ESSAYS ON THE WEALTHIEST AMERICANS

By

Kevin W. Capehart

Submitted to the

Faculty of the College of Arts and Sciences

of American University

in Partial Fulfillment of

the Requirements for the Degree

of Doctor of Philosophy

In

Economics

Chair: ______________________
Professor Alan G. Isaac

____
Professor Robert I. Lerman

____
Professor John A. Willoughby

____
Dean of the College

July 22, 2014
____
Date

2014
American University
Washington, D.C. 20016

Of all the classes, the rich are the most noticed and the least studied. So it was, and so it largely remains.

—John Kenneth Galbraith (1977, p. 44)

ESSAYS ON THE WEALTHIEST AMERICANS

by

Kevin W. Capehart

ABSTRACT

In every year since 1982, the popular magazine *Forbes Magazine* has published a list of the 400 wealthiest Americans. That list has attracted attention from the press and public, but it has been largely ignored by economists, at least in their professional capacities. Although a list published by a popular magazine may seem like a dubious source of data, the magazine's list is arguably the best source of data on the very top of the wealth distribution in the United States. This dissertation is a series of essays that use the magazine's list to study the wealthiest Americans. The essays study inequality between the wealthiest Americans and everyone else, inequality among the wealthiest Americans themselves, and mobility among the wealthiest Americans over time. Taken together, the essays offer insight into some basic empirical facts about a much noticed but little studied group.

ESSAYS ON THE WEALTHIEST AMERICANS

by
Kevin W. Capehart

ABSTRACT

In every year since 1982, the popular magazine *Forbes Magazine* has published a list of the 400 wealthiest Americans. That list has attracted attention from the press and public, but it has been largely ignored by economists, at least in their professional capacities. Although a list published by a popular magazine may seem like a dubious source of data, the magazine's list is arguably the best source of data on the very top of the wealth distribution in the United States. This dissertation is a series of essays that use the magazine's list to study the wealthiest Americans. The essays study inequality between the wealthiest Americans and everyone else, inequality among the wealthiest Americans themselves, and mobility among the wealthiest Americans over time. Taken together, the essays offer insight into some basic empirical facts about a much noticed but little studied group.

TABLE OF CONTENTS

ABSTRACT ii

LIST OF TABLES vii

LIST OF FIGURES viii

CHAPTER

1. INTRODUCTION 1

2. ON THE ESTIMATION OF THEIR WEALTH 12

2.1 Introduction 12

2.2 Top Wealths from Death Taxes 13

2.3 Top Wealths from Rich Lists 20

2.4 Possible Explanations for the Differences 24

2.4.1 Timing, Ownership, and Valuation Issues 25

2.4.2 A Person's Wealth Might Be Misestimated 29

2.4.3 The Magazine Might Miss Some People 29

2.4.4 Mortality Rates Might Be Misestimated 30

2.4.5 The Dead Might Not Be Representative of the Living 32

2.5 What If Forbes Was Right? 34

2.6 Extensions to the Exercise 41

2.6.1 The Probability of Substantially Smaller Estimates 41

2.6.2 The Source of the Substantially Smaller Estimates 49

2.7 Conclusion 57

2.8 A Postscript on the 2000s 60
3. ON THE DISTRIBUTION OF THEIR WEALTH 63
3.1 Introduction 63
3.2 Replication of Previous Work 64
3.3 Issues Identified During Replication 69
3.3.1 An Issue with Their Dataset 69
3.3.2 Other Issues with Their Dataset 71
3.3.3 An Issue with Ignoring Some of the Wealthiest Americans . . 74
3.3.4 An Issue with the Estimator 76
3.4 Extensions of Previous Work 78
3.4.1 Variation over Time in the Pareto Index 78
3.4.2 Deviations from a Pareto Distribution 80
3.4.3 Other Distributions besides an Untruncated Pareto 90
3.4.4 Deviations from a Truncated Pareto Distribution, Too 94
3.5 Conclusion 96
4. ON THE DURATION OF THEIR WEALTH 98
4.1 Introduction 98
4.2 A Magazine as a Data Source 100
4.3 A Brief Look at the Duration of Wealth 105
4.4 Factors Associated with Appearing Again 116
4.4.1 Model 116
4.4.2 Covariates 120
4.4.3 Results and Discussion 123
4.4.4 Some Quick Extensions 134
4.5 Conclusion 143
5. CONCLUSION 145

APPENDIX

A. APPENDIX TO THE FIRST ESSAY 147
A.1 Variances of the Estate-multiplier Estimators 147
A.2 The Pareto Extrapolation Method 148
A.3 Estimating Mortality Rates 149
A.4 Top Wealths from Surveys 153
A.5 A Similar Exercise with a Longer View 161
B. APPENDIX TO THE SECOND ESSAY 165
B.1 The Definition of a Pareto Distribution 165
B.2 An Interpretation of a Pareto Index 166
B.3 A Relationship between the Pareto and Gini Indexes 169
B.4 Parameter Estimates for Different Distributions 172
B.4.1 Pareto Distribution 172
B.4.2 Truncated Pareto Distribution 174
B.4.3 Truncated Log-Normal Distribution 176
B.4.4 Truncated Gamma Distribution 177
C. APPENDIX TO THE THIRD ESSAY 179
C.1 Movements Onto, Off of, and Throughout the List 179
C.2 Expressions for the Simple Estimates 185
C.3 Tests of the IIA Assumption 186
C.4 Estimates for the Extensions 189
C.5 A Thought Experiment on Wealth Deaccumulation 194
D. APPENDIX ON THE DATA 198
D.1 Collecting and Processing the Data 198
D.2 Replicating Bernstein and Swan's (2007) Table 200

D.3 Replicating Kilachand's (2012) List 208
D.4 Imputing the Wealth of One of Them 209
D.5 Fixing the Ages of Some of Them 209
D.6 Ignoring Human Capital 211
D.7 Ignoring Some Valuation Issues 213
D.8 Sharing the Data 214
D.9 Some Summary Statistics 215
D.9.1 Their Absolute Wealth 215
D.9.2 Their Relative Wealth 219
D.9.3 The Pro-Wealthiness of Their Accumulation 228
E. APPENDIX ON THE CODE 238
E.1 Code for Creating a Dataset 238
E.2 Code for Loading the Dataset 250
E.3 Code for the First Essay 253
E.4 Code for the Second Essay 277
E.5 Code for the Third Essay 307
REFERENCES 345

LIST OF TABLES

Table Page

2.1. Effect of Changes between 1990 and 2000 on Misestimation 54

3.1. Pareto Indexes for Some of the Wealthiest Americans, 1988–2003 . . 68

3.2. Effect of Errors in Klass et al.'s (2006) Dataset on the Pareto Index for Some of the Wealthiest Americans, 1988–2003 73

3.3. Number of the 400 Wealthiest Individuals in America Ignored by Klass et al.'s (2006) Study, 1988–2003 75

3.4. Goodness-of-fit Tests Against a Pareto Distribution, 1982–2013 85

3.5. Likelihood-ratio Tests, 1982–2013 93

3.6. Same Tests Against a Truncated Pareto Distributions, 1982–2013 . . 95

4.1. Multinomial Logit Model of the Competing Risks to Appearing Again 124

A.1. Effect of Changes between 1918 and 2000 on Misestimation 164

C.1. Number of Dropouts, 1982–2013 182

C.2. Number of Entrants, 1982–2013 183

C.3. Number of Incumbents, 1982–2013 184

C.4. Multinomial Logit Model with Philanthropic Dummies 191

C.5. Multinomial Logit Model with Self-made Dummies 192

C.6. Multinomial Logit Model with FIRE Industry Dummies 193

D.1. Decomposition of the Change in the Wealth of the 400 Wealthiest Americans, 1982–2013 237

LIST OF FIGURES

Figure Page

2.1. Direct and Estate-multiplier Estimates of the Wealth of the 400 Wealthiest Americans, 1982–2000 20

2.2. Mortality Rates, 1990 and 2000 32

2.3. Assumed and Estimated Wealth of the 400 Wealthiest Americans, 1982–2000 39

2.4. Misestimation of the Assumed Wealth of the 400 Wealthiest Americans with Different Assumptions about the Valuation of Wealth, 1982–2000 41

2.5. Probability of Underestimating the Assumed Wealth of the 400 Wealthiest Americans by at least 76 Percent with Different Assumptions about the Valuation of Wealth, 1982–2000 47

2.6. Median Misestimation of the Assumed Wealth of the 400 Wealthiest Americans with Different Assumptions about the Valuation of Wealth, 1982–2000 47

2.7. Probability of Underestimating the Assumed Wealth of the 400 Wealthiest Americans by at least 76 Percent with Different Assumptions about the Valuation of Wealth and the Rates at Which People Were Sampled, 1982–2000 48

2.8. Median Misestimation of the Assumed Wealth of the 400 Wealthiest Americans with Different Valuation and Sampling Assumptions, 1982–2000 48

2.9. Effect of Inequality on the Misestimation of the Assumed Wealth of the 400 Wealthiest Americans in 2000 55

2.10. Frequency Distribution of Estimates of the Assumed Wealth of the 400 Wealthiest Americans in 2000 Under Extreme Inequality 56

2.11. Assumed and Estimated Wealth of the 400 Wealthiest Americans, 1982–2010 61

3.1. Distribution of Wealth Among the 400 Wealthiest Americans in 2003 66

3.2. Pareto Index for the 400 Wealthiest Americans, 1982–2013 79

3.3. Ratios Between Wealth Estimates from *Bloomberg*'s and *Forbes Magazine*'s Lists, 2012 and 2013 88

4.1. Number of People who Appeared Again or Dropped Off Per Year, 1982–2013 107

4.2. Probability of Appearing Again After a Given Number of Years . . . 111

4.3. Probability of Appearing for at least a Given Number of Years 113

4.4. Effect of Age 127

4.5. Effect of Rank 128

4.6. Effect of the Calendar Year 129

4.7. Effect of Duration 130

4.8. Effect of Being Self-made 140

4.9. Effect of the FIRE Industry 142

A.1. Household Survey and Estate-multiplier Estimates of the Wealth of the Wealthiest One Percent of Households or People, 1989–2001 . . . 160

D.1. Wealth of the 400 Wealthiest Americans, 1982–2013 216

D.2. Percentage Change in the Wealth of the 400 Wealthiest Americans Over Select Years 218

D.3. Share of Wealth Held by the 400 Wealthiest Americans, 1982–2013 . 220

D.4. Shares of Wealth Held by the Wealthiest One Percent of American Households and the 400 Wealthiest Americans, 1989–2010 225

D.5. Shares of Wealth Held by the Least-wealthy Half of American Households and the 400 Wealthiest Americans, 1989–2010 226

D.6. Anonymous Changes in Wealth, 1982–2013 229

D.7. Non-anonymous Changes in Wealth, 1982–2013 230

CHAPTER 1

INTRODUCTION

The study of wealth is in the domain of economics. Indeed, economics has sometimes been called "the science of wealth" (Clark 2002, p. 415).[1] At least one reason why economists study wealth is because it can be a source of well-being (Davies 2009, p. 127). A person's wealth—or, equivalently, their net worth or their assets net of their debts—can be a source of well-being because it can be a source of consumption in either the present or future. Wealth may also improve a person's well-being in other ways. It may confer other powers besides its purchasing power over present or future goods and services, for example (Davies and Shorrocks 2000, pp. 606–7; Davies 2009, pp. 127–8).

This dissertation is a study of wealth. The dissertation is a series of essays that study a particular part of the personal distribution of wealth in a particular place over a particular period of time. The particular part of the wealth distribution is the very top, the particular place is the United States, and the particular period of time is recent decades. Each essay is self-contained, but this introductory chapter summarizes the essays and their relation to one another. This chapter also justifies the scope of the essays. The chapter starts with a justification for the scope of the essays and ends with a summary of the essays.

[1] Economics has been called other things besides the science of wealth, of course. For a discussion of alternative definitions of economics, including wealth-based definitions, see Backhouse and Medema (2009).

Part of the reason why the essays in this dissertation study the subject they study is that data on that subject has not been fully exploited by previous studies. Yet, if that somewhat practical reason is set aside for the moment, there are more principled reasons to study the subject studied by the essays in this dissertation. Those reasons can be organized into reasons to study the very top of the wealth distribution in the United States over recent decades, rather than *another part* of the wealth distribution in the same place over the same period of time, the same part of the wealth distribution in *another place* over the same period of time, or the same part of the wealth distribution in the same place over *another period of time*. Each of those sets of reasons can be taken up in turn.

Assuming for the moment that it seems reasonable to study the United States over recent decades, there are reasons to study *the very top* of the wealth distribution in that place over that period of time. The exact extent to which the distribution of wealth in the United States has been unequal at different points in time over recent decades can only be estimated, but, by all estimates, the distribution of wealth has been extremely unequal at every point in time. Wealth inequality has been more extreme than income or consumption inequality, for example (Davies and Shorrocks 2000, p. 607; Davies 2009, p. 127). Such extreme inequality implies that a small proportion of people own a disproportionately large share of wealth. The large share of wealth held by people at the very top of the wealth distribution is a reason to study that part of the distribution, even if the proportion of people at the very top might seem too small to be worthy of study.

Extreme inequality in the distribution of wealth also implies that, if an empirical study is not designed to try to capture the small proportion of people who own a large share of wealth, then the study will almost surely fail to capture the very top of the wealth distribution (Davies and Shorrocks 2000, pp. 629–30). The intentional or inadvertent failure of some empirical studies to capture the very top of the wealth distribution is another reason to study that part of the distribution, in particular, at least as a supplement to the study of other parts.

A final reason to study the very top of the wealth distribution relates to the failure of some theoretical models to predict or explain the large share of wealth held by the small number of people at the very top of the wealth distribution. The life-cycle model fails in that regard, for example, at least in its simplest form (Cagetti and De Nardi 2008). Dissertations like this one have been devoted to modifying such models so that they might generate levels of wealth inequality that match certain summary statistics more closely (Francis 2007), but actually studying the empirical realities of the part of the wealth distribution that some models fail to predict or explain could suggest which models and modifications thereof have a basis in reality and which are baseless.

A reason to study the very top of the wealth distribution in *the United States* over recent decades, rather than the same part of the wealth distribution in another country over the same period of time, is that many of the wealthiest people in the world have been Americans over recent decades. According to the inaugural year of *Forbes Magazine*'s list of the world's billionaires (which started as an extension to the magazine's list of the 400 wealthiest Americans), the United States was the country with the most billionaires in 1987 (Seneker 1987). By the most-recent year of the list, the United States was still the country with the most billionaires (Kroll 2014). In both years, Americans accounted for about one third of the world's billionaires (40 out of 140 and 492 out of 1,645, respectively). The wealth of those at the very top of the wealth distribution in the United States has therefore been relatively large, not only relative to other people in the same country, but also relative to everyone in the world. Again, the large share of wealth held by people at the very top of the wealth distribution in the United States over recent decades is a reason to study that part of the wealth distribution in that place over that period of time.

It should perhaps be admitted that the wealthiest people in *the entire world*, rather than in any one particular country, would also seem to be worthy of study for similar reasons. By studying the wealthiest Americans over recent decades, this dissertation studies some of the wealthiest people in the world over the same period of time, although obviously not a random sample of them. Studying a random or complete sample of the wealthiest people in the world is outside the scope of the essays in this dissertation. Of note, however, *Forbes Magazine*'s list of the world's billionaires has been used to study aspects related to the wealthiest people in the world, including the size or distribution of their wealth (Atkinson 2008; Davies et al. 2008; Ogwang 2013; Piketty 2014), their prevalence in particular countries at particular points in time, as well as the possible causes or consequences of their prevalence (Neumayer 2004; Sanandaji and Leeson 2013; Torgler and Piatti 2013), and their movements between countries over time (Sanandaji 2012).

Reasons to study the very top of the wealth distribution in the United States *over recent decades*, rather than the same part of the wealth distribution in the same place over another period of time, are as follows. Recent decades have seen a series of booms and busts in the markets for stocks, homes, and other assets. Those at the very top of the wealth distribution own a disproportionately large share of assets almost by definition, so they should be disproportionately affected by asset-market booms and busts, although some assets like stocks are more unequally distributed than other assets like homes (Wolff 2012, sec. 7.2), so maybe not. The extent to which different groups have been affected by recent asset-market booms and busts, especially by the most-recent recent bust in the housing market and its broader effects, is an active area of research (Kennickell 2011; Wolff 2012).

Another active area of research is on longer-term trends in inequality in the United States and the possible causes or consequences of any trends. While it is relatively well-established that those at the very top of the income distribution took home a larger share of income over recent decades (Piketty and Saez 2003 and the updates to that article), whether those at the very top of the wealth distribution took home a larger share of wealth over recent decades is less established (Kennickell 2003; Kopczuk and Saez 2004a). Despite the need for further research on whether there were actually any trends in inequality with respect to the distributions of income, wealth, or other sources of economic well-being, researchers have not been reticent about suggesting possible explanations for any such trends (Kaplan and Rauh 2013; Piketty 2014) or suggesting that increased inequality may have contributed to recent crises (Krugman 2010; Stiglitz 2012).

There are a number of reasons to study the very top of the wealth distribution in the United States over recent decades, therefore. The particular subject studied by the essays in this dissertation is also at least partly driven by the fact that data on that subject is available and, moreover, that the available data has not been fully exploited by previous studies. The primary source of data used by each essay is a list of the 400 wealthiest Americans that has been published annually since 1982 by *Forbes Magazine*. Lists of named wealth-holders like the magazine's are one source of data on wealth, as Davies and Shorrocks (2000, pp. 628–43) point out in their review of the theory and empirics of wealth and its distribution. Such lists have been largely neglected by academic researchers, however, as Davies and Shorrocks (2000) also point out in their review.

The neglect of such lists of named wealth-holders by academic researchers is understandable to some extent, as discussed below, but the neglect is nevertheless difficult to understand. The neglect of *Forbes Magazine*'s list of the 400 wealthiest

Americans is especially puzzling because it is arguably the best source of data on the very top of the wealth distribution in the United States over recent decades. Other sources of data like surveys or estate-tax records offer only incomplete accounts of that part of the wealth distribution in that place over that period of time. Consider surveys. Surveys that are not designed to try to over-sample the very top of the wealth distribution often fail to capture that part of the wealth distribution (Davies and Shorrocks 2000, pp. 629–30). The wealth survey conducted as part of the Panel Survey of Income Dynamics (PSID) is an example of a survey that often fails in that regard (Juster et al. 1999). Even if such surveys did try to over-sample the very top of the wealth distribution, the surveys could still be inaccurate accounts due to the non-response and misreporting problems that plague surveys of the wealthy (Davies 2009, pp. 129–30; Davies and Shorrocks 2000, pp. 630–1). *Forbes Magazine* may miss some of the 400 wealthiest Americans or misestimate what some of them are worth, of course, but there is no guarantee that the wealthy would response accurately or respond at all to a survey.

The Federal Reserve's Survey of Consumer Finances (SCF) is generally seen as the best survey on wealth in the United States at least partly because it is designed to try to over-sample the very top of the wealth distribution (Davies and Shorrocks 2000, p. 632). The SCF is still an incomplete account of the very top of the wealth distribution, however, given that the survey is explicitly designed to exclude the people (or, to be more precise, the households of the people) who appear on *Forbes Magazine*'s list of the 400 wealthiest Americans (Kennickell 2006, p. 84). Some of the reasons that have been given for excluding the (households of the) people who appear on the magazine's list are that it would be too difficult to get them to respond and, even if they did respond, it would be too difficult to protect their confidentiality (Kennickell 2007, p. 2). Yet, even if the SCF was not designed to exclude part of the very top of the wealth distribution, the survey might still be an inaccurate account due to non-response or misreporting errors.[2]

Another source of data on the very top of the wealth distribution is estate-tax records from the International Revenue Service (IRS). Estate-tax records provide data on some of the wealthiest Americans, but only those who die and pay the estate or so-called "death" tax. The wealth of the wealthiest living Americans must be inferred by using the dead as a sample for the living. Such inferences have potential problems, even if tax avoidance and evasion are assumed away (Atkinson 2008, sec. 2.3; Davies and Shorrocks 2000, sec. 3.3). *Forbes Magazine*'s list, in contrast, is reserved for Americans who are still alive.

The magazine's list of the 400 wealthiest Americans is therefore arguably the best source of data on the very top of the wealth distribution, at least in a cross-sectional sense. The list is also arguably the best source of data for studying mobility throughout that part of the wealth distribution over time. People only die once, so estate-tax records cannot be used to study wealth mobility, except at an inter-generational level. The SCF is typically only a cross-sectional survey conducted every three years, so it cannot be used to study wealth mobility, either. The 2009 SCF was atypical because it was conducted only two years after the previous survey and it tried to re-survey the same households from before. Again, however, the SCF is typically only a triennial, cross-sectional survey. The wealth survey conducted as part of the PSID is a panel survey, so it can be used to study mobility throughout some parts of the wealth distribution (Diaz-Gimenez et al. 2011, pp. 27–28), but that survey cannot be used to study mobility throughout the very top of the distribution, given that it often fails to capture that part of the distribution.

---

[2] Due to its design, the SCF can adjust for non-response errors (Kennickell 2007, p. 7), but there is no guarantee that those adjustments eliminate non-response errors.

While *Forbes Magazine*'s list is arguably the best source of cross-sectional and panel data on the very top of the wealth distribution, it is not an ideal source of data. The most notable and serious limitation of the list is that *Forbes Magazine* only provides limited details about its methods for identifying the 400 wealthiest Americans and estimating what each of them is worth. The details provided by the magazine are discussed in this dissertation (see, especially, sec. 2.3), but the details are limited. A closely related limitation is that, except in rare instances like Fitch's (2006) breakdown of Donald Trump's assets, the magazine does not provide details about the assets owned and debts owed by the people on its list. The failure to provide such details severely limits the ability of researchers to assess the magazine's methods and the accuracy of its estimates.

That said, no source of data on wealth is ideal, and other sources of data have similar limitations. In terms of the SCF, that survey tries to over-sample the very top of the wealth distribution, but the Federal Reserve only provides limited details about how it does so. Income-tax records are apparently used to try to identify wealthy households, but the exact manner in which those records are used is not revealed because of concerns over inadvertently violating the confidentiality of income-tax records (Kennickell 1999, 2001). In terms of estate-tax records, such records are also confidential, so access to them is severely limited. Analyzing such records requires befriending someone at the IRS and having them analyze the records for you (Kopczuk and Saez 2004a, p. 484; Kopczuk and Saez 2004b, p. 47).[3]

Other limitations of *Forbes Magazine*'s list are notable but less concerning. An obvious limitation is that, even if researchers accepted that the magazine's list was an

---

[3]The author's attempts to befriend someone at *Forbes Magazine* who can provide access to any additional details behind its wealth estimates have been unsuccessful, but gaining access to such data could open up new lines of research. Without trying to disparage the work of Emmanuel Saez and his colleagues, it can be noted that one of the great innovations of their work has simply been gaining access to confidential income- and estate-tax records.

accurate account of the 400 wealthiest Americans, it would still only be an account of a small number of people. The notion that a group of people can be too small to be worthy of study seems ultimately indefensible, but, even if the 400 wealthiest Americans were seen as insignificant in human terms, they would still be significant in economic terms. Again, extreme inequality in the distribution of wealth implies that small proportions of people own large shares of wealth. Information on a larger group of the wealthiest Americans could provide more insights, of course, but that does not imply that any insights offered by the magazine's list should be ignored.

Such limitations may explain why academic researchers have largely neglected *Forbes Magazine*'s list of the 400 wealthiest Americans as a source of data, but it should be emphasized that the list has not been entirely neglected. Previous studies that have used the magazine's list to study the 400 wealthiest Americans include studies on the size of their wealth (Mishel et al. 2012, pp. 382–4), the distribution of their wealth (Brzezinski 2012; Clauset et al. 2009; Klass et al. 2006; Levy and Solomon 1997), their mobility throughout the distribution of wealth over time (Castaldi and Milakovic 2007; Choi 2002; Hertz 2008; Keister 2005; Kennickell 2006), the sources of their wealth (Blitz and Siegfried 1992; Broom and Shay 2000; Canterbery and Nosari 1985; Foster and Holleman 2010; Kaplan and Rauh 2013; Petras and Davenport 1990; Potts 2006), which political parties they tend to contribute to (Republicans, according to Burris 2000), whether they are celebrities (only a few of them are, according to Cagetti and De Nardi 2006, p. 841), whether they are happier than their less-wealthy counterparts (they are, according to Diener et al. 1985), and whether they tend to have more children than their less-wealthy counterparts (they do, according to Essock-Vitale 1984, which should make Gregory Clark optimistic). The ability to add to that literature is yet another reason to study the particular subject studied by the essays in this dissertation.

Each of the essays can be summarized as follows. The first essay is on the distribution of wealth between the wealthiest Americans and everyone else. The essay compares what different data sources, including *Forbes Magazine*'s list of the 400 wealthiest Americans, suggest about recent trends in wealth inequality. Different data sources suggest different trends over certain periods. The essay attempts to reconcile the conflicting trends by considering potential problems with the different data sources. The essay argues that, while it may be unclear whether the magazine's list is an accurate account, it is clear that other sources of data would fail to capture the 400 wealthiest Americans, if the magazine's list was an accurate account.

The second essay is on the distribution of wealth among the wealthiest Americans themselves. The essay replicates and extends previous studies that used *Forbes Magazine*'s list to claim that the wealth of the 400 wealthiest Americans follows a particular distribution called a Pareto distribution. Although it would be captivating if their wealth followed that distribution, the essay argues that such a claim is not supported by the magazine's list. The magazine's list suggests that the distribution of wealth among the 400 wealthiest Americans deviates from a Pareto distribution to a statistically and substantively significant extent at almost every point in time.

The third and final essay is on mobility among the wealthiest Americans over time. The essay studies the amount of time that people who appear on the magazine's list remain there. The essay finds that most people remain for a relatively short amount of time, while some people remain for a relatively long amount of time, which can perhaps reconciling conflicting opinions about the persistence of wealth. The essay also finds that a person is less likely to remain on the list, if he or she is older and therefore closer to death or poorer and therefore closer to poverty. The recent crisis appears to have had little if any effect on mobility, however, even for those in the finance, insurance, or real estate industries.

Taken together, the essays offer insight into some basic empirical facts about the very top of the wealth distribution in the United States over recent decades. While the same source of data could surely be used to discover other empirical facts about the same part of the wealth distribution in the same place over the same period of time, this dissertation leaves the discovery of such facts as a direction for future research. Indeed, if nothing else, this dissertation has hopefully laid the foundation for such research by collecting data that was spread across several decades of a magazine, assembling that disparate data in a manner that is amenable to analysis, and sharing the data for others to study.

CHAPTER 2

ON THE ESTIMATION OF THEIR WEALTH

2.1 Introduction

For almost every year in the twentieth century, the economists Wojciech Kopczuk and Emmanuel Saez have estimated the shares of wealth held by some of the wealthiest Americans. They have estimated the shares of wealth held by groups ranging from the wealthiest two percent of Americans to the wealthiest 0.01 percent of Americans (Kopczuk and Saez 2004a,b). Kopczuk and Saez estimated the shares of wealth held by those different groups by applying the estate-multiplier method to estate-tax records, as discussed below. According to their estimates, the wealth shares of the different groups followed qualitatively similar trends over the twentieth century. The wealth shares were relatively high before the Great Depression of the 1930s, fell after the Depression, and then remained relatively stable for the rest of the century. The wealth shares were "remarkably stable" in the 1990s, in particular (Kopczuk and Saez 2004a, p. 453).

One reason why that stability in the 1990s was remarkable was because other estimates suggest that the share of wealth held by the 400 wealthiest Americans increased dramatically during that decade (Kopczuk and Saez 2004a, pp. 479–83). Those estimates were not made by applying the estate-multiplier method to estate-tax records, but rather by directly estimating the wealth of the 400 wealthiest Americans from whatever information was available, as discussed below. The estimates were also not made by economists as part of their academic work, but rather by the popular magazine *Forbes Magazine* as part of an annual list. Although one might presume that the direct estimates made by the magazine are erroneous, Kopczuk and Saez recognized that their estate-multiplier estimates might have failed to capture a dramatic increase in the share of wealth held by those at the very top of the wealth distribution during the 1990s (Kopczuk and Saez 2004a, pp. 480, 482).

This essay begins by comparing direct estimates of the wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list, on the one hand, to estate-multiplier estimates of their wealth based on the Kopczuk-Saez estimates, on the other hand. The essay shows that the estate-multiplier estimates were smaller than the direct estimates, especially by the end of the 1990s. Possible explanations for why the estate-multiplier estimates might have been smaller are then discussed. Next, the essay performs a simple exercise. The exercise involves applying the estate-multiplier method, not to estate-tax records, but to the magazine's list by only using information about people who died shortly after appearing on the list. The exercise suggests that it is possible and arguably probable that the Kopczuk-Saez estimates failed to capture a dramatic increase in the share of wealth held by the 400 wealthiest Americans during the 1990s.

2.2 Top Wealths from Death Taxes

As part of levying the estate tax, a person's wealth is estimated and recorded, so estate-tax records are a source of data on personal wealth (Davies and Shorrocks 2000, p 628). Estate-tax records are only a source of data on the wealth of people who died, however, given that the estate tax or so-called "death tax" is a tax levied on a person's wealth when he or she dies. Moreover, at least in the Untied States, estate-tax records are only a source of data on the wealth of people who died while relatively wealthy. The rules and regulations related to the estate tax are always changing, but only about one or two percent of the Americans who die in a given year are wealthy enough to be subject to the tax (Graetz 2011).

Although estate-tax records are only a source of data on the wealth of people who died while wealthy, the wealth of people who were alive and wealthy can be estimated by applying the "estate-multiplier method" to those records. That method assumes that the people who died while wealthy were a random sample of the people who were alive and wealthy. A wealthy person's probability of being sampled was his or her probability of dying. A person only dies once, so the sampling is assumed to be without replacement (Kopczuk and Saez 2004a and references therein).[1]

Given those assumptions, the total wealth of the people who were alive and wealthy can be estimated by multiplying the wealth of each person who died while wealthy by a multiple—the inverse of his or her mortality rate—and summing over those wealths. Formally, the estate-multiplier estimate of the total wealth of people who were alive and wealthy is $\sum_{i \in D} w_i/m_i$, where i indexes wealthy people, D denotes the set of wealthy people who died, w_i is the wealth of the i-th wealthy person, and m_i is his or her mortality rate.

Similarly, the total number of people who were alive and wealthy can be estimated by multiplying each wealthy person who died by a multiple—again, the inverse of his or her mortality rate—and summing over those people. The estate-multiplier estimate of the total number of people who were alive and wealthy is $\sum_{i \in D} 1/m_i$, where the variables have the same meanings as before.[2]

In principal, when the estate-multiplier method is applied to someone's estate-tax record, that method should only be used to estimate the wealth of people who

[1]On the history of the estate-multiplier method, see Atkinson and Harrison (1978, pp. 7–11).

[2]The variances of these estimators are given in section A.1 of this essay's appendix.

were alive and wealthy at the point in time just before the person died. In practice, however, too few people die at any given point in time to make it tenable to look at given points in time or even relatively short periods of time. Estate-tax records over longer periods of time like a year are therefore often used. Estate-tax records from over the course of a year were used by Kopczuk and Saez (2004a,b), in particular.

Ignoring any biases that might arise from using estate-tax records from a period of time that is longer than a point in time, the estate-multiplier estimates of the total number and total wealth of people who were alive and wealthy are unbiased, if each wealthy person dies with a non-zero probability, and if mortality rates are known with certainty (Horvitz and Thompson 1952).

Mortality rates are not known with certainty, so estimates must be used. Kopczuk and Saez used nationally representative age-, gender-, and year-specific mortality rates (Kopczuk and Saez 2004a, p. 448–9; Kopczuk and Saez 2004b, p. 37). Those estimates were made by the Social Security Administration, and the estimates are available from Wilmoth (1997).

A wealthy person is perhaps less likely to die in any given year than an average American, even after controlling for age and gender. As such, Kopczuk and Saez deflated the nationally representative mortality rates by age- and gender- (but not year-) specific social-differential factors (Kopczuk and Saez 2004a, p. 471). Social-differential factors for wealthy people were "not available," according to Kopczuk and Saez, so they used factors for a group with high socioeconomic status, specifically, white college graduates (Kopczuk and Saez 2004a, p. 449; Kopczuk and Saez 2004b, p. 39). Those factors are available from Brown et al. (2002).

Kopczuk and Saez assumed that the social-differential factors were constant over the twentieth century. They made that assumption because evidence for year-specific factors was "very sketchy," and also because the sketchy evidence did not suggest that there was any discernible trend over time in the social-differential factors for groups with high socioeconomic status (Kopczuk and Saez 2004b, p. 39).

Using those social-differential factors and nationally representative mortality rates, Kopczuk and Saez applied the estate-multiplier method to estate-tax records from various year between 1916 and 2000. The estate-tax records that they used were not available to the public, and those records are still not available to the public. The records were not even available to Kopczuk and Saez. Instead, Kopczuk and Saez submitted their desired calculations to an employee at the International Revenue Service (IRS) named Barry Johnson who made the calculations and reported his results (Kopczuk and Saez 2004a, p. 484; Kopczuk and Saez 2004b, p. 47).

Applying the estate-multiplier method to estate-tax records yields an estimate of the total number and total wealth of the people who were alive and wealthy, yet Kopczuk and Saez wanted to estimate the wealth of groups ranging from the wealthiest two percent of Americans to the wealthiest 0.01 percent of Americans. For years in which the estate-multiplier estimate of the total number of wealthy people was larger than the largest group for whom Kopczuk and Saez wanted a wealth estimate, the wealth of each group could simply be interpolated. Instead of summing over all of the people who were estimated to be alive and wealthy, Kopczuk and Saez (and Johnson) only needed to sum over some of them.[3]

For some years, the estate-multiplier estimates of the total number of wealthy people was smaller than largest group for whom Kopczuk and Saez wanted a wealth estimate. In order to estimate the wealth of a larger group, Kopczuk and Saez extrapolated the wealth of a smaller group by assuming that wealth followed a Pareto distribution (Kopczuk and Saez 2004a, p. 452; Kopczuk and Saez 2004b, p. 52). That extrapolation method is relatively straightforward, but a discussion of the method

---

[3] An example of this interpolation method is given in a subsequent section.

is relegated to the appendix (see section A.2). Kopczuk and Saez claimed that the extrapolation method could be "checked," and they did not indicate that the method failed any of their checks (see Kopczuk and Saez 2004b, p. 52).

After estimating the wealth of a given group, the wealth of the given group as a share of the total wealth of all Americans can be estimated. The total wealth of all Americans cannot be (or, at least, should not be) estimated from estate-tax records. Only a small percentage of the Americans who die in a given year are wealthy enough to be subject to the estate tax, as noted above, and estate-multiplier estimates suggest that only a similarly small percentage of the population would be wealthy enough to be subject to the tax if they died (Kopczuk and Saez 2004a, p. 450). The wealth of all Americans could be extrapolated, perhaps, but it is unclear what distributional form, if any, might be an accurate approximation across the entire distribution of wealth. To the extent that wealth is thought to follow a Pareto distribution, it is only thought to follow that particular distribution at the very top of the wealth distribution (Klass et al. 2006).

Instead of using estate-tax records, Kopczuk and Saez used another source of data to estimate the total wealth of all Americans. (They used a measure based on the total wealth of households and non-profit organizations, according to the Federal Reserve's Flow of Funds Accounts; Kopczuk and Saez 2004b, pp. 44–47.)

Note that, given an estimate of the wealth of some of the wealthiest Americans, estimating their share of wealth is a distinct issue that depends on estimating the total wealth of all Americans. As such, this essay focuses on estimates of *wealth*, rather than *shares* of wealth.

This essay also focuses on the wealth of the 400 wealthiest Americans, rather than a particular percentage of the population, in order to compare estate-multiplier and direct estimates of the wealth of that group. The 400 wealthiest Americans are a smaller group than the smallest group for whom Kopczuk and Saez estimated a wealth share. The smallest group for whom they estimated a wealth share was the wealthiest 0.01 percent of Americans. The number of people in that group increased over time as the population grew, but the number of people in that group was much larger than the 400 wealthiest Americans, even at the beginning of the century, and especially by the end of the century. In the year 2000, for example, the 400 wealthiest Americans were only about the wealthiest 0.0002 percent of Americans, if their share of the population is calculated by using the measure of population used by Kopczuk and Saez. (They used the number of people who were 20 years of age or older, according to the Census Bureau; Kopczuk and Saez 2004a, p. 448.)

Kopczuk and Saez could have estimated the wealth of the 400 wealthiest Americans, but they did not. It is somewhat strange that they did not, given that they made some attempt to compare their estate-multiplier estimates to *Forbes Magazine*'s direct estimates of the wealth of that group (see Kopczuk and Saez 2004a, pp. 479–83). Nevertheless, estate-multiplier estimates of the wealth of the 400 wealthiest Americans can be constructed as follows. Using the estimates that Kopczuk and Saez (2004a,b) made for a larger group and using the same extrapolation method that they used, the wealth of the 400 wealthiest Americans can be extrapolated from the wealth of the larger group. The wealth of the 400 wealthiest Americans can be extrapolated from the wealth of the smallest group for whom Kopczuk and Saez (2004a,b) estimated a wealth share, in particular.[4]

Part of figure 2.1 shows estate-multiplier estimates of the wealth of the 400 wealthiest Americans based on the Kopczuk-Saez estimates. Again, the wealth of

[4]If Kopczuk and Saez (2004a,b) had estimated the wealth of the 400 wealthiest Americans, then we could simply use those estimates without using their extrapolation method, but, again, they did not make such estimates. Estate-multiplier estimates could also be obtained by befriending an employee at the IRS and having him or her make the same calculations that Johnson made for Kopczuk and Saez, but that is beyond the scope of this essay.

the 400 wealthiest Americans was extrapolated from the wealth of the smallest group for whom Kopczuk and Saez (2004a,b) estimated a wealth share by using the same extrapolation method that they used. Estimates are shown for as recently as the year 2000, which is the most-recent year for which Kopczuk and Saez (2004a,b) estimated wealth shares. Estimates are only shown for as far back as the year 1982 for a reason that will become obvious. (The direct estimates shown as another part of the same figure are discussed in the next section.)

As shown in the figure: The estate-multiplier estimates suggest that, between the years 1990 and 2000, the wealth of the 400 wealthiest Americans increased from about 154 to 285 billion constant dollars, if current dollars are converted to constant dollars by using a version of the price index used by Kopczuk and Saez with the year 2000 as the base year. (The price index that they used was the Consumer Price Index for All Urban Consumers—the CPI-U—but since the 2008 update of Piketty and Saez 2003, Saez has used the Consumer Price Index Research Series Using Current Method—the CPI-U-RS—so this essay uses the latter index, although the results are quantitatively similar with either index.)

It can be noted that the estate-multiplier estimates suggest that the share of wealth held by the 400 wealthiest Americans increased from about 0.74 to 0.86 percent over the same period of time, if their wealth share is calculated by using the measure of total wealth used by Kopczuk and Saez.

The wealth of the 400 wealthiest Americans therefore increased by about 131 billion constant dollars and their share of wealth increased by about one tenth of a percentage point during the 1990s, at least according to estimates based on applying the estate-multiplier method to estate-tax records.

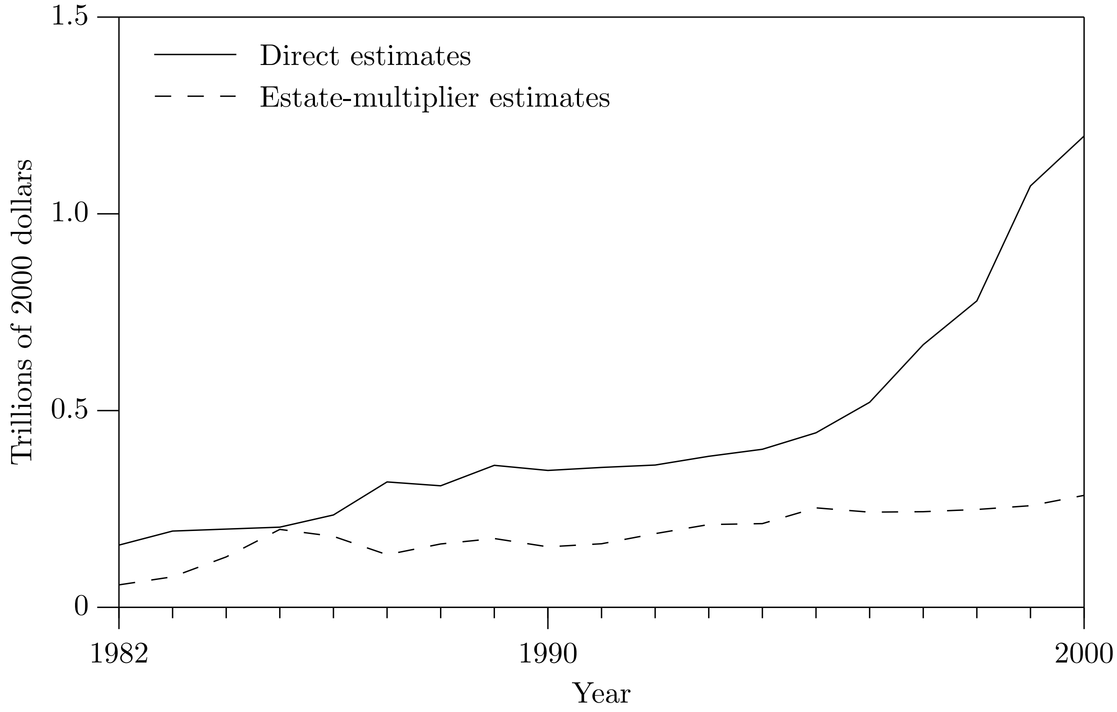

Figure 2.1. Direct and Estate-multiplier Estimates of the Wealth of the 400 Wealthiest Americans, 1982–2000

Source: Data adapted from *Forbes Magazine* (1982–2000); Kopczuk and Saez (2004a,b).

Note: This figure shows, for each year from 1982 to 2000, direct estimates of the wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list and estate-multiplier estimates of their wealth based on the Kopczuk-Saez estimates.

2.3 Top Wealths from Rich Lists

Estate-tax records are only one source of data on personal wealth. Lists of named wealth-holders or so-called "rich lists" are another source of data (Davies and Shorrocks 2000, p. 628). Perhaps the most-notable example of such a list is *Forbes Magazine*'s annual list of the 400 wealthiest Americans. Unlike estate-tax records, which are only supposed to provide information about people who died while wealthy enough to be subject to the estate tax, the magazine's list is supposed to provide information about people who were alive and wealthy enough to be one of the 400 wealthiest living Americans when the magazine made its list.

In order to make its list of the 400 wealthiest Americans, the magazine apparently uses whatever information is available. For example: As the magazine said on the 30th anniversary of its inaugural list:

> While we've been at it a long time, it is never an easy task [to make our list of the 400 wealthiest Americans]. Our reporters dig deep. [. . .] When possible we met with [people who we thought might be one of the 400 wealthiest Americans] in person [. . .] We also interviewed their employees, handlers, rivals, peers, and attorneys. We poured over hundreds of Security and Exchange Commission documents, court records, probate records, federal financial disclosures, and Web and print stories. (*Forbes Magazine* 2012, p. 226)

Other years of the magazine's list have offered similar descriptions of the varied sources that were used, although Web stories were obviously not a source of information before the advent of the World Wide Web.

When estimating a person's wealth, the magazine apparently takes into account all the assets and debts it can identify. "We took into account all assets: stakes in public and private companies, real estate, art, yachts, planes, ranches, vineyards, jewelry, car collections, and more," the magazine said on the 30th anniversary of its list, again, for example (*Forbes Magazine* 2012, p. 226). "We also factored in debt," the magazine said, while admitting that "we don't pretend to know what is listed on each [person's] private balance sheet" (ibid.).

The wealth of family members is also apparently taken into account when estimating a person's wealth. According to the inaugural year of the magazine's list, a person's wealth "generally" includes the wealth of his or her spouse, children, or other family members, "especially if family ties are manifestly close or they all share interests in an ongoing business" (*Forbes Magazine* 1982, p. 101). If "family or business ties" are "broken or at least notably frayed," then a person's wealth does not include the wealth of family members (ibid.). A person's wealth also excludes the wealth of a spouse, a child, or other family member, if the family member is wealthy enough to be one of the 400 wealthiest Americans themselves. In that case, the member is listed as a separate person on the magazine's list (ibid.). That way of assigning the ownership of wealth has not changed in any obvious way over time (cf. *Forbes Magazine* 1982, p. 101, and *Forbes Magazine* 2012, p. 226).

The magazine also offers some information about how it values assets. Publicly traded stocks are valued at the close of the stock market on a particular day of the year, while privately held companies are valued based on the value of similar public companies, for example (*Forbes Magazine* 2012, p. 226). The magazine's treatment of trusts is of particular interest to this essay. "Common sense," rather than trust law, is apparently applied to determine whether a trust should count towards someone's wealth (*Forbes Magazine* 1982, p. 101). "Most trusts are plainly set up to carry out a normal pattern of inheritance and exist to minimize inheritance taxes," the magazine said in one year. Trusts that are set up as such are "generally attributed to the person who created the wealth [...] or else the principal controlling family member" (ibid.). Some trusts like irrevocable charitable trusts are not attributed to the person who created the wealth, however (ibid.).

The magazine's estimates for the wealth of some of the wealthiest Americans are obviously made in a different way than estate-multiplier estimates. The estimates are even made in a different way than survey estimates. The magazine talks to people about how much they are worth, but the magazine also relies on other sources, and does not necessarily rely on what a person says he or she is worth. (The magazine does not rely on what Donald Trump says he is worth, for example; Fitch 2006.) The magazine's estimates can be said to be "direct" estimates (following the terminology of Davies and Shorrocks 2000, p. 642), rather than estate-multiplier, survey, or any other sort of estimates, given that the magazine tries to directly estimate a person's wealth from whatever information is available.

Along with estate-multiplier estimates based on the Kopczuk-Saez estimates, figure 2.1 also shows direct estimates of the wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list. For each year, except for a few years, the direct estimate of the wealth of the 400 wealthiest Americans is simply the sum of what the magazine estimated each person on its list was worth.

The exceptions are that, for the years 1982 to 1989, *Forbes Magazine* did not report an estimate of the wealth of one person on its list, Malcolm Stevenson Forbes, who was the editor-in-chief of the magazine. "People would have assumed that the printed figure [for my wealth] was real, not an estimate, as all the rest are," he explained (quoted in *Forbes Magazine* 1983, p. 168; also see *Forbes Magazine* 1982, p. 170). This essay imputed his wealth in a given year as the median wealth of the other 399 people on the list in the given year.

The direct estimates suggest that, between 1990 and 2000, the wealth of the 400 wealthiest Americans increased by almost a trillion constant dollars from about 360 to 1,198 billion constant dollars, if current dollars are converted to constant dollars by using the same price index that was used above.

It can be noted that the direct estimates suggest that, over the same period of time, the share of wealth held by the 400 wealthiest Americans increased by almost two percentage points from about 1.7 to 3.6 percent, if their share of wealth is calculated by using the same measure of the total wealth of all Americans that was used by Kopczuk and Saez.

The largest difference between the estate-multiplier and direct estimates occurs in the year 2000. In that year, *Forbes Magazine* estimated that the 400 wealthiest Americans were worth about 1,198 billion dollars in total, while the estate-multiplier estimates suggest that they were only worth about 285 billion dollars. The difference between those two estimates is about 913 billion dollars.

It is unclear whether that difference is statistically significant, given that neither *Forbes Magazine* nor Kopczuk and Saez (2004a,b) quantified the uncertainty associated with their estimates, but the difference is clearly substantively significant. The difference is almost a trillion dollars, which about 2.8 percent of the total wealth of all Americans, if the measure of total wealth used by Kopczuk and Saez is used. The difference is also equal to about 76 percent of the total wealth of the 400 wealthiest Americans, if *Forbes Magazine*'s direct estimates are to be believed. Or, if the estate-multiplier estimates are to be believed, then the difference is equal to over four times (about 321 percent of) the wealth of the 400 wealthiest Americans, about 71 percent of the wealth of the smallest group for whom Kopczuk and Saez estimated a wealth share (the wealthiest 0.01 percent of Americans), and about 10 percent of the wealth of the largest group for whom they estimated a wealth share (the wealthiest two percent of Americans).

2.4 Possible Explanations for the Differences

Possible explanations for why direct estimates of the total wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list might be substantially larger or simply different than estate-multiplier estimates of their wealth based on the Kopczuk-Saez estimates are as follows. A person's wealth as it is reported by the magazine and the person's wealth as it is reported for estate-tax purposes may be different because the magazine and estate taxes might not be measuring the same concept of wealth. Even if they were trying to measure the same concept, they might mismeasure it. Moreover, even if their estimates of a person's wealth were always the same and always equal to what a person was actually worth, direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans might still be different. The direct estimates could be wrong because the magazine might miss some people. The estate-multiplier estimates could also be wrong because mortality rates might be misestimted and, even if they were not, the dead might be too poor or too wealthy to be representative of the living. These reasons are expanded upon below. Although one might hope that other sources of data could suggest which estimates are more accurate, other sources like surveys are not helpful in that regard, but a discussion of that point is relegated to section A.4 of this essay's appendix.

2.4.1 Timing, Ownership, and Valuation Issues

The differences between the direct and estate-multiplier estimates may ultimately be due to differences between a person's wealth as it is reported by the magazine and the person's wealth as it is reported for estate-tax purposes. There are a number of issues that may generate the latter differences and that may, in turn, generate the former differences. First of all, there may be a timing issue (McCubbin 1994, p. 368). Whereas *Forbes Magazine*'s list is supposed to be snapshot of wealth at the close of the stock market on one day of the year, people die all year long. For estate-tax purposes, a person's wealth can be valued on either the day that he or she died or an alternative date, whichever date would result in a lower valuation. The alternative date for valuation has changed over time, but, in recent years, the alternative date has been six months after the day that a person died (Kopczuk and Saez 2004b, p. 43; Raub et al. 2010, p. 4). Thus, if a person died after the close of the stock market on the day that the magazine made its list, and if the person's wealth did not fall over the next six months, then his or her wealth would be valued at the same point in time. Otherwise, a person's wealth would be valued at different points in time. That timing issue may explain differences between the direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans, although it probably cannot explain large and ever-larger differences. There are other issues that may explain differences between the estimates, however.

Even if wealth was valued at the same point in time, there may be issues related to the way that the ownership of wealth is assigned (McCubbin 1994, p. 368). Whereas a person's wealth as it is reported by *Forbes Magazine* can include the wealth of the person's family members, a person's wealth for estate-tax purposes does not include the wealth of family members. Assets that were solely owned by a person's spouse or other family member are excluded from the person's estate, for example (Raub et al. 2010, p. 12).Also, whereas a trust can count towards a person's wealth as it is reported by the magazine, trusts do not count towards a person's wealth as it is reported for estate-tax purpose, which is the purpose of most trusts, at least according to *Forbes Magazine* (1982, p. 101). Those ownership issues may explain why the direct estimates of the wealth of the 400 wealthiest Americans are, not only different, but larger than the estate-multiplier estimates of their wealth.

Finally, even if wealth was valued at the same point in time and assigned in the same way, there may be issues related to the way that wealth is valued. Whereas the magazine apparently tries to estimate the market value of every asset, some assets do not need to be reported at their market values for estate-tax purposes (McCubbin 1994, p. 368). The assets that can be discounted and the amounts by which they can be discounted have changed over time (Kopczuk and Saez 2004b, p. 43; Raub et al. 2010, p. 13), but an example of a valuation discount is as follows. If a person owned a large amount of a certain type of asset like a company's stock, then the value of that type of asset can be discounted, given that selling off a large amount of the asset may drive its price down (Kopczuk and Saez 2004b, p. 42). Those valuation issues may also explain why the direct estimates of the wealth of the 400 wealthiest Americans are larger than the estate-multiplier estimates.

Interestingly, for people who appeared on *Forbes Magazine*'s list of the 400 wealthiest Americans while they were alive and then died while they were still wealthy enough to be subject to the estate tax, the extent to which a person's wealth as it was reported by the magazine was different than the person's wealth as it was reported for estate-tax purposes has actually been studied by a group of employees at the IRS (Raub et al. 2010). The group, which included the employee who helped Kopzuk and Saez construct their estate-multiplier estimates, were updating and extending an earlier study by another IRS employee (McCubbin 1994). For one part of their study, the group collected the estate-tax records of 181 people who died shortly after appearing on the magazine's list in any year between 1982 and 2008. Those people died in either the same year they were on the list or in the year after they were on the list (Raub et al. 2010, p. 9). Note that up to about a year and a half could have past between the time when the magazine estimated a person's wealth and the time when his or her wealth was estimated for estate-tax purposes. There may be a timing issue when comparing the wealth estimates, therefore, but the comparison still seems interesting.

Using those estate-tax returns, the group of IRS employees calculated the ratio between a person's wealth as it was reported for estate-tax purposes and the person's wealth as it was reported by the magazine. They called that ratio the "Forbes ratio" (Raub et al. 2010, p. 11). They found that, across the 181 people, the Forbes ratio was 50 percent on average (ibid.). So, on average, a person's wealth as it was reported for estate-tax purposes was only half of the person's wealth as it was reported by the magazine. For 28 of the 181 people, their estate-tax records said they were worth more than what the magazine said they were worth (ibid., pp. 11–12). However, for most of the 181 people, their estate-tax records said they were worth less. That could explain why the direct estimates of the wealth of the 400 wealthiest Americans were larger. The Forbes ratio also apparently fell slightly over time (ibid., p. 11), which may explain why the direct estimates were ever-larger.

The group of IRS employees found that the difference between what the magazine and the estate-tax records said someone was worth could be partially explained by an ownership issue. For people who were married, when a person's wealth for estate-tax purposes was recalculated to include the (other half of the) assets that the person owned jointly with his or her spouse, the Forbes ratio rose, albeit only slightly from 46 to 53 percent on average (Raub et al. 2010, p. 12). Other ownership issues may also explain the difference. The Forbes ratio might move closer to unity, if a person's wealth for estate-tax purposes was recalculated to include assets solely owned by his or her spouse or other family members. The group at the IRS did not have any data to study the effect of assigning wealth in that way, however. The only ownership issue that they could study was the effect of including assets jointly owned with a spouse.

The difference could also be partially explained by a valuation issue. Data on valuation discounts has apparently only been collected since 2004 (Raub et al. 2010, p. 13), so the effect of such discounts can only be studied for people who died since then. Yet, at least for people who died since then and claimed such discounts, when a person's wealth was recalculated without the discounts, the Forbes ratio rose slightly. The ratio rose from 47 to 54 percent on average. The ratio rose slightly more, to 58 percent on average, if a person's wealth was recalculated to also include assets that the person owned jointly with his or her spouse (ibid. pp. 12–13).

Thus, for people who died shortly after appearing on the magazine's list, the difference between their wealth as it was reported by the magazine and their wealth as it was reported for estate-tax purposes could be partially, but only partially, explained by certain ownership and valuation issues. The other ownership and valuation issues discussed above, as well as the timing issue discussed above, could explain the rest of the difference, perhaps.

2.4.2 A Person's Wealth Might Be Misestimated

Yet, even if the above-discussed issues did not exist, a person's wealth as it was reported by the magazine might still not be the same as the person's wealth as it was reported for estate-tax purposes. Audits suggest that estate taxes are evaded to at least some extent (Kopczuk and Saez 2004a, p. 470), so a person's wealth as it is reported for estate-tax purposes may underestimate whatever concept of wealth it is that the estate tax is trying to tax.

A person's wealth as it is reported by *Forbes Magazine* may also misestimate whatever concept of wealth the magazine is trying to capture, if some assets or debts are difficult for the magazine to either identify or value (Atkinson 2008, p. 69; Blitz and Siegfried 1992, p. 5). The magazine may overestimate a person's wealth, in particular, if assets are easier to identify than debts (Atkinson 2008, p. 70).

2.4.3 The Magazine Might Miss Some People

Even if a person's wealth as it was reported by the magazine was always the same as the person's wealth as it was reported for estate-tax purposes, and even if that amount was what the person was actually worth, direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans might still be different. The estimates might still be different because the magazine might miss some of the 400 wealthiest Americans. According to the group of IRS employees, between 1982 and 2008, there were 26 people whose estate-tax records suggested that they should have been on the magazine's list in the year that they died, yet they did not appear on the list in any year (Raub et al. 2010, pp. 6–7). The magazine might have missed some people in some years, therefore. Of course, the direct estimates of the wealth of the 400 wealthiest Americans would actually be too small, if the people the magazine missed were wealthier than the people it identified.

2.4.4 Mortality Rates Might Be Misestimated

There are other reasons why direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans might still be different, even if a person's wealth as it was reported by the magazine was always the same as the person's wealth as it was reported for estate-tax purposes, and even if that amount was what the person was actually worth.

Like other scholars who have applied the estate-multiplier method (e.g., Lampman 1962, p. 14), Kopzuk and Saez recognized that their estate-multiplier estimates could be wrong, if the mortality rates they used were wrong (Kopczuk and Saez 2004a, pp. 449, 471). Their estate-multiplier estimates could have been too small, in particular, if the mortality rates they used were too high.

Whether the Kopzuk-Saez mortality rates were too high, too low, or just right is unknown, given that the rates at which wealthy people died can only be estimated. The mortality rates used by Kopzuk and Saez were apparently the best estimates that were available. There is at least one other source of data that can be used to make such estimates, however. *Forbes Magazine*'s list can be used.

For each person who appears on the magazine's list in any given year, the magazine reports an estimate of his or her age. The magazine also reports the names of people who appeared on its list in the previous year but died since then. Using that information, we can estimate the probability that a person on the list in a given year at a given age would die by the time the magazine made its list in the next year.

A non-parametric approach to estimating those mortality rates is ultimately unappealing because of the relatively small number of people on the magazine's list in any given year and the even smaller number of people who died, so we estimated the mortality rates by assuming that people lived and died according to the Gompertz-Makeham law of mortality (Gompertz 1825; Makeham 1860). That assumption is the same assumption that Brown et al. (2002) made in order to construct the social-differential factors used by Kopczuk and Saez (2004a,b).[5]

We estimated year- and age- but not gender-specific mortality rates. We did not estimate gender-specific mortality rates for two reasons. One reason is that the number of women on the list has been small and shrinking since 1982. Only about one eighth of the people on the list were women by the year 2000, for example (Edlund and Kopczuk 2009, p. 164, table 4). Another reason we did not try to estimate gender-specific mortality rates is because, if we did try to do that, then we would not be able to estimate year-specific mortality rates for some years. In some years, only men died. To the extent that women have lower mortality rates than men of the same age, ignoring gender differences would bias our estimates downward, although any bias should shrink over time as the number of women on the list shrunk.

Estimated mortality rates for two years, 1990 and 2000, are shown in figure 2.2 of this essay. For comparison, the Kopzuk-Saez mortality rates for men in the same years are also shown in the figure. The Kopzuk-Saez mortality rates for women were lower, but, again, most of the people on the magazine's list in those years were men. As shown in the figure, in both of the years and for each age past about 50 years of age, the mortality rates estimated from the magazine's list are lower than the mortality rates used by Kopzuk and Saez for men. Most of the people on the magazine's list in those years were at least 50 years of age. (To be more specific, about 87 and 76 percent of them were over the hill in 1990 and 2000, respectively.) The Kopzuk-Saez mortality rates may have been too high, therefore, at least for most of the 400 wealthiest Americans. If so, then estate-multiplier estimates of the wealth of the 400 wealthiest Americans may have been too small.

---

[5]See section A.3 of the appendix for more details on the way in which we estimated the mortality rates and the way in which Brown et al. (2002) estimated the social-differential factors.

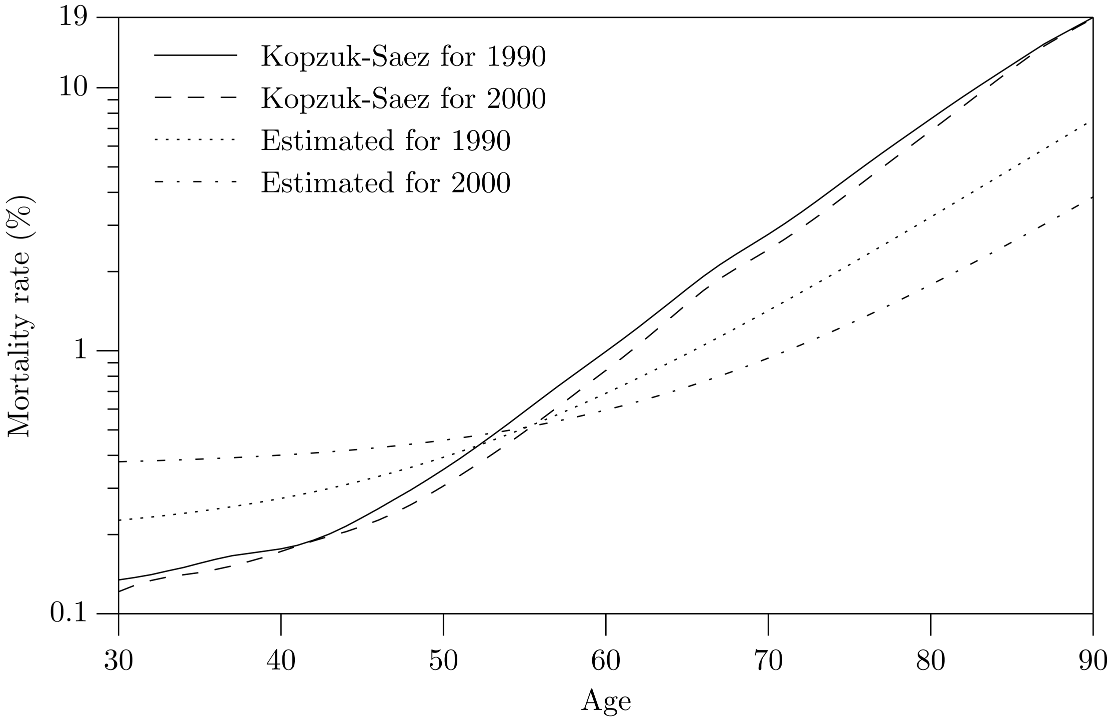

Figure 2.2. Mortality Rates, 1990 and 2000

Source: Data adapted from Brown et al. (2002); *Forbes Magazine* (1990, 1991, 2000, 2001); Wilmoth (1997).

Note: This figure shows, for the years 1990 and 2000, the mortality rates for males that were used by (Kopczuk and Saez 2004a,b). The figure also shows mortality rates that were estimated by using information about the people who appeared on *Forbes Magazine*'s list in a given year and then were either alive or dead by the next year.

2.4.5 The Dead Might Not Be Representative of the Living

There is at least one other reason why the estate-multiplier estimates may have been too small, even if estate-tax records always reported what a person was actually worth, and even if the mortality rates used by Kopzuk and Saez were actually the rates at which people died. The dead may have been too poor to be representative of the living (Atkinson 2008, p. 71). Scholars who have applied the estate-multiplier method have often recognized that people may deaccumulate wealth before dying for a variety of reasons, including bequest motives, tax avoidance, and end-of-life care (ibid.; Davies and Shorrocks 2000, p. 638; Kopczuk and Saez 2004a, pp. 471–2). Yet, even if people do not deaccumulate wealth before dying, the dead could still be too poor to representative of the living for the same reason that many of the living are too poor to be representative. The distribution of wealth is highly skewed, so most people are too poor to be representative (which is exactly why the best surveys on wealth try to over-sample the wealthy; Kennickell 2007).

There are all sorts of possible explanations for why the direct estimates of the wealth of the 400 wealthiest Americans might be different than the estate-multiplier estimates of their wealth, therefore. Not all of the reasons imply that one set of estimates was right and the other wrong. They may have both been right because they may have been measuring different things. Of course, both sets of estimates could have been wrong. Or, one set could have been wrong and the other right.

Kopczuk and Saez (2004a,b) did not dismiss the possibility that their estimates might have been wrong and the magazine's estimates might have been right. After pointing out that the increase in *Forbes Magazine*'s direct estimate of the wealth of the 400 wealthiest Americans in the late 1990s was largely due to an increase in the wealth of an even smaller number of people (the 100 wealthiest or so), they noted in a footnote that, "It seems possible that a few-year long surge of [the] wealth of a few individuals can remain unnoticed [by estate-multiplier estimates]" (Kopczuk and Saez 2004a, p. 480; Kopczuk and Saez 2004b, p. 31).

The possibility or impossibility, as well as probability or improbability, that the estate-multiplier estimates may have been right or wrong can be explored through a simple exercise implemented in the next section. The exercise involves assuming that *Forbes Magazine*'s list was an accurate account of the 400 wealthiest Americans and then trying to estimate what the 400 wealthiest Americans were assumed to be worth by only using information about people on the list who died.

2.5 What If Forbes Was Right?

In every year since 1982, *Forbes Magazine* has published a list of the 400 wealthiest Americans. If we look at the magazine's list in any given year, then we can see the name, estimated wealth, and estimated age of each person who was on the list in that year. And if we look at the list in any given year, except the inaugural year, then we can also see the name of each person who was on the list in the previous year but died. The names of the people who were on the list in the previous year but died are reported by the magazine, regardless of whether a person would have made it onto the list again if only he or she had lived. The magazine does not report any sort of estimate of what a person was worth at the point in time when he or she died, although we can look at the previous year's list in order to see what the magazine said he or she was worth at that point in time.

Whether *Forbes Magazine*'s list of the 400 wealthiest Americans in any given year was an accurate account of the 400 wealthiest Americans when the magazine made its list in that year is unknown. Yet, for the sake of an exercise to evaluate the estate-multiplier method, suppose that the magazine's list was an accurate account in every year. That is to say, suppose that the 400 people on the list were actually the 400 wealthiest Americans, suppose that they were actually worth what the magazine said they were worth, and suppose that they were actually as young or old as the magazine said they were.

For the sake of the same exercise, suppose we want to know the total wealth of the 400 wealthiest Americans in any given year at the point in time when the magazine made its list in that year. If we could see what the magazine said each person was worth, then the task would be trivial. We could simply sum up what the magazine said each person was worth, and we would arrive at what we assumed the 400 wealthiest Americans were worth in total.

Suppose, however, that we cannot see information about each person on the magazine's list. Suppose we can only see information about some of them. Suppose we can only see information about the people who were on the list in a given year but died by the time the magazine made its list in the next year. Estate-tax records only provide information about the people who died while wealthy enough to be subject to the estate tax. So, we are similarly assuming that we can only see information about people who died shortly after appearing on the magazine's list.

Suppose, moreover, that we do not know that the people for whom we see information must have been one of the 400 wealthiest Americans. Suppose we only know that, in order to see the name, wealth, and age of a person, he or she must have been worth at least a given amount of wealth. The given amount of wealth corresponds to the minimum wealth of the 400 wealthiest Americans, but we will assume that we only know the amount, not what it corresponds to. When working with estate-tax records, we do not know the number of people who would have been subject to the estate tax in the event of their death. We only know that a decedent must have been worth at least a given amount in order to be subject to the tax.

If that is all the information we have, then we do not know what the 400 wealthiest Americans were assumed to be worth in total, but we can estimate their total wealth by applying the estate-multiplier method to that information in the same way that Kopczuk and Saez (2004a,b) applied the method to estate-tax records. The estate-multiplier method as it was applied by Kopczuk and Saez can then be evaluated by dropping the pretense that we do not know what the 400 wealthiest Americans were assumed to be worth, and comparing the estate-multiplier estimate of their wealth to their assumed wealth.

This exercise cannot tell us whether *Forbes Magazine*'s list in any given year was an accurate account of the 400 wealthiest Americans when the magazine made its list in that year, of course. That said, the exercise can tell us whether the estate-multiplier method would have correctly estimated the total wealth of the 400 wealthiest Americans, if the magazine's list was an accurate account. Any differences cannot be attributed to any timing, ownership, or valuation issues, given that we are taking everyone's information directly from the magazine's list, although we will explore the effect of misvaluing wealth in just a moment.

The magazine's list from the year 2000 can used to illustrate this exercise. When its list for the next year came out, *Forbes Magazine* informed its readers that, out of the 400 people who were on its list in 2000, three people had died. The three people who died were William Redington Hewlett, Randolph Apperson Hearst, and Michael Chowdry. It can be noted that Hewlett was one of the co-founders of the Hewlett-Packard Company, Hearst was one of the sons of the newspaper mogul William Randolph Hearst, and Chowdry was the founder of a cargo airline company.

In order to estimate the total wealth of the 400 wealthiest Americans in 2000, we would only use information about these three people who died. According to the magazine's list from 2000, the ages, sexes, and wealths of the three people were as follows. Hewlett was an 87-year-old male who was worth nine billion dollars, Hearst was an 84-year-old male who was worth 1.8 billion dollars, and Chowdry was a 45-year-old male who was worth 920 million dollars (*Forbes Magazine* 2000). It can be noted that Hewlett and Hearst died of natural causes related to their advanced ages, while Chowdry died in a plane crash.

If we use the nationally representative mortality rates that Kopczuk and Saez used (again, Wilmoth 1997), as well as the social-differential factors that they used (again, Brown et al. 2002), then we would estimate that Hearst, Hewlett, and Chowdry should have died with about a 1-in-7, 1-in-9, and 1-in-464 chance, respectively. The estate-multiplier method therefore implies that there should have been about seven people like Hewlett, about nine people like Hearst, and about 464 people like Chowdry. That is to say, there should have been about seven 87-year-old males worth nine billion dollars, about nine 84-year-old males worth 1.8 billion dollars, and about 464 45-year-old males worth 920 million dollars. If that is all the information we have, then the estate-multiplier estimate of the total number and total wealth of the people who were alive and wealthier than the given amount of wealth would be about 481 people and 507 billion dollars, respectively.[6]

Four hundred and eighty one people is obviously greater than 400 people, so, if we knew that the given amount of wealth was the minimum wealth of the 400 wealthiest Americans, then we would know that something was amiss with our estate-multiplier estimates. Yet, again, when working with estate-tax records, we do not know how many people were wealthy enough to be subject to the estate tax in the event of their death. So, we are similarly assuming that we do not know that only 400 people should have been worth at least as much as the given amount.

Given that 481 people is greater than 400 people, the wealth of the 400 wealthiest people can be interpolated. We would estimate that the 400 wealthiest Americans were comprised of about seven people worth nine billion dollars, about nine people worth 1.8 billion dollars, and about 384 people worth 920 million dollars. Their total wealth would be estimated to be about 432 billion dollars.

That estate-multiplier estimate is wrong in the context of the exercise. The 400 wealthiest Americans were assumed to worth about 1,198 billion dollars in 2000. The estate-multiplier method therefore underestimates their wealth by about 765 billion dollars, which is over half (about 64 percent) of what the 400 wealthiest Americans were assumed to be worth in that year.

---

[6]Standard errors for those point estimates could be calculated by using the expression given in section A.1 of the appendix; but, again, the estate-multiplier method as it was applied by Kopczuk and Saez (2004a,b) did not account for any uncertainty associated with such estimates.

The same exercise can be repeated for the other years of *Forbes Magazine*'s list. It can be noted that the number of people who died shortly after appearing on the magazine's list in a given year varied by year. Between the years 1982 and 2000, as few as two and as many as 13 people died after appearing on the list. The percentage of the 400 wealthiest Americans who died varied from about one-half of one percent to 3.25 percent, therefore. Over the same period of time, the percentage of people who died while wealthy enough to be subject to the estate tax varied from about one to two percent, according to estimates made by Kopczuk and Saez (2004a, p. 450). So, the percentage of people who died while wealthy enough to be one of the 400 wealthiest Americans was similar to the percentage of people who died while wealthy enough to be subject to the estate-tax, if the magazine's list and Kopczuk and Saez's estimates are to be believed.

When repeating the exercise for other years, if the estate-multiplier estimate of the total number of people is greater than or equal to 400, like it is in the year 2000, then the total wealth of the 400 wealthiest Americans can be interpolated. In addition to the year 2000, for the years between 1982 and 2000, their wealth can be interpolated in the years 1986, 1989, 1992, and 1994. In the years in which the estate-multiplier estimate of the number of people is less than 400, the total wealth of the 400 wealthiest Americans cannot be interpolated, but it can be extrapolated by using the same extrapolation method used by Kopczuk and Saez, specifically, by assuming that wealth followed a Pareto distribution.

The results of this exercise for the years 1982 to 2000 are shown in figure 2.3 of this essay. As shown in the figure, in each year, the estate-multiplier method underestimates what the 400 wealthiest Americans were assumed to be worth. The smallest underestimation occurs in 1994. In that year, the estate-multiplier method underestimates their wealth by only about one billion constant dollars or less than

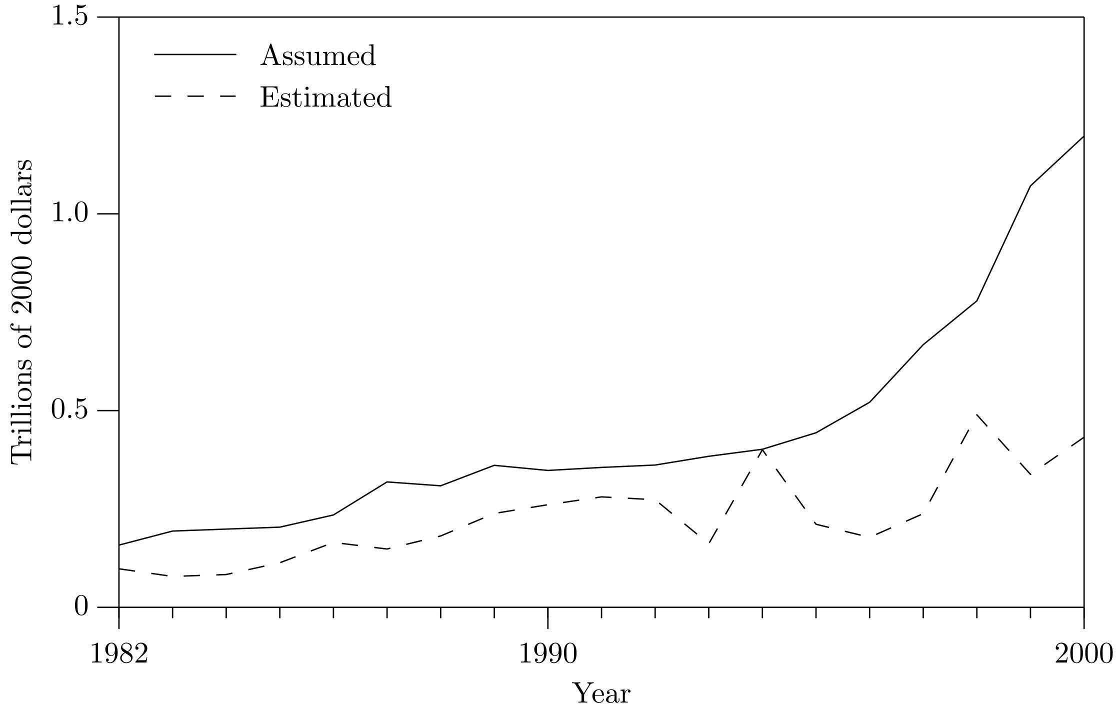

Figure 2.3. Assumed and Estimated Wealth of the 400 Wealthiest Americans, 1982–2000

Source: Data adapted from Brown et al. (2002); *Forbes Magazine* (1982–2001); Wilmoth (1997).

Note: This figure shows what the 400 wealthiest Americans were assumed to be worth for the sake of the exercise discussed in the text and what the 400 wealthiest Americans were estimated to be worth as part of the same exercise.

one percent of their wealth. The underestimation is more substantial in other years. In 1990, the estate-multiplier method underestimates their wealth by about 87 billion or 25 percent, and the method underestimates their wealth by about 765 billion or 64 percent in 2000, as noted above.

This figure of the assumed and estimated wealth of the 400 wealthiest Americans (again, fig. 2.3) is similar to the earlier figure of the direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans (fig. 2.1). Such similarity is suggestive. The similarity suggests that, if *Forbes Magazine*'s list was an accurate account of 400 wealthiest Americans, then the estate-multiplier method as it was applied by Kopczuk and Saez (2004a,b) may have substantially underestimated the wealth of the 400 wealthiest Americans, especially by the end of the 1990s.

The differences that were generated by the exercise are not actually quite as large as the differences that were observed between the direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans. In the year 2000, the exercise generates a difference of only about 765 billion dollars, rather than about 913 billion dollars. A difference of 765 billion dollars still seems substantial, but it is obviously smaller than 913 billion dollars.

However, if we make a realistic assumption about how wealth is valued during the exercise, then the differences generated by the exercise are at least as large the differences that were observed. The assumption is that, for each person who dies, his or her wealth is undervalued by half. That assumption is erroneous in the context of the exercise, but it is realistic for the reason noted above. As noted above, for people who died shortly after appearing on *Forbes Magazine*'s list, when the magazine's estimate of a person's wealth was compared to the wealth that was reported on his or her estate-tax records, the wealth reported on the estate-tax records was about half as much, at least on average.

Figure 2.4 shows the results of performing the exercise again while undervaluing what the people who died were worth by half. For comparison, the figure also shows the results from the earlier exercise, in which the wealth of the people who died was valued in full. As shown in the figure, the differences generated by the exercise are larger when wealth is undervalued by half. The differences generated by the exercise when wealth is undervalued by half are also at least as large as the differences that were observed. In the year 2000, again for example, the estate-multiplier method underestimate what the 400 wealthiest Americans were assumed to be worth by about 982 billion dollars, which is obviously greater than 913 billion dollars.

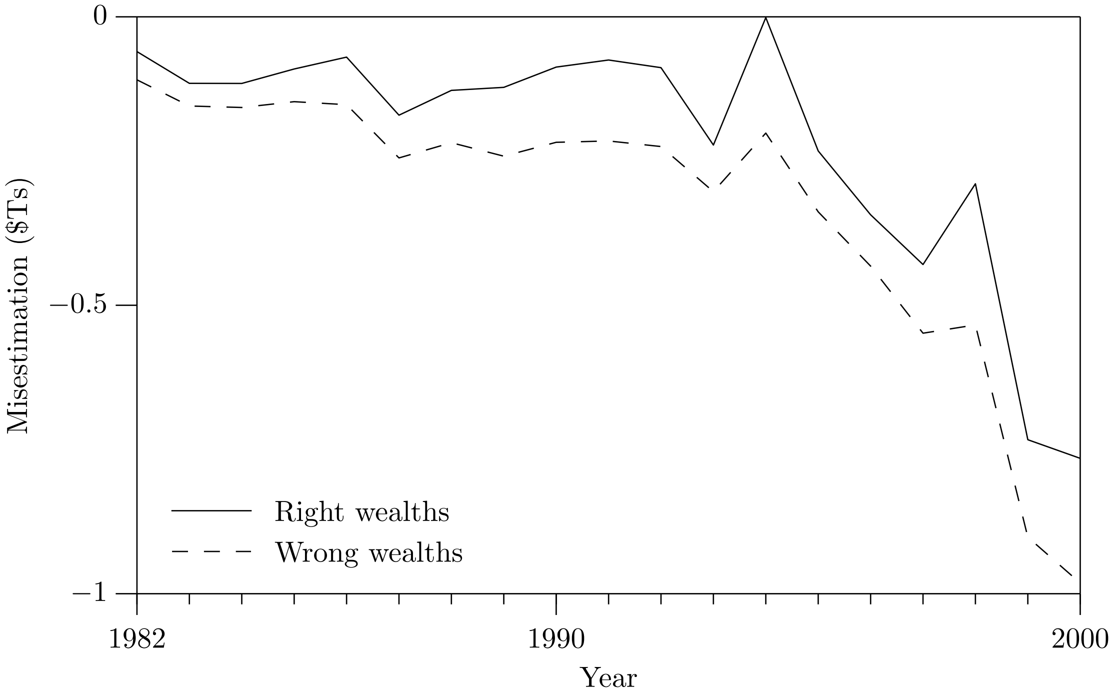

Figure 2.4. Misestimation of the Assumed Wealth of the 400 Wealthiest Americans with Different Assumptions about the Valuation of Wealth, 1982–2000

Source: Data adapted from Brown et al. (2002); *Forbes Magazine* (1982–2001); Wilmoth (1997).

Note: This figure shows, for different assumptions about the valuation of wealth, the difference between what the 400 wealthiest Americans were estimated to be worth as part of the exercise discussed in the text and what they were assumed to be worth as part of the same exercise. The wealth of someone who was sampled was either valued in full ("right wealths") or undervalued by half ("wrong wealths").

2.6 Extensions to the Exercise

2.6.1 The Probability of Substantially Smaller Estimates

In the exercise discussed above, we assumed that *Forbes Magazine*'s list was an accurate account of the 400 wealthiest Americans. We then applied the estate-multiplier method to the magazine's list by only using information about the people who died shortly after appearing on the list. The estate-multiplier method was applied to that information in the same way that Kopczuk and Saez (2004a,b) applied the method to estate-tax records. We used the mortality rates that they used, in particular. The main result of the exercise was that the estate-multiplier method underestimated what the 400 wealthiest Americans were assumed to be worth by a substantial amount, especially by the end of the 1990s. Such underestimation cannot be attributed to any timing, ownership, or valuation issues related to wealth because we took everyone's information directly from the magazine's list, although we did see that undervaluing wealths by a realistic proportion (namely, one half) led to even more severe underestimation.

The source of the underestimation in the exercised discussed above could be that the people who died were less likely to die than the Kopczuk-Saez mortality rates implied. Unfortunately, whether those mortality rates were the source of the misestimation or not is difficult to determine in the above-discussed exercise, given that we do not know how likely it was that a given person on the magazine's list in a given year would die. The probability that a person would die can be estimated, as discussed above, but it cannot be known with certainty, either in the real world or in the exercise.

In order to study the source of the underestimation, as well as the probability of such underestimation, the exercise discussed above can be extended as follows. Instead of using information about the people who actually died shortly after appearing on the magazine's list in a given year, we can draw a different sample from the people who were on the list in the given year and might have died, and then only use their information to apply the estate-multiplier method. People on the list can be sampled without replacement according to the mortality rates estimated earlier. By drawing our own sample, we obviously know the rates at which people are sampled. We therefore do not need to use estimates of the rates at which people were sampled. We do not need to use Kopczuk-Saez mortality rates, in particular, although we will explore the effect of using those rates in just a moment.

An issue that arises when drawing our own sample is that we might not sample anyone. That issue did not arise before because, although it was presumably possible that everyone who appeared on the magazine's list in any given year could have lived until the magazine made its list in the next year, at least one person died in every year. (Indeed, at least two people died in every year, as noted above.) Yet when we draw our own sample, we might not sample anyone. For the year 2000, for example, we would expect to draw an empty sample with a probability of about six percent, given the ages of the people who were on the list in that year and the mortality rates estimated earlier. We will treat an empty sample in the following way. Recall that we assumed we knew the minimum amount that a decedent must be worth in order to be observed. If we do not sample anyone, then we will assume that the 400 wealthiest Americans were all worth exactly that amount. That assumption is erroneous in the context of the exercise, but it is less erroneous than assuming that no one was worth that much. Also, even if we correctly assumed that that the 400 wealthiest Americans were all worth at least that much, we would not have any basis for guessing what they were worth beyond that amount.

If a large number of samples (such as 10,000 samples) are repeatedly redrawn for any given year, and if the estate-multiplier method is repeatedly reapplied in the way discussed above, then the probability that the method would substantially underestimate what the 400 wealthiest Americans were assumed to be worth in that year can be estimated. This extension to the exercise cannot tell us how likely it was that *Forbes Magazine*'s list was an accurate account of the 400 wealthiest Americans, but it can tell us how likely it was that the estate-multiplier method would underestimate the wealth of the 400 wealthiest Americans, if the magazine's list was an accurate account of the 400 wealthiest Americans, and if they died at the rates at which we are sampling them. The source of any underestimation can also be studied by studying how certain changes that occurred to the 400 wealthiest Americans over the years affected the probability of underestimating their total wealth.

Part of figure 2.5 shows, for each year between 1982 and 2000, the probability that the estate-multiplier method would underestimate what the 400 wealthiest Americans were assumed to be worth by at least a specific percentage of their wealth. The results are similar for similar percentages, but, for the figure, the specific percentage is the same percentage by which the direct and estate-multiplier estimates of the wealth of the 400 wealthiest Americans differed in 2000. Again, that percentage was about 76 percent of the direct estimate of their wealth. As shown in the figure, the probability of underestimating what the 400 wealthiest Americans were assumed to be worth by at least 76 percent is small. The probability is only as large as about six percent in 1998, the probability is about a quarter of a percent in 2000, and the probability is almost zero (about 0.02 percent) in 1990.

Although the probability of underestimating their wealth by at least 76 percent is small, the median amount by which their wealth would be misestimated is arguably large, at least by the end of the 1990s, as shown as part of figure 2.6 of this essay. In 1990, the estate-multiplier method would underestimate what the 400 wealthiest Americans were assumed to be worth by at least about 80 billion constant dollars half of the time, while the method would underestimate their wealth by at least about half a trillion dollars (about 438 billion dollars) half of the time in 2000.

Also, as shown in figure 2.5, if we made the same assumption about the undervaluation of wealth that we made earlier, then the probability of underestimating what the 400 wealthiest Americans were assumed to be worth by at least 76 percent would be larger, especially by the end of the 1990s. The probability would be about 33 percent or more in every year between 1996 and 2000. In one of those years, underestimating their wealth by at least 76 percent would be more probable than not. The probability would be about 52 percent in 1999. The median misestimation would also be larger in every year, as shown in figure 2.6 of this essay.

The probability of underestimating their wealth by a substantial percentage or amount would be even larger, if we made another realistic assumption about what we know about the rates at which people are sampled. Although we obviously know the rates at which we are sampling people, we can pretend we do not know those rates, and use the mortality rates used by Kopczuk and Saez instead. When applying the estate-multiplier method to estate-tax records, we do not actually known the rates at which people were sampled by Death. We can only use estimates like the Kopczuk-Saez mortality rates. So, we are similarly assuming that our knowledge about the rates at which people are sampled is imperfect.

If wealth is undervalued by half instead of valued in full, and if the Kopczuk-Saez mortality rates are used instead of the actual sampling rates, then underestimating what the 400 wealthiest Americans were assumed to be worth by at least 76 percent would be about 50 percent or more in every year between 1996 and 2000, as shown as part of figure 2.7 of this essay. The median amount by which their wealth would be misestimated would also be at least about 423 billion dollars in every year between 1996 and 2000, as shown as part of figure 2.8 of this essay.[7]

Recall that 76 percent is the percentage by which, in the year 2000, the direct estimate of the wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list differed from the estate-multiplier estimate of their wealth based on the Kopczuk-Saez estimates. The percentages by which those estimates differed were different in different years. As a percentage of the direct estimate, the estimates differed by about 54, 64, 68, 76, and 76 percent in each year from 1996 to 2000. In the

[7]The Kopczuk-Saez mortality rates for males, rather than females, were used. The misestimation would be similar, but less severe, if the rates for females were used.

context of the exercise, the probability of underestimating what the 400 wealthiest Americans were assumed to be worth in each of those years by at least each of those percentages can be estimated. Undervaluing wealths by half and using the Kopczuk-Saez mortality rates, those probabilities would be estimated to be about 98, 93, 86, 73, and 48 percent. Of course, the probability of underestimating their wealth year after year would be smaller than the probability of underestimating their wealth in a given year. The probability would not be infinitesimal, however. The probability of underestimating their wealth by at least each of those percentages in each of the years between 1996 and 2000 would be almost 30 percent.

These results are suggestive. They suggest that, in the years towards the end of the 1990s when there were substantial differences between the direct estimates of the wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list and the estate-multiplier estimates of their wealth based on the Kopczuk-Saez estimates, it was possible and arguably probable that the estate-multiplier method as it was applied by Kopczuk and Saez (2004a,b) would have underestimated the wealth of the 400 wealthiest Americans by at least the amount by which the direct and estate-multiplier estimates differed, if *Forbes Magazine*'s list was an accurate account of 400 wealthiest Americans in those years.

Admittedly, the exercise also suggests that the probability of underestimating the wealth of the 400 wealthiest Americans by at least each of those amounts in each of those years would have been essentially zero, if the mortality rates used by Kopczuk and Saez (2004a,b) were accurate estimates of the rates at which people died, or if estate-tax records were accurate estimates of a decedent's wealth. Again, however, if *Forbes Magazine*'s list was an accurate account of the 400 wealthiest Americans, then it appears that the Kopczuk-Saez mortality rates were generally too high (sec. 2.4.4) and that the wealths of decedents were generally too low (sec. 2.4.1).

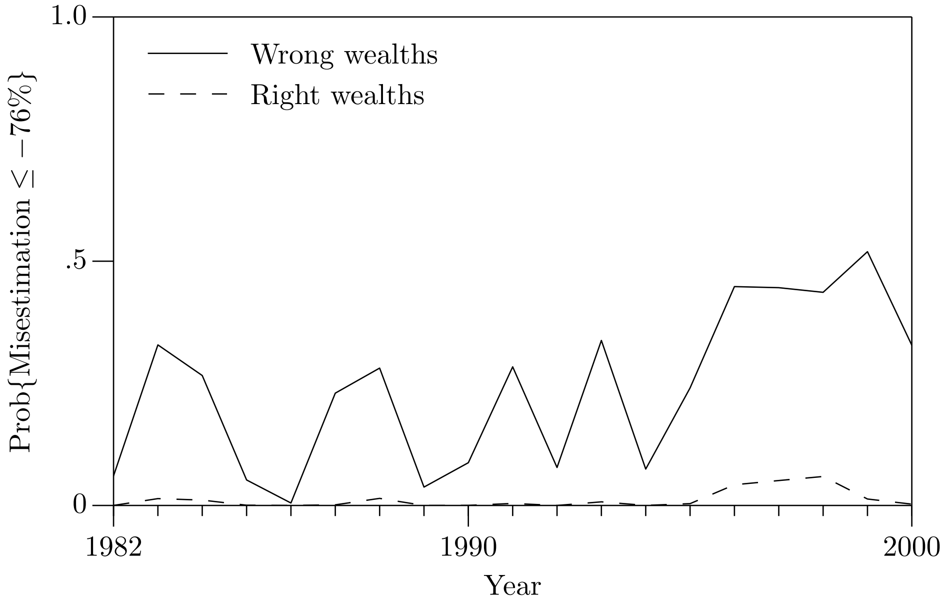

Figure 2.5. Probability of Underestimating the Assumed Wealth of the 400 Wealthiest Americans by at least 76 Percent with Different Assumptions about the Valuation of Wealth, 1982–2000

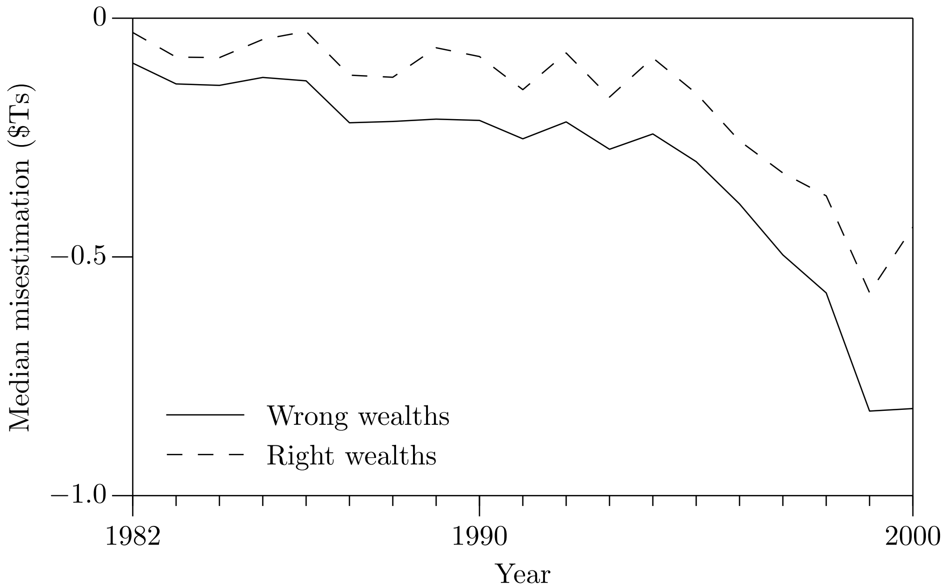

Figure 2.6. Median Misestimation of the Assumed Wealth of the 400 Wealthiest Americans with Different Assumptions about the Valuation of Wealth, 1982–2000

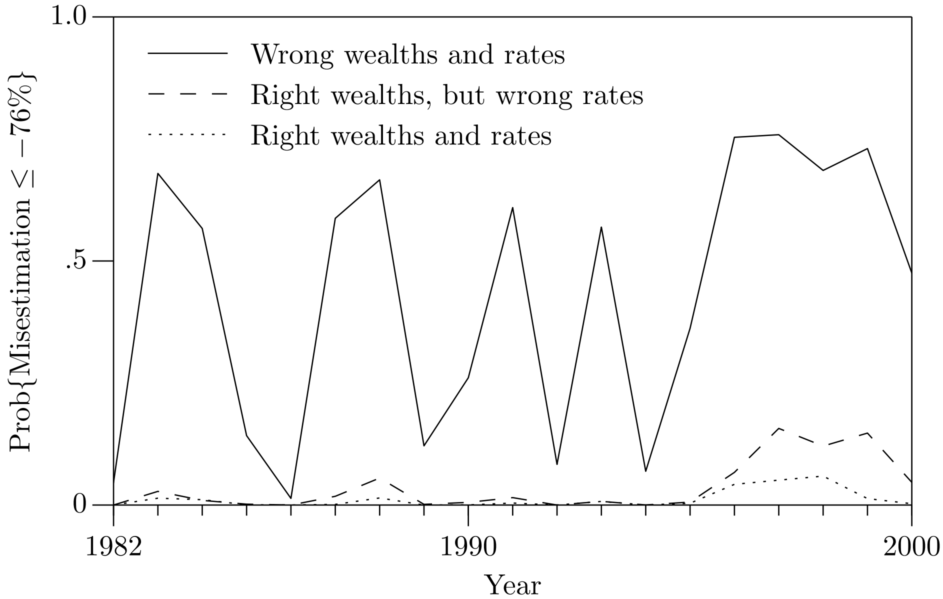

Figure 2.7. Probability of Underestimating the Assumed Wealth of the 400 Wealthiest Americans by at least 76 Percent with Different Assumptions about the Valuation of Wealth and the Rates at Which People Were Sampled, 1982–2000

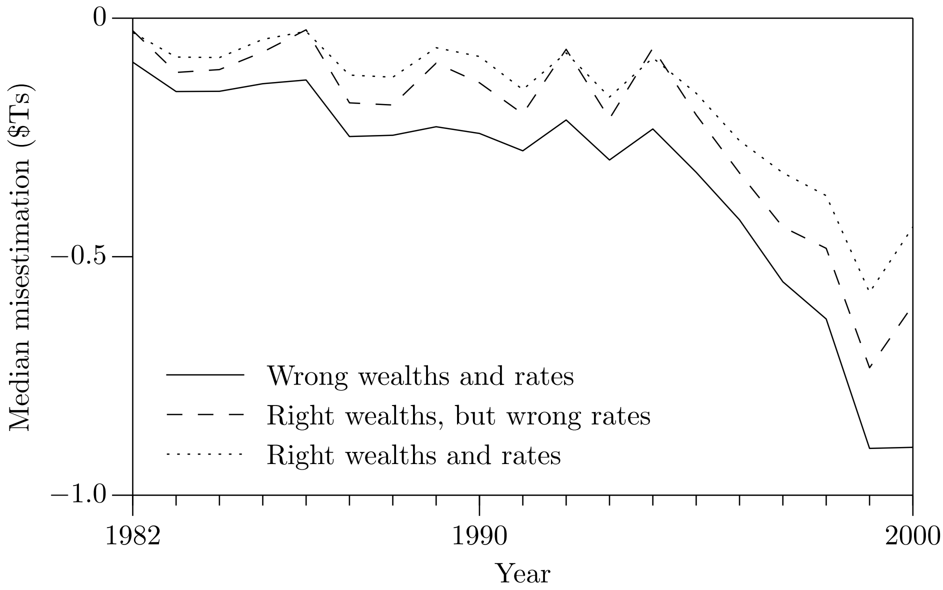

Figure 2.8. Median Misestimation of the Assumed Wealth of the 400 Wealthiest Americans with Different Valuation and Sampling Assumptions, 1982–2000

2.6.2 The Source of the Substantially Smaller Estimates

The probability of underestimating what the 400 wealthiest Americans were assumed to be worth varies by year in the exercise discussed above. The source of that underestimation can therefore be studied by studying how changes that occurred over those years would have affected the probability of underestimation. Changes that occurred between the years 1990 and 2000 can be studied, in particular. Maintaining our assumption that *Forbes Magazine*'s list was an accurate account of the 400 wealthiest Americans in every year, the relevant changes that occurred between the years 1990 and 2000 were the following. The ages of the 400 wealthiest Americans changed. They became younger on average as their average age fell from about 64 to 60 years of age, for example. (The standard deviation of their ages was about 13 years of age in both years.) Their wealths also changed, as discussed in more detail below. And the rates at which they died changed, at least according to the estimates we made earlier.

The net effect of all of those changes to the 400 wealthiest Americans—again, the changes to their ages, their wealths, and the rates at which they died—was that it was more likely that their total wealth would be substantially underestimated in our exercise. The probability of underestimating what the 400 wealthiest Americans were assumed to be worth by at least 76 percent would have almost doubled from about 26 to 48 percent, assuming that wealths were undervalued by half instead of valued in full, and also assuming that the Kopczuk-Saez mortality rates were used instead of the actual sampling rates, as shown as part of table 2.1 of this essay. We will focus on the implementation of the exercise in which wealths are undervalued and the Kopczuk-Saez mortality rates are used, which seems like the most realistic implementation, although some of the results for other implementations are reported in the same table for completeness.

The effect of each change can be studied by starting with the ages and wealths of the 400 wealthiest Americans in 1990, as well as our estimates of the rates at which they died in that year. Starting with those ages, wealths, and rates, the effect of changing the ages of the 400 wealthiest Americans can be studied by taking the ages of the wealthiest to the poorest of the 400 wealthiest Americans in 2000, and assigning those ages to the wealthiest to the poorest of the 400 wealthiest Americans in 1990. (When people have the same wealth, ties can be broken by the alphabetical order of a person's name, which should be a random way of breaking ties.) By doing that, the ages of the 400 wealthiest Americans change, and any association between a person's age and his or her rank in the distribution of wealth also changes, but everything else is held constant.

That change in the ages of the 400 wealthiest Americans has a relatively small effect on the probability of underestimating their wealth by at least 76 percent, as shown in the table. The probability is only about two percentage points lower. Note that, if a person with a given amount of wealth was less likely because he or she was younger, then the person's wealth would be inflated by a larger factor in the event of his or her death, which would tend to decrease the probability of underestimation.

Changing the rates at which people are sampled—either in isolation or in concert with a change in their ages—also has a relatively small effect on the probability of underestimating their wealth by at least 76 percent, as shown in the same table. The probability is about four percentage points higher. It seems that, at least for the set of ages associated with the people on the magazine's list in the year 1990, the difference between the estimated mortality rates and the Kopczuk-Saez mortality rates was greater in 2000 than 1990.

Changing the ages of the 400 wealthiest Americans or the rates at which they died has a relatively small effect, relative to the effect of changing the wealths of the

400 wealthiest Americans. The effect of that change can be studied by taking the wealths of the wealthiest to the poorest of the 400 wealthiest Americans in 2000, and assigning those wealths to the wealthiest to the poorest of the 400 wealthiest Americans in 1990. When the wealths of the 400 wealthiest Americans are changed in that way, the probability of underestimating their wealth by at least 76 percent almost doubles from about 26 to 53 percent. The net effect of all of the changes that occurred to the 400 wealthiest Americans between 1990 and 2000 is therefore about the same as the effect of just changing their wealths.

The change that occurred to their wealths between 1990 and 2000 could be characterized in different ways. In anticipation of another exercise, we can characterize the change as follows. If the wealth of the 400 wealthiest Americans follows a Pareto distribution, then there should be a percentage $p \in (0, 50)$ such that the wealthiest p percent of the 400 wealthiest Americans own $(100 - p)$ percent of the total wealth of the 400 wealthiest Americans (assuming that their mean wealth is finite; Hardy 2010, p. 42, proposition 2). Whether the wealth of the 400 wealthiest Americans actually follows a Pareto distribution is a subject of scholarly debate (Clauset et al. 2009), but the percentage of the wealth of the 400 wealthiest Americans owned by different percentages of the 400 wealthiest Americans can still be calculated. In 1990, the wealthiest 36 percent of the 400 wealthiest Americans owned about 64 percent of the wealth of the 400 wealthiest Americans (again, maintaining our assumption that the magazine's list was an accurate account). In 2000, in contrast, that same percentage of the 400 wealthiest Americans owned about 75 percent of the wealth of the 400 wealthiest Americans.

The effect of even more dramatic changes to the distribution of the wealth among the 400 wealthiest Americans on the probability of underestimating their total wealth can be illustrated as follows. Consider the 400 wealthiest Americans in

2000. Holding everything else constant—their total wealth, their ages, any association between ranks and ages, and the rates at which we are sampling them—we can change the distribution of their wealth. It is straightforward to distribute their wealth with different degrees of inequality, if we assume that their wealth follows a Pareto distribution. Note that the estate-multiplier estimates can only be improved by assuming that their wealth follows a Pareto distribution, given that the extrapolation method used by Kopczuk and Saez (2004a,b) assumes that wealth follows that particular distribution.

The probability of underestimating what the 400 wealthiest Americans were assumed to be worth by at least 76 percent can be calculated for Pareto distributions of wealth that range from almost perfect equality (i.e., slightly less than half of the 400 wealthiest Americans owning slightly more than half of the total wealth of the 400 wealthiest Americans) to almost perfect inequality (i.e., one person owning almost all of the wealth of the 400 wealthiest Americans). Those probabilities are shown in the top half of figure 2.9 of this essay. The median misestimation for each wealth distribution is shown in the bottom half of that figure.

As shown in the figure, if the distribution of wealth among the 400 wealthiest Americans was almost perfectly equal, then the estate-multiplier method would almost never overestimate their wealth by 76 percent. However, if the distribution of their wealth was almost perfectly unequal, then the method would almost always underestimate their wealth by at least that much. Like before, the estate-multiplier method was applied by undervaluing wealths and using the Kopczuk-Saez mortality rates, but, even under different assumptions about the valuation of wealth and rates at which people are sampled, the probability of underestimating their wealth by at least 76 percent would still approach certainty as the distribution of their wealth approached perfect inequality.

When the distribution of wealth among the 400 wealthiest Americans is almost perfectly unequal, their wealth will be substantially underestimated whenever the wealthiest person is not sampled. Given the age of the wealthiest person in 2000, and given the mortality rates estimated earlier, the wealthiest person should only be sampled with about a 1-in-240 chance. Yet, even in the unlikely event that the wealthiest person was sampled, the wealth of the 400 wealthiest Americans would still be substantially misestimated. Their wealth would be overestimated for the same reason that it would be underestimated whenever that person was not sampled. The wealth of the wealthiest person is unlike the wealth of anyone else.

Figure 2.10 shows a frequency distribution of estimates of the total wealth of the 400 wealthiest Americans when almost all of their total wealth is owned by just one person. The estimates form a bimodal distribution. At one mode, which reflects samples where the wealthiest person was not sampled, the estimates are at least one trillion dollars less than what the 400 wealthiest Americans were assumed to be worth. At the other mode, which reflects samples where the wealthiest person was sampled, the estimates are trillions of dollars too large. Again, the estate-multiplier method would over- or under-estimate what the 400 wealthiest Americans were assumed to be worth, depending on whether the wealthiest person was sampled or not, and the method would misestimate their wealth in either case.

The distribution of wealth among the 400 wealthiest Americans did not evolve from almost perfect equality to almost perfect inequality over the 1990s, of course, but the distribution of their wealth did become more unequal over that decade, at least according to *Forbes Magazine*'s list. If that list was an accurate account, then the increase in inequality among the 400 wealthiest Americans over the 1990s was apparently dramatic enough to have a relatively large effect on the probability that the estate-multiplier method would underestimate their total wealth.

Table 2.1. Effect of Changes between 1990 and 2000 on Misestimation

Changes	Probability of underestimating by at least 76%
	Right wealths and rates
No changes	0.02%
Wealths	0.79
All changes	0.27
	Right wealths, but wrong rates
No changes	1%
Wealths	7
All changes	5
	Wrong wealths, but right rates
No changes	9%
Wealths	34
All changes	33
	Wrong wealths and rates
No changes	26%
Ages	23
Sampling rates	22
Ages and sampling rates	22
Wealths	53
All changes	48

Sources: Data adapted from Brown et al. (2002); *Forbes Magazine* (1990, 2000); Wilmoth (1997).

Note: This table shows how certain changes that occurred between 1990 and 2000 would have affected the probability that the estate-multiplier method would underestimate what the 400 wealthiest Americans were assumed to be worth by at least 76 percent of their wealth, given different assumptions about the valuation of wealth and the rates at which people are sampled.

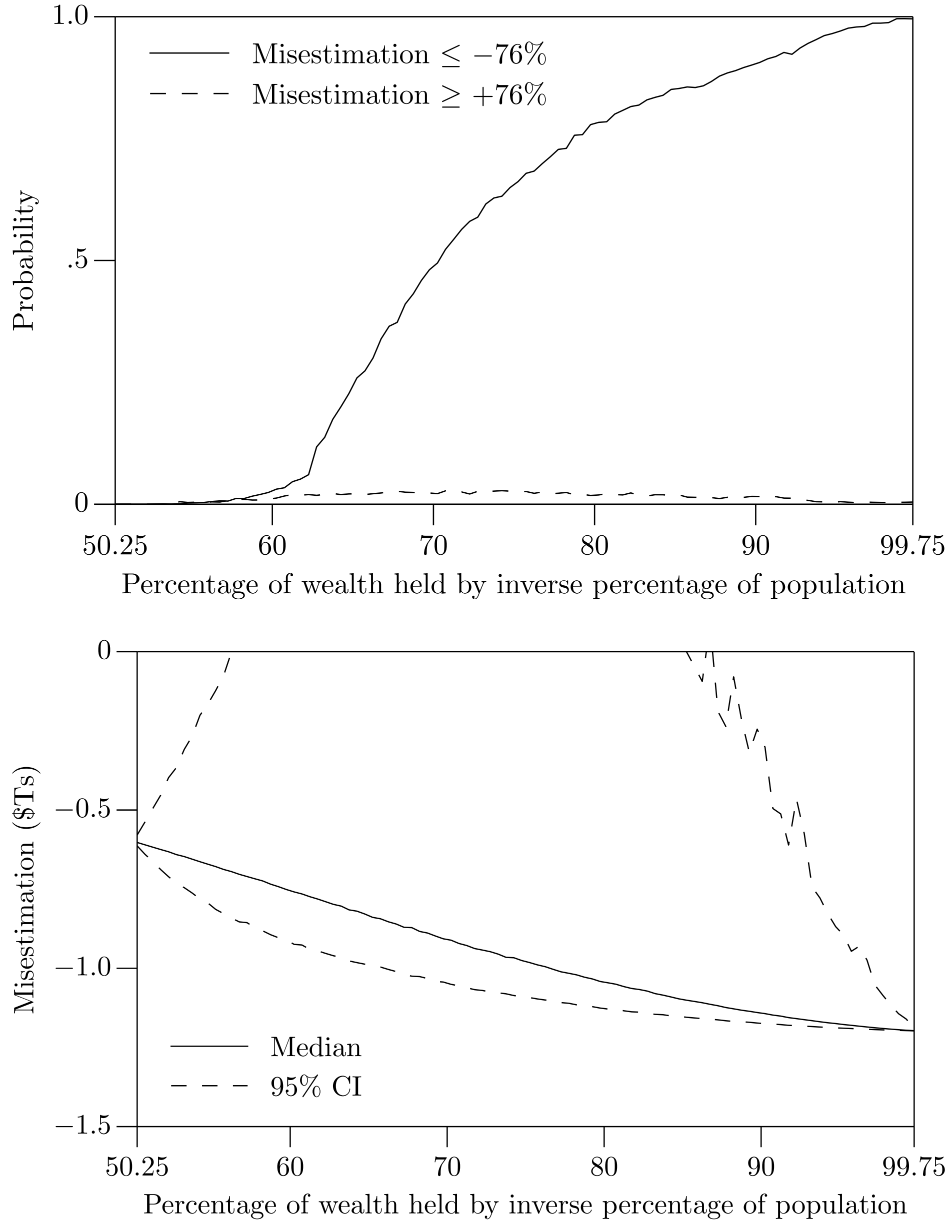

Figure 2.9. Effect of Inequality on the Misestimation of the Assumed Wealth of the 400 Wealthiest Americans in 2000

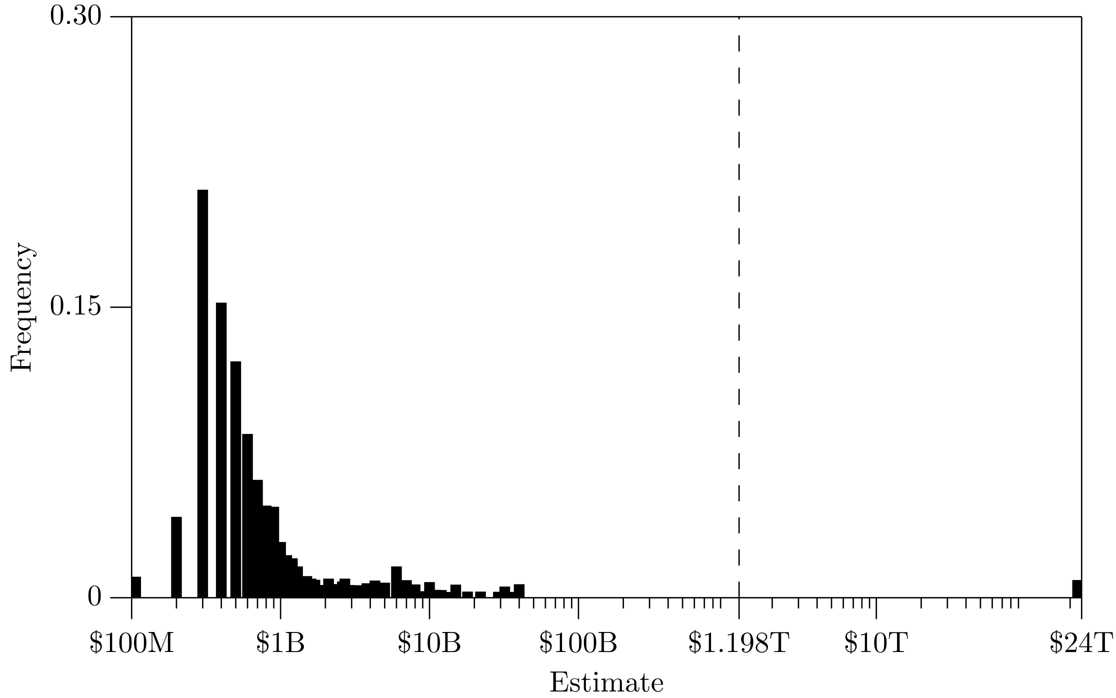

Figure 2.10. Frequency Distribution of Estimates of the Assumed Wealth of the 400 Wealthiest Americans in 2000 Under Extreme Inequality

Note: This figure shows the frequency distribution of estimates of the assumed wealth of the 400 wealthiest Americans in 2000, if their wealth was distributed so that one of them owned 99.75 percent of their total wealth. An estimate of their total wealth was rounded to the nearest one billion dollars, if it was greater than five billion, and the nearest 100 million, otherwise. The dashed line denotes what the 400 wealthiest Americans were assumed to be worth. The closest estimates are all at least one trillion dollars off.

2.7 Conclusion

The estate-multiplier estimator of the total wealth of people who were alive and wealthy enough to be subject to the estate tax is a specific type of estimator, specifically, a Horvitz-Thompson estimator (Scheuren and McCubbin 1987, p. 29). A fairly famous critique of that type of estimator comes in the form of a parable due to Basu ([1971] 2011, pp. 176–7). The parable can be paraphrased as follows.

An owner of a circus has a herd of 50 elephants. The owner wants to estimate the total weight of the herd because she wants to transport them, but weighing even one elephant is a laborious task, so she does not want to weigh each elephant. The owner did weigh each elephant in the not-too-distant past, however. When the owner weighed them in the past, one elephant named Sambo weighed the average amount, while another elephant named Jumbo weighed the most.

After consulting with the elephant trainer to make sure that Sambo still seemed to be an average-sized elephant, the owner was going to weigh Sambo again and estimate the total weight of the herd as 50 times his weight, but the owner decided to consult a statistician before doing anything.

The statistician suggested a different approach. The statistician suggested drawing a random sample of the elephants, weighing those elephants, and then using the Horvitz-Thompson estimator to estimate the total weight of the herd. The sampling scheme suggested by the statistician was to sample Sambo with a probability of 99 percent and sample the other 49 elephants with equal probabilities that summed to one percent (i.e., probabilities of 1/4,900).

After some discussion with the statistician, the owner realized that, in the likely event that Sambo and only Sambo was sampled, the estimate of the total weight of the herd would only be slightly larger than just Sambo's weight, specifically, just 100/99 times Sambo's weight. The owner also realized that, in the unlikely event that Jumbo and only Jumbo was sampled, the estimate would be 4,900 times Jumbo's weight. After realizing those things, the owner fired the statistician.

At least one lesson that can be drawn from this parable is that, although the Horvitz-Thompson estimator should be right on average across all possible samples (and although it is the "unique hyperadmissible estimator in the class of all generalized polynomial unbiased estimators," in the words of the statistician; ibid., p. 177), its estimates can be wildly inaccurate for any given sample (Welsh 2011).

If we extend the parable of Basu's elephants slightly, then a similar lesson can be drawn. Suppose that Jumbo is so heavy that he accounts for a substantial amount of the total weight of the herd. Then any random sample of the elephants will misestimate their total weight by a substantial amount, unless Jumbo is sampled with a probability close to 100 percent. To the extent that substantially misestimating the weight of the herd is something worth firing people over, it would be wise for the owner to either consult with the trainer or simply look at the herd for herself. Even a cursory look would surely reveal that Jumbo is a jumbo-sized elephant who must be weighed in order to come close to estimating the total weight of the herd.

To connect this parable about estimating the weight of a herd of pachyderms to the preceding study about estimating the wealth of a group of plutocrats, consider Bill Gates. There is almost no doubt that Gates has been blessed with extraordinary wealth. *Forbes Magazine* has taken note of that fact. Estate-tax records have been blind to that fact, however, given that Gates has also been blessed with longevity (and, even if he had died, estate-tax records may have still been blind to his extraordinary wealth, given that Gates has presumably been blessed with an estate planner). Anyone who uses estate-tax records to estimate the total wealth of the wealthiest Americans while overlooking the fact that Gates has been alive and wealthy is therefore ignoring an elephant in the room, so to speak.

The preceding study suggests that, when they applied the estate-multiplier method to estate-tax records, Kopczuk and Saez (2004a,b) may have overlooked some elephants in the room, especially by the end of the 1990s. The study suggests that, if *Forbes Magazine*'s list was an accurate account of the 400 wealthiest Americans, then, as the distribution of their wealth became increasingly unequal over the 1990s, it became increasingly likely that the estate-multiplier method as it was applied by Kopczuk and Saez (2004a,b) would have underestimated the wealth of the 400 wealthiest Americans by a substantial amount.

One direction for future research is to try to rescue the estate-multiplier method by accounting for the extraordinarily wealthy individuals who are the proverbial elephants in the room. *Forbes Magazine*'s list could perhaps be useful in that respect, although researchers will need to think carefully about how the information contained in that list might be incorporated into the estate-multiplier method. Indeed, a lesson that has been drawn from Basu's parable is that the statistician did not think carefully enough about how to incorporate information about the weights of the elephants during their earlier weigh in (Welsh 2011).

Instead of trying to improve the accuracy of estimates made by applying the estate-multiplier method to estate-tax records, an alternative direction for future research is to simply use direct estimates. To that extent that *Forbes Magazine*'s estimates are seen as somehow flawed, there is no reason why direct estimation of the wealth of the wealthiest Americans (400 or otherwise) must be left to a popular magazine. Such a project is well beyond the scope of this essay, but academic economists and others could devote time and resources to directly estimating the wealth of the wealthy in the same way that they devote time and resources to conducting surveys or creating other sources of data.

2.8 A Postscript on the 2000s

Unlike the top income share estimates that Saez and another colleague have continued to update on an annual basis since first publishing them (Piketty and Saez 2003), Saez and Kopczuk have not updated their top wealth share estimates since first publishing them. As such, the most-recent year for which their top wealth share estimates are available is the year 2000, which was over a decade ago, as of writing. Whether their top wealth share estimates would have remained as stable in the 2000s as they were in the 1990s is therefore unknown. Kopczuk and Saez or anyone else (or, at least, anyone else who befriends an employee at the IRS) could update their estimates, but that is beyond the scope of this essay.

Although top wealth estimates based on applying the estate-multiplier method to estate-tax records are not currently available for the years since 2000, the exercise that was performed in this essay for earlier years can also be performed for more-recent years because *Forbes Magazine* has continued to publish its list of the 400 wealthiest Americans. This postscript reports the results of performing the exercise for the 2000s. The results are reported mostly for completeness, although there is at least one interesting thing to note.

Figure 2.11 shows, for the years from 1982 to 2010, estate-multiplier estimates of what the 400 wealthiest Americans were assumed to be worth in a given year based on applying the estate-multiplier method to *Forbes Magazine*'s list from that year by only using information about the people who died by the time the magazine made its list in the next year. The figure also shows what the 400 wealthiest Americans were assumed to be worth in each year. This figure is the same as figure 2.3 for the years from 1982 to 2000. As shown in figure, in each year between 1982 and 2010, except one year, the estate-multiplier method underestimates what the 400 wealthiest Americans were assumed to be worth, often by a substantial amount.

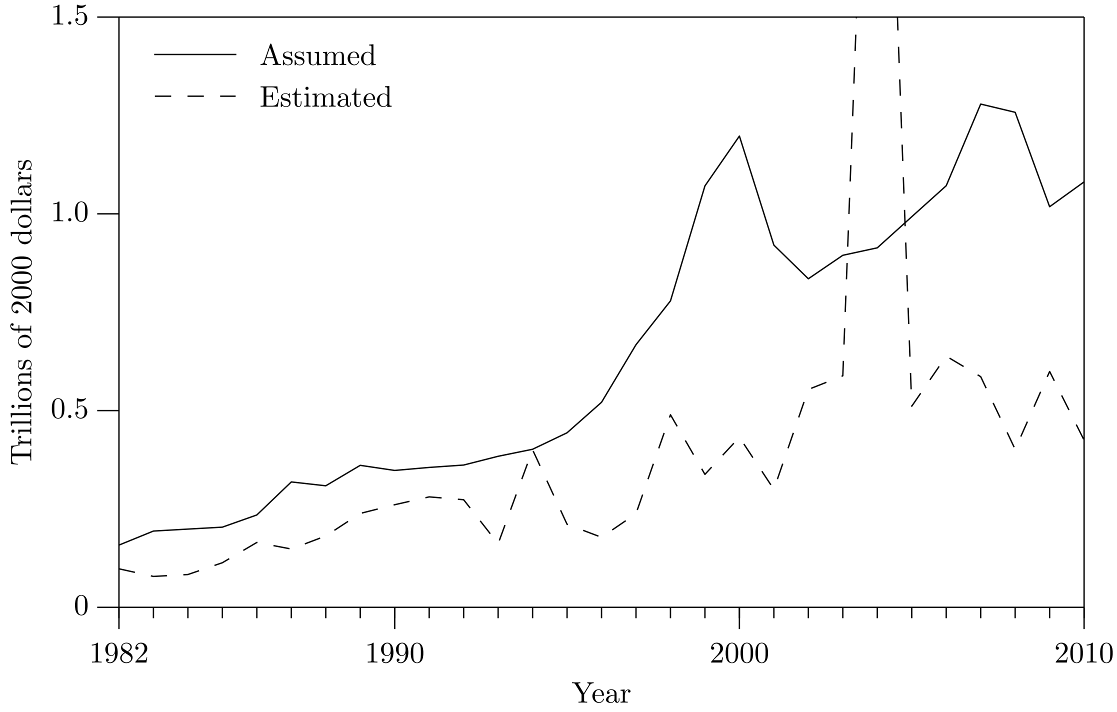

Figure 2.11. Assumed and Estimated Wealth of the 400 Wealthiest Americans, 1982–2010

Source: Data adapted from *Forbes Magazine* (1982–2011); Wilmoth (1997).

Note: This figure is identical to figure 2.3 except that it includes more-recent years.

The one exception is that, in the year 2004, the estate-multiplier method overestimates their wealth. The estimate is literally off the chart at about three trillion constant dollars. That overestimation occurred for the following reason. Out of the 400 people who appeared on the magazine's list in 2004, five of them died by the time the magazine made its list in the next year. One of those five decedents was the Walmart heir John Walton, who died after the experimental ultralight aircraft that he was piloting crashed (*Associated Press* 2005). In 2004, Walton was 58 years old and the fourth wealthiest American with a wealth of about 18 billion dollars, according to the magazine's list from the year. Using the mortality rates used by Kopczuk and Saez, Walton's probability of death was relatively small, given his relatively young age. He should have died with about a 1-in-158 chance. The estate-multiplier method therefore implies that there should have been about 158 people worth 18 billion dollars, but only a few people were worth at least that much. If Walton had not died in that year, then the estate-multiplier estimate would have underestimate what the 400 wealthiest Americans were assumed to be worth by about half a trillion constant dollars (specifically, about 536 billion constant dollars).

Scholars who have applied the estate-multiplier method to estate-tax records have often recognized that the death of someone like Walton could have a substantial affect on estate-multiplier estimates. For example: One noted scholar once noted that, “A wealthy young man crashing his sports car could add [a large number of people] to the estimate of the number of people with wealth over [a certain amount of wealth]” (Tony Atkinson quoted in Polanyi and Wood 1974, p. 231). For our exercise, Walton crashing his plane led to a substantial overestimation of what the 400 wealthiest Americans were assumed to be worth. A similar overestimation should occur when applying the estate-multiplier method to estate-tax records, unless Walton’s wealth for estate-tax purposes was much smaller than the magazine’s estimate.

CHAPTER 3

ON THE DISTRIBUTION OF THEIR WEALTH

3.1 Introduction

Over a century ago, Vilfredo Pareto discovered that the upper tail of the income distribution looked approximately like a distribution that is now named after him, namely, a "Pareto" distribution. Ever since Pareto's discovery, scholars have debated the extent to which income, wealth, and other variables follow Pareto distributions (Persky 1992). Relatively recently, Levy and Solomon (1997) studied whether wealth follows a Pareto distribution among the 400 wealthiest Americans. Their study used a list of the 400 wealthiest Americans published by *Forbes Magazine*. That magazine has published a list of the 400 wealthiest Americans in every year since 1982, but Levy and Solomon (1997) only used the magazine's list from one year, 1996. Based on their study of that year's list, they concluded that wealth appeared to follow a Pareto distribution among the 400 wealthiest Americans in at least that year.[1]

In a subsequent study, Klass et al. (2006), the same researchers and their colleagues used the magazine's list from each year between the years 1988 and 2003. They concluded that wealth appeared to follow a Pareto distribution among the 400 wealthiest Americans in each of those years. Other researchers have reached

[1] Wealth is a stock variable, so it should be measured at a point in time. According to *Forbes Magazine* (1996, p. 104), its 1996 list was a snapshot of wealth at the close of the stock market on one day of that year, August 23rd. When this essay speaks of the 400 wealthiest Americans on the magazine's list in a given year, the essay is speaking about the 400 wealthiest Americans at the point in time when the magazine made its list in the given year. We omit the point in time partly for readability and also partly because the magazine did not always report the point in time.

the same conclusion by studying the list from similar years (specifically, 1996 to 2004; Castaldi and Milakovic 2007). The conclusion has also been contradicted, however. Clauset et al. (2009, pp. 685–6) studied one year of the magazine's list and concluded that wealth almost certainly did not follow a Pareto distribution among the 400 wealthiest Americans in at least that year. Unlike previous studies, which based their conclusions on visual inspection or other forms of casual empiricism, the conclusion of Clauset et al. (2009) was based on a formal goodness-of-fit test.

This essay replicates and extends the study by Klass et al. (2006) in order to critically examine the extent to which wealth follows a Pareto distribution among the 400 wealthiest Americans. The essay begins by replicating their study. Issues identified during the replication are discussed next. The essay then extends their study. Using *Forbes Magazine*'s list but using every available year of that list, the essay examines the statistical and substantive significance of deviations between the distribution of wealth among the 400 wealthiest Americans and a Pareto distribution. A goodness-of-fit test like the one in Clauset et al. (2009) is applied as part of that examination. The essay argues that the distribution of wealth among the 400 wealthiest Americans appears to deviate from a Pareto distribution to a statistically and substantively significant extent at almost every point in time. Whether scholars should continue to see a Pareto distribution as even a crude approximation to the upper tail of the wealth distribution is debated to conclude.

3.2 Replication of Previous Work

What Pareto discovered over a century ago was that, at least in its upper tail, the empirical complementary cumulative distribution function (CCDF) for income looked approximately like a straight line on a double logarithmic scale (Pareto [1896] 2001). A Pareto distribution is simply a formalization of that sort of shape. The

CCDF for a Pareto distribution is $(x_{\min}/x)^{\alpha}$ for $x \geq x_{\min}$, where $x_{\min} > 0$ is a lower-bound parameter and α is a shape parameter (Arnold 1983). If that CCDF is drawn on a double logarithmic scale, then, for levels greater than the lower-bound parameter, it looks exactly like a straight line with a slope equal to the negative of the shape parameter. The empirical CCDF for a variable that follows a Pareto distribution should therefore tend to look like a straight line when it is drawn on a double logarithmic scale.[2]

Following a figure in Klass et al. (2006, p. 292, fig. 1), part of figure 3.1 of this essay shows the empirical CCDF for the wealth of the 400 wealthiest Americans on *Forbes Magazine*'s list in the year 2003. Note that the empirical CCDF evaluated at a given amount of wealth is drawn as the proportion of the 400 wealthiest Americans who were worth strictly more than that amount (as opposed to that amount or more), although the figure would look essentially the same if the alternative convention was used. Also note that the empirical CCDF is drawn on a double logarithmic scale. As seen in the figure, each of the 400 wealthiest Americans was worth at least 600 million current dollars in 2003, while the wealthiest one of them (Bill Gates) was worth 46 billion current dollars.

Based on a similar figure, Klass et al. (2006, p. 291) suggested that the empirical CCDF for the wealth of the 400 wealthiest Americans on the magazine's list in the same year looked "very close" to a straight line on a double logarithmic scale. The empirical CCDF for their wealth does look somewhat like a straight line on a double logarithmic scale, as the figure in this essay shows. Visual inspection might therefore suggest that wealth follows a Pareto distribution among the 400 wealthiest Americans on the magazine's list in at least that year (Klass et al. 2006, p. 291).

---

[2]Other distributions can also look like straight lines over certain ranges when they are drawn on a double logarithmic scale, as discussed below. The fact that a Pareto distribution looks like a straight line on that scale is shown formally in appendix B.1 and visually in the figure that follows.

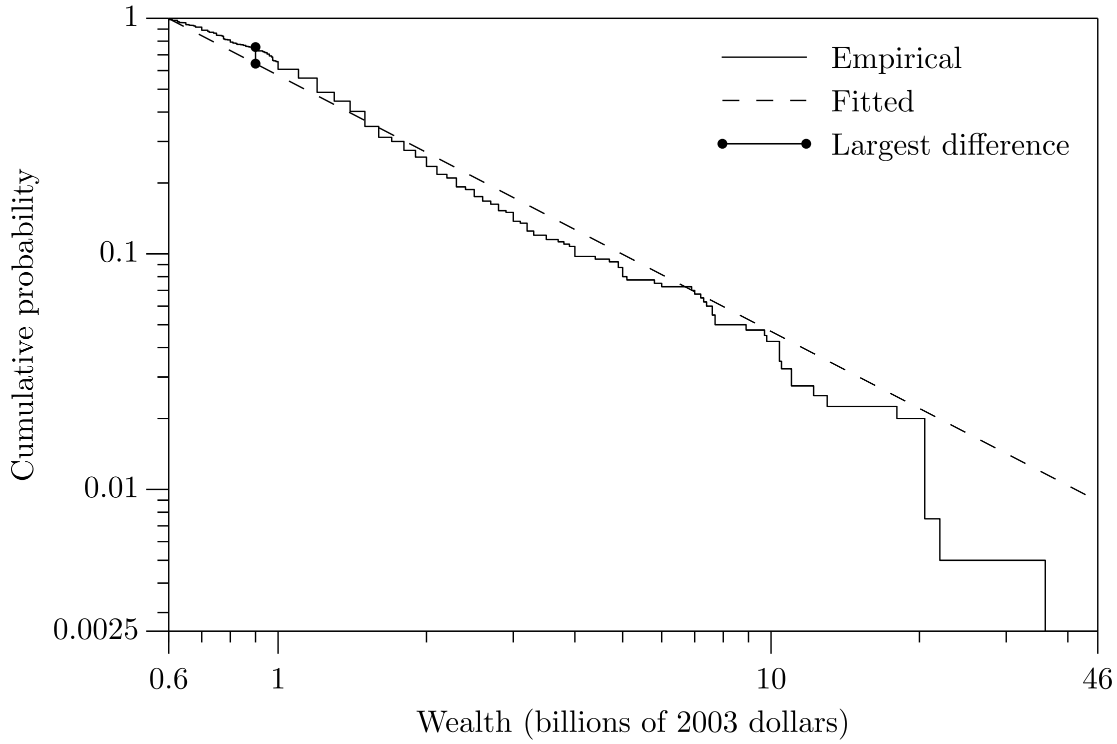

Figure 3.1. Distribution of Wealth Among the 400 Wealthiest Americans in 2003

Source: Data adapted from *Forbes Magazine* (2003).

Note: This figure shows the empirical CCDF for the wealth of the 400 wealthiest Americans on *Forbes Magazine*'s list in 2003, the CCDF for a Pareto distribution that was fit to the distribution of their wealth by maximum likelihood, and the largest absolute difference between the empirical and fitted CCDFs, which occurs once at a wealth of about 900 million current dollars. The difference between the empirical and fitted CDFs at that wealth is about 11 percent.

Regardless of whether or not someone sees something that looks like a straight line when he or she looks at the empirical CCDF for their wealth on a double logarithmic scale, and regardless of whether or not their wealth actually followed a Pareto distribution, the parameters of a Pareto distribution can be estimated from the distribution of their wealth. As part of their study, Klass et al. (2006) estimated the shape parameter of a Pareto distribution from the distribution of wealth among some of the wealthiest Americans on the magazine's list in 2003. They estimated the shape parameter in the following way.

Using the magazine's list from that year, they began by ranking the 400 wealthiest Americans from the wealthiest to the 400th wealthiest. The nine wealthiest were then ignored (Klass et al. 2006, p. 291). No explicit reason was given for why those nine were ignored, but it was perhaps because wealth did not appear to follow a Pareto distribution as closely among the nine wealthiest Americans as it did among the 391 Americans who were slightly less wealthy. As seen in figure 3.1 of this essay, although the empirical CCDF for the wealth of the 400 wealthiest Americans in 2003 looks somewhat like a straight line on a double logarithmic scale, it looks less like a straight line among the wealthiest of the 400 wealthiest.

Using the ranks and wealths of the 10th to 400th wealthiest Americans on the magazine's list in 2003, Klass et al. (2006) used a popular method for estimating the shape parameter of a Pareto distribution. They ran an ordinary least-squares regression of the logarithm of ranks against the logarithm of wealths, and they took the estimated slope parameter from that log-log rank-wealth regression as their estimate for the shape parameter of a Pareto distribution (Klass et al. 2006, p. 291). For more on that method and its popularity, see Gabaix and Ibragimov (2011).

In that way, Klass et al. (2006) estimated the shape parameter of a Pareto distribution from the distribution of wealth among some of the wealthiest Americans on the magazine's list in 2003. They reported that their point estimate for the shape parameter was about 1.22 (Klass et al. 2006, p. 291). We were able to replicate that result. The replication was aided by the corresponding author of Klass et al. (2006) sharing their dataset.

In addition to estimating the shape parameter of a Pareto distribution from the wealths of some of the wealthiest Americans on the magazine's list in 2003, Klass et al. (2006) also estimated the shape parameters of Pareto distributions from the distribution of wealth among some of the wealthiest Americans on the list in earlier

Table 3.1. Pareto Indexes for Some of the Wealthiest Americans, 1988–2003

Year	Pareto index
1988	1.60
1989	1.55
1990	1.54
1991	1.51
1992	1.50
1993	1.48
1994	1.44
1995	1.44
1996	1.49
1997	1.34
1998	1.29
1999	1.20
2000	1.13
2001	1.23
2002	1.20
2003	1.22

Source: Data adapted from Klass et al. (2006).

Note: This table shows, for the years 1988 to 2003, a point estimate for the shape parameter of a Pareto distribution fitted to the distribution of wealth among some of the wealthiest Americans in Klass et al.'s (2006) dataset in a given year. Following Klass et al. (2006), the estimates are based on log-log rank-wealth regressions for Americans who ranked between the 10th and 400th wealthiest in a given year.

years, specifically, each year between 1988 and 2003. For some of those years, they reported their point estimates for the shape parameters. They reported that, "The value of [the shape parameter of a Pareto distribution] gradually decrease[d] from about 1.6 in 1988, followed by a faster decline in the late 1990s down to about 1.1, after which it start[ed] to increase again" (Klass et al. 2006, p. 291).

We were able to replicate those results by using their dataset and their method for estimating the shape parameter of a Pareto distribution. Table 3.1 of this essay shows, for each year between 1988 and 2003, our point estimate for a Pareto index (i.e., the shape parameter of a Pareto distribution) estimated from the distribution of wealth among the 10th to 400th wealthiest Americans in their dataset in a given year.

As shown in the table, our estimate of the shape parameter of a Pareto distribution was about 1.6 in 1988, it decreased to about 1.1 by 2000, and it increased to about 1.22 by 2003.

3.3 Issues Identified During Replication

3.3.1 An Issue with Their Dataset

The dataset used by Klass et al. (2006) was constructed by the authors of that study and their research assistant from various issues of *Forbes Magazine* (Klass et al. 2006, p. 295). During the replication discussed in the previous section, we identified some relatively minor errors in their dataset. Those minor errors are discussed in a subsequent section. We also identified a more serious issue with their dataset. That issue is discussed in this section.

As noted in the introduction to this essay, in every year since 1982, *Forbes Magazine* has published a list of the 400 wealthiest Americans. The magazine's list in any given year is supposed to be a list of the 400 wealthiest *individuals* in America in that year. Although it is supposed to be a list of the 400 wealthiest individuals, the magazine's estimate of an individual's wealth can include the wealth of family members. According to the magazine, the wealth of an individual "generally" includes the wealth of his or her spouse, children, or other family members, "especially if family ties are manifestly close or they all share interests in an ongoing business" (*Forbes Magazine* 1982, p. 101). If "family or business ties" are "broken or at least noticeably frayed," however, then the wealth of an individual does not include the wealth of family members (ibid.). The wealth of an individual also excludes the wealth of a spouse, a child, or other family member, if the family member is wealthy enough to be one of the 400 wealthiest Americans themselves. In that case, the family member is listed as a distinct individual on the list.

In the years 1982 to 1999, alongside its list of the 400 wealthiest individuals in America, *Forbes Magazine* also published a list of some of the wealthiest families in America. Those families were so wealthy that many of the families were wealthier than many of the 400 wealthiest individuals (see below). Although the families were wealthy, the wealth of each family was so dispersed among the members of the family that none of the family members were wealthy enough to be one of the 400 wealthiest individuals. As the magazine said in one year, "Some of the largest fortunes in America are so divided among family members that no one individual qualifies for our rankings [of the 400 wealthiest individuals]" (*Forbes Magazine* 1999, p. 400). After the year 1999, the magazine stopped publishing a list of some of the wealthiest families alongside its list of the 400 wealthiest individuals. The magazine stopped publishing the list without explanation (at least to the best of our knowledge).

Forbes Magazine's list of the 400 wealthiest individuals in America is different than its discontinued list of some of the wealthiest families in America, therefore. Yet, in the dataset that they and their research assistant constructed, Klass et al. (2006) included the individuals on the magazine's list of the 400 wealthiest individuals as well as the families on its list of some of the wealthiest families. That seems like a mistake. It could be argued that it is difficult to draw a distinction between individuals and families. It could also be argued that the magazine does not draw a sharp enough distinction between individuals and families, given that the magazine's estimate of an individual's wealth can include the wealth of his or her family members. The magazine tried to draw such a distinction, however, so it seems like a mistake to ignore that distinction and conflate the two lists.[3]

---

[3]That mistake has been made by other studies, too. While it is unclear whether they constructed their own dataset or used Klass et al.'s (2006), Castaldi and Milakovic (2007) mistakenly claimed that "the size of the *Forbes* list [of the wealthiest individuals in America] was fixed at exactly 400 after the year 2000 and included more observations in the years before then" (p. 548).

Moreover, unlike its list of the 400 wealthiest individuals, the magazine's list of some of the wealthiest families was not intended to be a list of any particular number of the wealthiest families. As the magazine said in one year, its list of some of the wealthiest families was "extensive, but not intended to be as complete as our listing of the richest [i.e., wealthiest] individuals in America" (*Forbes Magazine* 1988, p. 285). The list was arguably extensive, therefore, but intentionally incomplete.

3.3.2 Other Issues with Their Dataset

As noted above, the dataset constructed by the authors of Klass et al. (2006) and their research assistant also contains some relatively minor errors. Those errors were identified by comparing their dataset to their data source. The errors are as follows, organized by year. All dollars values are reported in current dollars.

- Richard Alexander Manoogian was worth 625 million dollars in 1988, according to *Forbes Magazine* (1988, pp. 188, 340), not 885 million dollars, as in Klass et al.'s (2006) dataset.

- Klass et al.'s (2006) dataset includes the family of Charles E. Smith with a wealth of 290 million dollars in the year 1988, but Charles E. Smith and Robert H. Smith appeared on the magazine's list in that year, each with a wealth of 290 million dollars (*Forbes Magazine* 1988, pp. 258, 344–5).

- Henry Ross Perot Sr. was worth was 2,500 million dollars in 1989 (*Forbes Magazine* 1989, pp. 156, 352), not 500 million dollars, as in Klass et al.'s (2006) dataset.

- Klass et al.'s (2006) dataset does not include Shelby Cullom Davis for the year 1993, but he appeared on the magazine's list in that year with a wealth of 800 million dollars (*Forbes Magazine* 1993, p. 180).

- Klass et al.'s (2006) dataset includes Roy Michael Huffington with a wealth of 400 million dollars in the year 1994, but he was not wealthy enough to appear on the magazine's list in that year (*Forbes Magazine* 1994, pp. 331, 310).

- Leonard Samuel Skaggs Jr. was worth 950 million dollars in 1996, according to *Forbes Magazine* (1996, pp. 204, 354), not 945 million dollars, as in Klass et al.'s (2006) dataset.

- Klass et al.'s (2006) dataset includes a Frank Batten with a wealth of 1.6 billion dollars in the year 1999, but there was a Frank Batten Sr. who was worth 2.1 billion dollars and a Frank Batten Jr. who was worth 1.1 billion dollars on the magazine's list in that year (*Forbes Magazine* 1999, p. 242). The authors of Klass et al. (2006) and/or their research assistant apparently took the average wealth of the junior and senior Frank Batten and attributed that average to one individual.

Although studies like Brzezinski (2012) that have used Klass et al.'s (2006) dataset should probably still be aware of the errors noted above, this essay will not emphasize those errors because they do not seem to have either a statistically or substantively significant effect on the results reported by Klass et al. (2006). Table 3.2 shows, for each year in which errors were identified, the effect of the errors on the estimated Pareto index. Specifically, the table shows: the Pareto index estimated from a dataset with the errors; the Pareto index estimated from a dataset without the errors; and the difference between those two estimates, where the standard error of the difference was simply calculated as the squared sum of the variances of the estimated Pareto indexes (cf. Gelman and Stern 2006, p. 328). As seen in the table, none of the differences are significantly significant at conventional levels, nor would they be for almost any way in which their standard errors might be calculated, given that the differences are so small.

Table 3.2. Effect of Errors in Klass et al.'s (2006) Dataset on the Pareto Index for Some of the Wealthiest Americans, 1988–2003

	Pareto index		
Year	With errors	Without errors	Difference
1988	1.60 (0.11)	1.60 (0.11)	−0.00 (0.16)
1989	1.55 (0.11)	1.54 (0.11)	0.01 (0.16)
1990			
1991			
1992			
1993	1.48 (0.11)	1.49 (0.11)	−0.00 (0.15)
1994	1.44 (0.10)	1.44 (0.10)	0.00 (0.15)
1995			
1996	1.49 (0.11)	1.49 (0.11)	0.00 (0.15)
1997			
1998			
1999	1.20 (0.09)	1.21 (0.09)	−0.00 (0.12)
2000			
2001			
2002			
2003			

Sources: Data adapted from *Forbes Magazine* (1988–2003); Klass et al. (2006).

Note: This table shows, for years in which errors were identified in Klass et al.'s (2006) dataset, a Pareto index estimated from their dataset with the errors, a Pareto index estimated from a dataset without the errors, and the difference between the two estimated indexes. Following Klass et al. (2006), the point estimates for the Pareto indexes are based on log-log rank-wealth regressions for Americans who ranked between the 10th and 400th wealthiest in their dataset in a given year. Standard errors are reported in parentheses. The standard error of an estimated index was calculated in the manner suggested by Gabaix and Ibragimov (2011), while the standard error of the difference between two indexes was calculated as the squared sum of the variances of each index. None of the differences are statistically significant at conventional levels.

3.3.3 An Issue with Ignoring Some of the Wealthiest Americans

The other issues identified during the replication discussed above relate to the method used by Klass et al. (2006) to estimate the shape parameter of a Pareto distribution. Recall that, for any given year, they estimated the shape parameter of a Pareto distribution from the wealths of individuals or families who ranked between the 10th and 400th wealthiest in their dataset. In doing so, they ignored many of the 400 wealthiest individuals, as detailed in table 3.3 of this essay. They ignored most of those individuals because their dataset included a mix of individuals and families, and many of the families were wealthier than many of the individuals.

Some of the 400 wealthiest individuals were ignored by Klass et al. (2006) because the individuals were among the nine wealthiest individuals or families in their dataset (again, as detailed in table 3.3). Even if their dataset had only included individuals, rather than a mix of individuals and families, Klass et al. (2006) would have presumably still ignored the nine wealthiest individuals. Indeed, in the years after the magazine discontinued its list of some of the wealthiest families, they ignored the nine wealthiest individuals in each year.

It seems like a mistake to ignore the nine wealthiest Americans. Although wealth may not appear to follow a Pareto distribution as closely among the nine wealthiest Americans as it does among slightly less-wealthy Americans, it would be peculiar, if wealth followed a Pareto distribution among the 10th to 400th wealthiest Americans (and perhaps even less-wealthy Americans; Klass et al. 2006, p. 294), but not the nine wealthiest. Moreover, it is not clear that the worst fit between the distribution of their wealth and a Pareto distribution occurs among the nine wealthiest Americans. Thus, to the extent that wealth is thought to actually follow a Pareto distribution among the 400 wealthiest Americans, it seems like a mistake to ignore an arguably *ad hoc* number of the wealthiest of the 400 wealthiest.

Table 3.3. Number of the 400 Wealthiest Individuals in America Ignored by Klass et al.'s (2006) Study, 1988–2003

	Ignored because	
Year	Rank $<$ 10	Rank $>$ 400
1988[a]	6	100
1989[a]	5	95
1990	6	100
1991	7	94
1992	7	100
1993[b]	6	97
1994	6	98
1995	6	99
1996	6	91
1997	7	50
1998	8	50
1999	9	39
2000	9	0
2001	9	0
2002	9	0
2003	9	0

Sources: Data adapted from *Forbes Magazine* (1988–2003); Klass et al. (2006).

Note: Some of the 400 wealthiest individuals on *Forbes Magazine*'s list in the years 1988 to 2003 were ignored by Klass et al.'s (2006) study on some of the wealthiest individuals and families in the same years. This table shows the number of individuals ignored by their study. Some of the 400 wealthiest individuals were ignored because they were one of the nine wealthiest individuals or families in Klass et al.'s (2006) dataset in a given year. Other individuals were ignored because they were not one of the 400 wealthiest individuals or families in their dataset in the given year.

[a]For the years 1988 and 1989, Klass et al. (2006) also ignored one of the 400 wealthiest individuals on *Forbes Magazine*'s list in those years (namely, Malcolm Stevenson Forbes) presumably because the magazine did not report an estimate of that individual's wealth in any of those years.

[b]For the year 1993, Klass et al. (2006) also ignored one of the 400 wealthiest individuals on the magazine's list in that year (namely, Shelby Cullom Davis) presumably because of an error.

3.3.4 An Issue with the Estimator

Another issue identified during the replication discussed above also relates to the method used by Klass et al. (2006) to estimate the shape parameter of a Pareto distribution. Recall that they took the estimated slope parameter from a log-log rank-wealth regression as their estimate for the shape parameter of a Pareto distribution. That is a popular method, but undeservedly so, given that its estimates are "strongly" biased in small samples (Gabaix and Ibragimov 2011, p. 30).[4]

Of course, the small-sample bias of that least-squares estimator might not be an issue for a sample of the size that was of interest to Klass et al. (2006) and that is of interest to this essay, specifically, a sample of 400 observations. Four-hundred observations might be a large enough number to largely eliminate any small-sample bias. Yet Monte Carlo simulations suggest that, for sample sizes that are on the same order of magnitude as those that are of interest to this essay (specifically, samples of 200 observations), and even for samples that are larger than those that are of interest to this essay (specifically, samples of 500 observations), the least-squares estimator is still biased to a statistically significant extent (Gabaix and Ibragimov 2011, pp. 27–28, tables 1–2).

Moreover, even if the least-squares estimator of the shape parameter of a Pareto distribution was unbiased, an estimate of lower-bound parameter of that distribution would still be needed in order to test whether a variable actually follows a Pareto distribution. For studies like Ogwang (2013) that have used the least-squares estimate of the shape parameter and then tried to test whether a variable follows a Pareto distribution, they have used the biased version of the maximum-likelihood estimator for the lower-bound parameter, which is simply the smallest observation

[4] The bias is downwards. Gabaix and Ibragimov (2011) offer two proofs, several simulations, and zero intuition for why the bias is in that direction, in particular. The author of this essay is unable to offer any intuition, either.

(Arnold 1983, p. 194). The smallest observation is an obvious choice for the lower bound, but it nevertheless seems inconsistent to use different types of estimators for different parameters of the same distribution.

A different estimator than the least-squares estimator for the shape parameter of a Pareto distribution and the same type of estimator for the lower-bound parameter should be used, therefore, although there are a number of alternative estimators from which to choose. Rahman and Pearson (2003) discuss seven different estimators, for example. This essay adopts the (unbiased modification of the) maximum-likelihood estimator (Arnold 1983, pp. 196–7). Monte Carlo simulations suggest, for sample sizes on the same order of magnitude as those of interest to this essay or even smaller (specifically, samples as large as 100 and even as small as 10 observations), the maximum-likelihood estimator outperforms a number of other estimators in a number of areas (Rahman and Pearson 2003, p. 305, table 2).[5]

It should be noted that the a least-squares estimator for the shape parameter of a Pareto distribution might be more robust to deviations from a Pareto distribution than the maximum-likelihood estimator (in whatever way robustness might be measured), although it can also be noted that other estimators might be even more robust than either of those estimators (see, for example, Brazauskas and Serfling 2000). Yet, even if the maximum-likelihood estimator is less robust, if a more robust estimator is seen as somehow necessary due to deviations between the distribution of a variable and a Pareto distribution, then one should presumably try to test whether that variable actually follows a Pareto distribution.

[5]The performance of the (unbiased modification of the) maximum-likelihood estimator is especially good in terms of minimizing the differences between the true and fitted cumulative distributions, for example, at least relative to the performance of the other estimators considered by Rahman and Pearson (2003). For more details on the maximum-likelihood estimators, see section B.4.1 of this essay's appendix.

3.4 Extensions of Previous Work

3.4.1 Variation over Time in the Pareto Index

This essay now turns to extending the study by Klass et al. (2006). Like their study, this essay uses *Forbes Magazine*'s lists of the 400 wealthiest Americans, but this essay uses older lists from as far back as 1982 and more-recent lists from as recently as 2013. Aside from using older and newer data, this essay extends their study by critically examining the extent to which wealth follows a Pareto distribution among the 400 wealthiest Americans. The essay examines the statistical and substantive significance of variations over time in the shape parameter of a Pareto distribution fit to the distribution of wealth among the 400 wealthiest Americans on the magazine's list, as well as the statistical and substantive significance of the deviations from a Pareto distribution at each point in time.

The variation over time in the shape parameter of a Pareto distribution fit to the distribution of wealth among the 400 wealthiest Americans on the magazine's list is shown in figure 3.2 of this essay. As shown in the figure, the estimated shape parameter varied over time. That variation (or, really, the variation in the shape parameter of a Pareto distribution fit to the distribution of wealth among the 10th to 400th wealthiest individuals or families in their dataset) was recognized by Klass et al. (2006), but they did not examine the statistical or substantive significance of the variation. The estimated shape parameter was as large as 1.84 in the year 1984 and as small as about 1.06 in the year 2012 (with standard errors of about 0.09 and 0.05, respectively). The difference between those two estimates is statistically significant at less than the one percent level, at least when the standard error of the difference is calculated as the sum of squared variances of each estimated shape parameter, although the difference would be statistically significant at that level for almost any way in which its standard error might be calculated.

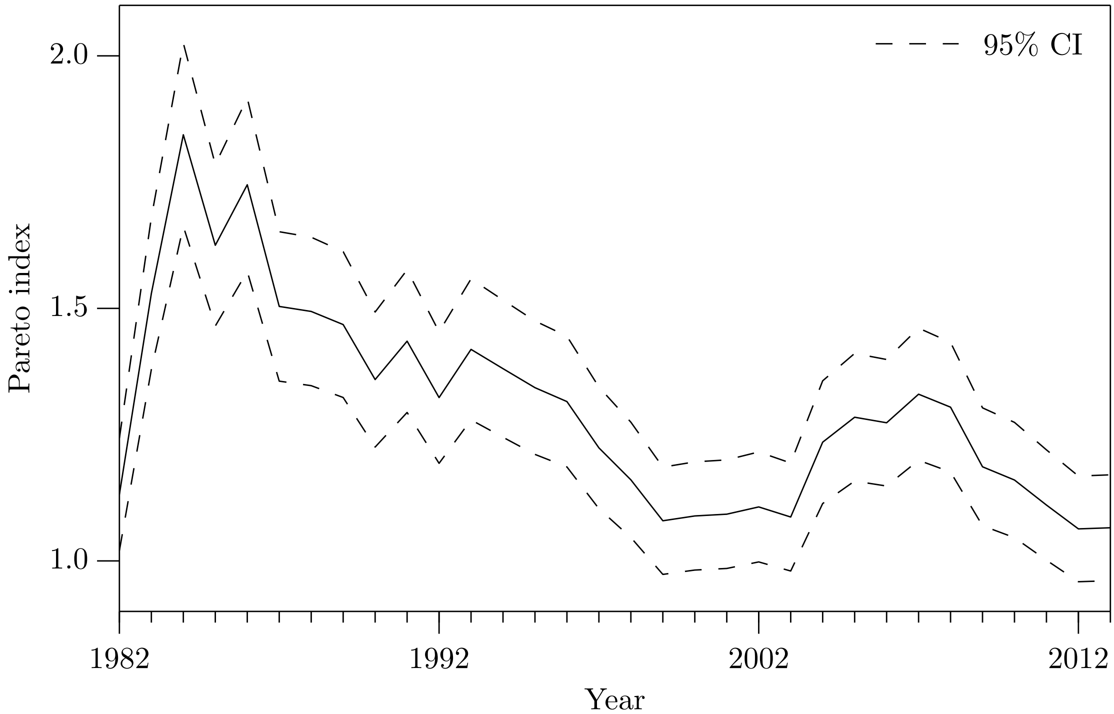

Figure 3.2. Pareto Index for the 400 Wealthiest Americans, 1982–2013

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows, for the years 1982 to 2013, the maximum-likelihood estimate for the shape parameter of a Pareto distribution fit to the distribution of wealth among the 400 wealthiest Americans on *Forbes Magazine*'s list in a given year.

The difference also seems to be substantively significant. Following the interpretation of Hardy (2010), a shape parameter of 1.84 would imply that the wealthiest 37 percent of the 400 wealthiest Americans should hold about 63 percent of the total wealth of the 400 wealthiest Americans. A shape parameter of 1.06, on the other hand, would imply that the same 37 percent should hold over 90 percent (about 94 percent) of the total wealth of the 400 wealthiest Americans. Thus, if wealth actually followed a Pareto distribution among the 400 wealthiest Americans on the magazine's list in those years, then the shape parameter of that distribution seems to have varied over time to a statistically and substantively significant extent.

3.4.2 Deviations from a Pareto Distribution

A standard test. After fitting a Pareto distribution to the distribution of wealth among the 400 wealthiest Americans on *Forbes Magazine*'s list in any given year, the goodness of that fit should be tested. A standard goodness-of-fit test is the Kolmogorov-Smirnov test (Hollander and Wolfe 1999, pp. 526–35). That test was applied by Clauset et al. (2009) in the same context and similar studies in similar contexts (see, for example, Ogwang 2013). In the current context, the intuition behind the test is that, if wealth actually followed a Pareto distribution among the 400 wealthiest Americans on the magazine's list in a given year, then there is only a small chance that there would be a large difference between the empirical CCDF for their wealth, on the one hand, and the CCDF for a Pareto distribution that was fit to the distribution of their wealth, on the other hand. The test statistic is simply the largest absolute difference between the empirical and fitted CCDFs (or, equivalently, the largest absolute difference between the empirical and fitted cumulative distribution functions). If the largest difference is larger than a critical value, then the null hypothesis that wealth follows a Pareto distribution is rejected. That test can be illustrated by applying it to the magazine's list from the year 2003.

Along with the empirical CCDF for the wealth of the 400 wealthiest Americans on *Forbes Magazine*'s list in 2003, figure 3.1 shows the CCDF for a Pareto distribution that was fit to the distribution of their wealth by maximum likelihood. The figure also shows the largest absolute difference between the empirical and fitted CCDFs. The largest absolute difference between those two CCDFs is about 11 percent. That difference occurs once at a wealth of about 900 million current dollars, as shown in the figure. The fitted CCDF implies that about 64 percent of the 400 wealthiest Americans should have been worth at least 900 million dollars, but the empirical CCDF implies that about 76 percent of them were worth at least that much. Whether that 11-percent difference between the empirical and fitted CCDFs is larger than a critical value or not is the critical question for the goodness-of-fit test. The critical value for the test can be found in the following way.

If we were trying to test whether their wealth followed a Pareto distribution with a lower-bound and shape parameter that had been specified independently of the data, then, in order to calculate the critical value for the test, we could perform a large number of Monte Carlo simulations like the following.[6] We could perform simulations in which, for each simulation, we calculate the largest absolute difference between the empirical CCDF for 400 random samples drawn from a Pareto distribution, on the one hand, and the CCDF for the hypothesized Pareto distribution, on the other hand. Those simulations could then be used to calculate the critical value for any desired level of statistical significance. If the desired level of statistical significance was one percent, for example, then the critical value would be the difference for which only one percent of the simulations had a difference that was at least as large as that difference. Of note, the critical values for such a test would be independent of the parameters of the Pareto distribution from which the random samples are drawn. That can be shown formally (see Lilliefors 1969, p. 387, and the reference therein) or through simulations (not shown). Also of note, the critical values would be independent of the parameters of the hypothesized Pareto distribution, again, assuming those parameters were specified independently of the data.

However, if the parameters of the hypothesized Pareto distribution are estimated from the data, rather than specified independently of the data, then critical

[6]Monte Carlo simulations are not actually necessary to calculate the critical value, if the parameters of the hypothesized distribution are specified independently of the data, and if the number of observations is large enough. The critical value should be approximately equal to $(\sqrt{-\ln\{\alpha/2\}/2})/\sqrt{n}$, where α denotes the desired level of statistical significance and n denotes the number of observations (see, for example, Miller 1956, p. 115, eq. 3). Monte Carlos simulations are necessary in other circumstances, however, as discussed in just a moment.

values based on the above-described simulations will generally be too large in the sense that, for any given level of statistical significance, they will generally fail to reject the null hypothesis frequently enough. They will generally be too large because, by fitting the hypothesized distribution to the data, the hypothesized distribution should tend to be a better fit to the empirical distribution than it would have been otherwise. To find the correct critical value, the Monte Carlo simulations described above should be performed in a slightly different way. We should perform simulations in which, for each simulation, we calculate the largest absolute difference between the empirical CCDF for 400 random samples drawn from a Pareto distribution, on the one hand, and the CCDF for a Pareto distribution whose parameters were estimated from the random samples, on the other hand. Like before, the critical values for such a test are independent of the parameters of the Pareto distribution from which the random samples are drawn.[7]

Accounting for the fact that the Pareto distribution in figure 3.1 was fit to the distribution of their wealth, a large number of Mote Carlo simulations (specifically, 10,000 simulations) like the above-described simulations suggest that the one-percent critical value would be about 6.5 percent. We would therefore expect to observe a difference at least as large as that difference with a probability of only one percent, if the wealth of the 400 wealthiest Americans on the magazine's list in 2003 actually followed a Pareto distribution. The 11-percent difference that was observed is obviously larger than 6.5 percent, so the null hypothesis that their wealth follows a Pareto distribution is rejected at the one-percent level of statistical significance.[8]

---

[7]It was therefore unnecessary when, in a similar context, Goldstein et al. (2004, p. 257) drew random samples from Pareto distributions with shape parameters that varied from one half to three. Any given shape parameter would have sufficed.

[8]If we follow Ogwang (2013) in a similar context and fail to account for the fact that the Pareto distribution was fit to the data, then the one-percent critical value would be erroneously calculated to be about 8.2 percent. As expected, that percentage is too large.

Accounting for rounding errors. The goodness-of-fit test applied above is a standard test, but it is not necessarily an appropriate test. The Kolmogorov-Smirnov goodness-of-fit test is only appropriate for continuous variables. Wealth is a continuous variable, so the test could be appropriate. Yet, the wealth estimates reported by *Forbes Magazine* are not reported down to fractions of pennies. In the year 2003, for example, it appears that the magazine rounded its estimate of someone's wealth to the nearest 100 million dollars, if he or she was a billionaire, and to the nearest five million dollars, otherwise. Such rounding can be incorporated into the simulations described above by simply rounding the random samples that we draw in the same way that the magazine appears to have rounded its wealth estimates.

Unfortunately, unlike before, the critical values for such a test are not independent of the parameters of the Pareto distribution from which the random samples are drawn, at least for a given rounding scheme. If the Pareto distribution is relatively equal, then it will be relatively more likely that more samples will be the same after rounding, and it will therefore be more likely that the largest absolute difference between the empirical and hypothesized distributions will be larger. An entirely different goodness-of-fit may be more appropriate, therefore. We leave that as a direction for future research, however, and we simply draw random samples from the Pareto distribution that was fit to the empirical data in order to get some sense of how rounding errors may affect conclusions about the goodness of the fit.

Accounting for rounding errors in that way, while also accounting for the fact that the Pareto distribution was fit to the data, the one-percent critical value would be estimated to be slightly larger than before. The one-percent critical value would be estimated to be about 7.7 percent. An 11-percent difference is still larger than that difference, so the null hypothesis is still rejected at the one-percent level of statistical significance.

Table 3.4 reports the results of applying that same test for each year between 1982 and 2013 while rounding the random samples in the same way the magazine seems to have rounded its wealth estimates in a given year. As shown in the table, for each year except two years (1994 and 2008), the test rejects the null hypothesis that wealth followed a Pareto distribution among the 400 wealthiest Americans on the magazine's list at the 10 percent level of statistical significance or less. The deviations from a Pareto distribution are statistically significant at conventional levels in 29 out of the 32 years, therefore.

In addition to being statistically significant at conventional levels, the deviations also seem to be substantively significant. For the year 2003, for example, the fitted CCDF implies that about 257 of the 400 wealthiest Americans (or, again, about 64 percent of them) should have been worth 900 million dollars, but there were almost 50 more Americans (or, again, about 11 percent more of the 400 wealthiest Americans) worth at least that much. There were 302 Americans (or, again, about 76 percent of the 400 wealthiest Americans) worth at least that much. A deviation of 50 multi-millionaires seems substantial. Similar deviations occur in the other years. Deviations from a Pareto distribution therefore seem to be statistically and substantively significant at almost every point in time.

It can be noted that Klass et al. (2006) did not draw much attention to deviations from a Pareto distribution, but Levy and Solomon (1997) did draw some attention to deviations. They suggested that the deviations might be due to rounding errors. "Most of the deviations from the theoretical fit occur at round values of wealth [which is] probably due to the rounding-off of estimated wealth values," they said (p. 92). However, the goodness-of-fit test we have applied made some attempt to account for rounding, and it still rejects a Pareto distribution at conventional levels of statistical significance.

Table 3.4. Goodness-of-fit Tests Against a Pareto Distribution, 1982–2013

Year	Test statistic	p-value accounting for: Rounding errors	p-value accounting for: Rounding and other measurement errors
1982	0.26	0.00	0.00
1983	0.10	0.00	0.00
1984	0.23	0.00	0.00
1985	0.11	0.00	0.00
1986	0.11	0.00	0.00
1987	0.10	0.00	0.00
1988	0.08	0.00	0.03
1989	0.08	0.00	0.00
1990	0.09	0.00	0.00
1991	0.06	0.06	0.18
1992	0.10	0.00	0.00
1993	0.09	0.00	0.00
1994	0.05	0.19	0.35
1995	0.06	0.04	0.04
1996	0.07	0.01	0.01
1997	0.07	0.01	0.00
1998	0.09	0.00	0.00
1999	0.07	0.06	0.04
2000	0.07	0.07	0.04
2001	0.10	0.00	0.00
2002	0.10	0.00	0.00
2003	0.11	0.00	0.00
2004	0.09	0.01	0.00
2005	0.11	0.00	0.00
2006	0.10	0.01	0.92
2007	0.09	0.01	0.80
2008	0.06	0.56	1.00
2009	0.10	0.00	0.00
2010	0.09	0.00	0.01
2011	0.11	0.00	0.00
2012	0.12	0.00	0.00
2013	0.12	0.00	0.00

Note: This table shows, for the years 1982 to 2013, the test statistic for a test of the goodness of the fit between the empirical CCDF for the wealth of the 400 wealthiest Americans in a given year and the CCDF for a Pareto distribution estimated from their wealths.

Accounting for other measurement errors. While some previous studies have suggested that rounding errors may explain deviations between the distribution of wealth among the 400 wealthiest Americans on *Forbes Magazine*'s list and a Pareto distribution, other studies seem to suggest that other measurement errors may also explain such deviations (Castaldi and Milakovic 2007, p. 544).[9]

To the extent that we know the ways (if any) in which *Forbes Magazine* misestimates what each of the 400 wealthiest Americans are worth, those measurement errors could perhaps be incorporated into the goodness-of-fit test discussed above in much the same way that rounding errors were incorporated into the test. Again, an entirely different goodness-of-fit test might be more appropriate, but, as part of the simulations, the random samples drawn from a Pareto distribution fit the empirical data could be subjected to such measurement errors.[10]

Aside from the way in which it appears to round its wealth estimates, we do not actually know the ways in which the magazine might misestimate wealth, but we can get some sense by making the following comparison between the magazine's list and another source of data on some of the wealthiest Americans.

Recently, the media corporation *Bloomberg* began publishing a daily list of some of the wealthiest people in the entire world. The list began as a list of the 20

[9]To quote Castaldi and Milakovic (2007) at length, they said the following in the context of discussing why wealth appeared to deviate from a Pareto distribution, especially among the least-wealthy of the 400 wealthiest Americans on *Forbes Magazine*'s list. They said, "One could speculate that measurement error is more pronounced for the lower part of the lists [i.e., among the least wealthy on the lists] because less effort might be devoted to the compilation of lower ranks and there is probably less publicly available information on lower-ranked agents" (p. 544).

[10]Except for Capehart (2014) who suggests exactly same thing in a slightly different context, we ares unaware of a goodness-of-fit test that accounts for measurement errors. If one is willing to assume that a variable follows a Pareto distribution, and if one is willing to assume that observations of the variable are subject to additive, normally distributed measurement errors, then Kondlo (2010) suggests how to fit the parameters of the Pareto distribution to the observations. Kondlo (2010) defends such assumptions on the basis that they are "common" (p. 50), but does not suggest how to test whether a variable actually follows a Pareto distribution, does not suggest whether measurement errors are actually additive or normally distributed, and thereby does not suggest whether it is actually appropriate for such assumptions to be so common.

wealthiest people in the world on March 5, 2012 (Miller and Newcomb 2012), and it has since expanded to a list of the 200 wealthiest. That list can be compared to *Forbes Magazine*'s annual list of the 400 wealthiest Americans for two days as of writing—August 24, 2012, and August 23, 2013.

Comparing those lists for the most-recent day on which they can be compared, all but one of the 64 Americans on *Bloomberg*'s list were also on *Forbes Magazine*'s list. The one exception was the grocery wholesaler Rick Cohen (on him and his wealth, see Coffey and Siraj 2013). *Forbes Magazine* apparently either failed to identify him or it estimated that he was worth much less than what *Bloomberg* estimated he was worth.

There were also 67 people on *Forbes Magazine*'s list who were wealthier by the magazine's account than the least-wealthy person on *Bloomberg*'s list but not on the latter list. One of those people was excluded from the list by design (namely, Michael Bloomberg), but *Bloomberg* must have either failed to identify the 66 other people on *Forbes Magazine*'s list or it must have estimated that those people were worth less than the least-wealthy person on its list.

Even for the 63 people who appeared on both lists, the lists disagreed about how much they were worth. Part of figure 3.3 shows kernel density estimates of the ratio between someone's wealth on August 23, 2013, according to *Bloomberg*'s list, and the same person's wealth on the same day, according to *Forbes Magazine*'s list. Only part of the range of ratios is shown in that figure. The ratios were as low as about 52 percent (Fidelty CEO Abigail Johnson was worth only 9.0 billion according to *Bloomberg* but 17.2 billion dollars according to *Forbes Magazine*) and as high as about 188 percent (Koch Industries shareholder Elaine Marshall was worth 15.6 billion according to *Bloomberg* but only 8.3 billion according to *Forbes Magazine*; on her and her wealth, see de Jong 2012).

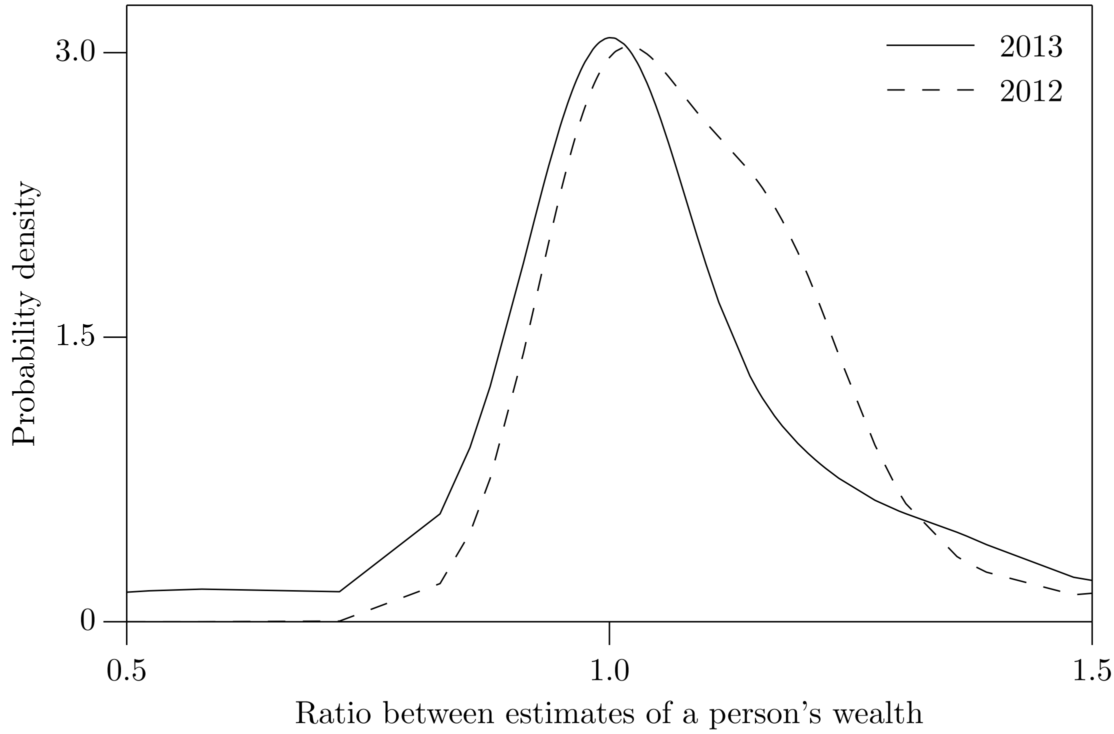

Figure 3.3. Ratios Between Wealth Estimates from *Bloomberg*'s and *Forbes Magazine*'s Lists, 2012 and 2013

Source: Data adapted from Bloomberg Billionaires Index (August 24, 2012; August 23, 2013); *Forbes Magazine* (2012, 2013).

Note: This figure shows, for the two days on which the lists can be compared in 2012 and 2013, and for people on both lists, kernel density estimates of the ratio between a person's wealth according to *Bloomberg*'s list of some of the world's wealthiest people and the same person's wealth according to *Forbes Magazine*'s list of the 400 wealthiest Americans. Normal kernels with bandwidths equal to the Silverman plug-in estimate were used. People who were not on both of the lists on a given day were ignored.

For the same reason that repeated, independent measurements of the same thing can reveal the imprecision of an instrument, the differences between *Forbes Magazine*'s and *Bloomberg*'s estimates of the wealth of the same people on the same day would seem to reveal the sorts of errors associated with trying to estimate the wealth of one of the wealthiest Americans. As such, those differences would seem to be our best estimate of any measurement errors associated with *Forbes Magazine*'s list (or *Bloomberg*'s list for that matter).

In order to account for such errors as part of the simulations discussed above, we could draw 400 samples from a Pareto distribution that was fit to the empirical data, mismeasure those samples by multiplying them by draws from one of the sets of kernel density estimates, and then do everything that we did before. Yet at least one problem with that approach is that we would essentially eliminate the possibility that someone who was not one of the 400 wealthiest Americans made it onto *Forbes Magazine*'s list because his or her wealth was overestimated.

So, instead, we will draw a larger number of samples from a Pareto distribution that extends over a wider range than the Pareto distribution that was fit to the empirical data, mismeasure those samples, and then take only the 400 largest. For simplicity, we will draw an order of magnitude more samples than 400 (i.e., 4,000 samples) from a Pareto distribution with a lower-bound parameter that is an order of magnitude lower than the lower-bound parameter of the Pareto distribution that was fit to the empirical data (about 60 million rather than about 600 million dollars in 2003, for example), and we will drawn from the most-recent set of kernel density estimates. Note that we are assuming that measurement errors are constant across people and time, but we have little evidence to suggest otherwise.

Part of table 3.4 reports the results of a goodness-of-fit test based on such simulations. As shown in the table, after accounting for rounding and other measurement errors in the manner discussed above, for 27 out of the 32 years, the test rejects the null hypothesis that wealth followed a Pareto distribution among the 400 wealthiest Americans on *Forbes Magazine*'s list at the 10 percent level of statistical significance or less. If an even more severe form of measurement errors was assumed, then the deviations between the distribution of wealth and a Pareto distribution might become insignificant, but there is no evidence to suggest assuming a more severe form of measurement error would be a reasonable assumption.

3.4.3 Other Distributions besides an Untruncated Pareto

The notion that wealth might follow a distribution as simple as a Pareto distribution is somewhat captivating (Schumpeter 1949, p. 156), so deviations from a Pareto distribution could perhaps be dismissed or deemphasized if similarly simple distributions cannot fit the distribution of wealth any better. The extent to which other distributions are a better or worse fit to the distribution of wealth among the 400 wealthiest Americans on *Forbes Magazine*'s list should therefore perhaps be explored. This essay will now turn to a brief exploration of a few alternative distributions, including the two-parameter versions of the gamma and log-normal distributions, as well as a slightly more general version of the Pareto distribution that only involves one additional parameter.

Gamma and log-normal. When its CCDF is graphed on a double logarithmic scale, the Pareto distributions looks like a straight line, but other distributions can look like straight lines, too, at least over certain ranges. The upper tails of the gamma and log-normal distributions can look like straight lines over certain ranges, for example. The two-parameter versions of those distributions (Cohen 1991, pp. 97, 114) are also as parsimonious as the Pareto distribution in terms of their number of parameters. Only looking at the upper tails of those distributions requires another parameter, a lower-truncation point, but the truncation point can be specified based on the wealth of the least-wealthy individual on the magazine's list in a given year and the way in which the magazine appears to round its wealth estimates.[11]

The lower-truncated (or even untruncated) gamma and log-normal distributions are not nested in the Pareto distribution or vice versa, so a conventional

---

[11] For the year 2003, for example, the truncation point can be specified as 597.5 million dollars, given that the minimum wealth in that year was 600 million dollars, and given that wealths less than one billion dollars were apparently rounded to the nearest five million. Formal expressions for the untrucuated and lower-truncated versions of the gamma and log-normal distributions can be found in appendices B.4.3 and B.4.4.

likelihood-ratio test of nested models cannot be applied, but a Vuong (1989) likelihood-ratio test of non-nested model can be applied (Cameron and Trivedi 2005, pp. 280–4; Clauset et al. 2009, pp. 679–80). Unlike a likelihood-ratio test of nested models, in which a nested model is specified as the "true" model, the true model is unspecified for a likelihood-ratio test of non-nested models, and the test is simply a test of which specified model is closer to the true, unknown model.

The first column of table 3.5 shows, for each year between 1982 and 2013, the results of a likelihood-ratio test for a Pareto distribution against a gamma distribution. If the test statistic is statistically significantly greater than zero at conventional levels, then the test favors the Pareto distribution; the test favors the gamma, if the test statistic is significantly less than zero; and neither distribution is favored over the other, if the test statistic is not significantly different than zero. As shown in the table, the Pareto distribution is favored in 23 of the years, the gamma is favored in one year (1982), and neither distribution is favored in eight of the years. Thus the Pareto distribution is favored over the gamma in most of the years.

The results are more ambiguous when a Pareto distribution is compared to a log-normal distribution. The Pareto is favored in eight years, the log-normal in four, and neither distribution is favored in 20 of the years (table 3.5, middle column). Neither distribution is favored over the other in most of the years, therefore.

Upper-truncated Pareto. The final alternative distribution this essay will explore is an upper-truncated Pareto distribution (Aban et al. 2006; Zhang 2013). That version of a Pareto distribution has one additional parameter, an upper-truncation parameter. If the truncation parameter is finite, then the truncated Pareto distribution looks like an untruncated one over most of its range, but it decays more rapidly in the upper tail. If the truncation parameter is infinite, on the other hand, then the truncated Pareto distribution is identical to the untruncated one. To the extent that wealth does not seem to follow an untruncated Pareto distribution as closely among the wealthiest (say, the nine wealthiest) of the 400 wealthiest Americans, the upper-truncated version of that distribution may offer a better fit.[12]

Given that the truncated Pareto distribution nests the untruncated one, a conventional likelihood-ratio test can be applied. The last column of table 3.5 shows the results of a likelihood-ratio test for each year between 1982 and 2013. For each year, except one year (1984), the test rejects the untruncated Pareto at the 10 percent level of statistical significance or less. An untruncated Pareto distribution was therefore almost always rejected in favor of an upper-truncated one.

Although there are an almost infinite number of other distributions that could be explored, the three distributions explored above reflect the range of ways in which other distributions might be a better or worse fit than the untruncated version of the Pareto distribution to the distribution of wealth among the 400 wealthiest Americans. The gamma seems to be a worse fit at almost every point in time; the log-normal seems to be neither better nor worse fit at most points in time; and the upper-truncated Pareto seems to be a better fit at almost every point in time.

Yet, even if another distribution is found to be a better fit than the untruncated version of the Pareto distribution (as appears to be the case with the upper-truncated version of that distribution), the extent to which it is actually a good fit should still be studied. For completeness, we will test the goodness of the fit of the upper-truncated version of the Pareto distribution before turning to some concluding comments about the extent to which the truncated version of that distribution is a good fit. That is, after all, the particular distribution that previous studies including Klass et al. (2006) have suggested is a good fit.

[12]For more on the upper-truncated Pareto distribution, see appendix B.4.2.

Table 3.5. Likelihood-ratio Tests, 1982–2013

Year	Distribution: Lower-truncated gamma	Lower-truncated log-normal	Upper-truncated Pareto
1982	-3.90^{***}	-6.32^{***}	24.01^{***}
1983	1.08	-0.85	11.18^{***}
1984	2.82^{***}	2.46^{**}	1.79
1985	1.28	-0.08	7.45^{***}
1986	3.04^{***}	2.01^{**}	2.97^{*}
1987	1.41	-0.72	3.49^{*}
1988	1.85^{*}	0.73	5.30^{**}
1989	0.96	-0.64	12.10^{***}
1990	-0.13	-0.68	14.21^{***}
1991	2.89^{***}	1.83^{*}	11.01^{***}
1992	2.02^{**}	0.84	13.87^{***}
1993	3.54^{***}	2.09^{**}	7.78^{***}
1994	3.23^{***}	1.90^{*}	7.89^{***}
1995	2.38^{**}	0.59	5.32^{**}
1996	1.56	0.75	5.75^{**}
1997	1.64	-1.00	3.65^{*}
1998	2.06^{**}	-1.08	3.27^{*}
1999	1.83^{*}	-2.10^{**}	4.13^{**}
2000	2.47^{**}	-0.96	6.63^{***}
2001	1.94^{*}	-0.43	6.26^{**}
2002	2.38^{**}	-0.96	6.91^{***}
2003	2.11^{**}	-3.02^{***}	7.77^{***}
2004	2.29^{**}	0.07	4.94^{**}
2005	2.28^{**}	0.27	4.69^{**}
2006	3.35^{***}	4.21^{***}	5.38^{**}
2007	3.40^{***}	8.85^{***}	5.28^{**}
2008	3.50^{***}	8.35^{***}	6.16^{**}
2009	1.68^{*}	-0.69	7.90^{***}
2010	2.06^{**}	0.87	8.56^{***}
2011	1.88^{*}	-0.22	10.14^{***}
2012	1.43	-1.82^{*}	11.59^{***}
2013	1.79^{*}	-0.75	12.62^{***}

Note: This table shows, for the years 1982 to 2013 and for different distributions, the test statistic for a likelihood-ratio test of a Pareto distribution against another distribution.

$^{*}p < .10$ $^{**}p < .05$ $^{***}p < 0.01$

3.4.4 Deviations from a Truncated Pareto Distribution, Too

Although the upper-truncated version of the Pareto distribution may be a better fit to the wealth of the 400 wealthiest Americans than the untruncated version of that distribution, the truncated version could still be a poor fit. The goodness-of-fit test that was applied before can be applied again in order to test the goodness of the fit between the empirical CCDF for the wealths of the 400 wealthiest Americans, on the one hand, and the CCDF for an upper-truncated Pareto distribution fit to the distribution of their wealth, on the other hand.

The goodness-of-fit tests discussed above—including the ones that account for rounding and other measurement errors, as well as the ones that only account for rounding errors—can be performed in exactly the same way as before, except that the upper-truncated version of the Pareto distribution should be used rather than the untruncated one.

The results of such goodness-of-fit tests are reported in table 3.6 of this essay. As seen in the table, in 28 out of the 32 years, the test that only accounts for rounding errors rejects, at the 10 percent level of statistical significance or less, the null hypothesis that wealth followed an upper-truncated Pareto distribution among the 400 wealthiest Americans on *Forbes Magazine*'s list. Similarly, the test that accounts for rounding and other measurement errors rejects that hypothesis in 25 out of the 32 years.

The substantive significance of the deviations from a truncated Pareto distribution are about the same as the deviations from an untruncated Pareto distribution discussed above. For the year 2003, again for example, the largest difference between the empirical and fitted CCDFs occurs at the same wealth as before (which is about 900 million dollars) and the difference is almost the same as before (a difference of about 39 multimillionaires).

Table 3.6. Same Tests Against a Truncated Pareto Distributions, 1982–2013

Year	Test statistic	p-value accounting for: Rounding errors	p-value accounting for: Rounding and other measurement errors
1982	0.24	0.00	0.00
1983	0.08	0.00	0.00
1984	0.23	0.00	0.00
1985	0.11	0.00	0.00
1986	0.12	0.00	0.00
1987	0.09	0.00	0.00
1988	0.07	0.00	0.04
1989	0.07	0.00	0.00
1990	0.08	0.00	0.00
1991	0.05	0.09	0.21
1992	0.09	0.00	0.00
1993	0.09	0.00	0.00
1994	0.04	0.51	0.66
1995	0.05	0.08	0.08
1996	0.06	0.01	0.02
1997	0.07	0.01	0.01
1998	0.08	0.00	0.00
1999	0.06	0.15	0.12
2000	0.06	0.22	0.19
2001	0.09	0.00	0.00
2002	0.10	0.00	0.00
2003	0.10	0.00	0.00
2004	0.08	0.03	0.00
2005	0.10	0.00	0.00
2006	0.09	0.02	0.77
2007	0.08	0.03	0.86
2008	0.06	0.47	1.00
2009	0.09	0.00	0.00
2010	0.08	0.00	0.01
2011	0.09	0.00	0.00
2012	0.10	0.00	0.00
2013	0.10	0.00	0.00

Note: This table shows, for the years 1982 to 2013, the test statistic for a test of the goodness of the fit between the empirical CCDF for the wealth of the 400 wealthiest Americans in a given year and the CCDF for a truncated Pareto distribution estimated from their wealths.

3.5 Conclusion

This essay was able to replicate the study by Klass et al. (2006), although the essay identified some issues with their study. The essay suggested that any future studies based on their data or methods should carefully consider whether individuals and families should be conflated, whether some of the wealthiest Americans should simply be ignored, and whether a strongly biased estimator should be used.

The essay also extended the study by Klass et al. (2006). The essay suggested that, to the extent that wealth actually follows a Pareto distribution among the 400 wealthiest Americans, the shape parameter of the Pareto distribution varies over time to a statistically and substantively significant extent. The essay also suggested that their wealth deviates from a Pareto distribution to a statistically and substantively significant extent at almost every point in time. The deviations would obviously be even worse (or, at least, not any better) if their wealth was thought to follow the exact same Pareto distribution at every point in time.

Deviations from a (time-varying or time-invariant) Pareto distribution could perhaps be dismissed by arguing that wealth truly follows a Pareto distribution, but *Forbes Magazine*'s lists are not accurate enough to reveal the true distribution. Some previous studies seem to have suggested as much (Castaldi and Milakovic 2007, p. 544; Levy and Solomon 1997, p. 92). Yet, this essay showed that, even after accounting for rounding errors and other possible measurement errors, the distribution of wealth among the 400 wealthiest Americans still seems to deviate from a Pareto distribution at almost every point in time.

Deviations from a Pareto distribution could also perhaps be dismissed by arguing that a Pareto distribution is nevertheless a close approximation to the upper tail of the distribution of wealth. Those who might make such an argument must be clear about what constitutes a close approximation. If the argument is that there are no distributions that are as simple as a Pareto distribution and that still offer as good a fit, then that argument would seem to be wrong, given that this essay identified at least one distribution that is similar in its simplicity and that offers a better fit (albeit still not an especially good fit).

Whether an approximation is adequate or not should ultimately depend on the purpose to which the approximation will be put, of course, yet empirical and theoretical studies that have assumed that wealth follows a Pareto distribution have not always seemed concerned with considering the adequacy of the approximation. Instead, they been content to appeal to previous studies that asserted it was adequate (see, for example, Kopczuk and Saez 2004b, p. 52). The results of this essay suggest that such studies should consider the adequacy of the approximation and not just appeal to previous studies that asserted that it was adequate.

A direction for future research is to consider whether previous empirical and theoretical studies that assumed that wealth follows a Pareto distribution would be affected if wealth did not actually follow that distribution in the real world. In terms of previous empirical studies, the effects may matter. Even in terms of something as simple as extrapolating or interpolating the number of people with a given amount of wealth, assuming a Pareto distribution can lead to significant misestimation, as the example in this essay showed.

In terms of previous theoretical studies that assumed that wealth follows a Pareto distribution, it may or may not matter whether wealth actually follows that distribution in the real world. Models are abstractions from reality, after all, and they are sometimes disconnected from reality altogether. Yet, if any attempt is made to apply conclusions drawn from such models to the real world, then the extent to which those conclusions rely on the assumption of a Pareto distribution of wealth should be considered.

CHAPTER 4

ON THE DURATION OF THEIR WEALTH

4.1 Introduction

Great wealth can persist, both within and across generations. The wealth accumulated by John D. Rockefeller during his lifetime (1839–1937) seems to have persisted for generations, for example, if we look at popular magazines that have kept track of such things. Rockefeller was the wealthiest American in 1918, according to a list of the 30 wealthiest Americans that was published by *Forbes Magazine* in that year (*Forbes Magazine* 1918). According to similar lists published by another magazine, some of Rockefeller's family members—including his sole son, his oldest grandson, and other relatives—were some of the wealthiest Americans decades after his death (*Fortune Magazine* 1957, 1968). And when *Forbes Magazine* published its first annual list of the 400 wealthiest Americans in 1982, 14 of the 400 wealthiest Americans were members of the Rockefeller family (Bernstein and Swan 2007, p. 231). One of them—John D. Rockefeller's oldest grandson, David Rockefeller—has appeared on the list in every year since then (*Forbes Magazine* 1982–2013).

Although great wealth can be persistent and perhaps permanent, it can also be more ephemeral. As pointed out by Steve Forbes, the current editor of *Forbes Magazine*, "When [our list of the 400 wealthiest Americans] appeared in 1982, it was populated by Rockefellers [but now] there is only one Rockefeller—ninety-four-year-old David" (Forbes and Ames [2009] 2011, p. 115). The wealthy are not a "fixed aristocracy," Forbes inferred. The fact that the current editor of *Forbes Magazine* is the son of the former editor, who was the son of the founding editor, might lead one to infer otherwise, but his point—that the wealthiest people and families are not always exactly the same people or families—is well taken. That point has also been made by other members of the staff at *Forbes Magazine*, perhaps at the behest of its editor. "For proof that America lacks a permanent overclass, look no further than [our list of the 400 wealthiest Americans]," a staff member wrote in one year (Lewis 2002). The proof was that, out of the 400 people who were on the list in 1982, "only" 58 of them were still on the list 20 years later.

The staff at *Forbes Magazine* are not the only ones who have used the magazine's list of the 400 wealthiest Americans to study wealth mobility, but the list has only been used in relatively anecdotal ways, as this essay elaborates upon below. Other sources of data like panel surveys have been used in relatively sophisticated ways to study inter- or intra-generational mobility throughout parts of the wealth distribution, but *Forbes Magazine*'s list is the only source of data that can be used to study inter- and intra-generational mobility throughout the part of the wealth distribution where great wealth is at its greatest, as this essay also elaborates upon below. This essay therefore studies what the magazine's list suggests about wealth mobility among the wealthiest Americans.

The rest of this essay is organized as follows. The essay begins by discussing the magazine's list as a source of data on wealth mobility. The list is then studied for what it suggests about wealth mobility, especially for what it suggests about the persistence of great wealth on an intra-generational time scale. The last section summarizes the results and suggests some directions for future research.

4.2 A Magazine as a Data Source

A popular magazine may seem like a dubious source of data on mobility throughout the distribution of wealth, but *Forbes Magazine*'s list of the 400 wealthiest Americans has already been used to study wealth mobility, at least in somewhat anecdotal ways. For example: As part of a wide-ranging defense of inequality in America, an academic economist by the name of Young Back Choi used the magazine's list from select years (1983, 1995, and 1999) to argue that mobility is "high" among the 10 wealthiest Americans (Choi 2002, p. 129). The economist pointed out that, "In 1983, only 12 years before [the year 1995], five out of the 10 [wealthiest Americans in 1995] were not even ranked among the [wealthiest] four hundred. By 1999, only four years later, some of them were already elbowed out by newcomers" (ibid.). All 10 of them were actually still ranked among the 400 wealthiest Americans in 1999, but six of them were no longer ranked among the 10 wealthiest, so in that sense they were elbowed out (ibid., p. 130, table 2).[1]

Another economist who has used *Forbes Magazine*'s list to study wealth mobility is Arthur Kennickell of the Federal Reserve. Kennickell (2006) studied whether people on the magazine's list in a select year (2001) were on the list in other, select years (1989, 1992, 1995, and 1998). He found that, out of the 400 people who appeared on the magazine's list in 2001, only about three-quarters of them (305 people or about 76 percent of them) were on the list three years before in 1998, and less than half of them (170 people or about 43 percent of them) were on the list 12 years before in 1989 (Kennickell 2006, p. 21, table 2.2). He also found that less-wealthy people were less likely to remain on the list. Whereas 68 people in the top quartile of the magazine's list in 1989 were still somewhere on the list by 2001, only 29 people

---

[1]The same economist made similar points (and a few of the exact same points) in an earlier paper by using the magazine's list from similar years (1983, 1989, and 1995; Choi 1999, pp. 246–7).

in the bottom quartile of the list in 1989 were still on the list by 2001.[2]

Even if *Forbes Magazine*'s list is assumed to be an accurate account of the 400 wealthiest Americans, the magazine's list can only provide a limited amount of information about wealth mobility. The list can only provide information about mobility into, out of, and throughout the part of the distribution of wealth that corresponds to the 400 wealthiest Americans. Yet, the information that the list can provide is greater than the information contained in the anecdotes discussed above. The list can provide more information than information about whether the 400 wealthiest Americans have always been exactly the same people, or about how many of the 400 wealthiest Americans in 1982 were still one of the 400 wealthiest Americans 20 years later, or about how an arbitrary number of the 400 wealthiest Americans (the 10 wealthiest, 100 wealthiest, etc.) in an arbitrary year (1995, 2001, etc.) were faring an arbitrary number of years before or after (12 years before, four years after, etc.), again, assuming that the list is an accurate account.

Whether *Forbes Magazine*'s list is an accurate account of the 400 wealthiest Americans is unknown. Several years of the Survey of Consumer Finances (SCF) conducted by the Federal Reserve have turned up people (or, really, households) who were apparently wealthy enough to appear on *Forbes Magazine*'s list in the same year but did not appear (see Kennickell 2007, p. 2, and the codebooks of the surveys since 1998). Estate-tax records have also been filed for people who were apparently wealthy enough to appear on *Forbes Magazine*'s list but did not appear

[2] The author of this essay and Arthur Kennickell both counted 305 people on the list in both 1998 and 2001, but the author of this essay counted only 169 people on the list in both 1989 and 2001, whereas Kennickell counted 170 people. That discrepancy is almost surely due to the misidentification of unique individuals, either by Kennickell or by the author of this essay. (Both the author of this essay and Kennickell counted 68 people who were in the top quartile in 1989 and still on the list by 2001, as well as 29 people who were in the bottom quartile in 1989 and still on the list by 2001.) It can be noted that, in another paper, Kennickell pointed out that, out of the 400 people on the list in 2007, 61 of them dropped off the list by 2009 (Kennickell 2011, p. 15). By the author's count, there were actually 62 people who dropped off the list between 2007 and 2009.

while they were still alive (Raub et al. 2010, pp. 11–12). The magazine's list may not be a completely accurate account of the 400 wealthiest Americans, therefore, although it may still be an approximately accurate account.

Of note, if a study is sophisticated enough that it accounts for sampling errors, then, in order for the magazine's list to be a useful source of data on the wealthiest Americans, all that is required of the list is that it be a representative sample of the those Americans. Of course, to the extent that the list is random sample, it may not be a representative sample. Scholars have suggested that, even if two people were worth the same amount, *Forbes Magazine*'s list might be biased towards including someone who has his or her wealth concentrated in a single or small number of assets rather than someone who has his or her wealth diversified across many assets (Blitz and Siegfried 1992, p. 7; Raub et al. 2010, p. 14). Scholars have also suggested that (again, even if two people were worth the same amount) the list might be biased towards including someone who owns more assets and owes more debt rather than someone who owns fewer assets and owes less debt (Atkinson 2008, p. 70).

Despite any systematic or unsystematic biases, the magazine's list has been used to study wealth mobility because it is the only source of data that can be used to study mobility at the very top of the wealth distribution in the United States on a frequency that is faster than an inter-generational time scale. Other sources of data like surveys and estate-tax records cannot be used. Estate-tax records could be used to study mobility at the very top of the wealth distribution, if those records were ever made available to the public; but, even if they were made publicly available, the records could only be used to study inter-generational wealth mobility, given that a person has to die before he or she has to pay estate taxes.

Surveys can be used—and, indeed, have been used—to study intra- or inter-generational wealth mobility; but, in order to study mobility at the very top of the wealth distribution, a survey would obviously need to be a panel survey that captures the very top of the wealth distribution. Unfortunately, no such survey exists for the United States. The wealth survey conducted as part of the Panel Survey of Income Dynamics (PSID) and similar panel surveys of wealth have been used to study wealth mobility (Diaz-Gimenez et al. 2011; Keister 2005). Those surveys were not designed to capture the very top of the wealth distribution, however, so they almost always fail to do so (Juster et al. 1999). The surveys cannot be used to study mobility at the very top of the wealth distribution, therefore.

A survey that might seem like it could be used to study mobility at the very top of the wealth distribution is the SCF, given that the SCF is designed to try to capture the very top of the wealth distribution, and given that a panel survey has sometimes been conducted as part of that survey. The SCF is typically only a cross-sectional survey, but, recently, a panel survey was conducted to study changes to household wealth between the years 2007 and 2009. That survey was conducted because those years were thought to be atypical (Bricker et al. 2011). Whether changes to household wealth between those years were atypical or not is unclear, however, given that the SCF's most-recent panel survey before that was back in the 1980s (ibid., p. 3).[3] Even if the panel survey had been conducted on a more-frequent basis, the SCF is designed to exclude the (households of the) people who appear on *Forbes Magazine*'s list of the 400 wealthiest Americans (Kennickell 2006, p. 84), so, to the extent that the magazine's list captures the very top of the wealth distribution, the SCF cannot be used to study mobility throughout that part of the distribution. The magazine's list therefore seems to be the best source of data, if only because it is the only source of data, on mobility at the very top of the wealth distribution.

---

[3]Studies of the SCF's earlier (1983–1989) panel survey and studies of its most-recent (2007–2009) panel survey include Kennickell and Starr (1997) and Bricker et al. (2011), respectively.

Although *Forbes Magazine*'s list seems to be the best source of data source on mobility at the very top of the wealth distribution in the United States over recent decades, the list does not seem to be a particularly good source of data on intra- or even inter-generational mobility among less-wealthy Americans, even though some studies have tried to use the list to study mobility from other parts of the wealth distribution into the very top (Blitz and Siegfried, 1992; Broom and Shay, 2000). A relatively recent and representative example of such a study is Kaplan and Rauh (2013). That study calculated (among other things) the percentage of people on *Forbes Magazine*'s list in select years (1982, 1992, 2001, and 2011) who grew up in families that were "wealthy," families that had "some wealth," or families that had "little or no wealth." *Forbes Magazine* does not systematically report such information, so the study apparently used the *Who's Who in America* book series and "internet searches" in order to classify people into those three categories (Kaplan and Rauh 2013, p. 44). The study found that the percentage of people who grew up in a wealthy family has fallen since 1982, which suggests that both intra- and inter-generational wealth mobility may have risen since then (Kaplan and Rauh 2013, p. 46). Such a finding could be correct, of course, but the problem with that study and similar studies that try to trace the origin of great wealth is that they require finding other sources of information that can supplement the information contained in *Forbes Magazine*'s list. Thus, while there is an obvious interest in studying how people are able to appear on *Forbes Magazine*'s list in the first place, the magazine's list is much better suited towards studying what happens once they get there. This essay will therefore focus on how long someone who appears on the list is able to remain there—the duration of their wealth, so to speak. We will now briefly look at some basic statistics about the duration of their wealth before looking at factors associated with longer or shorter durations.

4.3 A Brief Look at the Duration of Wealth

In every year since 1982, *Forbes Magazine* has published a list of the 400 wealthiest Americans. The list was published as recently as 2013, as of writing. For each person on the magazine's list in any given year, an interested reader can read about who the person is and how much he or she is estimated to be worth. The people on the magazine's list have not been the same people every year, so, for any given year except the inaugural year, the interested reader can also read about who came onto the list and who dropped off of it since the previous year.

Between the magazine's inaugural list in 1982 and its most-recent list in 2013, only about 1,500 (specifically, 1,474) unique individuals appeared on the list. That number of people is small relative to the number of people who would have appeared, if there had been complete turnover every year. If there had been complete turnover every year—with everyone coming onto the magazine's list in one year, dropping off by the next year, and then never reappearing again—then 12,800 people would have appeared on the list between 1982 and 2013.

The number of unique individuals who have been one of the 400 wealthiest Americans, according to *Forbes Magazine*'s list, is also small relative to the number of unique individuals who have been one of the 400 highest-income Americans, according to the International Revenue Service (IRS), at least over comparable periods of time. According to the IRS, almost 4,000 (specifically, 3,869) people were among the 400 highest-income Americans in any year between 1992 and 2009 (IRS 2012, p. 13). Over those same years, only about 1,000 (specifically, 1,027) people appeared on the magazine's list of the 400 wealthiest Americans, by comparison.[4]

[4]For the information about the 400 highest-income Americans released by the IRS, an individual could be either a primary or secondary taxpayer on either an individual or a joint tax return. Individuals were identified by their Social Security numbers. The author of this essay has been unable to find any account of why information about the 400 highest-income Americans was

The number of people who have appeared on the magazine's list over the years has been relatively small, therefore, but there has been turnover. Between 1982 and 2013, out of the 400 people on the list in any given year, as many as 376 and as few as 320 people appeared again in the next year, as part of figure 4.1 shows. About 352 people appeared again on average. So, as few as 24, as many as 80, and about 48 people on average have dropped off the list between any one year and the next.

The people who have dropped off the magazine's list have done so for different reasons. The reasons are as follows. One reason is that people have died. Although great wealth may afford a degree of immortality, even the wealthiest Americans eventually die. Out of the 1,474 people who appeared on the magazine's list in any year between 1982 and 2013, there were 219 people who dropped off the list because they died. The number of people who dropped off the list per year due to death is shown as a part of figure 4.1 of this essay. As shown in that figure, about seven people died on average per year, as many as 13 people died in some years (1982, 1992, and 1994), and as few as two people died in one year (1999).

Another reason why people have dropped off the list is that they have remained among the living but they have become too poor to remain among the 400 wealthiest Americans (at least by the magazine's account). That is to say, they dropped off the list because of an absolute or relative decline in their wealth. The number of people who dropped off per year due to such a decline in their wealth is shown as another part of figure 4.1 of this essay. About 41 people dropped off for that reason on average per year, as many as 72 people dropped off in one year (1986), and as few as 20 people dropped off in another year (2002).

---

released, rather than information about another number of the highest-income Americans, except that the information was released "in response to requests" (IRS 2003, p. 7). Information on the 400 highest-income Americans may have been requested because *Forbes Magazine* publishes a list of that number of the wealthiest Americans, but that could also just be a coincidence.

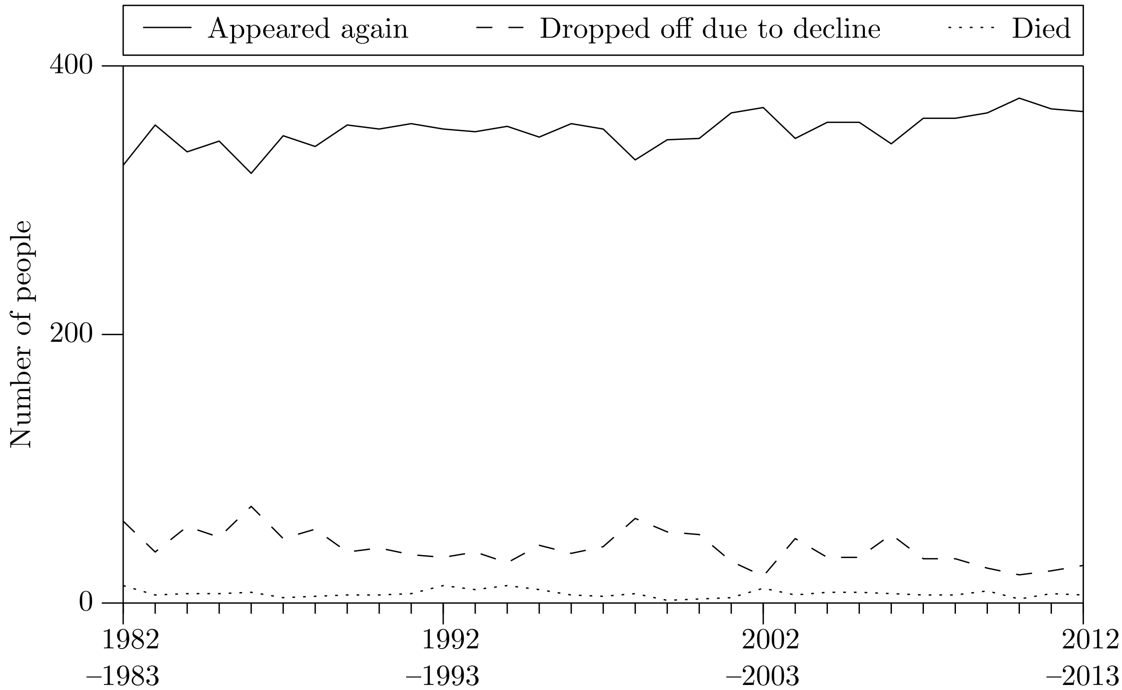

Figure 4.1. Number of People who Appeared Again or Dropped Off Per Year, 1982–2013

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows, for each year between 1982 and 2013, the number of people on *Forbes Magazine*'s list who appeared on the list again in the next year, dropped off due to an absolute or relative decline in their wealth, or dropped off due to death.

There is one other reason why people have dropped off of *Forbes Magazine*'s list of the 400 wealthiest Americans. A few people have dropped off the list because they have renounced their American citizenship. The Facebook co-founder Eduardo Saverin recently dropped off the list because he renounced his citizenship (*Forbes Magazine* 2012), but the Carnival Cruise Line founder Ted Arison (*Forbes Magazine* 1994), the Dart Container Company heir Robert Dart (*Forbes Magazine* 1995), and the Campbell Soup Company heir John Thompson Dorrance III (ibid.) also dropped off the list in earlier years because they renounced their citizenship.

Given that only four people have dropped off the list because they renounced their American citizenship, the rest of this essay largely ignores those people. The inferences that could be drawn from such a small number of people are severely limited. We could do little more than conclude that a person was more likely to drop off the list due to a renunciation of his or her American citizenship, if he or she was Edwardo Saverin or another one of the other people listed above. The fact that a few people dropped off the magazine's list because they renounced their citizenship is nevertheless of note.[5]

The 219 people who dropped off *Forbes Magazine*'s list due to death obviously never reappeared on the list again. They were dead. The four people who dropped off the magazine's list because they renounced their American citizenship also never reappeared again. None of them regained their American citizenship. Some of the people who dropped off the list due to an absolute or relative decline in their wealth did reappear again, however. Between 1982 and 2013, out of the 1,474 people who appeared at least once on the list, there were 344 people who reappeared on the list again after dropping off at least once before. Out of those people with several sets of consecutive appearances on the list, about half of them (177 people or about 51 percent of them) experienced a pattern wherein all of their sets of consecutive appearances on the list (of which the median number was two and the maximum number was four) only lasted one year.[6] Roughly a quarter of them (90 people or about 26 percent of them) experience a pattern wherein they appeared on the list twice in a row at least once, but, in between at least one of their sets of consecutive appearances, they were off the list for more than one year.[7] The remaining quarter or

[5]It is also perhaps of note that at least a few people who have appeared on the magazine's list were not born as American citizens. Edwardo Saverin only became an American citizen in 1998, for example. Almost all of the people who have appeared on the list were born as American citizens, however. Only 97 immigrants appeared in any year over the first 25 years of the magazine's list, for example, at least according to Bernstein and Swan (2007, pp. 34–37).

[6]That is to say, 177 people experienced an "occasional" pattern of appearing on the list, using the terminology used by Ashworth et al. (1994) in the context of repeated spells of poverty.

[7]I.e., 90 people experienced a "recurrent" pattern, using Ashworth et al.'s (1994) terminology.

so of them (77 people or about 22 percent of them) experience a pattern wherein they appeared on the list twice in a row at least once and they never disappeared from the list for more than one year in between their sets of consecutive appearances.[8]

Although some people reappeared on the list again after dropping off at least once before, most of the 1,474 people who appeared at least once on the list between 1982 and 2013 (specifically, 1,130 people or about 77 percent of them) came onto the list, appeared for some number of consecutive years, and then dropped off without ever reappearing again. Out of those people with a single set of consecutive appearances, a minority (254 people or about 22 percent of them) only appeared in one year, a majority (876 people or about 78 percent of them) appeared in more than one year, and a small fraction (25 people or about two percent of them) appeared in every single year since 1982 or 32 times in total.[9]

Some people have therefore appeared on *Forbes Magazine*'s list for longer than others. The probability that a person who appears on the magazine's list in a given year will appear on the list again in the next year (rather than drop off due to either death, decline, or a renunciation of his or her American citizenship) may depend on a number of factors. A host of other factors will be considered in the next section—including a person's age and their rank in the distribution of wealth—but one factor that may matter is the number of consecutive years that a person has already appeared on the list. There may be duration dependence such that, the longer a person has already appeared on the list, the more or less likely it is that he or she will appear again.

The longest time between any two sets of consecutive appearances was 23 years. The oilman Kenneth Stanley "Bud" Adams Jr. was off the list for that many years between his two sets of consecutive appearances in 1982 to 1985 and 2009 to 2012.

[8]I.e., 77 people experienced a "chronic" pattern, using Ashworth et al.'s (1994) terminology.

[9]I.e., 254, 876, and 25 people experienced "transient," "persistent," and "permanent" patterns of appearing on the magazine's list, respectively, using Ashworth et al.'s (1994) terminology.

A simple and perhaps overly simplistic way to estimate the probability that a person who has appeared on the magazine's list for a given number of consecutive years will appear again in the next year is as follows. That probability can be estimated as the total number of (possibly non-unique) people who appeared again after appearing for a given number of consecutive years, on the one hand, all over the total number of (again, possibly non-unique) people who either appeared again or dropped off after the given number of consecutive years, on the other hand. Such an estimate obviously assumes that everyone who appeared for a given number of years had the same probability of appearing again and, moreover, that the probability was equal to the proportion who appeared again rather than dropped off.[10]

Point estimates for those probabilities are shown as part of figure 4.2, although a point estimates for the probability of appearing again after either 30 or 31 consecutive years are not shown as part of that figure. At least for the data used by this essay, everyone who appeared on the list for either 30 or 31 consecutive years appeared again. The probability of appearing again after those many years would therefore simply be estimated to be 100 percent.

A point estimate for the probability of appearing again after 32 years is not shown as a part of that figure, either. Everyone who appeared on the magazine's list for 32 years was on the list in 2013 because they had appeared on the list in every year since 1982. Given that they were on the list in its most-recent year, it is unknown as of writing whether they will appear on the list again in the next year or not. Their sets of consecutive appearances are therefore said to be "right censored." The sets of consecutive appearances for everyone else on the list in its most-recent year are also right censored. Again, it is unknown whether they will appear again or not. Some people might appear again, but some might drop off due to a decline in

[10] A formal expression of the estimate is given as part of section C.2 of this essay's appendix.

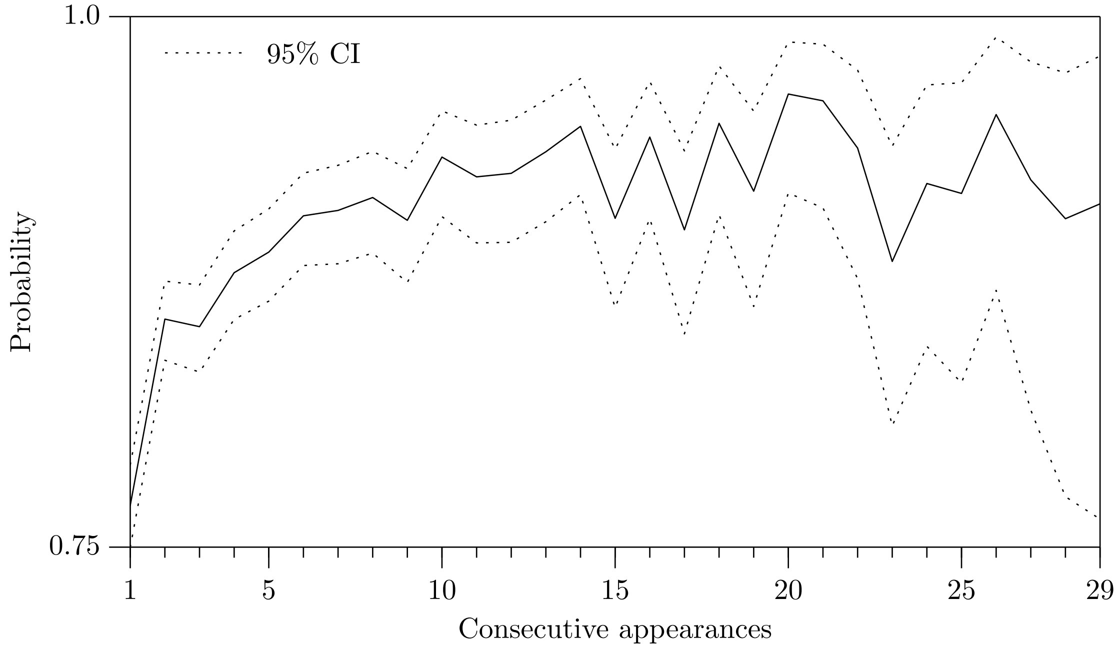

Figure 4.2. Probability of Appearing Again After a Given Number of Years

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows the estimated probability that a person who has appeared on *Forbes Magazine*'s list for a given number of consecutive years up to 29 years will appear again in the next year. A 95 percent confidence interval (CI) is also shown. Everyone who appeared for 30 or more consecutive years either appeared on the list again or was subject to right censoring.

their wealth, others might drop off due to death, and some might even drop off due to a renounced their American citizenship.

Ninety-five percent confidence intervals for each point estimate are also shown as part of figure 4.2 of this essay. At least for the data used by this essay, 95 percent Wald confidence intervals (i.e., confidence intervals based on the variance $p(1-p)/n$ where p is the estimated probability of appearing again and n is the number of people who either appeared again or dropped off) would include values greater than unity for some durations. The fact that that the Wald intervals for those years include values greater than unity is not surprising, given that the estimated probabilities in those years are close to unity. A value greater than unity is impossible for a proper probability, however, so we used another method (the Clopper-Pearson method) for calculating those intervals.[11]

Figure 4.2 suggests that, without controlling for any other factors, the longer someone has appeared on the list, the more likely it is that he or she will appear again, at least up until around 15 consecutive appearances. After that, the probability of appearing again is approximately constant, although there is greater uncertainty associated with those estimates, given that fewer people have appeared for that many consecutive years.

Using the estimates shown in figure 4.2, the probability that someone who appears at least once on the magazine's list will appear again and again for at least a given number of consecutive years can be estimated in a similarly simple way. That probability can be estimated as the product of the probabilities of appearing again after each year up to the given number of years. Figure 4.3 shows such estimates.[12]

As shown in that figure, the probability that someone who appears at least once on the list will appear at least once is 100 percent with perfect certainty. Most Americans will never appear on the magazine's list, of course, but, for those who do appear at least once, they will obviously appear at least once.

For the other probabilities, there is uncertainty. For those probabilities, figure 4.3 shows 95 percent confidence intervals along with each point estimate. The confidence intervals are based on the Greenwood formula for variance of the product of binomial proportions (Hollander and Wolfe 1999, pp. 541–42). The intervals for all the point estimates were relatively small at roughly one percentage point.

---

[11]Many methods for calculating such intervals have been suggested. Indeed, Pires and Amado (2008) compare 20 different methods. The method that we used is favored by Pires and Amado (2008, sec. 4.1) and Vos and Hudson (2005, sec. 5), at least in certain circumstances. That said, other methods yield similar intervals for our dataset. The Clopper-Pearson intervals are essentially the same as the Wald intervals, for example, except that there are no impossible values.

[12]Section C.2 of the appendix gives a formal expression for the estimates.

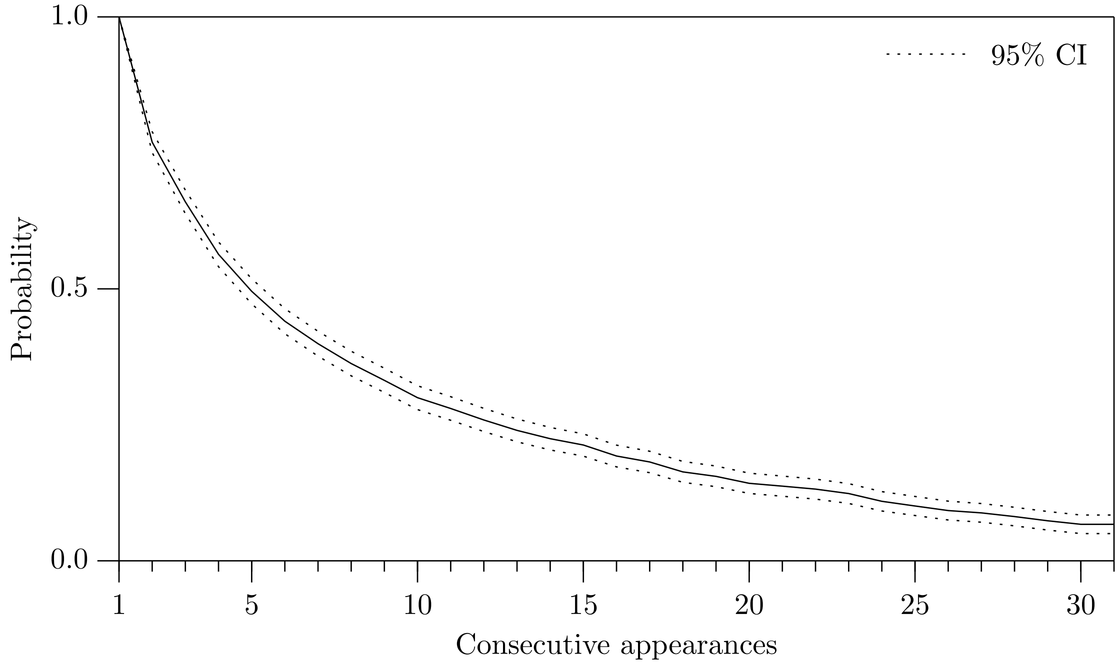

Figure 4.3. Probability of Appearing for at least a Given Number of Years

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows the estimated probability that a person who appears at least once on *Forbes Magazine*'s list would appear for at least a given number of consecutive years. A 95 percent confidence interval (CI) is also shown.

After appearing once, there are two extreme cases that could occur. In one extreme case, everyone who made it onto the list in a given year could appear again in the next year. In the other extreme case, everyone could disappear. The probability of appearing twice in a row was estimated to fall between those two extremes. The probability was estimated to be about 77 percent. That probability suggests that, if 400 people appeared on the list in a given year, then about 100 of them would drop off by the next year, while about 300 of them would appear again.

The probability of appearing at least five times in a row was estimated to be about 50 percent. So, again, if a group of 400 people appeared on the list in a given year, then only half of them would appear on the list at least five times in a row. Most people appear on the list for a relatively short amount of time, therefore.

Yet, the probability of appearing more than 10 times in a row was estimated to be about 30 percent. So, out of 400 people, about 120 would appear year after year for at least a decade. And the probability of appearing 30 times in a row was estimated to be about seven percent. So, out of 400 people, about 27 would appear year after year for at least three decades. Some people therefore stay on the list for a relatively long amount of time.

It can be noted that, if people who were on the list in 1982 were ignored—on the basis that their spells are "left-censored" because we do not know how long they might have already been the magazine's list if only it had made its list in earlier years—then the estimated probabilities of appearing on the list for a given number of consecutive years would be slightly different. The largest difference in the estimated probabilities would be associated with the probability of appearing at least three times in a row. The estimate would be about 2.7 percentage points lower. So, out of a group of 400 people, about 11 fewer people would be expected to appear year after year for three years. That difference is not statistically significant at the 10 percent level, however.

It can also be noted that another magazine published a list of the 66 wealthiest Americans back in the year 1968 (*Fortune Magazine* 1968) and a list of the 76 wealthiest Americans even further back in time in the year 1957 (*Fortune Magazine* 1957). One person, David Rockefeller, appeared on each of those lists in each of those years, as well as *Forbes Magazine*'s list in every year between 1982 and 2013. Thus, at least one person may have appeared on a list of the 400 wealthiest Americans in every year between 1957 and 2013, if such a list had existed over that entire period. As such, there may be at least a 0.25 percent chance that someone could appear on a list of the 400 wealthiest Americans year after year for 57 years or more. That chance is small, of course, but that period of time is also long. It is over half a century.

Figures 4.2 and 4.3 both suggest that the probability that a person who appears on *Forbes Magazine*'s list of the 400 wealthiest Americans will appear again may depend on the number of consecutive years that the person has already appeared. However, there may be other factors that affect the probability of appearing again, or, at least, there may be other factors that are associated with appearing again. The probability that a person will appear again is presumably associated with the person's age, for example. Younger people are presumably less likely to drop off the list due to death, given that they are presumably less likely to die, although being young and wealthy could be a lethal combination. Younger people may also be less likely to drop off of the list due to decline because they may be in an accumulation rather than a deaccumulation phase of life-cycle savings.

The probability that a person will appear again is also presumably associated with the person's rank in the distribution of wealth, as another example. Wealthier people are presumably less likely to drop off the list due to decline, given that they would need to experience a larger drop in their wealth in order to become the 401rst wealthiest American or worse. Wealthier people may also be less likely to die, given that there may be a social gradient to health even among the 400 wealthiest Americans, although more wealth may afford more ways to kill oneself (jet-setting around the world, flying into outer space, etc.).

Of course, young people become older over time. Wealthier people could also presumably become more or less wealthy over time. So, at least some of the factors that may be associated with appearing again may be time-varying factors. This essay therefore now turns to modeling, as a function of time-varying covariates, the number of consecutive years that a person who appears at least once on *Forbes Magazine*'s list of the 400 wealthiest Americans continues to appear.

4.4 Factors Associated with Appearing Again

4.4.1 Model

The number of consecutive years that a person appears on *Forbes Magazine*'s list can be modeled by using a discrete-time duration model with time-varying covariates and competing risks to appearing again. We assume that, for a person who appears on the list in any given year, there are $R = 2$ competing risks to appearing again. Let r index those risks. The risks are that he or she could drop off the list due to an absolute or relative decline in his or her wealth ($r = 1$) or drop off the list due to death ($r = 2$). For the reason noted above, we ignore the people who dropped off the list due to a renunciation of their American citizenship, and we therefore ignore the possibility of dropping off the list for that reason.

Suppose we observe N people who appear on the list for some number of consecutive years. Let $i = 1, 2, \ldots, N$ index those individuals. Some of the individuals may not be unique, given that some people reappeared on the list again after dropping off at least once before, but recall that most people (again, over three-quarters of the people who appeared at least once on the magazine's list) came onto the list, appeared for some number of consecutive years, and then dropped off without ever reappearing again.[13]

Let T denote a discrete random variable for the number of (uncensored but possibly unknown) consecutive years that a person appears on the list. Let $t = 1, 2, \ldots, T$ index those consecutive years. If a person came onto the list in one year but dropped off by the next year, then the person's realization of T would be one year, for example.

The conditional probability that the i-th person will drop off after his or her

[13]Modeling re-entry to the list is difficult, given that *Forbes Magazine* does not continue to cover people who drop off its list, except in cases when those people remain newsworthy.

t-th consecutive year on the list due to the r-th risk, conditional on the fact that he or she has appeared on the list for that many consecutive years, and conditional on any possibly time-varying covariates x_{it}, can be denoted by

$$P_{itr} \equiv P\left(T = t, R = r \mid T \geq t, x_{it}\right) \tag{4.1}$$

The probability of dropping off due to either death or decline can be denoted by $P_{it} \equiv P_{it1} + P_{it2}$. The probability of appearing on the list again in the next year is then the inverse of that probability (i.e., $1 - P_{it}$).

If it is assumed, for the moment, that the i-th person came onto the list, remained there for t consecutive years, and then dropped off due to either death or decline, then his or her contribution to the likelihood of the data would be

$$P_{itr} \prod_{k=1}^{t-1} \left(1 - P_{ik}\right) \tag{4.2}$$

which is the conditional probability that he or she drops off the list after his or her t-th year due to the r-th risk, conditional on the fact that he or she has appeared for that many years, multiplied by the probability that that he or she will appear for that many years.

For people who came onto the list, remained there for some number of years, and then dropped off due to either death or decline, the number of consecutive years that they appeared on the list is obviously known. However, for people who were still on the list in its most-recent year, we do not know whether they will appear again or drop off. We only know that they will appear for at least as many years as they have already appeared. Again, the observations associated with those people are said to be right-censored. Letting $d_{it} = 1$ if the i-th person was right-censored in his or her t-th year on the list, and letting $d_{it} = 0$ otherwise, the i-th person's contribution to the likelihood of the data would be

$$\left[P_{itr}\prod_{k=1}^{t-1}(1-P_{ik})\right]^{1-d_{it}}\left[\prod_{k=1}^{t}(1-P_{ik})\right]^{d_{it}} \tag{4.3}$$

which reduces to the previous equation, if his or her number of years on the list was not right-censored.

A convenient choice for the functional form of the conditional probabilities in the above-given equation given above is the logistic function

$$P_{itr} = \frac{\exp\{\beta_r x_{it}\}}{1+\exp\{\beta_r x_{it}\}} \tag{4.4}$$

where β_r are parameters that reflect the association between the covariates and the probability of dropping off due to the r-th competing risk. That functional form is convenient because then a discrete-time duration model of the number of years that a person appears on the magazine's list before dropping off due to either death or decline is identical to a multinomial logit model where each year that a person appears on the list is treated as a separate observation that could result in the person appearing again in the next year, dropping off due to decline, or dropping off due to death. That identity between a discrete-time duration model and a multinomial logit model was emphasized by Allison (1982, esp. p. 89), reemphasized by Jenkins (1995), and derived for the case of more than one competing risk by Lauer (2005, pp. 119–24). The identity has also been exploited by a range of studies on a range of topics (see Esteve-Perez et al. 2013 for a recent example).

As with any multinomial logit model, the relative probability of one alternative rather than another is a log-linear function of covariates. If the alternative of appearing again is taken as the base case (i.e., if the parameters associated with that outcome are normalized to zero for identification purposes), then the probability that the i-th person will drop off after his or her t-th year on the list due to the r-th risk (P_{itr}) relative to the probability that he or she will appear again in the next year

$(1 - P_{it})$ is a log-linear function of covariates such that

$$\ln \left\{ \frac{P_{itr}}{1 - P_{it}} \right\} = \beta_r x_{it} \tag{4.5}$$

The model that we will consider is therefore no different than a standard multinominal logit model. Such a model is perhaps the simplest model that allows for time-varying covariates to affect (or at least be associated with) the probability that someone will appear on *Forbes Magazine*'s list again, but even such a simple model has not been applied to the magazine's list before. Thus, while the model is simple, a study that relies on the model is still more sophisticated than previous studies. The model can be used to more carefully quantify associations that previous studies have only hinted at. Kennickell's (2006) finding that less-wealthy people seem to be less likely to remain on the list can be more carefully quantified, for example. The model can also hopefully be used to identify new associations, or, at least, it can be used to suggest associations that future studies could perhaps quantify more carefully with even more sophisticated models.

As a final note on our model, it can be noted that we have not explicitly modeled the competing risk of right censoring. We have not done so because, for a person on the list in a given year, whether the observation associated with that person in that year was right censored or not is completely determined by whether the year was the most-recent year or not. A dummy variable for whether someone was on the list in the most-recent year would therefore perfectly predict whether an observation was right censored. Some of the covariates we will consider are year-specific dummies, so will restrict our attention to the person-year observations associated with every year except the most-recent one, and we will thereby restrict our attention to the competing risk that cannot be completely explained by a single dummy.

4.4.2 Covariates

For our model of the probability that someone on the magazine's list in a given year will appear again, drop off due to a decline in wealth, or drop off due to death, the covariates we will consider include the person's age, his or her rank in the distribution of wealth, the calendar year, the number of consecutive years for which he or she has already appeared, and whether his or her number of consecutive years is left censored. Those covariates, a few of the other covariates we will consider, and some covariates we will not consider can be discussed in more detail as follows.

Age. One of the covariates we will consider is a person's age. Again, younger people are presumably less likely to drop off due to death. Younger people may also be less likely to drop off due to a decline in wealth. Out of all of the people who have appeared on *Forbes Magazine*'s list between 1982 and 2012, their minimum age was 22 years of age. Daniel Ziff was that young in 1994 when he came onto the list after inheriting part of his father's fortune (Wingfield 2007). He has been on the list ever since then. The maximum age of the people who appeared on the list was 99 years of age. The widow of an oil baron, Irene Wells Pennington, was that old in 1998. She dropped off the list in the next year due to a decline in her wealth. The median age of the people who appeared on the list was 64 years of age.

Rank. A person's rank in the distribution of wealth will also be considered as a covariate. Again, wealthier people are presumably less likely to drop off the list due to a decline in their wealth. They may also be less likely to drop off due to death. Out of the people who appeared on the magazine's list between 1982 and 2012, their ranks were obviously as high as the 1rst wealthiest. Their ranks were as low as the 399th wealthiest. The lowest rank was not the 400th wealthiest due to ties. Their median rank was the 194th wealthiest. The median rank was not the 200th wealthiest due to ties.

Year-specific dummies. A person's probability of appearing again after appearing in a given year could be associated with the given year in which he or she appeared. Some years may simply involve more or less turnover than others. A non-parametric approach to modeling any years-specific effects is to include dummy variables for each year as covariates. Dummies for each year (except an arbitrary year in order to avoid the well-known dummy variable trap) can be included because, after each year, some people appeared on the list again, some people dropped off due to a decline in wealth, and some people died.

Duration-specific dummies. A person's probability of appearing again could be associated with the number of years that the person has already appeared. A non-parametric approach to modeling any duration dependence would be to include dummy variables for each duration (again, except an arbitrary duration to avoid the dummy variable trap) as covariates. For the data used by this essay, duration dummies can be included for almost all of the durations, but not all of them. There would be a perfect prediction problem if dummy variables for durations of 30 or 31 years were included. Each person who appeared on the list for 30 or 31 years (of which there were 25 people) appeared again. None of them dropped off due to death or decline. There would also be a perfect prediction problem, if a dummy variable for a duration of 29 years was included. None of the people who appeared on the list for at least 29 consecutive years died by the next year. They all either appeared again or dropped off due to decline. In order to adopt a non-parametric approach to modeling any duration dependence while avoiding those perfect prediction problems, a dummy variable for a duration of 28 or more years can be included instead of separate dummies for 28 to 31 years. That approach seems preferable to adopting a parametric approach. The inferences that can be drawn about durations that are that long are also fairly limited anyway, as seen below.

Left-censored dummy. To the extent that the number of consecutive years that someone has appeared on the magazine's list affects the probability that he or she will appear again, a dummy variable for whether the person's number of consecutive years on the list is left censored or not should perhaps be considered as a covariate. If a person's number of consecutive years on the list are left censored, then he or she might have been on the list for many more consecutive years, if only the magazine had made its list in earlier years. We will consider such a left-censored dummy as a covariate.

Other covariates. Other covariates besides those discussed above could also be considered. Men may be more likely to die than women, other things being equal, for example. Unfortunately, *Forbes Magazine* has not systematically reported demographic information like the genders, races, or ethnicities of the people who have appeared on its list over the years. Such information could perhaps be found through other sources of data, but we leave that as a direction for future research, and do not include any demographic information as covariates.

It should also perhaps be noted that, even if such information was readily available, it would likely generate a number of perfect prediction problems for our model. According to Bernstein and Swan's (2007, p. 13) analysis of the first 25 years of *Forbes Magazine*'s list, out of the roughly 1,300 unique individuals who appeared on the magazine's list over those years, only about 200 were women, only about 30 were Asian Americans, only about 10 were Hispanic Americans, and only about 10 were African Americans with only one African American woman ever appearing (namely, Oprah Winfrey). That lack of diversity among the people on the magazine's list has fairly clear implications about who is more likely to move into the very top of the wealth distribution, but it could confound attempts to explain the probability of moving out of that part of the distribution.

While *Forbes Magazine* has not systematically reported demographic information about the people who have appeared on its list, the magazine has systematically reported other information, at least in the years since its list went online in 1996. For the people who have appeared on its list since that year, the magazine has reported the primary industry in which someone's fortune was made, as well as whether someone's fortune was self-made. We will initially ignore that information so that we can estimate the model across a larger number of years, but, after that, we will look at a smaller number of years in order to use that information.

Similarly, we will initially ignore the philanthropic activities of the people who have appeared on *Forbes Magazine*'s list so that we can estimate the model across a larger number of years, but we will consider those activities over a smaller number of year after that. In order to consider such activities, we must draw on other sources of data that are also limited to the years since 1996.

4.4.3 Results and Discussion

Table 4.1 reports the results of estimating the model discussed above with the covariates discussed above for people who appeared at least once on *Forbes Magazine*'s list in any year between the years 1982 and 2012. In total, there were a little less than 12,400 person-year observations (specifically, 12,379 observations) associated with 1,451 unique people. The remaining 21 observations were associated with the four people who dropped off the list because they renounced their American citizenship. One reason to ignore those observations is that any inferences drawn from such a small number of observations would be severely limited, as noted above. Another reason to ignore those observations, which was not noted above, is that they would create perfect prediction problems with some of the year and duration dummies, given that almost all of the people who appeared on the magazine's list never dropped off of it due to a renunciation of their citizenship.

Table 4.1. Multinomial Logit Model of the Competing Risks to Appearing Again

Covariates	Competing risks: Decline	Competing risks: Death
Constant	-3.58^{***}	-3.72^{*}
	(0.62)	(2.20)
Age	-0.04^{**}	-0.10^{*}
	(0.02)	(0.06)
Age squared	0.00^{**}	0.00^{***}
	(0.00)	(0.00)
Rank	0.01^{***}	0.00
	(0.00)	(0.00)
Left-censored dummy	-0.11	-0.02
	(0.12)	(0.21)
Year dummies	Yes	Yes
Duration dummies	Yes	Yes
Number of observations	12,379	
Percent correctly predicted	88.25	
Log-likelihood	$-3,965.22$	
Test for non-constant variables	$2,407.25^{***}$	
Test for age variables	321.09^{***}	
Test for rank variables	$1,416.29^{***}$	
Test for left-censored dummies	0.91	
Test for year dummies	181.79^{***}	
Test for duration dummies	69.54^{*}	

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows the maximum-likelihood estimates for the multinomial logit model described in the text, in which the unit of observation is a person on *Forbes Magazine*'s list in a given year who could either appear on the list again in the next year, drop off due to an absolute or relative decline in his or her wealth, or drop off due to death. Appearing again was taken as the base case. The covariates, the number of observations, the percentage of those observations that were correctly predicted, the likelihood-ratio tests, and the independence-of-irrelevant-alternatives tests are all described in the text.

$^{*}p < .10$ $^{**}p < .05$ $^{***}p < 0.01$

The results reported in that table are difficult to interpret in a direct manner, given that the model is a model of relative probabilities, and given that the relative probabilities are modeled as nonlinear functions of the covariates. The series of figures that follow aid in the interpretation. The figures that follow show the probability that a particular person would appear again, drop off due to decline, or drop off due to death. Unless otherwise stated, the person's age is assumed to be equal to the median age of 64 years of age. The person's rank is assumed to be equal to the median rank of the 194th wealthiest. The person is also be assumed to have come onto the list for the first time in 1996. Note that, if a person came onto the list in 1996, then he or she would obviously be on the list in that year, he or she would have only appeared on the list for one consecutive year, and his or her consecutive years on the list would not be left-censored (which would only be the case if he or she had been on the list in every year since 1982). This basis for comparison is somewhat arbitrary, but the results are similar with similar ages, ranks, years, and durations.

Ninety-five percent confidence intervals for the probabilities are shown as part of the figures that follow. Those intervals are Krinsky-Robb or "parametric bootstrap" intervals (Krinsky and Robb 1986). The intervals were constructed as follows. Given the point estimates for the parameters of the trinomial logit model, and given the estimated covariances for those parameters, the probabilities were repeatedly recalculated by repeatedly redrawing the parameters of the trinomial logit model from a multivariate normal distribution whose expected values were equal to the point estimates and whose covariances were equal to the estimated covariances. The parameters were redrawn and the probabilities were recalculated a large number of times (specifically, 10,000 times). The confidence intervals for the probabilities were then constructed by taking certain percentiles of those recalculated probabilities. Such intervals are computationally intensive to construct, but, unlike linearly approximated or "Delta-method" confidence intervals, the intervals cannot include impossible probabilities that are greater than unity or less than zero.

Effect of age. Holding a person's rank constant at the 194th wealthiest, holding his or her calendar year on the list constant at 1996, and holding his or her number of consecutive appearances on the list constant at one appearance, figure 4.4 shows the association between the person's age and his or her probability of appearing again, dropping off due to decline, or dropping off due to death. As seen in that figure, the probability that the person will appear again is approximately constant until about 70 years of age. After that, the probability of appearing again decreases at an accelerating rate as the person ages. The decrease in the probability of appearing again is almost entirely attributable to an increase in the probability of dropping off due to death. There is little, if any, association between the person's age and his or her probability of dropping off due to decline. Younger people are therefore more likely to remain on the list, but because they are less likely to die, and not because they are less likely to drop off due to decline.

Effect of rank. Holding the person's age constant at 64 years of age and holding the other covariates constant, figure 4.5 shows the association between the person's rank in the distribution of wealth and his or her probability of appearing again, dropping off due to decline, or dropping off due to death. As seen in that figure, if the person was the wealthiest person, then he or she would be expected to appear again with almost a 100 percent chance. If he or she was the least-wealthiest person among the 400 wealthiest Americans, on the other hand, then he or she would be expected to appear again with only about a 50 percent chance. That decrease in the probability of appearing again is almost entirely attributable to an increase in the probability of dropping off due to decline. There is therefore very little, if any, association between a person's wealth and his or her mortality rate.

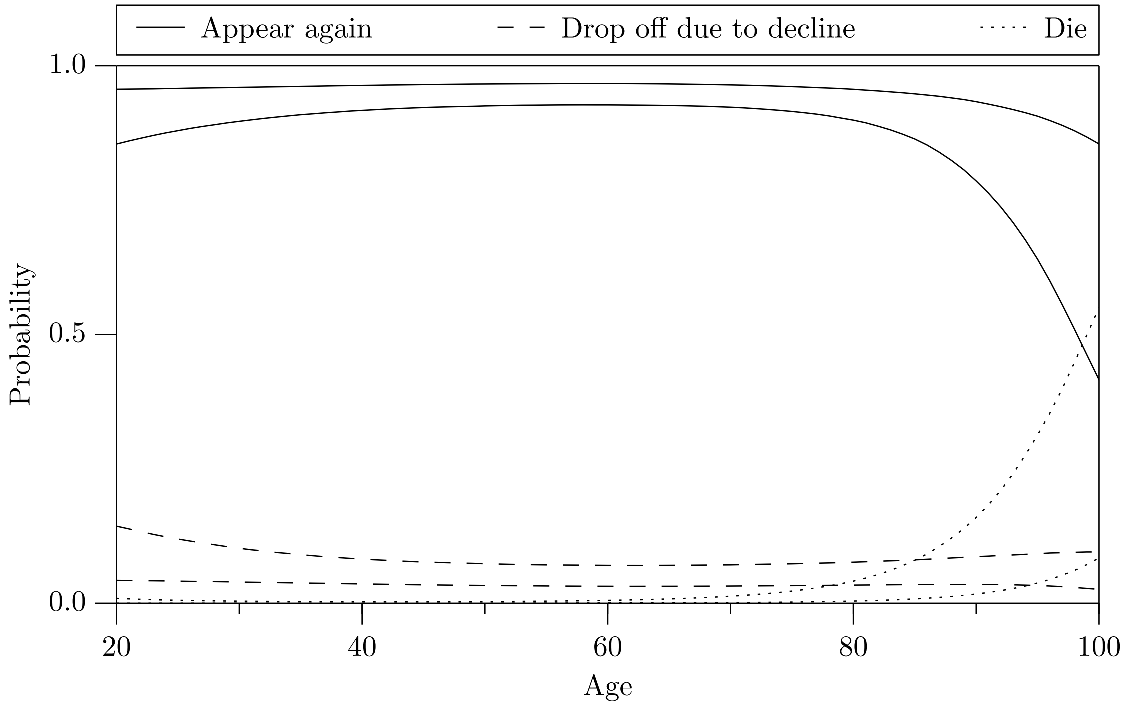

Figure 4.4. Effect of Age

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows 95 percent confidence intervals for the probability that a person on *Forbes Magazine*'s list in a given year would appear on the list again in the next year, drop off due to a decline in wealth, or drop off due to death as a function of the person's age in the given year.

Effect of the calendar year. Holding a person's age, rank, and consecutive appearances on the list constant, and also holding the left-censored dummy constant at zero for simplicity, figure 4.6 shows the association between the calendar year in which the person appears on the list and his or her probability of appearing again, dropping off due to decline, or dropping off due to death. The probability of appearing again is relatively high for each year, while the probabilities of dropping off due to death or decline are relatively low for each year, so the figure was split into two for the benefit of the reader. As seen in the figure, the probabilities of appearing again or dropping off due to either death or decline are approximately the same in each year, although there are some notable exceptions.

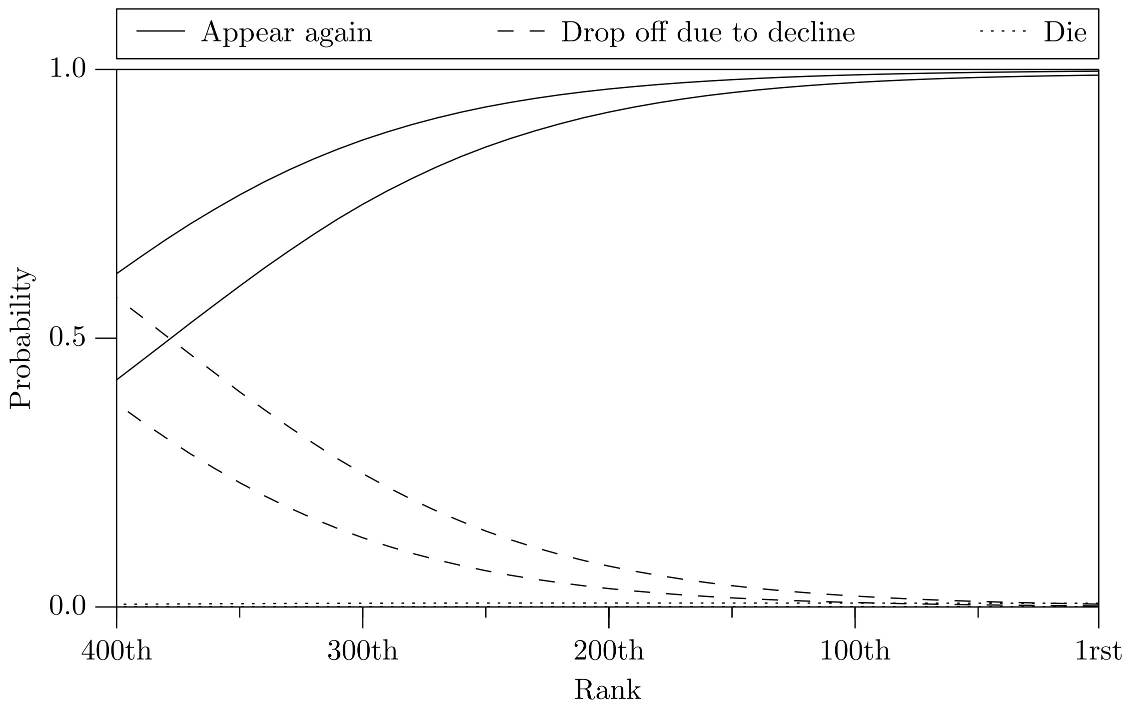

Figure 4.5. Effect of Rank

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows 95 percent confidence intervals for the probability that a person on *Forbes Magazine*'s list in a given year would appear on the list again in the next year, drop off due to a decline in wealth, or drop off due to death as a function of the person's rank in the distribution of wealth.

The probability of dropping off due to an absolute or relative decline in wealth appears to have been relatively high during the stock-market boom that collapsed with Black Monday in 1987, relatively high during the stock-market boom of the late 1990s that collapsed with the dot-com bubble, and relatively high during the recovery that followed the 2001 recession but preceded the crisis of recent years. Turnover therefore seems to have been higher during booms and lower during busts. Of course, certain booms or certain busts may be better or worse for certain people in certain industries. During the recent crisis, for example, people who made their fortunes in either finance, insurance, or real estate—the so-called FIRE industry—may not have fared as people well as people who made their fortunes in other industries.

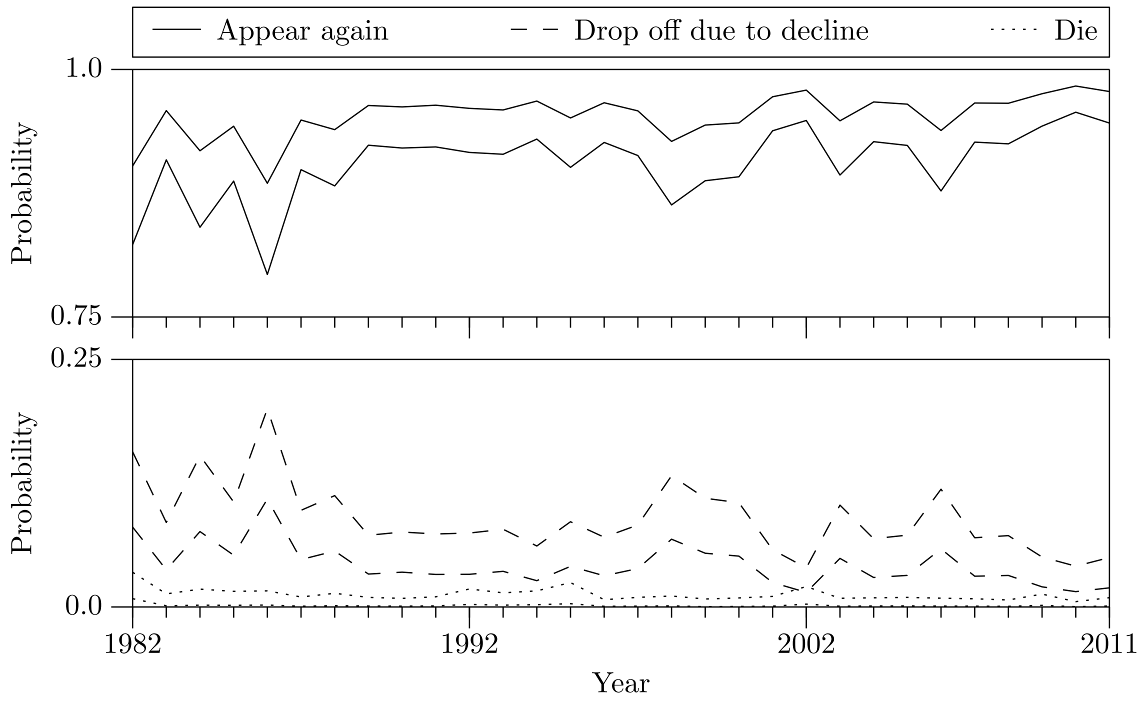

Figure 4.6. Effect of the Calendar Year

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows 95 percent confidence intervals for the probability that a person on *Forbes Magazine*'s list in a given calendar year would appear on the list again in the next year, drop off due to a decline in wealth, or drop off due to death as a function of the given year.

Effect of duration. Holding the other covariates constant, figure 4.7 shows the association between the number of consecutive years that a person has appeared on the list and his or her probability of appearing again, dropping off due to decline, or dropping off due to death. The figure was again split into two for the benefit of the reader. As seen in the figure, the probabilities are approximately the same for each duration. There is greater uncertainty about the effects of durations longer than about 20 years, but that uncertainty can be attributed to the relatively small number of people who appeared year after year for that many years. Thus, after controlling for covariates like a person's age and his or rank in the distribution of wealth, there does not appear to be any duration dependence. Figures 4.2 and 4.3

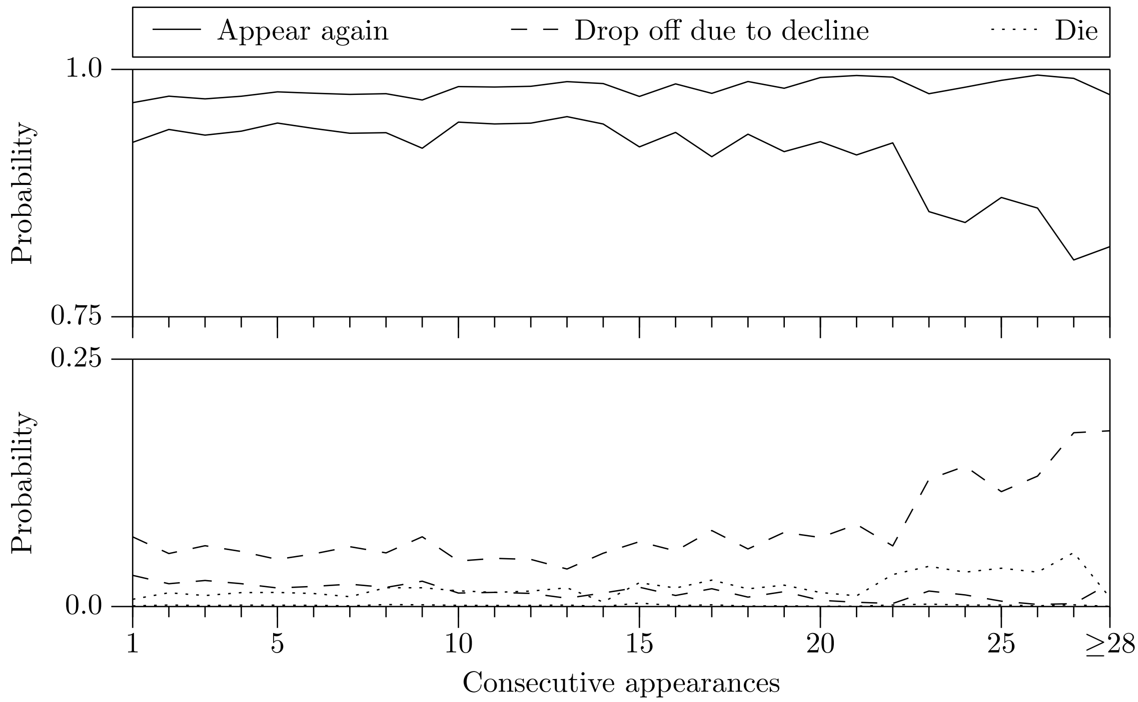

Figure 4.7. Effect of Duration

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows 95 percent confidence intervals for the probability that a person on *Forbes Magazine*'s list in a given year would appear on the list again in the next year, drop off due to a decline in wealth, or drop off due to death as a function of the number of consecutive years that the person has already appeared.

from before suggested otherwise, but recall that those figures did not control for any other covariates besides duration.

Effect of left-censoring. The only other covariate included in the model discussed above was a dummy variable for whether a person's consecutive appearances on the list were left-censored or not. That dummy variable was included as a covariate, but, as shown in the table, a likelihood-ratio test suggests that the likelihood of a model with the dummy variable is not statistically significantly different than a model without that dummy variable ($X_2^2 \approx 0.91$ $p \approx 0.64$). Such a finding seems to be consistent with the finding that, after controlling for covariates like a person's age and rank, there does not appear to be any duration dependence.

Correct and incorrect predictions. The extent to which the above-discussed model is a good model should be assessed. One measure of the goodness of the fit of the model to the data is the percentage of observations that were correctly predicted, in the sense that the outcome that actually occurred was the outcome that the model predicted was the most likely to occur. The percentage of observations that were correctly predicted by the model is relatively high at about 88 percent.[14]

Although many observations are correctly predicted by the model, the model incorrectly predicts some observations. Indeed, for 14 observations, the model even predicts that the outcome that actually occurred would occur with a probability of less than one half of one percent. For each of those 14 observations, the model predicts that a person would appear on the list again in the next year with a probability of more than 99.5 percent, but he or she did not appear again. Nine of the 14 people did not appear again because they died. The people who died all did so at relatively young ages in relatively unusual ways. One person, Tyson Foods heir Randal William Tyson, died at 34 years of age when he choked on a cookie (*Associated Press* 1986). Another person, Quad Graphics founder Harry Quadracci, died at 66 years of age when he drowned in a lake near his home (*New York Times* 2002). Three people died in plane crashes. The Heinz heir Henry John Heinz III died at 52 years of age when there was a mid-air collision between his plane and a helicopter (*New York Times* 1991); the Atlas Air founder Michael Chowdry died at 46 years of age when the demilitarized fighter jet that he was flying crashed (*Flying Magazine* 2001); and the Walmart heir John Walton died at 58 years of age when the experimental

[14] Admittedly, that percentage is only slightly higher than the percentage of observations that would have been correctly predicted by a constant-only model. Out of the almost 12,400 observations, a constant-only model would have correctly predict only 34 fewer observations. That said, the full model still performs better than a constant-only model in terms of the percentage of observations correctly predicted. The model also performs better in terms of other goodness-of-fit measures. The log-likelihood of the full model is statistically significantly greater than that of the constant-only model at less than the one percent level, as shown in the table.

ultralight aircraft that he was flying crashed (*Associated Press* 2005). Cancer killed four others. The real-estate magnate Robert Harris Lurie died at 48 years of age from colon cancer (*New York Times* 1990a); the Disney heir Sharon Disney Lund died at 56 from breast cancer (*Los Angeles Times* 1993); the Oracle co-founder Bob Miner died at 52 from lung cancer (*New York Times* 1994); and the Apple co-founder Steve Jobs died at 56 from pancreatic cancer (*New York Times* 2011).

Given that those nine people all died at relatively young ages in relatively unusual ways, it is not surprising that the model would incorrectly predict whether they appeared again. The people who died from cancer may have deaccumulated wealth, either because of medical expenses or bequest motives, but recall that the model found little if any association between wealth and mortality. Even if the model did find such an association, the model would almost surely still fail to predict the deaths of people who accidentally choked on cookies, drowned in lakes, or crashed their planes, given that such deaths do not afford the opportunity to deaccumulate wealth. Also, even though it seems like the wealthy might be at a higher risk of dying in demilitarized fighter jets or experimental ultralight aircrafts, such deaths presumably only occur with very small probabilities.

The other five people remained among the living, but they apparently became too poor to remain on the list. One person, William Bernard Ziff Jr., appears to have become too poor on his own accord. He was estimated to be worth about 1.5 billion dollars in 1993, but he fell off the list by the next year because he transfered ownership of Ziff Communications—a company that he himself had inherited after his father died in the 1950s—to his three sons. In the same year that William Jr. dropped off *Forbes Magazine*'s list, his three sons appear on the magazine's list for the first time with a combined wealth equal to their father's wealth in the previous year. Although William Jr. had a bout with prostate cancer about a decade earlier and eventually died from that form of cancer about a decade later, the transfer of ownership was apparently not in anticipation of his death. He transfered ownership because he was retiring (*Forbes Magazine* 1993, 1994; *New York Times* 2006). It can be noted as an aside that William Jr. appeared on the magazine's list in every year from 1982 to 1993, while his three sons—Daniel, Dirk, and Robert—have appeared on the list in every year from 1994 to the most-recent list.

The other four people who remained among the living but became too poor to remain on the list do not appear to have become too poor on their own accord. The real-estate and gambling mogul Donald Trump was estimated to be worth 1.7 billion current dollars in 1989, but he dropped off the list by the next year amid a downturn in the real-estate and gambling markets (*Forbes Magazine* 1989, 1990); the "King of Malls" Edward John DeBartolo Sr. was estimated to be worth 1.4 billion current dollars in 1991, but he dropped off the list by the next year amid the same downturn in the real-estate market (*Forbes Magazine* 1991, 1992); the Cablevision founder Charles Francis Dolan was estimated to be 2.6 billion current dollars in 2001, but he dropped off the list by the next year amid a dramatic decline in the price of his company's stock (specifically, a decline of over 90 percent over two years; *Forbes Magazine* 2001, 2002); and one of the members of the Pritzker family, Robert Alan Pritzker, was estimated to be worth 7.6 billion current dollars in 2003, but he fell off the list by the next year amid a legal fight with other family members (Bernstein and Swan 2007, pp. 259–61; *Forbes Magazine* 2003, 2004).In the case of Robert Alan Pritzker, it is somewhat unclear whether he actually became too poor amid the legal battles or, instead, whether he was already too poor and *Forbes Magazine*'s estimate of what he was worth was simply revised down when the legal battles revealed new information. Either way, 10 new members of the Pritzker family appeared on the list for the first time in the same year that Robert fell off of it (*Forbes Magazine* 2004).

We relate the details of those especially egregious prediction errors, not to suggest that a parsimonious model should necessarily be able to predict all of the outcomes, but rather to suggest some ways in which our model might be extended. For example, just like people in the FIRE industry were presumably more likely to drop off the list amid the recent crisis, it is not surprising that people like Donald Trump and the so-called King of Malls were apparently more likely to drop off the list amid a downturn in the real-estate market in the early 1990s (although there may have been other reasons why Trump was more likely to drop off; see Capehart forthcoming). Our model could therefore be extended to account for the possibility that people in certain industries at certain times may be more likely to drop off the list. We will now turn to considering a few such extensions.

4.4.4 Some Quick Extensions

The model discussed above can be extended to consider how other factors might affect the probability that a person on *Forbes Magazine*'s list in a given year will appear again. We will extend the model to consider the affect of whether someone was philanthropic, whether someone's wealth was self-made, and whether someone's wealth was made in certain industries rather than others. Each of those factors will be considered in turn.

Giving it away? One reason why someone might be more likely to drop off the magazine's list is that they may be more philanthropic. Other things being equal, a more philanthropic person would be more likely to give away more wealth and therefore more likely to drop off due to an absolute or relative decline in wealth. None of the especially egregious prediction errors discussed above could be attributed to philanthropy, unless the children of William Bernard Ziff Jr. or Robert Alan Pritzker are seen as charity cases, but wealthy people have been known to engage in philanthropy, so that might be something that matters.

While the philanthropy of the people who appear on its list has sometimes been the focus of *Forbes Magazine*'s coverage (see, for example, Adams 2014), the magazine has never systematically reported information about their philanthropy. Some relatively systematic information is available from other sources, however.

For every since 2000, the newspaper *The Chronicle of Philanthropy* has published the "Philanthropy 50," which is a list of the 50 Americans who were the most generous in terms of the total dollar value of their new commitments of money, stock, or other assets to charities and foundations over the course of the year. New commitments include donations that were paid out during the year, as well as pledges to pay out donations in the future, although pledges must be specific commitments to specific organizations in order to count towards the rankings (Di Mento 2014).

Note that, if someone is simply paying off a pledge made in an earlier year, then those payments do not count because they reflect old rather than new commitments. Also note that if someone makes a pledge to give away money but does not specify who they will give money to or when they will give it away, then that pledge does not count, either. Signing the so-called "Giving Pledge" without making any further commitments does not count, for example.[15]

The Philanthropy 50 only goes back to 2000, but, before that, *The Chronicle of Philanthropy* worked in conjunction with the magazine *Slate* to publish a similar list, the "Slate 60," which was an annual list of the 60 most-generous Americans where a person's generosity was measured in the same manner (or, at least, there is no indication that it was measured in a different manner). The inaugural year of that list included the 60 most-generous Americans of 1996 (Allen 1996; Plotz 2006).[16]

---

[15] On the Giving Pledge and some of its signatories who are and are not on the most-recent year of the Philanthropy 50, see Lewis (2014).

[16] As an aside, it can be noted that *Slate* was apparently inspired to start its list after a wealthy American suggested that, whereas *Forbes Magazine*'s list of the 400 wealthiest Americans discourages philanthropy, a list of the most-generous Americans could encourage it (Plotz 2006).

Thus, information is available on some of the most-generous Americans in every year since 1996. Over those years, some of the most-generous Americans according to the Slate 60 and Philanthropy 50 lists were also some of the wealthiest Americans according to *Forbes Magazine*'s list. Those people were presumably less likely to continue to appear on *Forbes Magazine*'s list. If someone made a donation that was generous enough to make them one of the most-generous Americans, then he or she was presumably more likely to drop off due to a decline in wealth.

In order to account for that possibility, we can incorporate the information contained in the Slate 60 and Philanthropy 50 lists into our model. There are various ways in which we might incorporate that information, but we will incorporate it as follows. We will include a dummy variable that reflects people who were one of the wealthiest Americans in a given year according to *Forbes Magazine*'s list and also one of the most-generous Americans in the same year or any earlier year according to either the Philanthropy 50 list for the years since 2000 or the Slate 60 for the years between 1996 and 1999. By defining the dummy in that way, we allow for the possibility that someone paying off a previous commitment would be more likely to drop off *Forbes Magazine*'s list even though those payments would not count towards either the Philanthropy 50 or Slate 60 lists.[17]

Including that philanthropic dummy obviously constrains the analysis to the years since 1996, but the model can otherwise be the same as it was before, except for one slight change. The slight change arises from the fact that a dummy variable for a duration of 14 years cannot be included in the model. For the people who appeared on any year of *Forbes Magazine*'s list since 1996 and who also appeared on

---

[17] Again, there are other ways in which we might incorporate the information contained in the Philanthropy 50 and Slate 60 lists. There is a bit of an inconsistency in looking at the 60 most-generous Americans in some years and only the 50 most generous in other years, but the results presented below are qualitatively similar if we only look at the 50 most generous on the Slate 60.

any year of the list since 1982 for at least 14 consecutive years, none of them died after their 14th consecutive year on the list. Including a dummy for a duration of 14 years would therefore create a perfect prediction problem. In order to maintain a non-parametric approach while avoiding that problem, a dummy variable for 13, 14, or 15 years can be included, rather than dummy variables for each of those years.

If the above-discussed model is re-estimated with that slight change and with the philanthropic dummy variable included as a covariate, then the results are as follows. A likelihood-ratio test suggests that the likelihood of a model with the philanthropic dummy is not significantly greater than a model without that dummy at conventional levels of statistical significance ($X_2^2 \approx 1.68$, $p \approx 0.43$). Moreover, the difference that the philanthropic dummy would make for the sort of typical person we considered above—again, a person who is 64 years of age, the 194th wealthiest American, and on the list for the first time in 1996—would not be statistically significant at conventional levels and, even if it was statistically significant, it would be arguably small. If the person ever appeared on the Philanthropy 50 or Slate 60 lists, then his or her probability of appearing on *Forbes Magazine*'s list again would differ by less than two tenths of a percentage point.

Given its lack of statistical or substantive significance, we should perhaps ignore the direction of that two tenths of a percentage point difference, but its direction can be noted. If our typical person ever appeared on the Philanthropy 50 or Slate 60 lists, then the person would actually be less likely to drop off *Forbes Magazine*'s list mostly because he or she would be less likely to drop off due to a decline in wealth. The philanthropic are those who can afford it the most, it seems. The direction of that effect is even more pronounced if we re-define the dummy variable so that it only reflects people who were on the Philanthropy 50 (or, in the years before 2000, the Slate 60) in the same year that they were on *Forbes Magazine*'s list. The likelihood of a model with that re-defined dummy is still not statistically significant greater than the likelihood of a model without it, and the effect for the typical person is also still not statistically significant, but the person would be about one percentage point less likely to drop off because he or she would be about one percentage point less likely to drop off due to decline.

It is of course possible that we have failed to find a negative effect of philanthropy on the probability of appearing again because our measure of philanthropy fails to reflect people who donated enough to drop off *Forbes Magazine*'s list but never donated enough to appear on the Philanthropy 50 or Slate 60 lists. Yet, if we look at the wealth of even the least-wealthy of the 400 wealthiest Americans in any given year and compare that to the largest amount that someone could have committed without appearing on the Philanthropy 50 or Slate 60 lists, then the commitments seem to be too small to have an effect that is discernible from any of the other varied forces that might affect the rise and fall of wealth. For the year 2013, for example, the least generous of the 50 most-generous Americans committed about 38 million current dollars over the course of that year according to the Philanthropy 50 list. That amount may seem like a hefty sum, but it is less than three percent of the wealth of even the least-wealthy of the 400 wealthiest Americans on *Forbes Magazine*'s list in the same year, who was worth about 1.3 billion current dollars.

Appearing on any of the Philanthropy 50 or Slate 60 lists does not seem to have a statistically or substantively significant effect on (or, at least, an association with) the probability of appearing on *Forbes Magazine*'s list again, therefore. If anything, a person is more likely to appear again because he or she is less likely to drop off due to a decline in wealth. Even though philanthropy does not seem to have an effect on the persistence of wealth, other factors might matter. The source of someone's wealth might matter, for example. We now turn to considering that possibility.

Profligate heirs? As discussed above, after its list went online in 1996, *Forbes Magazine* started systematically reporting whether someone's fortune was self-made. According to at least one year of the magazine's list, someone's fortune should be considered self-made if they "built [their] fortunes themselves," although someone's fortune could still be considered self-made, even if they did not build their fortunes "entirely from scratch" because they "borrowed money from in-laws or parents" or "started businesses with spouses or other relatives" (Kroll 2012). A dummy variable for whether someone's fortune was self-made can be included as a covariate in the above-discussed model. Including that dummy variable again constrains the analysis to the years since 1996, so the slight change discussed above must be made again.

If the above-discussed model is re-estimated with that slight change and with a self-made dummy variable included as a covariate, then a likelihood-ratio test suggests that the likelihood of a model with the self-made dummy is statistically significantly greater than a model without that dummy ($X_2^2 \approx 12$, $p < 0.01$). However, the difference that the dummy would make for the sort of typical person we have been considering would not be statistically significant at conventional levels. The difference would also seem to be small. If the person's wealth was self-made, then he or she would only be about one percentage point less likely to appear again because he or she would be about one percentage point more likely to drop off due to decline. Only if the person was barely one of the 400 wealthiest Americans would the difference be statistically and arguably substantively significant, as figure 4.8 shows.

Thus, to the extent that there is any association between whether someone's wealth is self-made or not and whether they appear again or not, it seems that the self-made are more likely to drop off. Those who inherit a fortune may actually be less profligate then those who make their own fortunes, perhaps. That said, the self-made may tend to hold their wealth in a single or small number of assets and, as

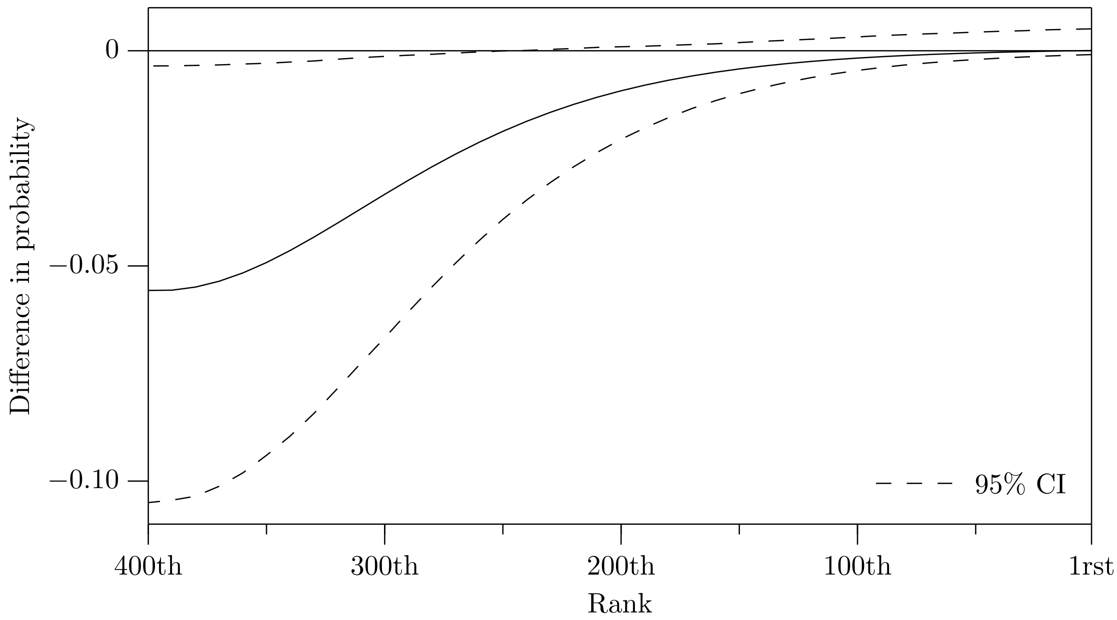

Figure 4.8. Effect of Being Self-made

Source: Data adapted from *Chronicle of Philanthropy*; *Forbes Magazine* (1982–2013); *Slate*.

Note: This figure shows, as a function of the person's rank in the distribution of wealth, point estimates and 95 percent confidence intervals for the difference between the probability that a person on *Forbes Magazine*'s list in a given year would appear on the list again in the next year if their wealth was self-made, on the one hand, and the probability that they would appear again if their wealth was not self-made, on the other hand. The probabilities correspond to a person who appeared on the list for the first time in 1996 at 63 years of age.

such, they may be more susceptible to a dramatic decline in their wealth. A person who became rich by building a company may hold almost all of his or her wealth in that single company instead of a well-diversified portfolio, for example.

A Great Recession? At the same time that *Forbes Magazine* began systematically reporting whether someone's wealth was self-made, the magazine also began systematically reporting the primary industry in which someone's wealth was made. Dummy variables for all the industries identified by the magazine could be included in the model discussed above, perhaps, but industries presumably rise and fall over time. The years around the recent crisis were years in which people who made their fortunes from the FIRE industry presumably saw their fortunes rise, fall, and then perhaps rise again, for example.

In order to study how people in the FIRE industry fared in the years around the recent crisis, the model discussed above with the slight changed discussed above can be re-estimated with a dummy variable for whether someone was in the FIRE industry or not and interaction terms between that industry-specific dummy and year-specific dummies. No one in the FIRE industry died after appearing on the list in either 1999, 2000, or 2012, so we will only include industry-year interaction terms for the years between 2001 and 2011.

Figure 4.9 shows, for a person who is 64 years of age, who is the 194th wealthiest American, and who came onto the list for the first time in a given year, 95 percent confidence intervals for the difference between—on the one hand—the probability that the person would drop off due to an absolute or relative decline in his or her wealth, if his or her fortune was made in the FIRE industry, and—on the other hand—the probability that the person would drop off due to decline, if his or her fortune was made in any other industry besides the FIRE industry.

As shown in that figure, the model suggests that, if a person's fortune was made in the FIRE industry rather than another industry, then the person would have been slightly less likely to drop off due to decline in the years leading up to the crisis (with the year 2005 as the only exception). The model also suggests that he or she would have been slightly more likely to drop off due to decline in each year between 2007 to 2010. By 2011, he or she would have again been slightly less likely to drop off due to decline.

The differences are so slight that there is almost no difference, however. The differences are not statistically significant at the five percent level, for example, and the differences are all on the order of only about one percentage point. Only back

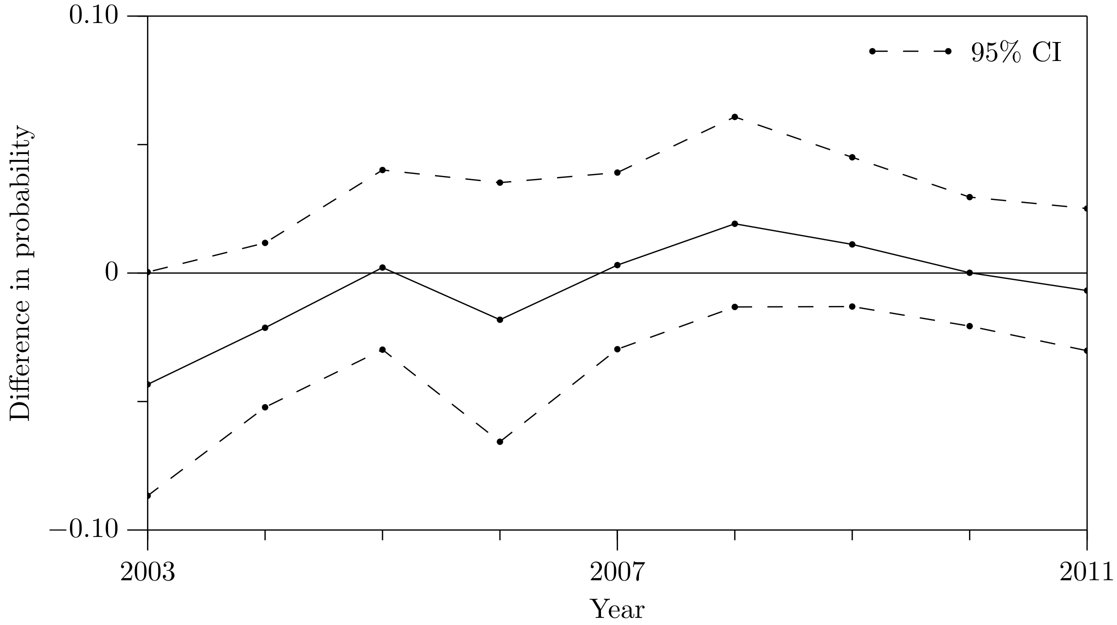

Figure 4.9. Effect of the FIRE Industry

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows, for select years, point estimates and 95 percent confidence intervals for the difference between the probability that a person on *Forbes Magazine*'s list in a given year would drop off by the next year due to a decline in wealth if he or she was in the FIRE industry, on the one hand, and the probability that the person would drop off due to a decline in wealth if he or she was in any other industry, on the other hand. The probabilities for a given year correspond to a person with the median age and rank who appeared on the list for the first time in the given year.

in the year 2003 was the difference statistically significant at that level; but, even in that year, the person would have only been about four percentage points less likely to drop off due to decline. In 2008, which appears to have been the worst year of the crisis for people who made their fortunes in the FIRE industry, a typical person would have been expected to drop off due to decline with a probability of about five percent rather than about three percent, if his or her fortune was made in the FIRE industry rather than another industry. Thus, even in its worst year, the Great Recession does not appear to have been much worse for the people who made their fortunes in the FIRE industry than any other industry, other things being equal.

4.5 Conclusion

This essay studied the extent to which wealth persists among some of the wealthiest Americans by using a data source that was well-suited to certain tasks. Specifically, this essay used *Forbes Magazine*'s annual list of the 400 wealthiest Americans in order to study the extent to which people who appear on that list continue to appear. The study found that, while most people have remained on the magazine's list for a relatively short amount of time, some have remained on the list for longer amounts of time, and a few have even remained on the list in every year since it started over 30 years ago. Thus, while the wealth of most people seems to be more ephemeral, the wealth of a few seems to be almost permanent.

This essay also studied some of the factors associated with the continuing to appear on the list. Some of the findings were hardly surprising. Younger, wealthier people were found to be less likely to drop off the magazine's list compared to their older and poorer counterparts. Wealthier people were less likely to drop off the list because they were less likely to become too poor stay there, while younger people were less likely to drop off because they were less likely to die. Although those findings are not surprising, there is nevertheless value in identifying or at least confirming those associations, as well as quantifying them.

Some findings were more surprising. The finding that older people were not any more or less likely to become too poor to stay on the list is surprising, perhaps, although that finding should only be surprising to someone who is familiar with the life-cycle theory of savings but unfamiliar with all of its empirical failures. The more surprising results were related to whether appearing on the magazine's list for more years helps one's chances of appearing again for yet another year (it seems to neither help nor hurt, after controlling for other factors), whether philanthropy hurts one's chances of appearing again (it seems to help if anything, perhaps because the philanthropic are those who can afford it the most), whether building rather than inheriting a fortune helps one's chances of appearing again (it seems to hurt if anything, perhaps because the portfolios of the self-made are riskier), and whether being in the wrong industry at the wrong time hurts one's chances of appearing again (it seems to hurt, but not as much as one might expect in at least some cases).

Directions for future research to which *Forbes Magazine*'s list may be well-suited include more fully exploring possible explanations for this essay's findings, especially its more surprising ones. The fact that the more surprising findings are all findings of little or no association suggests one possible explanation. It could be the case that the experiences of the wealthiest Americans are too varied for a simple model to identify strong associations. This essay argued that recent crises presumably hurt people in the FIRE industry, and that industry was indeed the wrong industry for some people during the crises; yet that industry was also the right industry for other people during that time. John Paulson saw his wealth rise by betting against the housing market, for example (*Forbes Magazine* 2007).

Some other directions for future research into the persistence of wealth involve tasks to which the magazine's list is not especially well-suited, but which are of interest. There is an interest in studying, not just whether someone falls off the list, but also how far down the distribution of wealth they fall. Someone who becomes poor relative to the 400 wealthiest Americans could still be extraordinarily wealthy, although the fact that most people who drop off the list never return again suggests that their falls may be dramatic. There is also an interest in studying how people make it onto the magazine's list in the first place and, in particular, whether they tend to start from the bottom of the wealth distribution or from loftier heights. Studying mobility throughout other parts of the wealth distribution besides the very top would seem to require other sources of data besides the magazine's list, however.

CHAPTER 5

CONCLUSION

On the occasion of the 10th anniversary of its annual list of the 400 wealthiest Americans, *Forbes Magazine* asked the question, "What has [our list] accomplished in 10 years?" (*Forbes Magazine* 1991, p. 145). The magazine claimed that the accomplishment of its list was that it "fill[ed] what was once an important blank spot in the portrait of American society" (ibid., p. 146). The metaphorical blank spot that the list filled was the very top of the wealth distribution. According to the magazine, its list filled that blank spot by "showing what concentrated private wealth in this country is really like, as opposed to what ideologues and political opportunists of every stripe want people to believe it is like" (ibid.).

The magazine's list might not be an accurate representation of the part of the American portrait that it purports to fill, of course, either because of inadvertent mistakes or perhaps even ideology, opportunism, or some other nefarious force. Yet *Forbes Magazine*'s list is arguably the best source of data on the very top of the wealth distribution in the United States over recent decades, as this dissertation has tried to argue. The list would therefore seem to deserve the attention of social scientists and not just casual readers.

Using that list, the essays in this dissertation studied inequality and mobility among the wealthiest Americans. The contributions that each essay makes are admittedly modest with many directions for future research that remain open, but trying to understand some basic empirical facts about the wealthiest Americans would seem to be a meaningful contribution to the so-called science of wealth. If nothing else, by arguing that the information contained in *Forbes Magazine*'s list deserves serious study, and by compiling that information in a dataset that others can study, this dissertation has hopefully laid the foundation for further research.

The directions for future research vary widely from relatively narrow and technical questions like the ones considered by this dissertation—questions like whether wealth inequality and mobility have risen, fallen, or stayed the same—to grander questions like whether inequality and mobility are too low, too high, or just right. Answers to the narrower questions can offer insight into the world in which we live and, in doing so, inform debates about the world in which we wish to live.

APPENDIX A

APPENDIX TO THE FIRST ESSAY

A.1 Variances of the Estate-multiplier Estimators

Suppose there are N people who are worth W in total. Let $i = 1, 2, \ldots, N$ index those people, and let w_i denote the i-th person's wealth. Suppose that the i-th person dies with a non-zero probability m_i. Using information about people who die, the total wealth of the people who were alive can be estimated as

$$\hat{W} = \sum_{i=1}^{N} d_i \left(\frac{w_i}{m_i} \right) = \sum_{i \in D} \frac{w_i}{m_i} \tag{A.1}$$

where d_i is an indicator variable that the i-th person died and D is the set of people who died. Similarly, the total number of people who were alive can be estimated as

$$\hat{N} = \sum_{i=1}^{N} d_i \left(\frac{1}{m_i} \right) = \sum_{i \in D} \frac{1}{m_i} \tag{A.2}$$

The unbiasedness of those estimators is obvious once it is recognized that the indicator variable d_i should be unity with probability m_i.

If everyone's probability of dying is independent of everyone else's, then the variance of the estimator for the total wealth of the people who were alive is

$$\mathrm{Var}(\hat{W}) = \sum_{i \in D} m_i \left(1 - m_i\right) \left(\frac{w_i}{m_i} \right)^2 \tag{A.3}$$

(see, for example, Stehman and Overton 1994, p. 30, eq. 2). That variance can be derived in the same way that the variance of any binomial random variable can be derived. (If mortality rates are not independent, then the expression for the variance is much more complicated; see Stehman and Overton 1994.) Similarly, the variance of the estimator for the total number of people who were alive is

$$\mathrm{Var}(\hat{N}) = \sum_{i \in D} m_i \left(1 - m_i\right) \left(\frac{1}{m_i}\right)^2 \tag{A.4}$$

A.2 The Pareto Extrapolation Method

Suppose we know that the N-th wealthiest people are worth W in total and that each of them is worth at least w. Now suppose we want to know what the N'-th wealthiest people in the same population are worth in total. Let W' denote their unknown total wealth. If we are willing to assume that personal wealth follows a Pareto distribution with a lower-bound parameter that is lower than any level of interest and a shape parameter α that is greater than unity, then we can extrapolate the wealth of the latter group based on the wealth of the former group.

Given what we assumed we knew about the N-th wealthiest people, and given our assumption about the distribution of wealth, the total wealth W' of the N'-th wealthiest people can be extrapolated as

$$W' = W \left(\frac{N'}{N}\right)^{\left(1 - \frac{1}{\alpha}\right)} \tag{A.5}$$

where

$$\alpha = \frac{W}{W - wN} \tag{A.6}$$

Both of those equations follow from the properties of a Pareto distribution. The latter equation (eq. A.6) follows from the fact that, if personal wealth follows a Pareto distribution with a shape parameter $\alpha > 1$, then the mean wealth of people with wealth above a given level is proportional to that given level of wealth, where the proportion is $\alpha/(\alpha-1)$. So, for the N-th wealthiest people whose minimum wealth is w and whose total wealth is W, their mean wealth can be expressed as

$$\frac{W}{N} = w\left(\frac{\alpha}{\alpha-1}\right) \tag{A.7}$$

That property of a Pareto distribution is derived below as part of a lemma in the appendix to the second essay.

The former equation (eq. A.5) follows from the fact that, if personal wealth follows a Pareto distribution with a shape parameter $\alpha > 1$, then the share of wealth held by the wealthiest $100 * p \in [0, 100]$ percent of people is $p^{1-\frac{1}{\alpha}}$. That property is also derived as a lemma in the appendix to the second essay.

A.3 Estimating Mortality Rates

As discussed in the text, the age-, gender-, and year-specific mortality rates used by Kopczuk and Saez (2004a,b) were constructed by taking nationally representative age-, gender-, and year-specific mortality rates from Wilmoth (1997) and deflating those mortality rates by age- and gender-specific social-differential factors for white college graduates taken from Brown et al. (2002). Those social-differential factors were estimated by Brown et al. (2002) as the ratio between age- and gender-specific mortality rates for white college graduates, on the one hand, and nationally representative age- and gender-specific mortality rates, on the other hand. The latter mortality rates were taken from somewhere else (although exactly where is unclear; see ibid., p. 453), while the former were estimated as follows.

Brown et al. (2002) started with non-parametric estimates of age- and gender-specific mortality rates for white college graduates. Those non-parametric estimates were constructed by an unnamed employee at the Census Bureau (ibid., p. 448).

Brown et al. (2002) could have used those estimates to directly construct social-differential factors, but they did not do that for various reasons (including a desire for smoother estimates; ibid., pp. 450–2). Instead, for each gender of white college graduates, they fit a parametric function to the non-parametric estimates. Specifically, they ran a non-linear regression of the form

$$q\{x\} = 1 - sg^{\left(c^{x+1} - c^{x}\right)} \tag{A.8}$$

where $q\{x\}$ is the mortality rate at age x and s, g, and c are reduced-form parameters (ibid., p. 453). That function was derived by assuming that white college graduates of a given gender lived and died according to the Gompertz-Makeham law of mortality (ibid., pp. 452–3.). The function can be derived from that law as follows.

Under the Gompertz-Makeham law, the probability of living to age x is

$$p\{x\} = \exp\{(\alpha/\beta)\,(1 - \exp\{\beta x\} - \gamma x)\} \tag{A.9}$$

where α is scaling parameter, β is a parameter that captures any age-dependent component to mortality, and γ is Makeham's (1860) parameter that captures any age-independent component to mortality. Or, in the notation of Brown et al. (2002),

$$p\{x\} = (1/g)s^{x}g^{c^{x}} \tag{A.10}$$

where $s \equiv \exp\{-\gamma\}$, $g \equiv \exp\{-\alpha/\beta\}$, and $c \equiv \exp\{\beta\}$.

The probability of dying by age $x+1$, conditional on living to age x, is then

$$q\{x\} = \left(p\{x\} - p\{x+1\}\right)/p\{x\} \tag{A.11}$$

(ibid., p. 453), or, substituting in the expression for $p\{x\}$ and simplifying,

$$q\{x\} = 1 - \exp\{(\alpha/\beta)\,(\exp\{\beta x\} - c\beta\,(x+1)\}) - \gamma\} \tag{A.12}$$

which is equivalent to the function estimated by Brown et al. (2002) when s, g, and c are defined in the same way as before.

In order to estimate the mortality rates for people who appeared on *Forbes Magazine*'s list of the 400 wealthiest Americans, we made the same assumption about them that Brown et al. (2002) made about white college graduates. We assumed that they lived and died according to the Gompertz-Makeham law of mortality. Although we made the same assumption that Brown et al. (2002) made, we did not use the same method that they used to estimate the parameters associated with that law. Instead of minimizing the squared differences between non-parametric estimates, on the one hand, and fitted values from a parametric function fit to those non-parametric estimates, on the other hand, we simply maximized the probability of observing the deaths that were observed. By assuming that the people who appeared on *Forbes Magazine*'s list in any given year lived and died according to the Gompertz-Makeham law of mortality, the probability of observing the deaths that were observed between any one year of the magazine's list and the next can be expressed as

$$\prod_{i=1}^{400} \left(1 - q\{x_i\}\right)^{1-d_i} q\{x_i\}^{d_i} \tag{A.13}$$

where x_i is the age of the i-th person on the list in a given year, d_i is an indicator variable that the i-th person died by the next year, and $q\{\cdot\}$ is defined above. That expression is simply the sum of the probabilities that each person on the list would either live, if he or she lived, or die, if he or she died. The sum of those probabilities is the probability of observing all the deaths that were observed. The parameters that maximize that probability can be found by using numerical methods.

Thus, while we used a different estimation method, we still made the same assumption about mortality rates that Brown et al. (2002) made. Aside from the fact that they also made that assumption, we can defend the assumption as follows.

The Gompertz law of mortality is the prototypical example of a law of mortality (Olshansky and Carnes 1997). Other laws of mortality, including Makeham's (1860) modification of that law, were suggested in order to account for ways in which that law fails. Makeham suggested his modification to the Gompertz law because that law fails to account for any age-independent component to mortality.

While other laws that account for other ways in which the Gompertz-Makeham law fails could be used, other ways in which that law fails can be ignored for this study. One way in which the law fails is that it fails to account for the complex relationship between age and mortality at very young ages. The Heligman-Pollard law of mortality modified the Gompertz-Makeham law to try to account for that complex relationship (Heligman and Pollard 1980). The complexity of the relationship is reflected in the fact that the Heligman-Pollard law involves eight parameters. The failure of the Gompertz-Makeham law to account for the complex relationship between age and mortality at very young ages can be ignored for this study, however, given that the youngest person to appear on *Forbes Magazine*'s list since 1982 was relatively old at 22 years of age.

Another way in which the law fails is that it fails to account for the deceleration of mortality at very old ages. Other laws that account for that deceleration like a logistic law have been suggested (see Olshansky and Carnes 1997, p. 10). The failure of the Gompertz-Makeham law to account for the deceleration of mortality at very old ages can be ignored for this study, however, given that any failure to account for that deceleration should bias the estate-multiplier estimates upwards. The essay in this dissertation suggests that the Kopczuk-Saez estimates may be biased in the other direction. The assumption that people on *Forbes Magazine*'s list lived and died according to the Gompertz-Makeham law of mortality seems defensible, therefore.

A.4 Top Wealths from Surveys

Although one might hope that other sources of data could suggest whether the direct or estate-multiplier estimates of the total wealth of the 400 wealthiest Americans were more accurate, other sources like surveys are not helpful in that regard. Some surveys are especially unhelpful. The exact extent to which the distribution of wealth is skewed can only be estimated, but, by all estimates, the distribution is highly skewed. Given that the distribution of wealth is highly skewed, surveys that do not try to over-sample the top of the wealth distribution often fail to capture that part of the distribution (Davies and Shorrocks 2000, pp. 629–35; Juster et al. 1999). For example, in a recent year of a household survey of wealth that does not try to over-sample wealthy households—the wealth survey conducted as part of the Panel Study of Income Dynamics (PSID)—the wealthiest household in the survey was estimated to be worth only about 50 million current dollars (PSID 2007). A household worth 50 million dollars would have been worth more than most households in that year. According to estimates based on a household survey of wealth that does try to over-sample wealthy households—the Survey of Consumer Finances (SCF)—a wealth of 50 million dollars would have placed a household in the wealthiest one percent of households in that year (Kennickell 2011, p. 27, table A1). Yet the wealthiest household in the SCF in the same year was worth much more than a mere 50 million. The wealthiest household in the SCF was worth over one billion dollars (SCF 2007). Surveys like the PSID that do not try to over-sample the wealthy are therefore especially unhelpful for estimating the wealth of relatively wealthy groups, let alone the 400 wealthiest Americans.

The SCF is generally recognized as the best survey on wealth in the United States, at least partly because it tries to over-sample wealthy households (Davies and Shorrocks 2000, p. 632). The survey uses two sampling mechanisms. For one sampling mechanism, data from the Census Bureau is used to sample households with equal probabilities (conditional on their geographic area in order to capture certain areas of the country; Bricker et al. 2012, p. 78). It is unlikely that this sampling mechanism would sample any of the relatively small number of households that own a relatively large amount of wealth, but the mechanism should sample a large number of households with small amounts of wealth.

For the other sampling mechanism, which generates the so-called "supplemental" or "list" sample, income-tax returns are used to try to sample wealthy households (Bricker et al. 2012, p. 78). A flow of income is obviously not a stock of wealth, and a tax-filing unit may not be a household, but the staff at the SCF apparently use the income-tax returns to predict how much the household of a tax-filing unit may be worth. The staff then samples households that are predicted to be wealthy (Kennickell 1999, 2001). Interestingly, some of the households with the most wealth are apparently sometimes some of the households with the smallest incomes, perhaps because the wealthy can afford to minimize their income, or perhaps because of variability in their income (Kennickell 1999, pp. 7–8).

Despite the fact that the SCF tries to over-sample wealthy households, the survey is still not especially helpful for estimating the wealth of the 400 wealthiest Americans for the following reason. By design, the SCF excludes the households of people on *Forbes Magazine*'s list of the 400 wealthiest Americans (Kennickell 2006, p. 84). To belabor this point: One of the directors of the SCF in some of his publications (ibid.), other members of the staff at the SCF in some of their publications (Bertaut and Starr 2002, p. 214; Bricker et al. 2012, p. 78; etc.), and the codebooks for the publicly available datasets (see, for example, SCF 2010a) all emphasize that the survey is designed to exclude the households of people who appear on the magazine's list. The income-tax returns that are used to generate the supplement or list sample specifically exclude the income-tax returns of people who appear on the magazine's list, so their households can never be part of that sample. Also, in the unlikely event that one of the households of the people who appear on the magazine's list was ever sampled as part of the equal-probability sample, that household would presumably be excluded, although it is somewhat unclear what would happen, perhaps because it has never happened.

It is somewhat strange that the SCF would be designed to try to over-sample wealthy households, while also designed to exclude households that may have been some of the wealthiest households, but that is how the survey was designed. Some reasons for excluding those households have been offered. One reason is that attempting to survey them would be prohibitively expensive and frequently unsuccessful. To quote the director of the SCF at length:

> The argument for excluding this group [i.e., the households of people who appear on the magazine's list] from the SCF sample is that because such people are typically surrounded with levels of staff intended to keep other people away, they would be extraordinarily expensive to attempt to interview, and the success rate could reasonably be expected to be quite low. (Kennickell 2007, p. 2)

Another reason is that it would be difficult to protect their confidentiality. To continue the quote from above:

> Moreover, because these people are so well known, it would be almost impossible to protect their confidentiality without destroying the statistical utility of the data they would provide. (ibid.)

Whether it is reasonable to exclude them or not, the survey excludes them. As such, the SCF is not especially helpful for determining the accuracy of direct or estate-multiplier estimates of the wealth of the 400 wealthiest Americans.

That said, according to the codebook for the most-recent year of the survey (which was the year 2010, as of writing), there were 10 households that were estimated to be worth more than the minimum wealth that it took to make it onto

Forbes Magazine's list in the same year (SCF 2010a). In earlier years of the survey, there were a similar number of such households. (There were four, three, seven, and four such households in 2007, 2004, 2001, and 1998, respectively, according to the codebooks from those years. The numbers were not reported for other years of the triennial survey, to the author's knowledge, perhaps because the number was zero.) Exactly how much those households were estimated to be worth is not known, at least to the public, because the households were removed from the public dataset. They were removed because "it would be very difficult to obscure sufficiently the identity of [those households] without rendering their data virtually useless" (see, for example, SCF 2010a).

The fact that some households in the SCF were estimated to be worth more than the minimum wealth that it took to make it onto *Forbes Magazine*'s list suggests that the magazine may have missed some of the 400 wealthiest Americans (Kennickell 2007, p. 2). If so, then the magazine would have at least not overestimated the total wealth of the 400 wealthiest Americans, and would have underestimated their wealth, if the people it missed were wealthier than the people it identified. Thus, even if the SCF suggests that the magazine missed some of the 400 wealthiest Americans, this does not suggest that the magazine's direct estimates of the total wealth of the 400 wealthiest Americans were too large relative to the estate-multiplier estimates. The fact that the survey found some households that were credibly estimated to be worth at least as much as what the magazine estimated the 400th wealthiest American was worth also suggests that the magazine's estimate for the wealth of at least the 400th wealthiest American was credible or at least not incredibly large.

Even if the SCF did somehow capture the 400 wealthiest households in America without sampling any of the households of people on the magazine's list, survey estimates of the wealth of 400 wealthiest households in America based on the SCF may not be directly comparable to either the direct estimates of the wealth of the 400 wealthiest Americans based on *Forbes Magazine*'s list or the estate-multiplier estimates of their wealth based on the Kopczuk-Saez estimates.

Especially when trying to compare the survey and estate-multiplier estimates, there is an issue about the assignment of ownership. For any given person, his or her wealth for estate-tax purposes excludes the wealth of any family members, as discussed above, so if a person had family members who were worth something, then his or her wealth according to estate-tax records would be different than his or her wealth according to the survey, even if any other issues were assumed away. The magazine and the survey both account for the wealth of family members, so the person's wealth according to the magazine should be relatively similar to the wealth of the person's household according to the survey (relative to the similarity between the estimates according to estate-tax records and the survey), although perhaps not the same. There may also be other issues.

Even if wealth was assigned in the same way, there may be a timing issue. For the most-recent year of the SCF, the interviews for that survey were "largely" conducted between the months of May and December (Bricker et al. 2012, p. 79), but the magazine's list in the same year was suppose to be a snapshot of wealth at the close of the stock market on one day of that year (August 25th; *Forbes Magazine* 2010), and people died all throughout that year (although, again, the valuation date for estate-tax records could have been either the day a person died or six months thereafter; Kopczuk and Saez 2004b, p. 43). Thus, the survey, direct, and estate-multiplier estimates may not reflect the same point in time.

And even if wealth was assigned in the same way and valued at the same time, there may be a valuation issue. Although misreporting may plague the SCF and other surveys of wealth (Davies 2009, pp. 129–30), even if households accurately responded to the questions that they were asked, the questions asked by the SCF limit the value of certain types of assets and debts that a household can report.

In terms of assets: As of the most-recent year of the survey, a household can only include up to two "investment real estate" or "second (or vacation) homes" (sec. E of SCF 2010b), only up to two businesses in which they have an "active management interest" (sec. F), only up to four "cars, trucks, vans, minivans, [or] jeep-type (sport-utility) vehicles" (sec. G), only up to two "motorhomes, RVs, motorcycles, boats, airplanes, [or] helicopters" (sec. G), only up to six checking accounts, which would only add up to 1.5 million dollars, assuming that each account held the maximum amount that was insured by the Federal Deposit Insurance Corporation in that year (sec. N), only up to five "saving or money-market accounts," which would only add another 1.25 million under the same assumption (sec. N), only up to three other "important" financial assets that are not covered by another part of the survey (sec. N), only up to 10 pension plans (up to two that are owned, up to four that are currently paying out or being drawn from, and up to four others; sec. R), only up to three "inheritance[s], substantial gift[s], or trust[s]" that have already been received (sec. X), apparently only up to one "substantial inheritance or transfer" that is expected in the future (sec. X), and apparently only up to one "personal trust or foundation" (so possibly no personal trusts if the household would rather speak about a foundation; sec. X).

In terms of debts: A household can only include up to three mortgages on homes that are currently owned (sec. D), only up to six lines of credit (including home-equity lines of credit; sec. D), only up to two loans on homes that were previously owned (sec. E), only up to six loans for "educational purposes" (sec. H), and only up to six other loans not covered by another part of the survey (sec. I).

Those limits on the number of assets and debts may not be binding for most households, of course, but if some of the limits bind for some wealthy households, then survey, direct, and estate-multiplier estimates may not be comparable. The total value of real estate is recorded for estate-tax purposes, for example, rather than just the value of up to two homes, as even a cursory glance at an estate-tax form would show. Likewise, the magazine would apparently account for all the real estate it could identify, rather than just up to two homes.

Having said all of that, survey and estate-multiplier estimates for the wealth of a slightly less-wealthy group than the 400 wealthiest Americans can be compared. Figure A.1 shows the wealth of the wealthiest one percent of households, according to survey estimates made by Kennickell (2011, p. 12, table 5) based on the triennial SCF.[1] The wealthiest one percent of households was the smallest group that Kennickell (2011) considered. The figure also shows the wealth of the wealthiest one percent of people, according to the estate-multiplier estimates made by Kopczuk and Saez (2004a,b). As shown in the figure (which is loosely based on fig. 11 of Kopczuk and Saez 2004a), the survey estimates were larger and increasingly larger than the estate-multiplier estimates in the years in which those estimates can be compared. Yet, again, it is unclear what conclusions could be drawn from such a comparison. The discrepancy may simply be due to people increasingly pairing up along class lines to form households, for example (as pointed out by Kopczuk and Saez 2004a, p. 476, among others).

In summary, even the SCF is not especially helpful for suggesting whether the direct or estate-multiplier estimates of the wealth of the 400 wealthiest Americans were right or wrong.

---

[1] Kennickell (2011) only reports the share of wealth held by the wealthiest one percent of households. He does not report their total wealth or the total wealth of all households. However, the mean wealth of all households is reported elsewhere (Kennickell 2012, p. 4, table 2) and the number of households represented by the SCF is also reported elsewhere (Bricker et al. 2012, p. 78, table A.3), so the total wealth of the wealthiest one percent of households can be calculated.

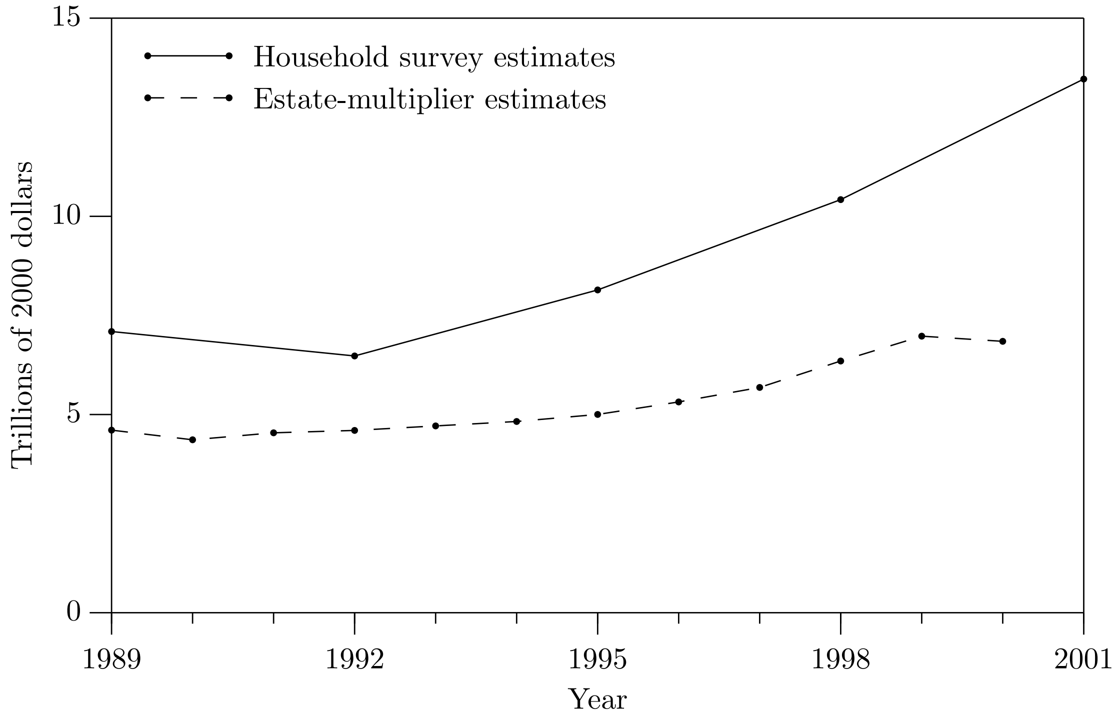

Figure A.1. Household Survey and Estate-multiplier Estimates of the Wealth of the Wealthiest One Percent of Households or People, 1989–2001

Source: Data adapted from Bricker et al. (2012); Kennickell (2011, 2012); Kennickell and Starr (1994); Kennickell et al. (2000); Kopczuk and Saez (2004a,b).

Note: This figure shows, for every three years from 1989 to 2001, estimates of the wealth of the wealthiest one percent of households based on the Survey of Consumer Finances. The figure also shows, for every year from 1989 to 2000, estimates of the wealth of the wealthiest one percent of people based on the Kopczuk-Saez estimates.

A.5 A Similar Exercise with a Longer View

An exercise that is similar to the exercise discussed in the text can be performed in order to study how changes that occurred over the course of the 20th century, and not just the last decade of that century, may have affected the probability that the estate-multiplier method would underestimate what some of the wealthiest Americans were worth. *Forbes Magazine*'s list of the 400 wealthiest Americans only goes back to 1982, so we cannot perform the same exercise for years before that year. We can perform a similar exercise for one year towards the start of the 20th century, however. In 1918, *Forbes Magazine* published what would eventually prove to be the precursor to its list of the 400 wealthiest Americans. In that year, the magazine published a list of the 30 wealthiest Americans (*Forbes Magazine* 1918).

Using that list, we can perform an exercise that is almost identical the exercise discussed in the text. We can assume that the list was an accurate account of the 30 wealthiest Americans in 1918. We can then try to estimate the total wealth of the 30 wealthiest Americans by only using information about certain people on the list.[2] For simplicity, we can sample people on the list at the mortality rates used by Kopzcuk and Saez for that year, and we can inflate the people who are sampled by the rates at which they were sampled. A person's wealth can also be valued in full, given that we do not have a basis for assuming otherwise for that year.

Such an exercise suggests that the estate-multiplier method would underestimate what the 30 wealthiest Americans were assumed to be worth by at least about 43 percent of their wealth half of the time, as shown as part of table A.1 of this appendix. The same exercise suggests that this median misestimation would be even

[2]The magazine's list of the 30 wealthiest Americans in 1918 did not report their ages, but a person's age in that year can be calculated as difference between 1918 and the year in which they were born. Their birth years were reported by *Forbes Magazine* (1983, pp. 56–64).

more severe by the end of the 20th century. If we use information about just the 30 wealthiest Americans from *Forbes Magazine*'s list of the 400 wealthiest Americans in 2000, and if we use the Kopzcuk-Saez mortality rates from the same year, then the same exercise suggests that the estate-multiplier method would underestimate what the 30 wealthiest Americans were assumed to be worth by at least about 55 percent half of the time, as shown in the same table.

Unlike the exercise discussed in the text, the more severe underestimation in more-recent years cannot be attributed to a change in their wealths. If we take the wealths of the wealthiest to the poorest of the 30 wealthiest Americans in 2000, assign those wealths to the wealthiest to the poorest of the 30 wealthiest Americans in 1918, and repeat the exercise again for 1918, then the median misestimation would actually be less severe. The estate-multiplier method would only underestimate their wealth by at least about 36 percent half of the time, as shown in the table. The underestimation would be less severe because, although inequality among the wealthiest Americans may have been greater in 2000 than it was in 1990, inequality among the 30 wealthiest Americans was greater in 1918 than it was in 2000, at least according to *Forbes Magazine*'s lists from those years. The wealthiest person in 1918, John D. Rockefeller, was much wealthier relative to his peers than the wealthiest person in 2000, Bill Gates. Rockefeller was over five times wealthier (about 433 percent wealthier) than the next-wealthiest American, while Gates was only slightly wealthier (about eight percent wealthier). Rockefeller's wealth also accounted for about one third of the total wealth of the 30 wealthiest Americans in 1918, while Gates' wealth only accounted for about 12 percent in 2000.

The increase in the severity of the underestimation also cannot be attributed to a change in the ages of the 30 wealthiest Americans or a change in any association between a person's age and his or her rank in the distribution of wealth. The median misestimation would be about the same, if we took the ages of the wealthiest to the poorest of the 30 wealthiest Americans in 2000, assigned those ages to the wealthiest to the poorest of the 30 wealthiest Americans in 1918, and repeated the exercise again for 1918. The median misestimation would be about the same because, somewhat surprisingly, the ages of the 30 wealthiest Americans were about the same in both years. In 1918 and 2000, the average age of the 30 wealthiest Americans was about 59 and 60 years of age, respectively, while the standard deviation of their ages was about 16 and 15 years of age, respectively. There was also no obvious change in any association between a person's age and his or her rank in the distribution of wealth.

The more severe underestimation can instead be attributed to the change in the rates at which people were sampled. If we repeat the exercise for 1918, but sample people at the Kopzcuk-Saez mortality rates for 2000, then the estate-multiplier method would underestimate what the 30 wealthiest Americans were assumed to be worth by at least about 59 percent half of the time. The net effect of all of the changes that occurred to the 30 wealthiest Americans between 1918 and 2000 is therefore about the same as the effect of just changing the rates at which they are sampled. To the extent that the change in the Kopzcuk-Saez mortality rates between 1918 and 2000 captures the change in the rates at which people died, it seems that people were simply less likely to die at any given age by the end of the century.

Thus, between the start and end of the twentieth century, unlike between the start and end of the last decade of that century, the increase in the probability of underestimating what the some of the wealthiest Americans were assumed to be worth was not driven by greater inequality in the distribution of their wealth. Again, however, the increase in inequality among the 400 wealthiest Americans over the 1990s was apparently dramatic enough to have a relatively large effect on the probability that the estate-multiplier method would underestimate their total wealth.

Table A.1. Effect of Changes between 1918 and 2000 on Misestimation

Changes	Median misestimation
No changes	−43%
Ages	−43
Sampling rates	−59
Ages and sampling rates	−59
Wealths	−36
All changes	−55

Sources: Data adapted from Brown et al. (2002); *Forbes Magazine* (1918, 1983); Wilmoth (1997).

Note: This table shows how certain changes that occurred between 1918 and 2000 would have affected the median amount by which the estate-multiplier method would underestimate what the 30 wealthiest Americans were assumed to be worth. The median amount is reported as a percentage of their wealth. For both years, people were sampled at the Kopzuk-Saez mortality rates, the people who were sampled were inflated by the rates at which they were sampled, and their fully valued wealths were also inflated by those rates.

APPENDIX B

APPENDIX TO THE SECOND ESSAY

B.1 The Definition of a Pareto Distribution

Over a century ago, Vilfredo Pareto discovered that, if the number of people with income greater than a given level of income was plotted against income on a double logarithmic scale, then the plot looked approximately like a straight line, at least at the highest levels of income (see, for example, Pareto [1896] 2001). What is now called a (type I) Pareto distribution has the complementary cumulative distribution function

$$P(X > x) = \left(\frac{x_{\min}}{x}\right)^{\alpha} \tag{B.1}$$

for $x \geq x_{\min}$, where $x_{\min} > 0$ is a lower-bound parameter and α is a shape parameter (Arnold 1983). This function is is simply a formalization of the sort of distribution that Pareto discovered for the distribution of income. If the function was drawn on a double logarithmic scale, then it would be a straight line. The slope of the line would be equal to the negative of the shape parameter, given that

$$\frac{\partial \ln\{P(X > x)\}}{\partial \alpha} = -\alpha \tag{B.2}$$

for $x > x_{\min}$.

B.2 An Interpretation of a Pareto Index

A lemma and a theorem on interpreting the shape parameter of a Pareto distribution (or, equivalently, a Pareto index) are as follows.

Lemma 1. *If personal wealth follows a Pareto distribution with a shape parameter* $\alpha > 1$*, then the share of wealth held by the wealthiest* $100 * p \in [0, 100]$ *percent of people is*

$$p^{1-\frac{1}{\alpha}} \tag{B.3}$$

Proof. Suppose that personal wealth follows a Pareto distribution. The probability density function for that distribution evaluated at a level of wealth w would be

$$\alpha w_{\min}^{\alpha} w^{-1-\alpha} \tag{B.4}$$

for $w \geq w_{\min}$, where $w_{\min} > 0$ is a lower-bound parameter and $\alpha > 0$ is a shape parameter. The percentage of people with wealth greater than a level of wealth w_p would be

$$p \equiv \int_{w_p}^{\infty} \alpha w_{\min}^{\alpha} w^{-1-\alpha} dw = \alpha w_{\min}^{\alpha} \left[\frac{1}{-\alpha} w^{-\alpha} \right]_{w_p}^{\infty} = w_{\min}^{\alpha} w_p^{-\alpha} \tag{B.5}$$

where the shape parameter α was assumed to be strictly greater than unity. So, for a percentage of people $100 * p$, the wealth of the least-wealthy person in the wealthiest $100 * p$ percent of people would be

$$w_p = w_{\min} p^{-1/\alpha} \tag{B.6}$$

The total wealth of everyone—or, at least, the total wealth of the people whose wealth follows the Pareto distribution—would be

$$W \equiv \int_{w_{\min}}^{\infty} wN\left(\alpha w_{\min}^{\alpha} w^{-1-\alpha}\right) dw = N\alpha w_{\min}^{\alpha}\left[\frac{1}{1-\alpha} w^{1-\alpha}\right]_{w_{\min}}^{\infty} = N\frac{\alpha}{\alpha-1} w_{\min} \tag{B.7}$$

where N is the total number of people. As before, the shape parameter α was assumed to be strictly greater than unity. It can be noted, tangentially, that this equation can be rewritten to show that the mean wealth of people with wealth above the minimum level of wealth is proportional to that minimum level of wealth, where the proportion is $\alpha/(\alpha-1)$.

Similarly, the wealth of the wealthiest $100 * p$ percent of people would be

$$W_p \equiv \int_{w_p}^{\infty} wN\left(\alpha w_{\min}^{\alpha} w^{-1-\alpha}\right) dw = N\alpha w_{\min}^{\alpha}\left[\frac{1}{1-\alpha} w^{1-\alpha}\right]_{w_p}^{\infty} = N\frac{\alpha}{\alpha-1} w_{\min} p^{1-\frac{1}{\alpha}} \tag{B.8}$$

where the expression for w_p was substituted into the equation.

The share of wealth held by the wealthiest $100 * p$ percent of people is then

$$W_p/W = p^{1-\frac{1}{\alpha}} \tag{B.9}$$

$\square$

Theorem 1. *(Hardy 2010, proposition 2) If personal wealth follows a Pareto distribution with a shape parameter* $\alpha > 1$ *and the wealthiest* $100 * p \in (0, 50]$ *percent of people own* $(1-p) \in [50, 100)$ *percent of the total wealth of the population, then*

$$\alpha = \log_{p/(1-p)} p \tag{B.10}$$

Proof. If personal wealth follows a Pareto distribution with a shape parameter $\alpha > 1$, then the share of wealth held by the wealthiest $100 * p \in [0, 100]$ percent of people is

$$p^{1-\frac{1}{\alpha}} \tag{B.11}$$

by lemma 1. So, if the wealthiest $100 * p$ percent of people own $100 * (1 - p)$ percent of the population's wealth, then

$$p^{1-\frac{1}{\alpha}} = 1 - p \tag{B.12}$$

Taking the logarithm of both sides and solving for the shape parameter,

$$\alpha = \frac{\ln p}{\ln \{p/(1-p)\}} \tag{B.13}$$

where it was assumed that p is not equal to zero. The right-hand side of this equation can be expressed as the logarithm of p to base $p/(1-p)$ by using the change-of-base rule. $\square$

Note that, as a Pareto distribution of wealth approaches perfect equality with the wealthiest 50 percent of people owning almost half of the population's wealth, the shape parameter approaches positive infinity. Also note that, as a Pareto distribution of wealth approaches perfect inequality with a infinitesimally small percentage of people owning almost all of the population's wealth, the shape parameter approaches unity from above.

The shape parameters for other situations besides perfect equality and perfect inequality include the following. The wealthiest 49 percent of people owning 51 percent of the population's wealth corresponds to a shape parameter of about 17.83; the wealthiest 40 percent owning 60 percent corresponds to about 2.26; the wealthiest 30 percent owning 70 percent corresponds to about 1.42; the wealthiest 20 percent owning 80 percent corresponds to about 1.16; the wealthiest 10 percent owning 90 percent corresponds to about 1.05; and, as a final example, the wealthiest one percent of people owning 99 percent of the population's wealth corresponds to a shape parameter that is almost unity (about 1.002).

B.3 A Relationship between the Pareto and Gini Indexes

For those who prefer to interpret inequality in terms of the Gini index, it can be noted that there is an inverse relationship between the shape parameter of a Pareto distribution (or, again, a Pareto index) and the Gini index.

Theorem 2. *If personal wealth follows a Pareto distribution with a shape parameter $\alpha > 1$, then the Gini index is*

$$\frac{1}{2\alpha - 1} \tag{B.14}$$

Proof. The Gini index can be expressed in different ways, but one expression for the Gini index is as unity minus two times the area under a Lorenz curve (Cowell 2000, p. 112, eq. 26). In the context of the personal distribution of wealth, the Lorenz curve shows the proportions of wealth held by the least-wealthy proportions of the population. If personal wealth follows a Pareto distribution with a shape parameter $\alpha > 1$, then the percentage of wealth held by the poorest $100 * p$ percent of the population is

$$1 - (1 - p)^{\left(1 - \frac{1}{\alpha}\right)} \tag{B.15}$$

by lemma 1 with a slight change of notation. The Gini index is then

$$1 - 2\int_0^1 \left(1 - (1 - p)^{\left(1 - \frac{1}{\alpha}\right)}\right) dp = 1 - 2\left(\frac{\alpha - 1}{2\alpha - 1}\right) = \frac{1}{2\alpha - 1} \tag{B.16}$$

□

Note that this relationship between the Pareto and Gini indexes is an inverse relationship. As a Pareto distribution approaches perfect equality with its shape parameter approaching unity from above, the Gini index approaches unity from below.

Conversely, as the Pareto distribution approaches perfect inequality with its shape parameter approaching infinity from below, the Gini index approaches zero from above. The same relationship can be seen by rewriting the expression given above as $\alpha = (1+G)/2G$ where G denotes the Gini index.

It should be emphasized that this relationship only holds if wealth (or whatever variable is under consideration) follows a Pareto distribution. However, to the extent that wealth is thought to follow a Pareto distribution, it is only thought to follow that particular distribution at the very top of the wealth distribution. Assuming that wealth actually follows a Pareto distribution at the top of the wealth distribution, but assuming that wealth does not follow that distribution elsewhere, the relationship between the Pareto and Gini indexes would only be a relationship between the Pareto index and the within-group Gini index for wealth inequality among those at the top of the wealth distribution.

Interestingly, as shown by Alvaredo (2011), the overall Gini index for the distribution of wealth among a given population can be decomposed into the contributions of inequality among the wealthiest $100{*}p$ percent of the population, inequality among the least-wealthy $100{*}(1-p)$ percent of the population, and inequality between those two groups as

$$(Gps) + [G'(1-p)(1-s)] + (s-p) \tag{B.17}$$

respectively, where G is the within-group Gini index for the distribution of wealth among the former group, s is their share of wealth, and G' is the within-group Gini index for the distribution of wealth among the latter group (ibid., p. 275, eq. 4). Thus, if the wealth of the wealthiest $100 * p$ percent of the population follows a Pareto distribution with a shape parameter $\alpha > 1$, then $G = [1/(2\alpha - 1)]$ and $s = p^{1-\frac{1}{\alpha}}$ by theorem 2 and lemma 1, respectively.

Although it is somewhat off topic, it can be noted that the Gini index for the distribution of wealth among all American households increased from about 83 to 87 percent between the years 1989 and 2010, according to estimates that Edward Wolff made by using the Survey of Consumer Finances (Wolff 2012, p. 50, table 2). That survey is designed to exclude the (households of the) people who appear on *Forbes Magazine*'s list of the 400 wealthiest Americans, however (Kennickell 2006, p. 84). The estimates made by Wolff are therefore actually estimates for the within-group Gini index for the distribution of wealth among all American households, except the 400 wealthiest, if it is assumed that the magazine's list captures the 400 wealthiest households (which may be an incorrect assumption, even if the list captures the 400 wealthiest Americans). Under the same assumption, the Gini index for the distribution of wealth among all American households can be estimated by using the estimates from Wolff (2012), information from the magazine's list, and equation B.17 from above, as well as information on the total number of households, which can be taken from various reports on the Survey of Consumer Finances (Bricker et al. 2012, p. 78, table A.3; Kennickell and Starr 1994, p. 880; Kennickell et al. 2000, p. 27). Making those estimates, the Gini index would be slightly larger in each year between 1989 and 2010, but it would still only increase from about 83 to 87 percent between those years (Author's calculations).

That change in the Gini index can be interpreted as follows. In the context of the household distribution of wealth, twice the Gini index is the average absolute difference between the wealth of each household as a proportion of the average wealth of all households. This interpretation has been pointed out by at least Tony Atkinson (quoted in Raskall and Matheson 1992, p. 11), and the interpretation is obvious once the Gini index is expressed in terms of an average absolute difference (Cowell 2000, p. 111, eq. 25). The mean wealth of all American households was about 464 thousand current dollars in 2010, according to Wolff (2012, p. 49, table 1). If that is taken as their mean wealth, then a Gini index of 87 percent in the year 2010 boils down to an average absolute difference of about 36 thousand dollars more than it would have been if the Gini index was only 83 percent. Thirty-six thousand dollars may seem like a small amount of wealth, but just that amount of wealth would have placed a household somewhere between the wealthiest 20 to 40 percent of households in the year 2010 (based on the figures reported by Kennickell 2012, p. 4, table 2).

B.4 Parameter Estimates for Different Distributions

B.4.1 Pareto Distribution

The probability density function for a Pareto distribution is

$$f(x;\alpha,x_{\min}) \equiv \alpha x_{\min}^{\alpha} x^{-1-\alpha} \tag{B.18}$$

for $x \geq x_{\min}$, where $x_{\min} > 0$ is a lower-bound parameter and $\alpha > 0$ is a shape parameter. Let $x_1, x_2, \ldots, x_n$ be independent samples from the same Pareto distribution. Then the likelihood function is

$$\begin{aligned} L(\alpha, x_{\min}; x_1, x_2, \ldots, x_n) &\equiv \prod_{i=1}^{n} f(x_i;\alpha,x_{\min}) \\ &= \alpha x_{\min}^{\alpha} x_i^{-1-\alpha} \\ &= \alpha^n x_{\min}^{\alpha n} \left(\prod_{i=1}^{n} x_i \right)^{-1-\alpha} \end{aligned} \tag{B.19}$$

(Arnold 1983, p. 194, eq. 5.2.1). The logarithm of that likelihood function is

$$\begin{aligned} \ln L(\alpha, x_{\min}; x_1, x_2, \ldots, x_n) &= \sum_{i=1}^{n} \ln f(x_i;\alpha,x_{\min}) \\ &= n\ln\{\alpha\} + n\alpha\ln\{x_{\min}\} + (-1-\alpha)\sum_{i=1}^{n}\ln\{x_i\} \end{aligned} \tag{B.20}$$

(Arnold 1983, p. 194, eq. 5.2.2). That log-likelihood function is obviously an increasing function of the lower-bound parameter, but the lower-bound parameter cannot be larger than the smallest of the n samples, so the maximum-likelihood estimator for the lower-bound parameter is

$$\hat{x}_{\min} \equiv x_{(n)} \tag{B.21}$$

where $x_{(n)}$ denotes the smallest of the n samples (Arnold 1983, p. 194, eq. 5.2.3). Substituting that estimator into the likelihood function and using a little bit of calculus, the maximum-likelihood estimator for the shape parameter is

$$\hat{\alpha} \equiv n \bigg/ \sum_{i=1}^{n} \ln \left\{ \frac{x_i}{x_{(n)}} \right\} \tag{B.22}$$

(ibid.). It can be noted, tangentially, that the reciprocal of this maximum-likelihood estimator for the shape parameter of a Pareto distribution falls into the class of concave "richness" measures suggested by Peichl et al. (2010), if the "richness line" is defined as the smallest of the n samples.

The maximum-likelihood estimators for the parameters of a Pareto distribution are consistent, but biased (Arnold 1983, pp. 194–5). Unbiased estimators for the lower-bound and shape parameters are

$$\left(1 - \frac{1}{\hat{\alpha}(n-1)}\right) \hat{x}_{\min} \tag{B.23}$$

(Arnold 1983, p. 197, eq. 5.2.18) and

$$\left(\frac{n-2}{n}\right) \hat{\alpha} \tag{B.24}$$

(Arnold 1983, p. 196, eq. 5.2.13), respectively. Monte Carlo simulations suggests that these biased-corrected maximum-likelihood estimators perform better along a range of dimensions than a range of other estimators, even for samples that are an order of magnitude smaller than those considered in this dissertation (i.e., sample sizes closer to 40 than 400; Rahman and Pearson 2003).

The variances of the bias-corrected maximum-likelihood estimators for the lower-bound and shape parameters are

$$\frac{x_{\min}^2}{\alpha(n-1)(\alpha n-2)} \tag{B.25}$$

(Arnold 1983, p. 197, eq. 5.2.19) and

$$\frac{\alpha^2}{n-3} \tag{B.26}$$

(Arnold 1983, p. 196, eq. 5.2.16), respectively.

B.4.2 Truncated Pareto Distribution

The probability density function for an upper-truncated Pareto distribution is

$$f(x;\alpha,x_{\min},x_{\max}) \equiv \frac{\alpha x_{\min}^{\alpha} x^{-1-\alpha}}{1-(x_{\min}/x_{\max})^{\alpha}} \tag{B.27}$$

for $x \geq x_{\min}$, where $x_{\min} > 0$ is a lower-bound parameter, $x_{\max} > x_{\min}$ is an upper-truncation parameter, and $\alpha > 0$ is a shape parameter (Aban et al. 2006, p. 271, eq. 2). Note that, as the upper-truncation parameter approaches infinity, the probability density function for a truncated Pareto distribution (again, eq. B.27) approaches the probability density function for an untruncated Pareto distribution (eq. B.18). The maximum-likelihood estimators for the lower-bound and truncation parameters are

$$\hat{x}_{\min} \equiv x_{(n)} \tag{B.28}$$

and

$$\hat{x}_{\max} \equiv x_{(1)} \tag{B.29}$$

respectively, where $x_{(n)}$ denotes the smallest of the n samples and $x_{(1)}$ denotes the largest (Aban et al. 2006, p. 271, theorem 2). The maximum-likelihood estimator for the shape parameter, which can be denoted by $\hat{\alpha}$, must be found by using numerical methods, but it can be found by solving the following equation.

$$\frac{n}{\hat{\alpha}} + \frac{n(x_{(n)}/x_{(1)})^{\hat{\alpha}} \ln\{x_{(n)}/x_{(1)}\}}{1-(x_{(n)}/x_{(1)})^{\hat{\alpha}}} - \sum_{i=1}^{n} \ln\left\{\frac{x_{(i)}}{x_{(n)}}\right\} = 0 \tag{B.30}$$

(ibid.). The maximum-likelihood estimators for the lower-bound and truncation parameters are consistent (Aban et al. 2006, p. 271), but biased (Zhang 2013, theorem 1). Unbiased estimators for the lower-bound and truncation parameters were derived analytically by Zhang (2013) as

$$x_{(n)}\left(1+\frac{(x_{(n)}/x_{(1)})^{\hat{\alpha}}-1}{n\hat{\alpha}}\right) \tag{B.31}$$

and

$$x_{(1)}\left(1+\frac{\ln\{x_{(1)}/x_{(n)}\}}{n}\right) \tag{B.32}$$

respectively (Zhang 2013, eq. 3). Given that the maximum-likelihood estimator for the shape parameter must be found numerically, a correction for any asymptotic bias cannot be derived analytically, but Maschberger and Kroupa (2009) suggest simply using

$$\left(\frac{n-3}{n}\right)\hat{\alpha} \tag{B.33}$$

where, again, $\hat{\alpha}$ is the solution to equation B.30. This adjusted for any bias was suggested because it is similar to the correction for the bias of the maximum-likelihood estimator for the shape parameter of an untrucated Pareto distribution (where three rather than two is subtracted to account for the additional parameter in the truncated case; Maschberger and Kroupa 2009, p. 3).

B.4.3 Truncated Log-Normal Distribution

If the logarithm of a random variable is normally distributed, then the random variable is log-normally distributed. The probability density function for a (two-parameter) log-normal distribution is

$$f(x;\mu,\sigma) \equiv \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-\left(\ln\{x\}-\mu\right)^2}{2\sigma^2}\right\} \tag{B.34}$$

for $x > 0$, where $\mu \in (-\infty, \infty)$ and $\sigma > 0$ are the mean and standard deviation of the logarithm of the random variable, respectively. The cumulative distribution function for the log-normal distribution is

$$F(x;\mu,\sigma) \equiv \frac{1}{2}\left(1 + \operatorname{erf}\left\{\frac{\ln\{x\}-\mu}{\sqrt{2\sigma^2}}\right\}\right) \tag{B.35}$$

where

$$\operatorname{erf}\{z\} \equiv \frac{2}{\sqrt{\pi}} \int_0^z \exp\left\{-t^2\right\} dt \tag{B.36}$$

is the error function.

If a log-normal distribution is truncated from below at a known truncation point τ, then the probability density function for the truncated log-normal distribution is

$$f(x;\mu,\sigma|x>\tau) = \frac{f(x;\mu,\sigma)}{1-F(\tau;\mu,\sigma)} \tag{B.37}$$

Let $x_1, x_2, \ldots, x_n > \tau$ be independent samples from the same truncated log-normal distribution. Then the likelihood function is

$$L(\mu, \sigma; x_1, x_2, \ldots, x_n) \equiv \prod_{i=1}^{n} \frac{f(x_i; \mu, \sigma)}{1 - F(\tau; \mu, \sigma)} \tag{B.38}$$

The logarithm of that likelihood function is

$$\ln L(\mu, \sigma; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \left[\ln f(x_i; \mu, \sigma) - \ln \{1 - F(\tau; \mu, \sigma)\}\right] \tag{B.39}$$

where

$$\ln f(x_i; \mu, \sigma) = \ln \left\{ \frac{1}{x_i \sqrt{2\pi\sigma^2}} \right\} - \frac{\left(\ln\{x_i\} - \mu\right)^2}{2\sigma^2} \tag{B.40}$$

The maximum-likelihood estimators for the mean and standard-deviation parameters can be found by using numerical methods.

B.4.4 Truncated Gamma Distribution

The probability density function for a (two-parameter) gamma distribution is

$$f(x; \alpha, \beta) \equiv \frac{(x/\beta)^{\alpha-1} \exp\{-x/\beta\}}{\beta \Gamma\{\alpha\}} \tag{B.41}$$

for $x > 0$, where $\alpha > 0$ is a shape parameter, $\beta > 0$ is a scale parameter, and $\Gamma\{\cdot\}$ is the gamma function. The cumulative distribution function for the gamma distribution is

$$F(x; \alpha, \beta) \equiv \frac{\gamma\{\alpha, x/\beta\}}{\Gamma\{\alpha\}} \tag{B.42}$$

where $\gamma\{\cdot\}$ is the lower incomplete gamma function.

If a gamma distribution is truncated from below at a known truncation point τ, then the probability density function for the truncated gamma distribution is

$$f(x; \alpha, \beta | x > \tau) = \frac{f(x; \alpha, \beta)}{1 - F(\tau; \alpha, \beta)} \tag{B.43}$$

Let $x_1, x_2, \ldots, x_n > \tau$ be independent samples from the same truncated gamma distribution. Then the likelihood function is

$$L(\alpha, \beta; x_1, x_2, \ldots, x_n) \equiv \prod_{i=1}^{n} \frac{f(x_i; \alpha, \beta)}{1 - F(\tau; \alpha, \beta)} \tag{B.44}$$

The logarithm of that likelihood function is

$$\ln L(\alpha, \beta; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} [\ln f(x_i; \alpha, \beta) - \ln \{1 - F(\tau; \alpha, \beta)\}] \tag{B.45}$$

where

$$\ln f(x_i; \alpha, \beta) = (\alpha - 1) \ln \{x_i\} - \alpha \ln \{\beta\} - (x_i/\beta) - \ln \{\Gamma \{\alpha\}\} \tag{B.46}$$

The maximum-likelihood estimators for the shape and scale parameters can be found by using numerical methods.

APPENDIX C

APPENDIX TO THE THIRD ESSAY

C.1 Movements Onto, Off of, and Throughout the List

Figure 4.1 in the text of the essay shows, for each year between 1982 to 2013, the number of people on *Forbes Magazine*'s list in one year who appeared again or dropped off due to either death or decline by the next year. The tables that follow show the same information in numerical rather than graphical form, as well as some additional information about the number of people who moved onto, off of, or throughout the list over time.

Table C.1 shows, for each year between 1982 and 2013, the number of people on the magazine's list in one year who dropped off the list by the next year ("dropouts"). The table also shows the number of those people who dropped off because they died ("decedents"), renounced their American citizenship ("renunciants"), or experienced an absolute or relative decline their wealth ("other dropouts").

Table C.2 shows, for the same years, the number of people who came onto the list between one year and the next ("entrants"). The people who came onto the list were either appearing for the first time ("first-timers") or reappearing after dropping off in an earlier year ("returnees"), so the table also shows the number of people in those two groups. Everyone who appeared on the first year of the magazine's list was new to the list and appearing for the first time, of course, while everyone who was new in its second year must not have been on the list in its first.

Finally, table C.3 shows, for the same years, the number of people who were on the list in a given year and still on the list in the next year ("incumbents"). The table also shows the number of incumbents who moved up the wealth rankings by becoming wealthier in absolute or at least relative terms ("up in rank"), the number who became poorer and moved down the rankings ("down in rank"), and the number who stayed where they were in the rankings between the given year and the next ("same rank"). Note that someone could have only moved so far down the rankings before dropping off the list altogether.

The tables in this section suggest that, for years associated with asset-market booms like the stock market boom in the late 1990s, there tended to be more people who appeared on the list for the first time, more people who dropped off, and more incumbents who moved up in rank. For years associated with asset-market busts like the dot-com crash and the recent recession, on the other hand, there tended to be fewer first-timers and fewer incumbents dropping off, but also fewer incumbents moving up in rank. Whether someone on the list is more or less likely to drop off in certain years after controlling for their characteristics is one of the central questions explored by the essay.

The tables also suggest that, aside from years associated with such asset-market booms and busts, mobility among the wealthiest Americans has been fairly stable and arguably low over recent decades. Between any one year and the next, most people have remained on the list. On average, only about 10 percent of the 400 people have dropped off per year. Most dropped off due to decline, while some dropped off due to death, and only rarely did anyone drop off because they renounced their American citizenship. Among the people who came onto the list to replace the dropouts, most came onto the list for the first time, but some of them were also returning after dropping off in an earlier year.

It should be noted that it is difficult to assess whether mobility among the 400 wealthiest Americans, according to *Forbes Magazine*'s list, is consistent with mobility among other wealthy groups, according to panel surveys. A panel survey conducted as part of the Survey of Consumer Finances (SCF) suggests that, between 2007 to 2009, the amount of turnover among the wealthiest one percent of households was consistent with the amount of turnover among the people on *Forbes Magazine*'s list between the same years (Kennickell 2011, pp. 10, 15). It is unclear whether turnover between other years is consistent, however, given that the SCF is typically only a cross-sectional survey.

Evidence from the Panel Survey of Income Dynamics (PSID) suggests that the amount of turnover among the wealthiest groups in its survey is not inconsistent with the amount of turnover among the people on the magazine's list. The survey suggests that most of the wealthy remain wealthy. (See Diaz-Gimenez et al. 2011, pp. 27–28, on inter-quintile movements between 2001 and 2007; also see Hurst et al., 1998, pp. 282–87, on inter-decile movements between 1984 and 1994). Yet the wealthiest groups in the PSID are often relatively poor (Juster et al. 1999), so it is unclear whether turnover among those groups should be compared to turnover among the people on the magazine's list.

Table C.1. Number of Dropouts, 1982–2013

Year	Decedents	Renunciants	Other dropouts	Total
1982–1983	13	0	61	74
1983–1984	6	0	38	44
1984–1985	7	0	57	64
1985–1986	7	0	49	56
1986–1987	8	0	72	80
1987–1988	4	0	48	52
1988–1989	5	0	55	60
1989–1990	6	0	38	44
1990–1991	6	0	41	47
1991–1992	7	0	36	43
1992–1993	13	0	34	47
1993–1994	10	1	38	49
1994–1995	13	2	30	45
1995–1996	10	0	43	53
1996–1997	6	0	37	43
1997–1998	5	0	42	47
1998–1999	7	0	63	70
1999–2000	2	0	53	55
2000–2001	3	0	51	54
2001–2002	4	0	31	35
2002–2003	11	0	20	31
2003–2004	6	0	48	54
2004–2005	8	0	34	42
2005–2006	8	0	34	42
2006–2007	7	0	51	58
2007–2008	6	0	33	39
2008–2009	6	0	33	39
2009–2010	9	0	26	35
2010–2011	3	0	21	24
2011–2012	7	1	24	32
2012–2013	6	0	28	34
Average	7	0	41	48

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows, for the years 1982 to 2013 and as an average across those years, the number of people who dropped off *Forbes Magazine*'s list of the 400 wealthiest Americans between one year of that list and the next year because they died ("decedents"), renounced their American citizenship ("renunciants"), or experienced a decline in their wealth ("other dropouts"). The total number of people who dropped off between one year and the next is also shown.

Table C.2. Number of Entrants, 1982–2013

Year	First-timers	Returnees	Total
1982–1983	74	0	74
1983–1984	35	9	44
1984–1985	41	23	64
1985–1986	51	5	56
1986–1987	60	20	80
1987–1988	36	16	52
1988–1989	35	25	60
1989–1990	27	17	44
1990–1991	27	20	47
1991–1992	30	13	43
1992–1993	37	10	47
1993–1994	34	15	49
1994–1995	39	6	45
1995–1996	40	13	53
1996–1997	31	12	43
1997–1998	36	11	47
1998–1999	60	10	70
1999–2000	46	9	55
2000–2001	25	29	54
2001–2002	16	19	35
2002–2003	15	16	31
2003–2004	45	9	54
2004–2005	33	9	42
2005–2006	28	14	42
2006–2007	46	12	58
2007–2008	32	7	39
2008–2009	19	20	39
2009–2010	17	18	35
2010–2011	18	6	24
2011–2012	21	11	32
2012–2013	20	14	34
Average	35	13	48

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows, for the years 1982 to 2013 and as an average across those years, the number of people who came onto *Forbes Magazine*'s list of the 400 wealthiest Americans between one year of that list and the next year for the first time since its inception in 1982 ("first-time") or after an absence ("returnees"). The total number of people who came onto the list between one year and the next is also shown.

Table C.3. Number of Incumbents, 1982–2013

Year	Up in rank	Down in rank	Same rank	Total
1982–1983	184	141	1	326
1983–1984	215	138	3	356
1984–1985	217	116	3	336
1985–1986	184	153	7	344
1986–1987	187	128	5	320
1987–1988	174	165	9	348
1988–1989	224	113	3	340
1989–1990	151	198	7	356
1990–1991	163	187	3	353
1991–1992	155	196	6	357
1992–1993	204	142	7	353
1993–1994	192	151	8	351
1994–1995	181	167	7	355
1995–1996	221	120	6	347
1996–1997	191	162	4	357
1997–1998	176	172	5	353
1998–1999	217	106	7	330
1999–2000	227	117	1	345
2000–2001	76	260	10	346
2001–2002	120	237	8	365
2002–2003	229	130	10	369
2003–2004	201	131	14	346
2004–2005	231	121	6	358
2005–2006	202	149	7	358
2006–2007	207	124	11	342
2007–2008	171	180	10	361
2008–2009	165	187	9	361
2009–2010	207	149	9	365
2010–2011	206	166	4	376
2011–2012	193	162	13	368
2012–2013	222	128	16	366
Average	190	155	7	352

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows, for the years 1982 to 2013 and as an average across those years, the number of people who stayed on *Forbes Magazine*'s list of the 400 wealthiest Americans between one year of that list and the next year while seeing their rank in the distribution of wealth go up, down, or stay the same. The total number of people who stayed on the list between one year and the next is also shown.

C.2 Expressions for the Simple Estimates

The text of the essay discusses a simple estimate for the probability that someone would appear on *Forbes Magazine*'s list again after appearing for a given number of consecutive years. The text also discussed a related estimate for the probability that someone would appear again and again for a given number of consecutive years. Those estimates can be rediscussed with more notation as follows.

Looking at any given year of *Forbes Magazine*'s list, each person on the list in the given year will have been on the list for some number of consecutive years. Some will appear again in the next year, while others will not. Looking across each year of the magazine's list, we can add up the total number of (possibly non-unique) people who appeared for some number of consecutive years t. Let n_t denote that number of people. For people on any year of the list except its most-recent year, we know whether they appeared again or not. Let d_t denote the number of people who dropped off the list after their t-th consecutive year on it. For people on the most-recent year of the list, we do not yet know whether they will appear again or not. Let c_t denote the number of those right-censored people.

A simple estimate of the probability of that someone will appear again in the next year after appearing for a given number of consecutive years t is

$$p_t \equiv 1 - \left(\frac{d_t}{n_t - c_t} \right) \tag{C.1}$$

which is one less the number of people who appeared again after appearing for a given number of consecutive years all over the number of people who either appeared again or dropped off after the given number of years. That estimate is simplistic in the sense that it assumes that, conditional on the number of years that someone has already appeared, everyone appears again or drops off with the same probability.

Using those estimates for the probability of appearing again after a given number of consecutive years, the probability that someone will appear again and again for given number of consecutive years T can be estimated as the product

$$\prod_{t=1}^{T} p_t \tag{C.2}$$

That estimate is simplistic in the sense that it assumes that everyone has the same probability of appearing on the magazine's list again and again for a given number of consecutive years. The estimate is akin to the popular Kaplan-Meier or "product-limit" estimator (Hollander and Wolfe 1999, pp. 535–50) and similarly simple non-parametric estimators for survival functions.

C.3 Tests of the IIA Assumption

In the model discussed in the text, a person's probability of dropping off *Forbes Magazine*'s list due to a decline in his or wealth rather than appearing again, his or her probability of dropping off due to death rather than appearing again, and his or her probability of dropping off due to decline rather than death were all assumed to be independent of the possibility of the remaining alternative. That is the independence of irrelevant alternatives (IIA) assumption of any multinomial logit model.

The IIA assumption can be tested by applying a popular test suggested by Hausman and McFadden (1984). The intuition behind the test is that, if the IIA assumption is correct, then the estimated effects of the covariates on the relative probability of any two alternatives should not change too much if the observations associated with the remaining, supposedly irrelevant alternative were ignored. So, for example, the estimates from a binomial logit model for the probability of dropping off due to decline rather than appearing again should not be too different from the comparable estimates from the trinomial logit model discussed in the text that also models the probability of the supposedly irrelevant alternative of dropping off due to death rather than appearing again.

The statistic for the test is as follows. If $\hat{\beta}^{full}$ and $\hat{\beta}^{restricted}$ denote the estimates of the comparable coefficients from a full, trinomial logit model that models all of the alternatives and a restricted, binomial logit model that ignores one of the alternatives, respectively, and if $\hat{V}^{full}$ and $\hat{V}^{restricted}$ denote the estimated variance-covariance matrices of the comparable coefficients, then the test statistic is

$$\left(\hat{\beta}^{restricted} - \hat{\beta}^{full}\right)' \left(\hat{V}^{restricted} - \hat{V}^{full}\right)^{-1} \left(\hat{\beta}^{restricted} - \hat{\beta}^{full}\right) \tag{C.3}$$

which should be asymptotically distributed as a Chi-squared distribution with degrees of freedom equal to number of comparable coefficients (Hausman and McFadden 1984). That statistic will obviously be zero if the estimated coefficients do not change when one of the alternatives is ignored. If the statistic is too large, then the IIA assumption is rejected.

Given that our model has three alternatives, the test can be performed in three different ways by ignoring different alternatives, although, at least asymptotically, the test should yield the same result, regardless of the way in which it is performed. One way to perform the test is to ignore the observations associated with dropping off due to decline. When those observations are ignored and a binomial logit model is estimated, the change in the estimated parameters yields a test statistic that is a negative number (specifically, about -5.22).

Although the test statistic should be non-negative asymptotically, given that a Chi-squared distribution does not take on negative values, negative values are not atypical with either real or simulated data (Cheng and Long 2007; Vijverberg 2011; etc.). Such values are typically interpreted as a failure to reject the IIA assumption (Vijverberg 2011, p. 39, table 1). Most notably, Hausman and McFadden (1984,

p. 1226) interpret negative values in that way. Interpreted in that way, the test suggests that the probability of dropping off due to death rather than appearing again is independent of the possibility of dropping off due to decline.

A similar conclusion is drawn if the observations associated with dropping off due to death are ignored. If those observations are ignored, then the test statistic is negative again (specifically, about −7.91), which again suggests that the probability of dropping off due to decline rather than appearing again is independent of the possibility of dropping off due to death, if a negative test statistic is interpreted in the same way as before.

The test statistic is also negative (specifically, about −710.35) if the observations associated with appearing again are ignored. Thus the probability of dropping off due to death rather than decline appears to be independent of the possibility of appearing again. It should be noted that, by ignoring the observations in which a person appears again, the number of observations is reduced by almost an order of magnitude from about 12,400 person-year observations to 1,488 observations. With a smaller number of observations, the uncertainty associated with any given estimate should tend to be larger, so we would not expect to find statistically significant differences between the estimates of comparable coefficients. All three versions of the test therefore imply that the IIA assumption cannot be rejected.

There is however an important caveat to that conclusion. We have applied the Hausman-McFadden test because the possible violation of the IAA assumption is so well known and the popularity of that test is so widespread that any well-trained economist would expect the test to be applied. Yet scholars who have used Monte Carlo simulations to study whether the Hausman-McFadden test actually tests what it claims to test conclude that any inferences drawn from that test should be completely ignored (Cheng and Long 2007; Vijverberg 2011).

The intuition behind the test is nevertheless appealing, so it can be noted that, even when a supposedly irrelevant alternative is ignored, the relative risk of one alternative rather than another does not seem to change by a substantial amount. For a person with a median age and median rank who comes onto the list for the first time in 1996, for example, the trinomial model suggests that he or she is about 418 times more likely to appear on the list again rather than drop off due to death, while a binomial model that ignores the possibility of dropping off due to decline suggests that the person is about 426 times more likely to do so. The trinomial model also suggests that the person is about 20.2 times more likely to appear again than drop off due to decline, while a binomial model that ignores the possibility of dropping off due to death suggests that he or she is about 20.3 times more likely to do so. And the person is also much more likely to drop off the list due to decline than death, according to both the trinomial model and a binomial model that ignores the possibility of appearing on the list again in the next year, although the trinomial model suggests that he or she is about 21 times more likely to drop off due to decline, while the binomial model suggests that he or she is about 129 times more likely to do so. In each case, the relative risks are qualitatively similar and, except for the latter case, quantitatively similar.

C.4 Estimates for the Extensions

Table 4.1 in the text shows estimates for the model that was subsequently extended, but similar tables were not shown for the extensions to that model. The following tables show estimates for the extensions to the model. Recall that one extension included a dummy variable for whether someone was philanthropic. Another extension included a dummy variable for whether someone's wealth was self-made. And yet another extension included a dummy variable for whether someone's wealth was made in the finance, insurance, or real estate (FIRE) industry, as well as interaction terms between that dummy and year-specific dummies. All of the extensions were restricted to the years since 1996 because we do not have information on those dummy variables in earlier years. *Forbes Magazine* did not start systematically reporting the source of someone's wealth until 1996. That was also the first year for which *Slate* constructed its list of the most-generous Americans.

Table C.4. Multinomial Logit Model with Philanthropic Dummies

	Competing risks	
Covariates	Decline	Death
Constant	-4.24^{***}	-6.56^{*}
	(0.92)	(3.47)
Age	-0.06^{**}	-0.10
	(0.03)	(0.09)
Age squared	0.00^{**}	0.00^{**}
	(0.00)	(0.00)
Rank	0.02^{***}	-0.00
	(0.00)	(0.00)
Left-censored dummy	-0.57^{**}	0.26
	(0.28)	(0.34)
Philanthropic dummy	-0.04	-0.34
	(0.15)	(0.27)
Year dummies	Yes	Yes
Duration dummies	Yes	Yes
Number of observations	6,798	
Percent correctly predicted	89.19	
Log-likelihood	$-1,927.46$	
Test for non-constant variables	$1,392.83^{***}$	
Test for philanthropic dummies	1.68	

Source: Data adapted from *Chronicle of Philanthropy*; *Forbes Magazine* (1982–2013); and *Slate*.

Note: This table shows maximum-likelihood estimates for the model discussed in the text that was essentially the same as the model estimated in table 4.1, except that it included a dummy variable that reflects people who were one of the wealthiest Americans in a given year according to *Forbes Magazine*'s list and also one of the most-generous Americans in that or any other earlier year according to either the Philanthropy 50 list (for the years since 2000) or the Slate 60 list (for the years between 1996 and 1999).

$^{*}p < .10$ $^{**}p < .05$ $^{***}p < 0.01$

Table C.5. Multinomial Logit Model with Self-made Dummies

	Competing risks	
Covariates	Decline	Death
Constant	-4.44^{***}	-6.05^{*}
	(0.93)	(3.45)
Age	-0.06^{**}	-0.11
	(0.03)	(0.09)
Age squared	0.00^{*}	0.00^{***}
	(0.00)	(0.00)
Rank	0.02^{***}	-0.00
	(0.00)	(0.00)
Left-censored dummy	-0.57^{**}	0.09
	(0.28)	(0.34)
Self-made dummy	0.22^{**}	-0.57^{***}
	(0.11)	(0.22)
Year dummies	Yes	Yes
Duration dummies	Yes	Yes
Number of observations	6,798	
Percent correctly predicted	89.19	
Log-likelihood	$-1,922.51$	
Test for non-constant variables	$1,402.74^{***}$	
Test for self-made dummies	11.60^{***}	

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows maximum-likelihood estimates for the model discussed in the text that was essentially the same as the model estimated in table 4.1, except that it included a dummy variable for whether someone's wealth was self-made or not.

$^{*}p < .10$ $^{**}p < .05$ $^{***}p < 0.01$

Table C.6. Multinomial Logit Model with FIRE Industry Dummies

	Competing risks	
Covariates	Decline	Death
Constant	-4.35^{***}	-6.56^{*}
	(0.94)	(3.51)
Age	-0.05^{*}	-0.10
	(0.03)	(0.09)
Age squared	0.00^{*}	0.00^{**}
	(0.00)	(0.00)
Rank	0.02^{***}	-0.00
	(0.00)	(0.00)
Left-censored dummy	-0.56^{*}	0.28
	(0.28)	(0.34)
FIRE industry dummy	-0.81^{***}	-0.52
	(0.19)	(0.51)
Year dummies	Yes	Yes
Duration dummies	Yes	Yes
Industry-year interaction terms	Yes	Yes
Number of observations	6,798	
Percent correctly predicted	89.50	
Log-likelihood	$-1,908.35$	
Test for non-constant variables	$1,431.05^{***}$	
Test for FIRE industry dummies and industry-year interaction terms	39.91^{**}	

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows maximum-likelihood estimates for the model discussed in the text that was essentially the same as the model estimated in table 4.1, except that it included a dummy variable for whether someone's wealth was made in the FIRE industry or not and, also, interaction terms between that dummy and year dummies.

$^{*}p < .10$ $^{**}p < .05$ $^{***}p < 0.01$

C.5 A Thought Experiment on Wealth Deaccumulation

While we might simply assume that a wealthy person could become less wealthy if he or she wanted to, we should think about the ways in which someone could actually do so if they wanted. To that end, suppose that someone has somehow accumulated enough wealth to make himself or herself wealthier than anyone else. That wealthy person can be called Mr. Moneybags. It has sometimes been suggested that the wealthy accumulate wealth as an end unto itself, but suppose that Mr. Moneybag wants to deaccumulate wealth as an end unto itself. Reasons why someone might want to deaccumulate wealth are noted below, but, for now, suppose that Mr. Moneybags merely wants to deaccumulate wealth as an end unto itself.

There are ways in which he could try to reduce his wealth. Some assets can be destroyed. Cash can be burned, for example. So, if at least some of his assets were already held in a destructible form like cash, then Mr. Moneybags could destroy those assets and thereby reduce his wealth. Yet, if Mr. Moneybags destroyed some of his cash, then the rate of interest on his remaining cash should rise, at least from a loanable funds perspective. Burning a dollar would therefore decrease Mr. Moneybag's wealth by a dollar, all other things being equal, but it might not decrease his wealth by a full dollar if the interest rate rose, unless he burned any additional interest income, although he would also need to burn any additional interest income that was generated by burning that dollar, and so forth (and likewise for any other destroyable assets that yield a return determined by supply and demand). The destruction of assets would nevertheless seem to be an effective way for Mr. Moneybags to reduce his assets and thereby reduce his wealth.

Other ways in which Mr. Moneybags could try to reduce his wealth may also be effective. Suppose that instead of destroying assets, Mr. Moneybags simply transfers them to some other individual or organization. If at least some of his assets were already held as cash, then he could drop his cash out of a helicopter, for example. Such a helicopter drop of cash would reduce his wealth, at least immediately. Yet, if Mr. Moneybags owned a business that manufactured bags for holding money, and if there was an increase in the demand for such bags after he started dropping cash out of a helicopter, then a fraction of each dollar he dropped might return to him. The same thing would be true, if he owned any sort of business that benefited from broadly shared wealth.[1] In order to reduce his wealth by a dollar after dropping a dollar out of a helicopter, he would also need to drop any cents that return to him out of a helicopter, and so forth. Again, however, transferring assets would seem to be an effective way for Mr. Moneybags to reduce his wealth.

Another seemingly effective way to reduce his wealth would be to take on debt. Taking on debt and then eventually repaying that debt would be equivalent to simply transferring whatever the interest payments would have been (usurious or otherwise), at least up until the point where the wealth of Mr. Moneybags would turn negative with his debts exceeding his assets. Whereas Mr. Moneybags could never reduce his wealth below zero by destroying or transferring his assets, he could reduce his wealth below zero by taking on enough debt.

Taking on debt, transferring assets, or destroying assets would therefore seem to be effective ways for Mr. Moneybags to reduce his wealth. Of course, except when they are at their most euphoric, it is uncommon for bankers to extend a loan simply so that someone can take on debt. The intentional destruction of one's assets is also uncommon, except perhaps among those who either attend potlatch rituals or reenact the actions of protagonists from Ayn Rand novels. Transferring assets is more common. The wealthy transfer assets to their children and other charity cases all

[1]To take a more realistic example, Bill Gates might get some of his money back if he bought a personal computer for every child in Africa and then some of those children or their parents went out and bought Microsoft products.

the time, for example. A "warm glow" or some other form of utility associated with such transfers may be at least one reason why someone would want to deaccumulate wealth as more than an end unto itself (Andreoni 1990).

Other ways in which Mr. Moneybags could try to reduce his wealth may be less effective. Assuming that at least some of his assets were already held in cash or another liquid form, Mr. Moneybags could use those assets to buy goods and services. Unlike a pure transfer of wealth, in which he would exchange his wealth for nothing in return except perhaps a warm glow or some other form of utility, Mr. Moneybags would get a good or service back in exchange for his wealth. If a good or service was non-durable, then Mr. Moneybags could consume it, and his wealth would be reduced by the cost of the item. If he merely wanted to deaccumulate wealth as an end unto itself, then there would presumably be no limit to the amount of non-durable goods and services that he could buy. Yet, as pointed out by Carroll (2000, pp. 476–7), if someone actually wanted to consume what they bought, and if someone was only willing to pay so much for a particular good and service, then there would presumably be a limit to the amount non-durable goods and services that anyone could buy and consume.

Durable goods could also be bought, but such goods are assets that add to one's wealth (or, at least, they could be counted as such). If Mr. Moneybags bought a helicopter, for example, then his wealth might not decline by the full cost of that helicopter, depending on its resale value. The wealth of Mr. Moneybags might even increase if the mere fact that he owned a helicopter increased the demand for and thereby the value of the helicopter. That said, Mr. Moneybags might be able to reduce his wealth through either durable or non-durable consumption. Such consumption would be another reason to deaccumulate wealth as more than an end unto itself.

Some of the above-discussed ways in which Mr. Moneybags could reduce his wealth depended on him already having some of his assets in cash or another liquid or destructible form. Transforming his wealth into that form might be difficult, however. Suppose that most of his wealth was initially held in the stock of a specific company. One difficulty that might arise is that, even if there were market makers or other market participants who wanted to buy the stock, trying to sell too much stock too quickly might cause the price of the stock—or maybe even the entire market for the stock—to collapse (as Einstein 2000 pointed out in the context of Bill Gates trying to sell off all of his stock in Microsoft).

Also, if market participants did not know that Mr. Moneybags was selling his stock simply so that he could deaccumulate wealth, then the market participants might presume that Mr. Moneybags had insider information (as Einstein 2000 also pointed out). Again, the price of the stock might collapse. If Mr. Moneybags merely wanted to deaccumulate wealth, such a collapse would be welcomed; but, if someone was actually trying to transform those assets into another form that could be used for consumption, charity, or what not, then he or she would obviously wish to avoid such a collapse. Someone trying to sell off their stock in a specific company would therefore need to convince market participants that he or she was selling off stock, but not at any price and only for reasons unrelated to the company. A new-found desire to help the less fortunate might be one reason.

APPENDIX D

APPENDIX ON THE DATA

D.1 Collecting and Processing the Data

The primary source of data used by this dissertation is *Forbes Magazine*'s annual list of the 400 wealthiest Americans. That list was first published in 1982, and it was published as recently as 2013, as of writing (*Forbes Magazine* 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013). The author of this dissertation was informed that the magazine does not share its data on the 400 wealthiest Americans (personal communication on June 17, 2009, with Tanya Prive, circulation assistant at *Forbes Magazine*). The magazine obviously does share its data through its print publication and its online presence. The data has also been shared as part of a book (Bernstein and Swan 2007), as discussed below. The magazine apparently does not share its data in a less-dispersed form, however. Thus, in order to study the magazine's list, the data must be collected and processed from the print or online versions of the magazine. The author of this dissertation collected and processed the data himself.

Previous researchers who have studied the magazine's list have also apparently had to collect and process the data themselves. For example: An economist at the Federal Reserve, Arthur Kennickell, apparently collected and processed data on the magazine's list for the years 1989, 1992, and 1995, as well as for each year between the years 1998 and 2002 (Kennickell 2006). Kennickell appears to have collected and processed the data himself because, while presenting at a conference, he said, "This [i.e., the years around 1989, 1992, and 1995] is when the data was only available in the magazine, [so] I had to type them into the computer, and I was not going to do [that for] every year; but, since then, I can get [the data] off the internet, so I have an annual series beyond that [i.e., in the years 1998 to 2002]" (Kennickell 2003).

As another example: The authors of Klass et al. (2006) and their research assistant apparently collected and processed the data from print or online versions of the magazine for the years 1988 to 2003. The corresponding author of that paper was kind enough to share their dataset with the author of this dissertation (personal communication on June 19, 2009, with Moshe Levy, corresponding author of Klass et al. 2006), although the author of this dissertation independently collected and processed data for those years, as discussed in one of the essays in this dissertation, and some errors were identified in their dataset, as discussed in the same essay.

The main difficulty with collecting and processing the magazine's data is that the same people have been listed under slightly different names in different years. Bill Gates was "William Henry Gates III" when he first appeared on the list in 1986 (*Forbes Magazine* 1986, p. 173), but he was just "Bill Gates" by 2013 (*Forbes Magazine* 2013, p. 124), to take but one example. For any longitudinal analysis, identifying unique individuals across time is obviously important. The author of this dissertation tried to ensure that he identified individuals across time by attempting to replicate a list of all the unique individuals who appeared on the magazine's list in any year between 1982 and 2006 (Bernstein and Swan 2007, pp. 331–60) and in both 1982 and 2012 (Kilachand 2012). Those attempted replications are discussed in the next two sections, before discussing a few other issues that should be addressed before trying to study the data.

D.2 Replicating Bernstein and Swan's (2007) Table

There is a book entitled *All the Money in the World: How the Forbes Four Hundred Make and Spend Their Fortunes* (Bernstein and Swan 2007). That book was published to coincide with the 25th anniversary of *Forbes Magazine*'s annual list of the 400 wealthiest Americans, and it was created in collaboration with the magazine (see Bernstein and Swan 2007, p. vii, on the collaboration).

As an appendix to the book, there is a table entitled "The Forbes Four Hundred, 1982–2006" (Bernstein and Swan 2007, pp. 331–60). That table is supposed to include all of the unique individuals who appeared on the magazine's list between 1982 and 2006 (ibid., p. 331). The table also includes, for each supposedly unique individual, his or her peak wealth (in millions of current dollars) and the year of his or her peak wealth (or, if there were multiple years at the peak wealth, the most-recent year of the peak wealth).

The author of this dissertation tried to replicate that book's table. The following discrepancies were identified during the attempted replication.

One discrepancy was in the number of unique individuals who appeared on the magazine's list over its first 25 years. According to the book's table, there were 1,302 unique individuals who appeared. The author identified only 1,301 unique individuals, or one fewer.

This discrepancy arose because the book inadvertently included nine individuals who should have been excluded, and inadvertently excluded eight individuals who should have been included.

One of the nine individuals who should have been excluded, Stephen C. Hilbert, should have been excluded because he did not appear on the magazine's list in any year. According to the book's table, he appeared once in 1998, but he was a near miss in that year, according to the magazine (*Forbes Magazine* 1998, p. 384).

The other eight individuals should have been excluded because the supposedly unique individuals were not actually unique individuals. Eights pairs of names are given below. For each pair of names, each name in the pair was included in the book's table as a unique individual, but both names refer to the same individual.

- "Crocker, Ruth Chandler" and "Von Platen, Ruth Chandler"
- "Dedman, Robert H." and "Dedman, Robert Henry, Sr."
- "Getty, J. Paul, Jr." and "Getty, Eugene Paul (J. Paul), Jr."
- "Harris, Ann Clark Rockefeller" and "Roberts, Ann Clark Rockefeller"
- "Hendrix, Helen Hunt" and "Hunt, Helen"
- "O'Neill, Abby M. Rockefeller" and "O'Neill, Abby Mitton"
- "O'Neill, Laura Simpson" and "Thorn, Laura Simpson"
- "Resenberg [*sic*], Henry A., Jr." and "Rosenberg, Henry A., Jr."

For most of those pairs, each name refers to a woman who changed her name after marriage or divorce. For the other pairs of names, one name in the pair is either a misspelling of the other name or a slight variation on the other name. Eight non-unique individuals were therefore inadvertently included in the book's table.

The eight individuals who were inadvertently excluded from the book's table are as follows.

- Robert Dart appeared on the magazine's list once in 1994, but he is not included in the book's table.
- Mark Getty appeared once in 1994, but he is not included.

- David Whitmire Hearst Sr. appeared four times, but he is not included. The book seems to attribute three of his appearances to his son, David Whitmire Hearst Jr., who appeared on the magazine's list for 19 years, not the 22 years reported in the book's table.

- Stanley Stub Hubbard appeared for 20 years, but he is not included. The book seems to attribute his appearances to his father, Stanley E. Hubbard, who appeared on the magazine's list for two years, not the 22 years reported in the book's table.

- Glenn Robert Jones appeared once in 1994, but he is not included.

- William Myron Keck Jr. appeared once in 1982, but he is not included. The book seems to attribute his appearances to William Myron Keck II, who appeared on the magazine's list for eight years, not the nine years reported in the book's table.

- Kim Magness appeared four times, but he is not included. The book seems to attribute his appearances to his father, Bob John Magness, who appeared on the magazine's list for 12 years, not the 16 years reported in the book's table.

- Ariadne Getty Williams appeared once in 1994, but she is not included.

If those eight individuals who should have been included are included, and if the nine individuals who should have been excluded are excluded, then the book's table would have identified the same number of unique individuals as the author.

Other discrepancies besides a discrepancy in the total number of unique individuals were also identified during the attempted replication. Those discrepancies related to an individual's years on the magazine's list, his or her peak wealth, or the most-recent year of his or her peak wealth. The other discrepancies were as follows. All the discrepancies are due to errors in the book's table, at least to the best of the author's knowledge. All dollar values are in current dollars.

- John Edward Anderson appeared on the magazine's list for 19 years in total (he was a new entry in the year 1988 and appeared through the year 2006), not the 20 years that was reported by the book's table.

- Robert Orville Anderson appeared for seven years (he appeared from 1982 to 1988), not six years.

- George Leon Argyros appeared for 10 years, not 11 years, although he was a near miss in 1992.

- Kenneth Eugene Behring appeared for eight years (he appeared from 1989 to 1995 and in 1997), not nine years, and his peak wealth was in 1997, not 1998 (he was not on the list in that year).

- Arthur Bejer Belfer appeared for 11 years (he appeared from 1982 to 1992), not 10 years.

- Virginia McKnight Binger appeared for 17 years, not 18 years, although she was a near miss in 1998.

- Octavia Mary du Pont Bredin appeared for 15 years (she appeared from 1982 to 1987, dropped out in 1988, appeared from 1989 to 1997, and dropped out in 1998), not 14 years.

- August Anheuser Busch Jr. appeared for eight years (he appeared from 1982 to 1989), not seven years.

- James H. Clark appeared for nine years, not 10 years, although he was a near miss in 1998.

- Tristram C. Colket Jr. appeared for eight years, not nine years, although he was a near miss in 1992.

- Andrea B. Currier appeared for five years (she appeared in 1986 and from 1990 to 1993), not three years.

- Constance Simons du Pont Darden appeared for 16 years (she appeared from 1982 to 1987, dropped out in 1988, and appeared from 1989 to 1999), not 15 years, and her peak wealth was 525 million dollars in 1998, not 500 million dollars in 1997.

- Robert Henry Dedman Sr. (or, equivalently, Robert H. Dedman; see above) appeared for 17 years, his peak wealth was 1.2 billion dollars, and the most-recent year of his peak wealth was 1999.

- Barry Diller appeared for nine years, not 10 years, although he was a near miss in 1998.

- Thomas Henry Dittmer appeared for four years (he appeared from 1986 to 1989), not three years.

- John Thompson Dorrance II appeared for seven years (he appeared from 1982 to 1988 and dropped out due to death in 1989), not eight years.

- John Thompson Dorrance III appeared for seven years (he appeared in 1982, dropped out in 1983, returned in 1989, and appeared until 1994), not five years.

- Helena Allaire Crozer du Pont appeared for five years (she appeared from 1984 to 1988), not four years.

- The peak wealth of Pierre Samuel du Pont III was 300 million dollars. According to the book's table, the most-recent year he was worth that much was 1982; but the most-recent year was 1987, according to the magazine's list.

- Malcolm Stevenson Forbes appeared for eight years (he appeared from 1982 to 1989), not six years.

- Samuel Joseph Frankino appeared for four years (he appeared in 1985 and from 1987 to 1989), not three years.

- John Brooks Fuqua appeared for five years (he appeared from 1982 to 1984 and in 1988 and 1992), not four years.

- Eugene Paul Getty Jr. (or, equivalently, J. Paul Getty Jr.; see above) appeared for 11 years with a peak wealth of one billion dollars in 1997.

- Guilford Glazer appeared for 18 years, not 19 years, although he was a near miss in 1998.

- John Murdoch Harbert III appeared for 11 years (he was a new addition in 1984, appeared from 1984 to 1994, and dropped due to death in 1995), not 12 years, and his peak wealth was 710 million dollars in 1993, not 1.5 billion dollars in 2006 (he was not on the list in that year).

- Marguerite Harbert appeared for 11 years, not 10 years. The book seems to attribute her 2006 appearance to her deceased husband, John Murdoch Harbert III.

- David Whitmire Hearst Jr. appeared for 19 years, not 22 years, as noted above.

- Helen Hunt Hendrix (or, equivalently, Helen Hunt; see above) appeared for five years, her peak wealth was 200 million dollars, and the most-recent year of her peak wealth was 1986.

- Teresa F. Heinz appeared for 10 years, not 11 years, although she was a near miss in 1999.

- Leon Hess appeared for 17 years (he appeared from 1982 to 1998, and dropped out in 1999), not 18 years, and his peak wealth was in 1993, not 1999 (he was not on the list in that year).

- Samuel J. Heyman appeared for eight years (he was a new entry in 1991, appeared until 1998, and dropped out in 1999), not seven years.

- Harry Howard Hoiles appeared for nine years (he was a new entry in 1987, appeared until 1995, and dropped out in 1996), not 10 years.

- Ronald Holden appeared for only one year, not two years, although he was a near miss in 1998.

- Amos Barr Hostetter Jr. appeared for 20 years (he was a new entry in 1987 and appeared through 2006), not 21 years.

- Stanley E. Hubbard appeared for two years, not 21 years, as noted above. Also, his peak wealth was 175 million dollars in 1983, not 1.8 billion dollars in 1996 (he was not on the list in that year).

- Muriel Kauffman appeared for two years (she appeared once in 1993 and again in 1994), not one year.

- William Myron Keck II appeared for eight years, not nine years, as noted above.

- Randolph D. Lerner appeared for four years (he was a new entry in 2003 and appeared through 2006), not seven years.

- Theodore Nathan Lerner appeared for four years, not one year. The book seems to attribute some of his appearances to Randolph D. Lerner, who is a different (and apparently unrelated) individual.

- Bob John Magness appeared for 12 years (from 1985 to 1996), not 16 years, as noted above. Also, his peak wealth was 1.2 billion dollars in 1994, not 2.3 billion dollars in 2000 (he was not on the list in that year).

- Morton Leon Mandel appeared for 12 years (he appeared from 1982 to 1984, in 1987, and from 1989 to 1996), not 11 years.

- Jacqueline Mars appeared for 20 years (she was a new entry in 1987 and appeared through 2006), not 21 years.

- Craig O. McCaw appeared for 20 years (he was a new entry in 1987 and appeared through 2006), not 21 years.

- Wendy McCaw appeared for two years, not three years, although she was a near miss in 1998.

- Alice Francis du Pont Mills appeared for 17 years (she appeared from 1982 to 1998, and dropped out in 1999), not 16 years.

- Abby Mitton Rockefeller O'Neill (or, equivalently, Abby M. Rockefeller O'Neill; see above) appeared for five years with a peak wealth of 425 million dollars in 1985.

- Milton Jack Petrie appeared for 13 years (he appeared from 1982 to 1994), not 12 years.

- Ann Clark Rockefeller Roberts (or, equivalently, Ann Clark Rockefeller Harris; see above) appear for three years, her peak wealth was 150 million dollars, and the most-recent year of her peak wealth was 1984.

- Marvin Maynard Schwan appeared for five years (he appeared from 1988 to 1992), not four years.

- Edward Wyllis Scripps appeared for four years (he appeared from 1982 to 1985), not five years, and his peak wealth was 160 million dollars in 1985, not 1.6 billion dollars in 1993 (he was not on the list in that year). The book seems to confuse the dispersed wealth of the family of E.W. Scripps (who died in 1926) with the wealth of Edward Wyllis Scripps.

- Ollen Bruton Smith appeared for 11 years, not 12 years, although he was a near miss in 1998.

- George Soros appeared for 20 years (he was a new entry in 1987 and appeared through 2006), not 21 years.

- Laura Simpson Thorn (or, equivalently, Laura Simpson O'Neill; see above) appeared for four years with a peak wealth of 200 million dollars in 1985.

- Ruth Chandler Von Platen (or, equivalently, Ruth Chandler Crocker; see above) appeared for five years with a peak wealth of 300 million dollars in 1987.

Those discrepancies relate to 54 unique individuals, which is a relatively small number of people, given that over one thousand people appeared on the magazine's list between 1982 and 2006.

D.3 Replicating Kilachand's (2012) List

For the 30th anniversary of its inaugural list of the 400 wealthiest Americans, a staff member at *Forbes Magazine* published a list of the names and wealths of the unique individuals who appeared on the list in both 1982 and 2012 (Kilachand 2012). That list included 36 people. Each of those 36 people did indeed appear on the magazine's list in both 1982 and 2012, but three other people—Phoebe Hearst Cooke, George Phydias Mitchell, and Leslie Herbert Wexner—also appeared on the list in both of those years. There were 39 people who appeared on the list in both 1982 and 2012, therefore. It can also be noted that Kenneth Stanley "Bud" Jr. was worth 150 million current dollars in 1982, according to the magazine, not 152 million current dollars, as Kilachand (2012) suggests.

D.4 Imputing the Wealth of One of Them

In each year, except in a few years, *Forbes Magazine* has reported an estimate of the wealth of each person who has appeared on its list of the 400 wealthiest Americans. The exceptions are that, in the years 1982 to 1989, the magazine did not report an estimate for the wealth of one person on its list, Malcolm Stevenson Forbes, who was the editor-in-chief of *Forbes Magazine*. "People would have assumed that the printed figure [for my wealth] was real, not an estimate, as all the rest are," he explained (*Forbes Magazine* 1983, p. 168; also see *Forbes Magazine* 1982, p. 170). This dissertation imputed his wealth in a given year as the median wealth of the other 399 people on the list in the given year. The wealths of one person (corresponding to eight observations) were therefore imputed.

It can be noted that, by using that imputation method, Malcolm Stevenson Forbes would be estimated to be worth 450 million current dollars in 1989. When he died a year later in 1990, the *New York Times* reported that, "Mr. Forbes' worth was estimated at 400 million to one billion [current] dollars" (*New York Times* 1990b, p. D1). So, an estimated wealth of 450 million current dollars in 1989 is not inconsistent with other estimates of his wealth from around that time.

D.5 Fixing the Ages of Some of Them

In addition to reporting an estimate of the wealth of each person who appears on its list of the 400 wealthiest Americans, *Forbes Magazine* also reports the age of each person, although there have been a few exceptions over the years.

For some people in some years, the magazine did not report their ages or only reported a range for their ages like "70s." Some of those people appeared on the list in subsequent years along with their ages, however, so their ages can be extrapolated backward. This dissertation extrapolated the ages of 22 people (corresponding to

61 observations). The fact that *Forbes Magazine* was sometimes able to estimate a person's wealth but not their age may make one wonder about the accuracy of the wealth estimates, although the magazine's primary interest was obviously in estimating people's wealths rather than their ages.

For other people, their ages cannot be extrapolated backwards because they never appeared again in subsequent years or, at least, they never appeared again along with their ages. For those people, except for one person, this dissertation imputed their ages as the average conditional age of the other people on the magazine's list in the same year. So, for example, if the magazine reported that a person was somewhere in his or her seventies, then this dissertation imputed his or her age as the average age of everyone else on the list in the same year who was in their seventies. The ages of 15 people (34 observations) were imputed in that way.

The one person who is the exception is David S. Oros, who appeared on the magazine's list once in the year 2000. A 2001 article in *Forbes Magazine* reported that Oros was 41 years of age (Brown 2001), so this dissertation assumed that he was 40 in 2000. That person was the only person for which the magazine reported his or her age in another context, at least to the author's knowledge.

A final issue with the ages reported by the magazine is the following. If the ages of unique people are compared from list to list, then it is obvious that at least some of the ages reported by the magazine were incorrect. If the ages were correct, then the difference in age for a given person between any two lists should never be less than zero or greater than two years of age. The difference does not necessarily need to be one year of age because the magazine's list has not been a snapshot of wealth on the exact same day every year, so a person could be the same age or two years older between any two lists. Some of the age differences are greater than two, however, in which case the person would have had too many birthdays. Some of the age difference are also negative, in which case the person would have grown younger. This dissertation corrected those obviously incorrect ages by assuming that the most-recent age reported by the magazine was correct—which seems like a reasonable assumption, given that more information should come to light over time—and then extrapolating backwards. The ages of 53 people (111 observations) were corrected in that way.

In total, the ages of 90 people corresponding to 206 observations were either extrapolated, imputed, or corrected. That represents a relatively small number of people (only about six percent of the total number of people who appeared on the magazine's list) and a relatively small number of observations (only about two percent of the total number of observations).

D.6 Ignoring Human Capital

When it estimates what someone is worth, *Forbes Magazine* apparently tries to account for all sorts of real and financial assets ("stakes in public and private companies, real estate, art, yachts, planes, ranches, vineyards, jewelry, car collections, and more;" *Forbes Magazine* 2012, p. 262), but there is no indication that it tries to account for "human capital" or any of the other sorts of capital identified by social scientists such as "personal capital," "mental capital," "social capital," "erotic capital," or "natural capital."[1]

Human capital is the only sort of capital that economists commonly complain about whenever someone fails to account for it, so not trying to account for other sorts of capital can perhaps pass without comment, but a few comments should be made about the magazine not trying to account for human capital.

---

[1]The sorts of capital listed above are borrowed from Fine's (2013, p. 8) extensive but not exhaustive list of a plethora of capitals. Other sorts of capital identified by social scientists include "cultural capital" and "symbolic capital" (Fine 2007, pp. 48–49).

It should first be recognized that, even if the magazine does not try to account for human capital, its wealth estimates may account for such capital after all, at least partially, and at least in some cases. Some of the wealthiest Americans according to *Forbes Magazine* have been people who created companies, operated those companies, and held some of their wealth in the companies. If we take the view that markets are information processors, then part of the price of such a company should reflect something akin to the human capital of its creator and operator. When the price of Apple's stock rose and fell along with news about the health of Steve Jobs, those prices swings presumably reflected new expectations about how long Jobs would live and how much value he would contribute over the rest of his life, for example. To the extent that the price of a company reflects a person's human capital, the person's wealth held in that company would also reflect his or her human capital. Thus, while it is important to note that the magazine does not try to account for human capital, that does not imply that it completely fails to do so.

It should also be noted that, even if *Forbes Magazine* wanted to try to account for human capital, it is unclear how it could do so. Economists often measure a stock of human capital by a level of education, but many of the Americans who have been the wealthiest Americans according to *Forbes Magazine* have also been college or even high school dropouts. The man who has been the wealthiest American for several years according to the magazine (namely, Bill Gates) is a college dropout, for example (Bernstein and Swan 2007, p. 8). Selection effects are an obvious reason why someone who became wealthy would forgo higher levels of education, but it is nevertheless the case that education may be a poor proxy for human capital. Another typical measure of a stock of human capital is a flow of income, but some of the wealthiest Americans according to *Forbes Magazine* have worked for salaries of only one dollar per. Steve Jobs, Larry Page, and others have worked for an annual salary of only one dollar, for example (Coffey 2011). Stock options are an obvious reason why someone would work for such a seemingly low amount, but, if we account for stock options, then those options would reflect the person's human capital only to the extent that the price of a stock reflects such capital.

Finally, it should be recognized that, even if there was a clear way to try to account for human capital, there are reasons to study non-human assets separately from human ones. Unlike non-human assets, human assets cannot be bought or sold, so they cannot be spent down or pledged as collateral (Davies 2009, pp. 127–8).

In summary, the fact that *Forbes Magazine* does not try to account for human capital is noteworthy but not indefensible.

D.7 Ignoring Some Valuation Issues

Human capital is not the only asset for which it is somewhat unclear how it should be valued when trying to estimate someone's wealth. Even publicly traded stocks could reasonably be valued in different manners. Such stocks may seem straightforward to value because they change hands almost constantly at prices that are publicly available. Simply using market prices is an obvious and not unreasonable approach to valuing a stock, and that approach is the one that *Forbes Magazine* apparently adopts. Again, however, there are other reasonable manners in which publicly traded stocks might be valued.

If a person owns a certain amount of a stock that is trading at a certain price, then he or she would only be worth the current value of that stock to the extent that wealth is defined in terms of market prices. If the person actually sold the stock, then trying to sell it might push down its price and, even if its price was not pushed down, capital-gain taxes and other expenses might be incurred, as emphasized by Einstein (2000). Thus, in some sense, the person was never worth as much as the market price of his or her stock might suggest. Any taxes or transactions costs associated with selling a stock should therefore perhaps be accounted for when valuing the stock.

Of course, for other assets besides publicly traded stocks, just estimating their market prices can be difficult. Estimating any taxes and transaction costs that someone might face when selling an asset is an additional difficulty. That difficulty may also be unnecessary, if it is reasonable to simply value assets at market prices. It is therefore not surprising that the magazine does not try to account for the costs associated with selling an asset.[2]

D.8 Sharing the Data

As mentioned above, *Forbes Magazine* declined to directly share its data on the 400 wealthiest Americans with the author of this dissertation, so the author went through each year of the magazine's list and collected the data himself. The magazine also declined the author's request to republish that raw data as part of this dissertation (personal communication on May 8, 2014, with Elena Coster, sales coordinator at *Forbes Magazine*'s content management firm, PARS International). We will therefore refrain from sharing that raw data as part of this dissertation. Instead, we will share the dataset we created by collecting data from each year of the magazine's list, identifying unique individuals in the manner discussed above, imputing some wealth estimates in the manner discussed above, and fixing some of the ages reported by the magazine in the manner discussed above. That dataset is too large to include here, but it is included as a supplementary file to this dissertation. That file is available through ProQuest/UMI along with the full text of the dissertation.

---

[2] An implication of valuing assets at market prices that is perhaps noteworthy is as follows. While a fortune with substantial unrealized capital gains may appear bigger than a fortune for which capital gains have already been realized, the former may actually be smaller than the latter after accounting for capital-gains taxes.

D.9 Some Summary Statistics

D.9.1 Their Absolute Wealth

According to the magazine's list, the total wealth of the 400 wealthiest Americans increased dramatically between the inaugural year of its list in 1982 and the most-recent year of its list, which was the year 2013, as of writing. The increase in the nominal wealth of the 400 wealthiest Americans was especially dramatic. Their nominal wealth increased over 20-fold from about 92 billion to two trillion current dollars between 1982 and 2013 (*Forbes Magazine* 1982, 2013).[3] That increase in their nominal wealth is shown as part of figure D.1 in this section.

Of course, for comparisons over time, nominal values should be adjusted to account for changes to the value of a dollar. Previous studies based on *Forbes Magazine*'s list of the 400 wealthiest Americans have either erroneously ignored the distinction between nominal and real wealth (Klass et al. 2006) or used a version of the Consumer Price Index (CPI) made by the Bureau of Labor Statistics (BLS) in order to deflate nominal wealth into real wealth (Broom and Shay 2000; Kennickell 2006; Kopczuk and Saez 2004a; Mishel et al. 2012).

It could be argued that the CPI is an inappropriate deflator for at least two reasons. First: If wealth is accumulated for future consumption (either by one's self or one's beneficiaries), rather than current consumption, then wealth should perhaps be measured in terms of future rather than current prices (Alchian and Klein 1973). Second: If wealth is accumulated for some power it confers over people or institutions, rather than for its purchasing power over current or future goods and services, then

[3] Again, in the years 1982 to 1989, *Forbes Magazine* did not report an estimate of the wealth of one of the 400 wealthiest Americans on its list (the editor-in-chief of the magazine, Malcolm Stevenson Forbes). His wealth in a given year was imputed as the median wealth of the other 399 wealthiest Americans in that year. Previous studies that have used the magazine's list to study the wealth of the 400 wealthiest Americans do not appear to have used any imputation (Broom and Shay 2000; Kennickell 2006; Mishel et al. 2012). Those studies were actually only studies of the wealth of 399 out of the 400 wealthiest Americans up until the year 1990, therefore.

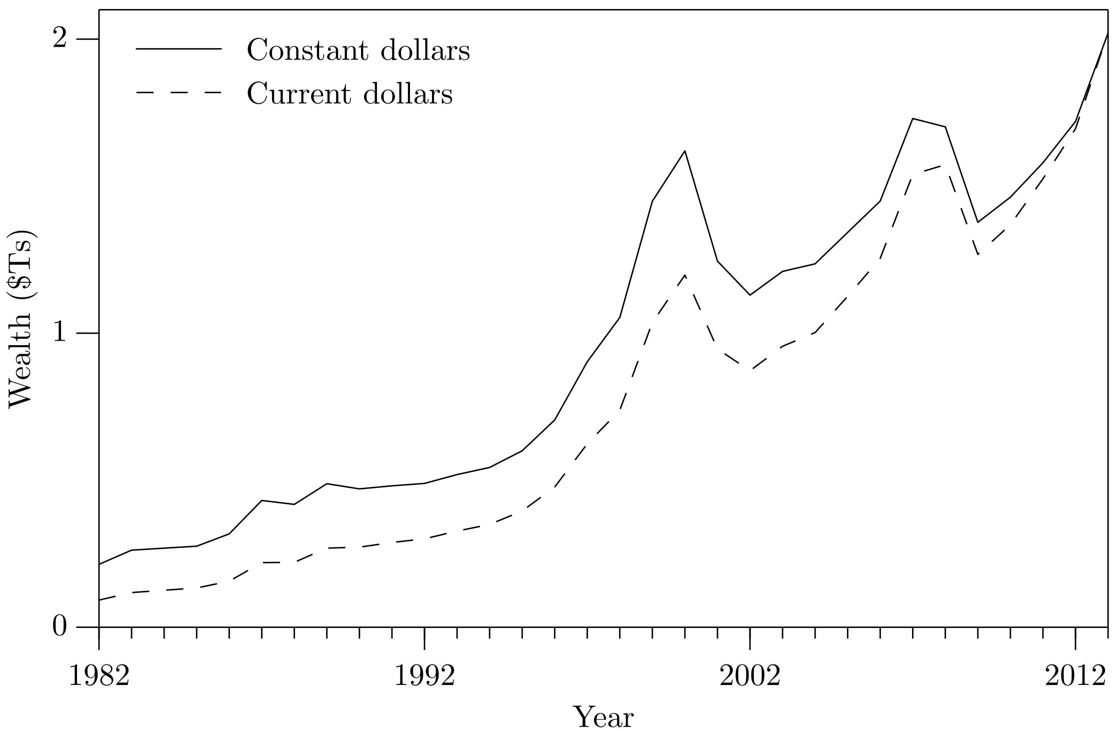

Figure D.1. Wealth of the 400 Wealthiest Americans, 1982–2013

Source: Data adapted from BLS (2013); *Forbes Magazine* (1982–2013).

Note: This figure shows the wealth of the 400 wealthiest Americans, based on *Forbes Magazine*'s list from the years 1982 to 2013. Their wealth is shown in current dollars and constant (i.e., CPI-deflated, 2013) dollars.

wealth should perhaps be measured in terms of the social power it confers (Officer and Williamson 2006). Unfortunately, a measure of future prices is not readily available and a measure of social power is not well established. So, if only out of necessity, we will simply use a version of the CPI as a deflator.[4]

When their nominal wealth is deflated by the CPI, the increase in the wealth of the 400 wealthiest Americans between 1982 and 2013 is somewhat less dramatic. Their wealth in CPI-deflated, 2013 dollars increased a little less than 10-fold from a little over 200 billion to about two trillion constant dollars.

[4]Although the results would be qualitatively similar with different versions of that index, we will use the CPI-U-RS, in particular (BLS 2013).

As seen in figure D.1, the wealth of the 400 wealthiest Americans—whether it is measured in current or constant dollars—did not increase in every year between 1982 and 2013. Their wealth decreased in some years. The most-dramatic decrease in their wealth occurred after their wealth peaked in the year 2000 and then fell over the next two years. Their wealth peaked in the year 2000 at about 1.2 trillion current or 1.6 trillion constant dollars. Their wealth then fell between 2000 and 2002 by about 325 current or 484 constant dollars. The decline in their constant-dollar wealth after the year 2000 is shown as part of figure D.2. The run-up and decline in their wealth around those years can be associated with the stock-market boom of the late 1990s that went bust in the early 2000s (*Forbes Magazine* 2001, 2002).

By the year 2007, the wealth of the 400 wealthiest Americans had exceeded its previous peak in terms of both current and constant dollars, but their wealth then fell again over the next two year. Their wealth fell by about 272 billion current or 348 billion constant dollars between 2007 and 2009. That decline in their wealth—which can be associated with the recent economic crisis that saw the collapse of a housing bubble, a panic in financial markets, a crash in the stock market, and a deep recession that rivaled the Great Depression (*Forbes Magazine* 2007–2009)—was relatively modest, relative to the decline in their wealth after the stock-market crash of the early 2000s, as figure D.2 shows.

Other declines in their wealth were even more modest. Between 1987 and 1988, as well as between 1989 and 1990, their wealth decreased slightly in current-dollar terms and it only increased slightly in constant-dollar terms. Their current-dollar wealth increased by amounts that were on the order of only about one billion dollars, while their constant-dollar wealth decreased by amounts that were on the order of only about 10 billion dollars. For the magazine's list from 1987, publicly traded stocks were valued on September 11th of that year (*Forbes Magazine* 1987, p. 112),

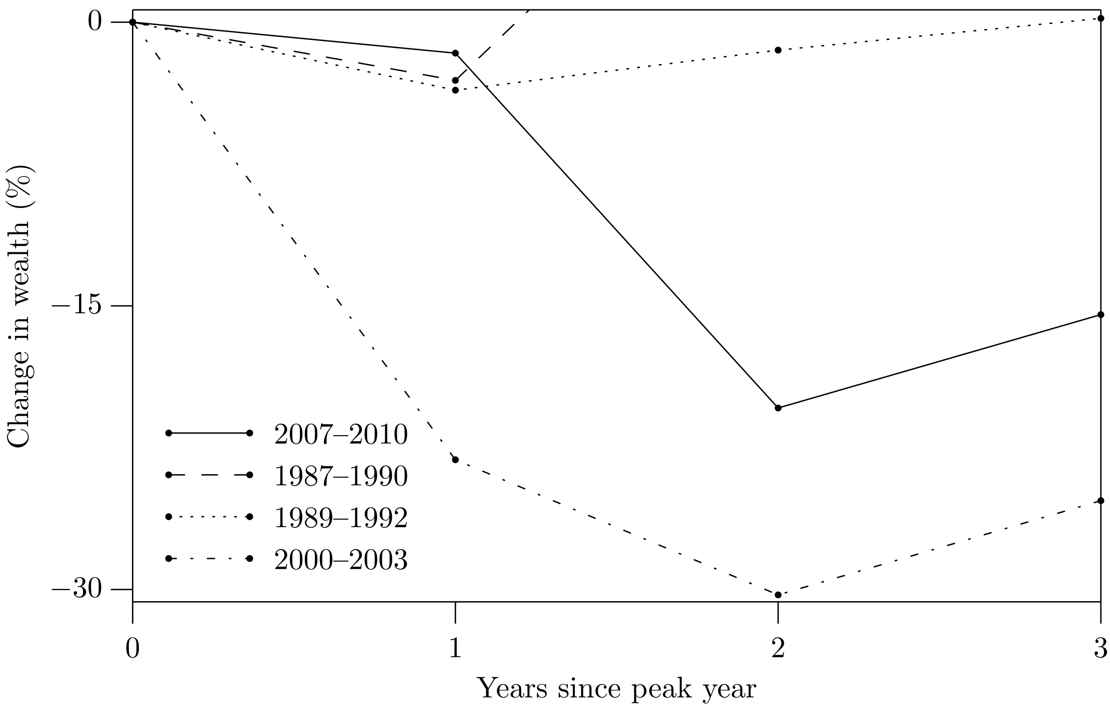

Figure D.2. Percentage Change in the Wealth of the 400 Wealthiest Americans Over Select Years

Source: Data adapted from BLS (2013); *Forbes Magazine* (1987–2010).

Note: This figure shows, for years after their wealth peaked, the percentage change in the wealth of the 400 wealthiest Americans, relative to their peak wealth. Dollar values are constant (i.e., CPI-deflated, 2013) dollars.

which was shortly before the "Black Monday" of October 19th when the stock market crashed. The decline in the wealth of the 400 wealthiest Americans between 1987 and 1988 can therefore be associated with that stock-market crash (*Forbes Magazine* 1988). The decline in their wealth between 1989 and 1990 can be associated with the downturn in the real-estate market, stock market, and economy more generally that occurred around those years (*Forbes Magazine* 1990).

Thus, while *Forbes Magazine*'s list does not suggest that the 400 wealthiest Americans became ever-wealthier in absolute terms since 1982, the list does suggest that they became wealthier in absolute terms since then.

D.9.2 Their Relative Wealth

As the wealth of the 400 wealthiest Americans increased in absolute terms over recent decades, their wealth also increased relative to everyone else's, although the exact extent to which their relative wealth increased obviously depends on how the wealth of everyone else is measured.

Part of figure D.3 shows, for the years 1982 to 2013, the wealth of the 400 wealthiest Americans as a share of one specific measure. The measure that we used for that figure was a weighted average of the end-of-the-second-quarter and end-of-the-third-quarter wealth of the household sector in the United States in a given year, as reported by the Federal Reserve's Flow of Funds Accounts, where the weights are discussed below. Using that measure, the share of household wealth held by the 400 wealthiest Americans roughly tripled from about 0.85 to 2.75 percent between 1982 and 2013, as shown as part of the figure.

One issue with using that specific measure is a timing issue. Wealth is a stock variable, so it should be measured at a point in time. The magazine's list in a given year is supposed to be a snapshot of wealth at the close of the stock market on a particular day of the given year, but the day of the year has varied over the years, and the day has never been at the end of any quarter. For years in which the magazine reported the day on which it took a snapshot of wealth, the day was as early as August 16th in 2002 and as late as September 12th in 1986 (*Forbes Magazine* 1982–2013). Those days fall between the end of the second quarter and the end of the third quarter, which is why a weighted average of the wealth of the household sector at the end of those two quarters was used above. The wealth of the household sector at the end of a quarter was weighted more heavily if the date on which the magazine took its snapshot of the wealth was closer (in terms of number of days) to the end of that quarter. Unfortunately, for some years, the magazine does not appear to have

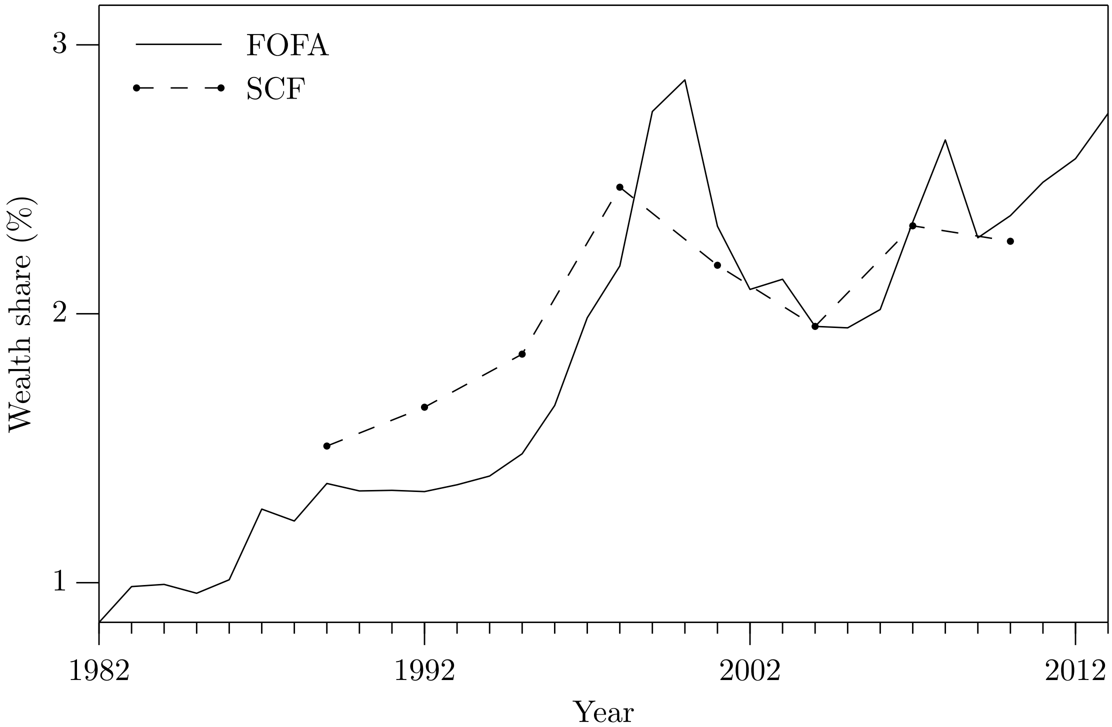

Figure D.3. Share of Wealth Held by the 400 Wealthiest Americans, 1982–2013

Source: Data adapted from Bricker et al. (2012); BLS (2013); Flow of Funds Accounts; *Forbes Magazine* (1982–2013); Kennickell (2011, 2012); Kennickell and Starr (1994); Kennickell et al. (2000).

Note: This figure shows the wealth of the 400 wealthiest Americans, based on *Forbes Magazine*'s list, as a share of household wealth, based on either the Flow of Funds Accounts (FOFA) or the Survey of Consumer Finances (SCF), for years between 1982 and 2013.

reported the day on which it took a snapshot of wealth. To the author's knowledge, the magazine did not report the day it took a snapshot of wealth for its lists in 1982, 1999, 2000, and 2003 to 2005, although the magazine did report that its 1982 list was a snapshot of wealth in "mid-August" (*Forbes Magazine* 1982, p. 101). For those years, the wealth of the household sector at the end of the second quarter was weighted equally with their wealth at the end of the third quarter.

Using a weighted average of the wealth of the household sector at the end of the second and third quarters does not necessarily resolve the timing issue. That would only resolve the issue if the wealth of the household sector changed in a linear fashion between the end of one quarter and the next. Yet, if we were to use the Flow-of-Funds estimates of the wealth of the household sector at the end of any quarter in a given year, then trends in the wealth of the 400 wealthiest Americans as a share of the wealth of the household sector would be quantitatively similar. To be more specific, if the Flow-of-Funds estimates at the end of any quarter in a given year were used, then their wealth share would be estimated to be as low as about 0.82 and as high as about 0.87 percent in 1982, and it would be estimated to be anywhere between about 2.62 and 2.85 percent in 2013. Their share of household wealth would have therefore increased between those years by a factor that was anywhere between about three to three-and-a-half percent—roughly tripling.

Although their share of wealth has roughly tripled over recent decades (at least according to *Forbes Magazine*'s list), some economists have suggested that, even at their largest, the share of wealth held by the 400 wealthiest Americans has still been small. For example: In a footnote to their 2002 article, two economists from the Federal Reserve noted that the SCF is designed to exclude the (households of the) people who appear on *Forbes Magazine*'s list (Bertaut and Starr 2002, p. 214, n. 10). They noted in the same footnote that the share of household wealth held by the people on the magazine's list was "on the order of 2.5 percent," which they suggested was merely a "modest" share (ibid.).

As another example: At a 2003 conference, an economist from the Federal Reserve presented a paper in which he used *Forbes Magazine*'s list to supplement the SCF (Kennickell 2003). He used that list to supplement the SCF because, as noted above, the SCF is designed to exclude the (households of the) people who appear on the magazine's list. The paper's discussant asked the audience the question, "What fraction of total wealth do you think that the [400 wealthiest Americans] hold?" (Smeeding 2003). "They hold two percent," he answered. "The lowest guess I've gotten from anyone is 10 percent. Everybody thinks it's much bigger [but] these rich guys hold two percent only, and that's it" (ibid.).

While it could perhaps be argued that the trillion or so dollars of wealth held by the 400 wealthiest Americans (about two trillion current dollars in 2013, according to the magazine's list from that year) is small relative to the tens of trillions of dollars of wealth held by all the households in America (roughly 75 trillion current dollars in 2013, according to the Flow-of-Funds estimate discussed above), it seems unreasonable to suggest that the share of wealth held by the 400 wealthiest Americans is small relative to either their share of the population or the shares of wealth held by other groups of Americans, as we will now argue.

Consider their share of wealth relative to their share of the population. At every point in time over recent decades, the 400 wealthiest Americans have been an infinitesimally small proportion of the population. In 1982, they accounted for about one five-thousandth of one percent (i.e., about 0.0002 percent) of the population using the Census Bureau's estimate of the resident population in that year. By 2013, the 400 wealthiest Americans only accounted for about one ten-thousandth of one percent (i.e., about 0.0001 percent) of the population (again, using the Census Bureau's population estimate). By comparison, the share of household wealth held by the 400 wealthiest Americans was almost one percent in 1982 and over two-and-a-half percent by 2013, according to the estimates discussed above. Their share of wealth has therefore been thousands of times larger than their share of the population. A share of wealth that is thousands of times larger than an infinitesimal share of the population is still small, perhaps, but the share of wealth held by the 400 wealthiest Americans has been at least disproportionately large.

Next, consider their share of wealth relative to the share of wealth held by another group of wealthy Americans, specifically, the wealthiest one percent of American households. Although there is no single source of data that can be used to compare the shares of wealth held by those two groups (Kennickell 2006, p. 84, n. 4), their wealth shares can be compared by using *Forbes Magazine*'s list and the Federal Reserve's SCF. One of the current directors of the SCF, Arthur B. Kennickell, has used the magazine's list to make similar comparisons (Kennickell 2003, 2006).[5] Those two data sources can be used to compare the shares of wealth held by the two groups in every three years between the years 1989 and 2010. Earlier years of the SCF are not comparable to latter years of that survey because the survey was redesigned (Kennickell 2011, p. 11, n. 10), and more-recent years of the survey are not yet available, but, again, the two data sources can be used to compare the shares of wealth held by the two groups in every three years between 1989 and 2010.

In 1989, the share of wealth held by the wealthiest one percent of American households was about 30.1 percent, according to estimates made by Kennickell (2012, p. 5, table 3) from the SCF. Kennickell (2012) reported that wealth share without reporting either an estimate of the total wealth of the wealthiest one percent of American households or an estimate of the total wealth of all the households in America, but he did report an estimate of the mean wealth of American households. He estimated that their mean wealth was about 188.9 thousand current dollars (Kennickell 2012, 4, table 2, if the constant-dollar figure he reports is converted to 1989 dollars by using the CPI-U-RS). The total wealth of American households must have therefore been estimated to be about 17.6 trillion current dollars, given that the number of households represented by the SCF in that year was apparently 93.1 million (Kennickell and Starr 1994, p. 880, n. 22).[6] That estimate of household wealth based on

[5]On Kennickell and his role in the SCF, see Rae and DeHaan (2012).

[6]The number of households represented by the SCF over the years has apparently been, in millions of households, 93.1 in 1989 (Kennickell and Starr 1994, p. 880, n. 22), 95.9 in 1992 (ibid.), 99.0 in 1995 (Kennickell et al. 2000, p. 27, n. 35), 102.6 in 1998 (ibid.), 106.5 in 2001 (Bricker et al. 2012, p. 78, table A.3), 112.1 in 2004 (ibid.), 116.1 in 2007 (ibid.), and 117.6 in 2010 (ibid.).

the 1989 SCF is similar to, although not exactly the same as, Flow-of-Funds estimates of household wealth at the end of any quarter in that year. According to the Flow of Funds Accounts, household wealth ranged from about 17.7 to 19.0 trillion current dollars over that year. If the wealthiest one percent of American households were estimated to be worth 30.1 percent of about 17.6 trillion current dollars, then they must have been estimated to be worth about 5.3 trillion current dollars. By comparison, the wealth of the 400 wealthiest Americans was about 270 billion current dollars in 1989, which would have been about 5.1 percent of that amount.

However, if we follow Kennickell (2006) and add the magazine's estimate of the wealth of the 400 wealthiest Americans to the SCF estimates, then the wealthiest one percent of households would have been worth about 31.2 percent of household wealth, and the 400 wealthiest Americans would have been worth about 4.8 percent of that. In a similar context, Kennickell (2006) adds the magazine's estimate because, again, the SCF is designed to exclude the (households of the) people who appear on the magazine's list. The wealth of the 400 wealthiest Americans as a share of the SCF estimate of household wealth with *Forbes Magazine*'s estimate of their wealth added to it (or, for that matter, without the magazine's estimate of their wealth added to it) has not been exactly the same as their wealth as a share of the Flow-of-Funds estimate of household wealth, which can be seen in figure D.3, but the shares have been similar, which can be seen in the same figure.

If the same calculations are made for subsequent years, then, by 2010, the share of wealth held by the wealthiest one percent of American households was about 34.5 percent and the 400 wealthiest Americans were worth about 6.7 percent of that. Or, if *Forbes Magazine*'s estimate of the wealth of the 400 wealthiest Americans is added to the SCF estimates, then share of wealth held by the wealthiest one percent of American households was about 36 percent and the 400 wealthiest Americans were

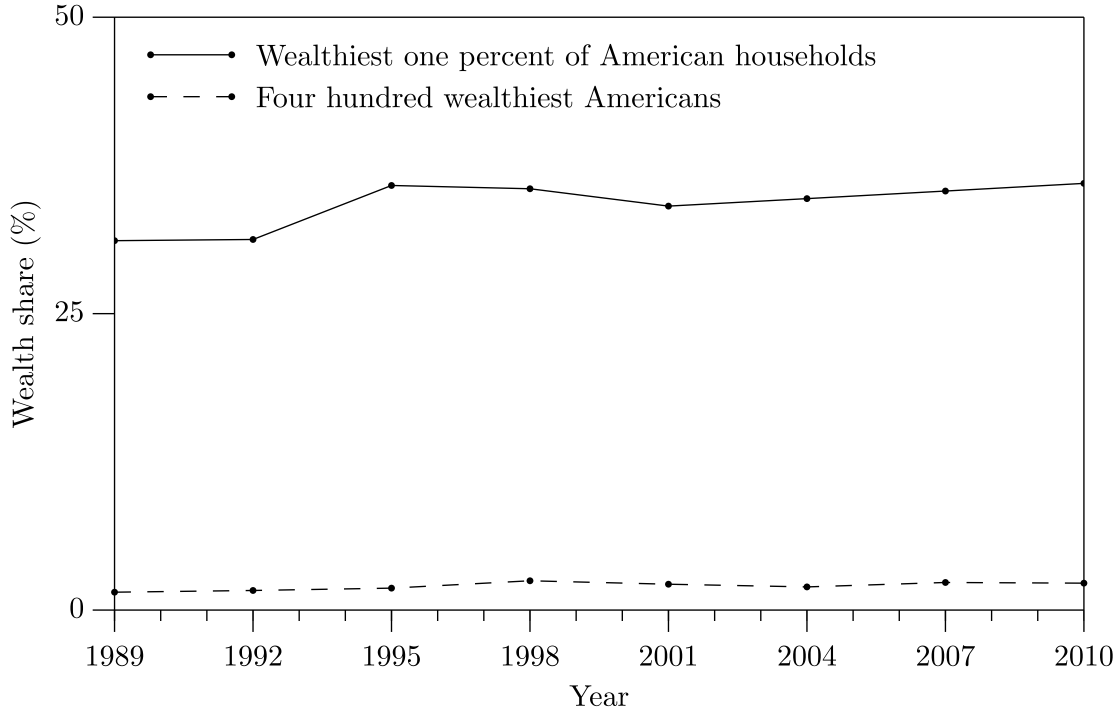

Figure D.4. Shares of Wealth Held by the Wealthiest One Percent of American Households and the 400 Wealthiest Americans, 1989–2010

Source: Data adapted from Bricker et al. (2012); BLS (2013); *Forbes Magazine* (1989–2010); Kennickell (2011, 2012); Kennickell and Starr (1994); Kennickell et al. (2000).

worth about 6.3 percent of that. The wealth shares in the other years between 1989 and 2010 are shown as part of figure D.4, where the magazine's estimate of the wealth of the 400 wealthiest Americans was added to the SCF estimates in each year.

As a final comparison, consider the share of wealth held by the 400 wealthiest Americans relative to the share of wealth held by a group of some of the least-wealthy Americans, specifically, the least-wealthy half of American households. Again, there is no single source of data that can be used to compare the shares of wealth held by those two groups, but their wealth shares can be compared for every three years between 1989 and 2010 by using *Forbes Magazine*'s list and the Federal Reserve's SCF. Figure D.5 shows, for each of those years, the shares of household wealth held by the least-wealthy half of American households, on the one hand, and the 400

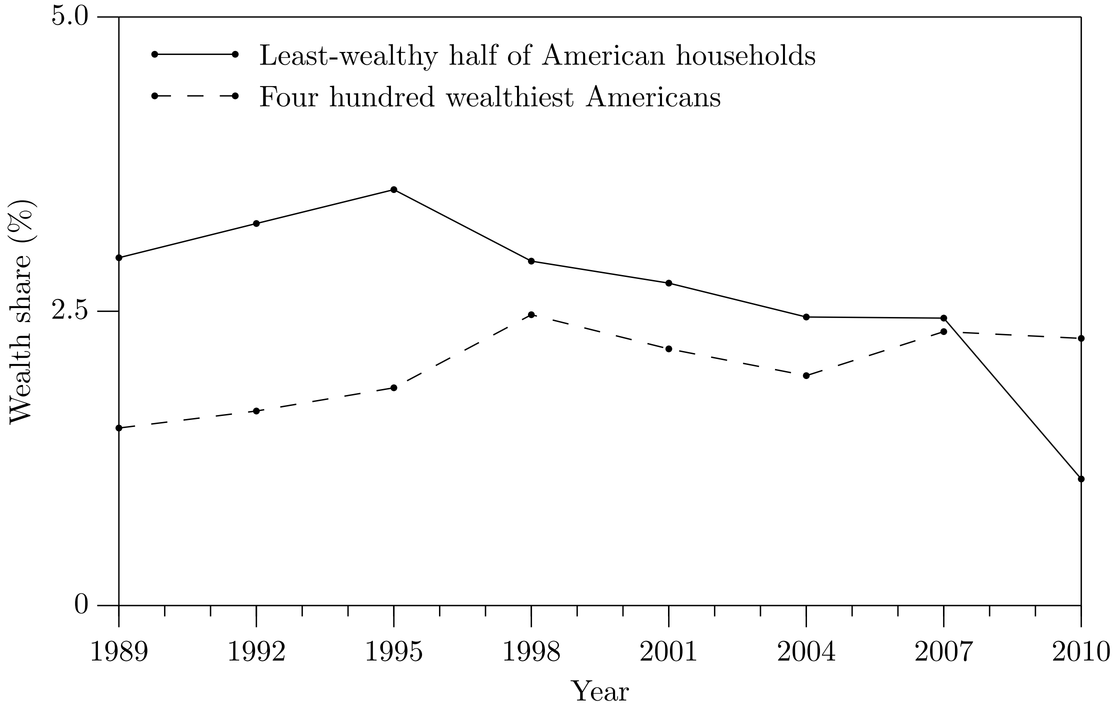

Figure D.5. Shares of Wealth Held by the Least-wealthy Half of American Households and the 400 Wealthiest Americans, 1989–2010

Source: Data adapted from the same sources as figure D.4.

wealthiest Americans, on the other hand, where the shares were estimated in the same way as in the earlier figure. As seen in the figure: In 1989, the least-wealthy half of American households owned about three percent of household wealth, while the 400 wealthiest Americans owned about 1.5 percent. By 2010, the 400 wealthiest Americans owned about 2.3 percent of household wealth, while the least-wealthy half of American households owned about 1.1 percent.

It can be noted that, as long as the magazine did not over-estimate what they were worth by more than about 720 billion current dollars or 53 percent, the 400 wealthiest Americans were worth at least as much as the least-wealthy half of American households in the 2010. The fact that the 400 wealthiest Americans were worth at least as much as the least-wealthy half of Americans households in that year was partially due to the fact that about 11.1 percent of households had negative wealth with their debts exceeding their assets (Kennickell 2012, p. 3, table 1), but it is nevertheless a striking statistic that a group that accounted for less than one thousandth of one percent of the American population was worth at least as much as half of the roughly 120 million households in America. That statistic, as well as the earlier statistic on the share of wealth held by the 400 wealthiest Americans relative to the share of wealth held by the wealthiest one percent of American households, has been emphasized by Foster and Holleman (2010) among others.[7]

The share of wealth held by the 400 wealthiest Americans over recent decades hsas therefore been disproportionately large relative to their share of the population, roughly equal to the share of wealth held by the least-wealthy half of American households, and arguably large relative to the share of wealth held by the wealthiest one percent of American households. Their share of wealth has also been larger in recent years than it was in earlier years, despite the fact they shrunk as a share of the population. Whether the wealth of the household sector is measured based on the SCF or the Flow of Funds Accounts, the wealth of the 400 wealthiest Americans increased relative to the wealth of everyone else over recent decades, as seen in figure D.3 of this section. Their share of wealth did not increase in every year, however, as seen in the same figure. Their wealth share fell by a relatively large amount amid the stock-market crash of the early 2000s, recovered to a large extent, and then fell by a smaller amount amid the recent economic crisis. The most-recent estimates suggest that their wealth share has all but recovered.

---

[7]Another striking statistic—which was estimated by Allegretto (2011), emphasized by Stiglitz (2012, p. 8), and expounded upon by Bivens (2012) and others—is that the share of household wealth held by the six heirs of Walmart in 2007 was roughly equal to the share of wealth held by the least-wealthy 30 percent of American households in that year.

D.9.3 The Pro-Wealthiness of Their Accumulation

As discussed in the previous section, *Forbes Magazine*'s list suggests that the wealth of the 400 wealthiest Americans increased relative to the wealth of everyone else over recent decades. The magazine's list also suggests that, even among the 400 wealthiest Americans, the wealth of wealthier people increased relative to the wealth of less-wealthy people, as seen in figure D.6 of this section.

That figure shows, for the wealthiest to the 400th wealthiest Americans in 1982 and 2013 (whoever they happened to be in those two years), the change in their wealth in billions of constant dollars. Note that the ordinate axis in the figure uses an inverse hyperbolic sine transformation.[8]

As seen in the figure: The increase in the wealth of the wealthiest American was relatively large at about 67 billion constant dollars; the increase in the wealth of the 100th wealthiest American was roughly equal to the average change of about five billion; the change in the wealth of the 200th wealthiest American was smaller at about three billion; and the change in the wealth of the 400th wealthiest American was smaller still at about one billion constant dollars.

Figures like figure D.6 are often used in the literature on poverty to study the distributional effects of income growth. In that context, such figures show changes in income between two periods of time as a function of relative income. The figures are used to study whether growth was "pro-poor" in the sense (among the other senses in which that term is used) that poorer people saw their incomes rise by more than

---

[8]That is to say, if y denotes the variable of interest, then that variable was transformed as $\ln\{\theta y + (\theta^2 y^2 + 1)^{1/2}\}/\theta$, where the scale parameter θ was somewhat arbitrarily set to one one-thousandth. An inverse hyperbolic sine or "arcsinh" transformation approximates a logarithmic transformation at extreme values, but the former is more popular than the latter in the literature on wealth because, while wealth can be zero or negative, the log of a non-positive number is infinite or imaginary (Kennickell 2006, p. 84; Pence 2006). (At values close to zero, an arcsinh transformation approximates a linear transformation or, equivalently, no transformation at all.) A semi-log scale could have been used for the figure, given that all of the changes in wealth were positive, but the semi-arcsinh scale was used in anticipation of a similar figure that follows.

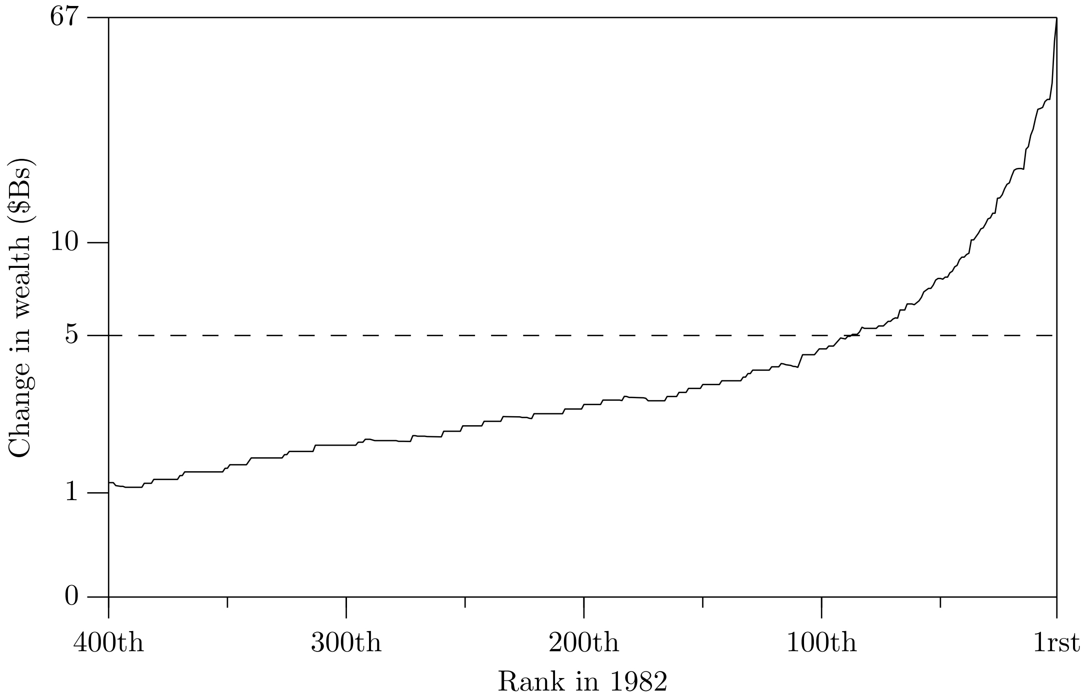

Figure D.6. Anonymous Changes in Wealth, 1982–2013

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This figure shows, on a semi-arcsinh scale, changes in the wealth of the wealthiest to the 400th wealthiest Americans between 1982 and 2013 (in billions of CPI-deflated, 2013 dollars), ignoring who they were as individuals in either of those years (solid line). The average change in their wealth is also shown (dashed line).

richer people, ignoring who those people might have been in either period of time. An obvious issue with such a figure is that it obscures the gains or losses experience by unique individuals (Grimm 2007). The same issue presents itself in our figure.

Figure D.7 extends our earlier figure by showing changes in the wealth of unique individuals who appeared on *Forbes Magazine*'s list in 1982, 2013, or both of those years. Only 38 people appeared in both years. For those 38 people, the figure shows the uncensored changes in their wealth. Considering just those changes for the moment, the figure suggests that some people fared better between 1982 and 2013 than figure D.6 would suggest. The investor Warren Buffett fared better, for example. In 1982, he was the 92nd wealthiest American with a wealth of about half

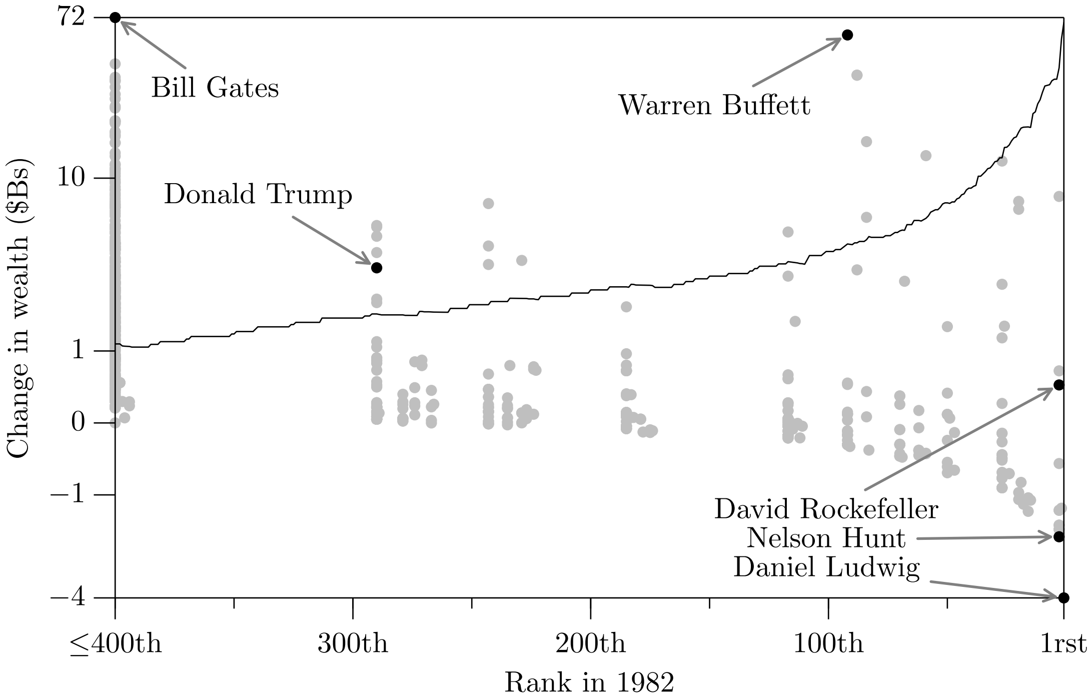

Figure D.7. Non-anonymous Changes in Wealth, 1982–2013

Source: Data adapted from the same sources as figure D.6.

Note: This figure shows, on a semi-arcsinh scale, changes in the wealth of the wealthiest to the 400th wealthiest Americans between 1982 and 2013, ignoring who they were as individuals (solid line). Uncensored or censored changes in the wealth of individuals who were one of the 400 wealthiest Americans in 1982 or 2013 are also shown (points).

a billion constant dollars. By 2013, the 92nd wealthiest American was worth about five billion dollars, but, by then, Buffett was the second wealthiest American with a wealth of over 50 billion dollars.

Other people who appeared on the list in both 1982 and 2013 fared worse than figure D.6 would suggest. The oldest grandson of John D. Rockefeller—David Rockefeller—fared worse, for example. David was the third wealthiest American in 1982 with a wealth of about two billion constant dollars. The third wealthiest American was worth 41 billion dollars by 2013, but David was not worth nearly that much by then. He still fared fairly well, of course, or else he would have fallen off the list. David was the 193rd wealthiest American with a wealth of 2.8 billion dollars.

An above-average change in wealth does not necessarily imply that someone fared as well in the intervening period as that change would suggest. Take Donald Trump, for example. According to *Forbes Magazine*, Trump was the 290th wealthiest American in 1982 with a wealth of about 200 million constant dollars, and he was the 134th wealthiest American with a wealth of over three billion dollars by 2013, but, for the first half of the 1990s, he was not one of the 400 wealthiest Americans. Indeed, his debts may have even exceeded his assets by hundreds of millions of dollars at one point (Barsky 1990). Again, however, an above-average (or below-average) change in wealth between 1982 and 2013 implies that someone ultimately fared better (or worse) than figure D.6 would suggest.

For everyone except the 38 people who appeared on the magazine's list in both 1982 and 2013, only censored changes in their wealth can be shown. For the 362 people who were on the list in 1982 but dropped off by 2013, the amount by which a person's wealth changed is shown as the difference between his or her wealth in 1982, on the one hand, and the minimum wealth of the 400 wealthiest Americans in any year after he or she dropped off the list, on the other hand. The change in a person's wealth would have been at least as large as that difference, if the person was still alive and still an American citizen by 2013.

Some of the people who dropped off the list were still alive and still Americans by then. After appearing on the list in 1982 with a wealth of over two billion constant dollars, and after dropping off the list in the 1980s amid an ill-fated attempt to corner the silver market with his brother, Nelson Hunt did not make it back on the list by 2013, but he was still alive and still an American in that year, for example.

Some of the people who dropped off the list were no longer alive or no longer Americans by 2013, however. Over a quarter (specially, 117) of the 400 people who appeared on the list in 1982 eventually dropped off because they died. Other people dropped off the list for other reasons, but at least some of them eventually died, too. At least some of the 243 people who dropped off the list because of a decline in their wealth eventually died. One of the sisters of the Hunt brothers—Margaret—eventually died after dropping off the list because of a decline in her wealth, for example. And one of the two people who dropped off the list because they renounced their American citizenship eventually died. Carnival Cruise Line founder Ted Arison, who dropped off the list in 1994 because he renounced his American citizenship, died in 1999. For the people who died, they obviously lost all their wealth and much more, but the change in their wealth between 1982 and 2013 was still at least as large as the difference that was described above, at least in some sense.[9]

For the 362 people who were not on the list in 1982 but came onto the list by 2013, the amount by which a person's wealth changed is shown as the difference between the minimum wealth of the 400 wealthiest Americans in any year between 1982 and the first year that the person appeared on the list, on the one hand, and his or her wealth in 2013, on the other hand. The change in a person's wealth would have been at least as large as that difference, if the person could have appeared, but did not appear, in the year in which the minimum wealth of the 400 wealthiest Americans was at its minimum.

Most of the people who came onto the list by 2013 could have appeared on the list in 1982 if only they were wealthy enough, but some could not have appeared, either because they had not been born yet or because they had not become American citizens yet. The minimum wealth of the 400 wealthiest Americans was at its minimum in 1982, so the three people who were born after 1982 (namely, oil-pipeline heir

[9]The one other person who was on the list in 1982 but eventually dropped off the list because he renounced his American citizenship was Campbell Soup Company heir John Thompson Dorrance III. He was still alive and still wealthy enough to be one of the 400 wealthiest Americans in 2013 if only he was still an American citizen (at least according to *Forbes Magazine*'s list of the world's billionaires in 2013), so his change in wealth was better than the change shown in figure D.7.

Scott Duncan and Facebook co-founders Dustin Moskovitz and Mark Zuckerberg) could not have appeared on the list back in that year, although the change in their wealth between 1982 and 2013 was still at least as large as the difference described above, at least in some sense.

The number of people who would have appeared on the list in 1982 if only they had already been American citizens is unknown, but the number might be zero. There were only 38 immigrants on the list in 2013 who were not already on the list in 1982 (Navarro 2013). The wealthiest one of them (namely, Google co-founder Sergey Brin) was almost surely not a multi-millionaire back in 1982, and the earliest year any of them appeared on the list was 1987. George Soros appeared on the list for the first time in that year, but he had been an American citizen since the 1960s. To the extent that any of those immigrants were already multi-millionaires back in 1982, the change in their wealth was worse than the change shown in our figure.

Although some of the changes shown in the figure may be misleading due to censoring, some of the greatest gains seem to have been experienced by people who were not on the list in 1982 but made it onto the list by 2013, while some of the greatest losses seem to have been experienced by people who were on the list in 1982 but dropped off by 2013. Bill Gates was not on the list in 1982, so he must have been worth less than the minimum wealth of the 400 wealthiest Americans in that year, which was about 174 million constant dollars. By 2013, Gates was the wealthiest American and worth 72 billion dollars. Gates fared well between 1982 and 2013, therefore. The wealthiest American in 1982—the shipping mogul Daniel Keith Ludwig—did not fare as well. Ludwig was worth over four billion constant dollars in 1982, but he eventually dropped off the list because he died.

Thus, the changes in wealth among the 400 wealthiest Americans between 1982 and 2013 were perhaps not as pro-wealthy as figure D.6 might suggest. Relatively wealthier members of the 400 wealthiest Americans became relatively wealthier between 1982 and 2013, but the 400 wealthiest Americans in 1982 were not exactly the same people—or even mostly the same people—as the 400 wealthiest Americans in 2013, and people who were the same moved up and down the ranks.

The change in the wealth of the 400 wealthiest Americans as a group between 1982 and 2013 was also perhaps not as pro-wealthy as the increase in their wealth relative to the wealth of everyone else might suggest. The change in the wealth of the 400 wealthiest Americans as a group between those two years can be decomposed into the contributions of the people who appeared on the list in both years, the people who came onto the list, and the people who dropped off. The contribution of people who appeared on the list in both years ("incumbents") is simply the sum of the changes in their wealth. If someone appeared on the magazine's list in both years, and if his or her wealth increased by a given amount, then the wealth of the 400 wealthiest Americans must have also increased by that amount.

The contributions of the people who came onto the list ("entrants") and the people who dropped off the list ("dropouts") can be calculated as the sum of the differences between a person's wealth in 2013 and the minimum wealth of the 400 wealthiest Americans in 1982, in the former case, and the negative of the sum of the differences between the minimum wealth of the 400 wealthiest Americans in 1982 and a person's wealth in 2013, in the latter case. If someone was on the list in 1982 but dropped off by 2013, and if all other things were equal, then the wealth of the 400 wealthiest Americans would have decreased by the difference between his or her initial wealth and the minimum wealth of the 400 wealthiest Americans. Likewise, if someone was not on the list in 1982 but came onto the list by 2013, and if all other things were equal, then the wealth of the 400 wealthiest Americans would have increased by the difference between his or her initial wealth and the minimum wealth of the 400 wealthiest Americans. All other things were not equal, of course. The minimum wealth of the 400 wealthiest increased between 1982 and 2013, in particular. Decomposing the change in the total wealth of the 400 wealthiest Americans in that way still seems sensible, however.

The contributions of the people who dropped off the list can be further decomposed into the contributions of people who dropped off the list because they died ("decedents"), people who dropped off the list because they renounced their American citizenship ("renunciants"), and people who dropped off the list because their wealth declined in either absolute or relative terms ("other dropouts").

Decomposing the 1,807-billion-constant-dollar increase in the wealth of the 400 wealthiest Americans between 1982 and 2013 in such a manner, incumbents added about 273 billion constant dollars (of which Buffett contributed about 21 percent all by himself) and entrants added about 1,652 billion (of which Gates contributed about four percent), while decedents subtracted about 57 billion (of which Ludwig subtracted about eight percent), other dropouts besides renunciants subtracted a similar dollar amount (about 61 billion, of which Nelson Hunt subtracted about four percent), and renunciants subtracted a negligible amount (less than one billion constant dollars). Thus, while some of the increase in the wealth of the 400 wealthiest Americans was due to the gains experienced by people who were one of the 400 wealthiest Americans in both of those years, most of the change was due to the gains experienced by people who were not yet one of the 400 wealthiest Americans back in 1982. There is no guarantee that those people started from a lowly station in life or that they ended up doing anything more than inheriting a fortune, of course, but they were at least not already one of the 400 wealthiest Americans in 1982.

The 1,652-billion-constant-dollar contribution of the people who became one of the 400 wealthiest Americans by 2013 can be decomposed further into the contributions of people who were self-made, at least according to *Forbes Magazine*, and people who were not. According to the magazine, 273 out of the 400 people on its list in 2013 were self-made, in the sense that they "built [their] fortunes themselves" (although they might not have built their fortunes "entirely from scratch" because they might have "borrowed money from in-laws or parents" or "started businesses with spouses or other relatives;" Kroll 2012). All but 17 of the 273 self-made people were not on the list in 1982. The self-made people who made it onto the list by 2013 contributed about 1,134 billion dollars. Seven of those people—including the Microsoft co-founder Bill Gates, the Orcale founder Larry Ellison, the Amazon founder Jeff Bezos, and the Google co-founders Larry Page and Sergey Brin, as well as the media mogul Michael Bloomberg and the casino owner Sheldon Anderson—contributed about 22 percent of that amount or about 248 billion dollars.

The other people who made it onto the list by 2013 but were not self-made contributed about 519 billion. The heirs of Walmart founder Sam Walton—oldest son Rob, youngest son Jim, daughter Alice, and daughter-in-law Christy—contributed about a quarter of that amount or about 135 billion dollars.

The magazine's list therefore suggests that, out of the 1,807-billion-constant-dollar increase in the wealth of the 400 wealthiest Americans between 1982 and 2013, about 792 billion can be attributed to people who were either already one of the 400 wealthiest Americans back in 1982 or became one of the 400 wealthiest Americans by 2013 without building their fortunes by themselves, while about 1,134 billion can be attributed to self-made people who became one of the 400 wealthiest Americans by 2013. The discrepancy is attributable to people who either died, experienced a decline in their wealth, or renounced their American citizenship. The contributions of each group are summarized in table D.1.

Table D.1. Decomposition of the Change in the Wealth of the 400 Wealthiest Americans, 1982–2013

Group	Contribution ($Bs)
Incumbents	273
Entrants	
Self-made entrants	1,134
Other entrants	519
Dropouts	
Decedents	−57
Renunciants	−1
Other dropouts	−61
All groups	1,807

Source: Data adapted from *Forbes Magazine* (1982–2013).

Note: This table shows how the change in the total wealth of the 400 wealthiest Americans between 1982 and 2013 can be decomposed into the contributions made by different groups of unique individuals. The decomposition is discussed in the text. The dollar values are billons of CPI-deflated, 2013 dollars.

APPENDIX E

APPENDIX ON THE CODE

E.1 Code for Creating a Dataset

This appendix includes code for a program that creates the dataset used by each essay in the dissertation, a program that loads that dataset, and programs that generate the results reported by the essays. All of the programs were written in the programming language Python (specifically, Python version 2.7). The programs should be readable without much knowledge of that particular language, although it can be noted that comments are set off by number signs or quotation marks.

This section of the appendix includes the program that creates the dataset used by each essay. That dataset is based on each year of *Forbes Magazine*'s list of the 400 wealthiest Americans between the inaugural year of its list in 1982 and the most-recent year, which is 2013, as of writing. The primary differences between the dataset created by the program and the information reported by the magazine are that unique individuals have the same names across years, some wealth estimates are imputed in the manner discussed in section D.4 of the data appendix, and some of the ages reported by the magazine are corrected in the manner discussed in section D.5 of the same appendix. The dataset is available through ProQuest/UMI along with the full text of this dissertation, as discussed in section D.8.

The next sections of this appendix include programs that load and analyze the dataset created by the program that follows.

‘‘‘
Filename: create_forbes_400_dataset.py
Python version: 2.7
Source: Capehart, Kevin W. Essays on the Wealthiest Americans. PhD dissertation, American University, Washington, DC, 2014.
Description: This program creates a dataset based on Forbes Magazine’s annual list of the 400 wealthiest Americans. The primary differences between the dataset created by this program and the information reported by the magazine are that unique individuals have the same names across years, some wealth estimates are imputed, and some of the ages reported by the magazine are corrected. See the dissertation cited above and the code below for more details on the manner in which unique individuals were identified, some wealths were imputed, and some ages were corrected.
’’’

IMPORT SOME STANDARD MODULES

from __future__ import division # define division as in Python v3
import numpy # import for its many useful functions
from os.path import isfile # import for checking whether a file exists

IMPORT RAW FORBES 400 DATA

The data to be imported is organized so that each row corresponds to an individual on Forbes Magazine’s list of the 400 wealthiest Americans in a given year. Some of the rows also correspond to a family on the magazine’s accompanying list of some of the wealthiest families in America the given year, but we will ignore those rows. (The distinction between those two lists is discussed in the dissertation.) For the rows of interest, the columns correspond to an individual’s name, age, wealth (in millions of current dollars), the specific source of his or her wealth, the general industry in which his or her wealth was made, and whether his or her wealth was self−made. That data is almost identical to the information reported by each year of the magazine’s list, except that some of the names were modified so that unique individuals have the same names across years. See the dissertation for a discussion of how the data was collected and how the unique individuals were identified. Of note, data on whether someone’s wealth was self−made, as well as data on the industry in which someone made his or her wealth, is only available for the years since 1996 when the magazine started making its list available online. Also of note, Forbes discontinued its list of some of the wealthiest families in America after 1999.

# define the filename for the raw data
filename = “./data/raw_forbes_400_data.txt”

assertion error if file does not exist
assert isfile(filename), “Forbes Magazine declined to directly share its data on the 400 wealthiest Americans with the author (personal communication on June 17, 2009, with Tanya Prive, circulation assistant at Forbes Magazine), so the author went through each year of the magazine’s list and collected the data himself. The magazine also declined the author’s request to republish that raw data as part of his dissertation (personal communication on May 8, 2014, with Elena Coster, sales coordinator at Forbes Magazine’s content management firm, PARS International). As such, a file with the raw data collected from the magazine is not included as a supplementary file to the dissertation. The code for processing the rawer version of the data (‘create_forbes_400_dataset.py’) cannot be executed without that version of the data (named ‘raw_forbes_400_data.txt’ by the author). The processed

version of the data ('forbes_400_dataset.txt') is included as a supplementary file to the dissertation, however, and the manner in which that version of the data was produced can be seen by looking at the code for processing the rawer version."

```
# get the header for the raw data
skiprows = 4 # number of rows to skip in file
delimiter = '\t' # file delimiter
f = open(filename, "r") # open file
header = f.readlines()[skiprows-1].replace('\n', '').split(delimiter)
f.close() # close file

# load the data with the headers
rawdata = numpy.loadtxt(filename, delimiter=delimiter, skiprows=skiprows,
    dtype={'names':(header), 'formats':(['<S99']*len(header))})

# get the most-recent year of the list
last_year = max([int(year) for year in set(rawdata['year'])])

## REMOVE ANY FAMILY MEMBERS FROM AN INDIVIDUAL'S NAME ##

# As discussed in the dissertation, an individual's wealth can include the wealth of family
    members in some cases. In such cases, an individual's name is usually given as his or name
    followed by the phrase ' and Family'. We'll remove that phrase.

for i in range(0, len(rawdata)):
  if ' and Family' in rawdata['name'][i]:
    rawdata['name'][i] = rawdata['name'][i].replace(' and Family', '')

# There are also a few cases that require special attention. In 2011, in three cases, a husband
    and wife were listed together, rather than as one individual and their family. The
    husband-wife teams were: Michael and Marian Ilitch, Stewart and Lynda Resnick, and Do
    Won and Jin Sook Chang. Michael Illitch had been on the list in the past, as either himself
    or 'Michael Illitch and Family', so we will remove ' and Marian Ilitch' from the name 'Ilitch,
    Michael and Marian Ilitch'. The other individuals were new to the Forbes 400. Forbes
    appears to report the age of Do Won Chang, so we will remove ' and Jin Sook Chang' from
    the name 'Sook, Jin and Do Won Chang'. Forbes does not report an age for the Resnicks, so
    we will arbitrarily remove ' and Lynda Resnick' from the name 'Resnick, Stewart and Lynda
    Resnick'. We'll do the same for all other years, as well. In 2012 and 2013, the brothers Hank
    and Doug Meijer were listed together, rather than as one individual and their family. Hank
    is the older brother, so we'll remove ' and Doug' from the name 'Meijer, Hank and Doug'.
    For the same years, we'll also arbitrarily remove ' and Judy' from the husband-wife team
    'Love, Tom and Judy' and remove ' and Peggy' from the husband-wife team 'Cherng,
    Andrew and Peggy'. Note that, if someone in any one of those pairs eventually dies while
    the other remains on the list, then, for any sort of longitudinal analysis, we should
    reconsider which family members (if any) should be ignored.

for i in range(0, len(rawdata)):
  if rawdata['name'][i] == 'Ilitch, Michael and Marian Ilitch':
    rawdata['name'][i] = rawdata['name'][i].replace(' and Marian Ilitch', '')
  if rawdata['name'][i] == 'Chang, Do Won and Jin Sook Chang':
    rawdata['name'][i] = rawdata['name'][i].replace(' and Jin Sook Chang', '')
  if rawdata['name'][i] == 'Resnick, Stewart and Lynda Resnick':
```

```
    rawdata['name'][i] = rawdata['name'][i].replace(' and Lynda Resnick', '')
  if rawdata['name'][i] == 'Meijer, Hank and Doug':
    rawdata['name'][i] = rawdata['name'][i].replace(' and Doug', '')
  if rawdata['name'][i] == 'Love, Tom and Judy':
    rawdata['name'][i] = rawdata['name'][i].replace(' and Judy', '')
  if rawdata['name'][i] == 'Cherng, Andrew and Peggy':
    rawdata['name'][i] = rawdata['name'][i].replace(' and Peggy', '')

## GET INDIVIDUALS' NAMES ##

# construct a list of non-unique names
names = dict(zip(['including families', 'excluding families'], \
  [list(), list()]))
for i in range(0, len(rawdata)):
  name = rawdata['name'][i]
  names['including families'].append(name)
  # the age associated with a family is 'n.a.' in the raw data in most cases
  if rawdata['age'][i] != 'n.a.':
    names['excluding families'].append(name)

# get unique names
for _ in ['including families', 'excluding families']:
  names[_] = list(set(names[_]))

# If excluding families, exclude family names that have associated ages that were not excluded
    by the code excuted above. It is unclear why these names have ages associated with them in
    the magazine, but they are families and not individuals.
family_names_to_remove = ['Alfond Family', 'Breed Family', 'Magness Family', 'Martin Family',
    'Miner Family', 'Ward Family']
for name in family_names_to_remove:
  names['excluding families'].remove(name)

# alphabetize names
for _ in ['including families', 'excluding families']:
  names[_].sort()

## EXPORT DATA ORGANIZED BY UNIQUE INDIVIDUALS ##

# if the file already exists, set the following to True to recreate it
re_create_file = False
filename = "./data/reorganized_forbes_400_data_1982_to_" + str(last_year) + ".txt"
if not isfile(filename) or re_create_file:
  # open file
  table = open(filename, 'w')
  # write first header for the table
  infos = ['wealth', 'age', 'source', 'industry', 'selfmade or inherited']
  years = [str(y) for y in range(1982, last_year+1, 1)]
  header = 'name\t'
  for info in infos:
    header += info + '\t'
    for year in years[1:]:
      if (info == infos[-1]) and (year == years[-1]):
```

```
            header += '\n'
        else:
            header += '\t' # if last info and last year, then break line
    table.write(header)
    # write second header for the table
    header = '/year\t'
    for info in infos:
        for year in years:
            header += year
            if (info == infos[−1]) and (year == years[−1]):
                header += '\n'
            else:
                header += '\t'
    table.write(header)
    # write data
    for name in names['excluding families']: # note: excluding families
        table.write(name + '\t') # write name
        for info in infos:
            for year in years:
                for i in range(0, len(rawdata)):
                    if (name == rawdata['name'][i]) and (year == rawdata['year'][i]):
                        table.write(rawdata[info][i])
                        break
                if (info == infos[−1]) and (year == years[−1]):
                    table.write('\n')
                else:
                    table.write('\t')
    # close file
    table.close()
else:
    print "Note that 're_create_file' must be set to True in order to recreate the file that
        reorganizes the raw data."

## IMPORT DATA THAT WAS JUST EXPORTED ##

# infos and years
infos = ['wealth', 'age', 'source', 'industry', 'selfmade or inherited']
years = [str(y) for y in range(1982, last_year+1, 1)]
# define header
header = ['name']
for info in infos:
    header.extend(info+year for year in years)
# import data
rawdata = numpy.loadtxt(filename, delimiter='\t', skiprows=2, \
    dtype={'names':(header), 'formats':(['<S99']*len(header))})
# create a dictionary of the data
data = dict(zip(rawdata['name'], [dict(zip(['wealth', 'age', 'source', 'industry', 'selfmade or
        inherited'], [dict(zip(years, [rawdata['wealth'+y][i] for y in years])), dict(zip(years,
        [rawdata['age'+y][i] for y in years])), dict(zip(years, [rawdata['source'+y][i] for y in years])),
        dict(zip(years, [rawdata['industry'+y][i] for y in years])), dict(zip(years, [rawdata['selfmade
        or inherited'+y][i] for y in years]))])) for i in range(0, len(rawdata['name']))]))
# get list of names sorted alphabetically
```

```
names = sorted(data.keys())

## USE OTHER INFO TO SOME FIX AGES ##

# The following age is missing, but it can be fixed with other information from forbes.
    According to http://www.forbes.com/global/2001/0709/030tab1.html, David Oros was 41 in
    2001, so he would've been about 40 in 2000.
data['Oros, David S']['age']['2000'] = '40'

## EXTRAPOLATE AND CORRECT SOME AGES ##

# some ages are "bad" in the sense that they're not reported or not reported as an integer
bad_ages = ['n.a.', 'blank', 'page23', 'died after press closing', \
  'unknown', '40s', '50s', 'early50s', '60s', 'late60s', '70s', '70+', 'early70s', \
  'mid70s', '75+', 'late70s', '80s', 'late80s']

# keep track of names with ages that are extrapolated or corrected
names_with_ages_extrapolated, names_with_corrected_ages = [], []

# fix the following ages by extrapolating forward
data['Lauder, Estee']['age']['1995'] = '87' # roll her age forward
data['Busch, August Anheuser Jr']['age']['1989'] = '90' # roll his age forward
names_with_ages_extrapolated.append('Lauder, Estee')
names_with_ages_extrapolated.append('Busch, August Anheuser Jr')

# define years in descending order
descending_years = [str(y) for y in range(last_year, 1982-1, -1)]

# fix the following ages by extrapolating backwards or correcting inconsistencies
# for each name
for name in names:
  # define a variable for the last age that was observed, extrapolated backwards, or corrected
  last_age = ''
  # look at each year in descending order
  for year in descending_years:
    # get the name's age in a given year
    age = data[name]['age'][year]
    # if an age is not observed in the given year and an age wasn't observed in later years, then
        pass
    if (age == '') and (last_age == ''):
      pass
    # if an age is not observed in the given year, but an age was observed in later years
    elif (age == '') and (last_age != ''):
      # continue to extrapolate the observed age backwards by subtracting unity
      last_age = str(int(last_age) - 1)
    # if a bad age is observed in the given year, but an age was observed in later years
    elif (age != '') and (age in bad_ages) and (last_age != ''):
      # extrapolate the observed age backwards by subtracting unity
      data[name]['age'][year] = str(int(last_age) - 1)
      # redefine the last age as this age
      last_age = str(int(last_age) - 1)
      # keep track of the names with extrapolated ages
```

```
    names_with_ages_extrapolated.append(name)
  # if a good age is observed in the given year and an age wasn't observed in later years
  elif (age != '') and (age not in bad_ages) and (last_age == ''):
    # redefine the last age as this age
    last_age = data[name]['age'][year]
  # if a good age is observed in the given year and an age was also observed in later years
  elif (age != '') and (age not in bad_ages) and (last_age != ''):
    # check to see if the age is consistent by calculating the difference between the age
        observed in later years extrapolated backwards, on the one hand, and the age
        observed in the given year, on the other hand.
    age_difference = int(last_age) - int(age)
    # depending on when their birthday occurs, the difference should be between zero and
        two years
    # if age is consistent, then redefine last age
    if 2 >= age_difference >= 0:
      last_age = age
    # if age is inconsistent
    else:
      # correct it
      data[name]['age'][year] = str(int(last_age) - 1)
      # redefine the last age as this age
      last_age = str(int(last_age) - 1)
      # keep track of the names with corrected ages
      names_with_corrected_ages.append(name)

# EXTRAPOLATIONS: If a person's age was missing in a given year, but she appeared on the
    list in a latter year with a non-missing age, then it was assumed that the most-recent
    non-missing age was correct, and the earlier missing age was extrapolated based on it. The
    ages of 22 people (61 observations) were extrapolated, in total.
#print "The ages of " + str(len(list(set(names_with_ages_extrapolated)))) + " people (" +
    str(len(names_with_ages_extrapolated)) + " observations) were extrapolated, in total. "

# CORRECTIONS: Note that, if the ages are correct, then the age difference for a given person
    between two lists should be no less than zero years (or else they would've grown younger)
    and at most two years (since the list is not constructed on the same date every year, it is
    possible that a person could have two birthdays between two lists). If those conditions did
    not hold, then we assumed that the most-recent age was correct while the latter age was
    incorrect, which is reasonable since new information presumably came to light. The ages of
    53 people (111 observations) were corrected, in total.
#print "The ages of " + str(len(list(set(names_with_corrected_ages)))) + " people (" +
    str(len(names_with_corrected_ages)) + " observations) were corrected, in total. "

## IMPUTE SOME AGES ##

# For each year, we'll impute unknown ages as the average age of a relevant group. M.S. Forbes'
    ages will be inclued in all of the imputations, as they should be, given that he was a
    member of the Forbes 400.

# create list of age categories
age_categories = ['all age categories']
rough_ages = ['40s', '60s', '70s', '80s', 'late70s', '70+', '75+']
age_categories.extend(rough_ages)
```

```
# create dict of ages by year by age category
ages_in = dict(zip(age_categories, [dict(zip(years, [[] for year in years])) for age_category in
    age_categories]))

# define a function that takes an age and returns an age group
def get_rough_age(age):
  if 40 <= age < 50:
    rough_age = rough_ages[0]
  elif 60 <= age < 70:
    rough_age = rough_ages[1]
  elif 70 <= age < 80:
    rough_age = rough_ages[2]
  elif 80 <= age < 90:
    rough_age = rough_ages[3]
  else:
    rough_age = 'ignore'
  return rough_age

# fill the dictionary
for name in names:
  for year in years:
    age = data[name]['age'][year]
    if (age != '') and (age not in bad_ages):
      ages_in['all age categories'][year].append(int(age))
      rough_age = get_rough_age(int(age))
      # age categories
      if rough_age != 'ignore':
        ages_in[rough_age][year].append(int(age))
      # late 70s
      if 75 <= int(age) < 80:
        ages_in['late70s'][year].append(int(age))
      # 70+
      if int(age) >= 70:
        ages_in['70+'][year].append(int(age))
      # 75 +
      if int(age) >= 75:
        ages_in['75+'][year].append(int(age))

# calculate average age by age category by year
avg_age_in = dict(zip(age_categories, [dict(zip(years,
    [str(int(round(numpy.mean(ages_in[age_category][year])))) for year in years])) for
    age_category in age_categories]))

# 'unknown' ages are imputed as the average age for the given year (3 people, 5 observations)
data['Pan, Jing Jong']['age']['2000'] = avg_age_in['all age categories']['2000']
data['Pan, Theresa']['age']['2000'] = avg_age_in['all age categories']['2000']
data['Wiskemann, Elizabeth S']['age']['2004'] = avg_age_in['all age categories']['2004']
data['Wiskemann, Elizabeth S']['age']['2005'] = avg_age_in['all age categories']['2005']
data['Wiskemann, Elizabeth S']['age']['2006'] = avg_age_in['all age categories']['2006']

# conditional ages are imputed as average conditional age (12 people, 29 observations)
data['Bettingen, Burton Green']['age']['1982'] = avg_age_in['70+']['1982']
```

data[‘Bettingen, Burton Green’][‘age’][‘1983’] = avg_age_in[‘70+’][‘1983’]
data[‘Bettingen, Burton Green’][‘age’][‘1984’] = avg_age_in[‘70s’][‘1984’]
data[‘Bettingen, Burton Green’][‘age’][‘1985’] = avg_age_in[‘70s’][‘1985’]
data[‘Bettingen, Burton Green’][‘age’][‘1986’] = avg_age_in[‘70s’][‘1986’]
data[‘Copeland, Gerret van Sweringen’][‘age’][‘1985’] = avg_age_in[‘40s’][‘1985’]
data[‘Copeland, Pamela Cunningham’][‘age’][‘1985’] = avg_age_in[‘70s’][‘1985’]
data[‘Franchetti, Anne’][‘age’][‘1991’] = avg_age_in[‘70s’][‘1991’]
data[‘Franchetti, Anne’][‘age’][‘1992’] = avg_age_in[‘70s’][‘1992’]
data[‘Franchetti, Anne’][‘age’][‘1993’] = avg_age_in[‘70s’][‘1993’]
data[‘Franchetti, Anne’][‘age’][‘1994’] = avg_age_in[‘70s’][‘1994’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1982’] = avg_age_in[‘70+’][‘1982’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1983’] = avg_age_in[‘70s’][‘1983’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1984’] = avg_age_in[‘70s’][‘1984’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1985’] = avg_age_in[‘70s’][‘1985’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1986’] = avg_age_in[‘70s’][‘1986’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1987’] = avg_age_in[‘late70s’][‘1987’]
data[‘Green, Dorothy (Dolly)’][‘age’][‘1988’] = avg_age_in[‘80s’][‘1988’]
data[‘Jones, Arthur’][‘age’][‘1983’] = avg_age_in[‘60s’][‘1983’]
data[‘Kauffman, Muriel’][‘age’][‘1994’] = avg_age_in[‘70s’][‘1994’]
data[‘Lauder, Joseph’][‘age’][‘1982’] = avg_age_in[‘70s’][‘1982’]
data[‘Rains, Liliore Green’][‘age’][‘1982’] = avg_age_in[‘70+’][‘1982’]
data[‘Rains, Liliore Green’][‘age’][‘1983’] = avg_age_in[‘70s’][‘1983’]
data[‘Rains, Liliore Green’][‘age’][‘1984’] = avg_age_in[‘70s’][‘1984’]
data[‘Rains, Liliore Green’][‘age’][‘1985’] = avg_age_in[‘70s’][‘1985’]
data[‘Rains, Liliore Green’][‘age’][‘1986’] = avg_age_in[‘70s’][‘1986’]
data[‘Whittier, Leland K’][‘age’][‘1982’] = avg_age_in[‘75+’][‘1982’]
data[‘Whittier, N Paul’][‘age’][‘1982’] = avg_age_in[‘75+’][‘1982’]
data[‘Woodward, Helen Whittier’][‘age’][‘1982’] = avg_age_in[‘75+’][‘1982’]

INTERPOLATIONS: The ages of 15 people (34 observations) were imputed, of which 3 people (5 observations) had unknown ages and 12 people (29 observations) had conditional ages (e.g., “70s”). Unknown ages were imputed as the year−specific average age. Conditional ages were imputed as the age−specific average conditional age; “70s” was imputed as the average age of people in their 70s, for example.

SUMMARY: The ages of 22 people (61 observations) were extrapolated. The ages of 53 people (111 observations) were corrected. And the ages of 15 people (34 observations) were imputed. So, in total, the ages of 90 people (206 observations) were extrapolated, corrected, or imputed. That represents a relatively small number of people: about six percent, specifically, 100. ∗ (90. / 1474.) percent, of the (unique) people in the sample. It also represents a relatively small number of the age observations: about two percent, specifically, 100. ∗ (206. / (400. ∗ 32.)) percent, of the age observations

IMPUTE THE WEALTH OF M.S. FORBES

M.S. Forbes was one of Forbes 400 in the 1982 to 1989 lists (dying by 1990 list), but the magazine did not report an estimate of his wealth. We’ll impute his wealth in a given year as the median wealth of the other 399 wealthiest Americans on the list in the given year

# remove M.S. Forbes from the list of names
names.remove(‘Forbes, Malcolm Stevenson’)
# calculate median wealth with M.S. Forbes

```
median_wealth_in = dict(zip(years, [[] for year in years]))
for year in years:
  for name in names:
    wealth = data[name]['wealth'][year]
    if wealth != '':
      median_wealth_in[year].append(float(wealth))
for year in years:
  median_wealth_in[year] = numpy.median(median_wealth_in[year])
# impute his wealth as the median wealth (1 person, 8 observations)
for year in [str(y) for y in range(1982, 1989+1)]:
  data['Forbes, Malcolm Stevenson']['wealth'][year] = str(int(median_wealth_in[year]))
# add M.S. Forbes back to list of names and sort the names alphabetically
names.append('Forbes, Malcolm Stevenson')
names.sort()

# SUMMARY: The wealths of one person (eight observations) were imputed as the
    year-specific median wealth.

## INDUSTRY DATA ##

# impute missing industries
data['Anderson, John Edward']['industry']['2000'] = 'Beverages'
data['Diller, Barry']['industry']['2000'] = 'Media/Entertainment'
data['Disney, Roy Edward']['industry']['2000'] = 'Media/Entertainment'
data['Eisner, Michael D']['industry']['2000'] = 'Media/Entertainment'
data['Gallo, Ernest']['industry']['2000'] = 'Beverages'
data['Gonda, Louis L']['industry']['2000'] = 'Investments' # source of wealth in 1999 was 'AIG
    stock', so industry should be investments. Industry was also investments in 1999
data['Kimmel, Sidney']['industry']['2000'] = 'Apparel'
data['Stowers, James Evans Jr']['industry']['2000'] = 'Finance'

# change 'Technology/Medicine' to 'Technology' for the following
data['Bezos, Jeffrey P']['industry']['2009'] = 'Technology'
data['Dangermond, Jack']['industry']['2009'] = 'Technology'
data['Dolby, Ray Milton']['industry']['2009'] = 'Technology'
data['Gates, William Henry III']['industry']['2009'] = 'Technology'
data['Kim, James']['industry']['2009'] = 'Technology'
data['Sun, David']['industry']['2009'] = 'Technology'
data['Zuckerberg, Mark']['industry']['2009'] = 'Technology'

# change 'Technology/Medicine' to 'Healthcare' for the following
data['Brown, John W']['industry']['2009'] = 'Healthcare'
data['Rahr, Stewart']['industry']['2009'] = 'Healthcare'

# should be 'Investments' for the following
data['Picower, Jeffry']['industry']['2009'] = 'Investments'

# change to 'Information Technology', 'Internet', and 'Software' to 'Technology' for simplicity
# also change 'Investments' and 'Finance/Investment' to 'Finance' for simplicity
for year in years:
  for name in names:
    wealth = data[name]['wealth'][year]
```

```
    if wealth != '':
      industry = data[name]['industry'][year]
      # change to 'Information Technology', 'Internet', and 'Software' to 'Technology'
      if industry in ['Information Technology', 'Internet', 'Software']:
        data[name]['industry'][year] = 'Technology'
      # change 'Investments' and 'Finance/Investment' to 'Finance'
      if industry in ['Investments', 'Finance/Investment']:
        data[name]['industry'][year] = 'Finance'

## SELF-MADE OR INHERITED DATA ##

# impute missing
data['McGlothlin, James']['selfmade or inherited']['1997'] = 'Self made' # this imputation can be
    justified in different ways, but based on earlier years of the Forbes 400 list and other
    information, he does not appear to have inherited his wealth
data['Corn, Elizabeth Turner']['selfmade or inherited']['1996'] = 'Inherited' # if someone's
    priminary source of wealth is 'Inheritance' and wealth is inherited in other years, then
    wealth shouldn't be 'Self made'
data['Boudjakdji, Millicent V']['selfmade or inherited']['1997'] = 'Inherited' # same justification
# it should be noted that 'Copley, Helen Kinney' has a priminary source of 'Inheritance' and
    her wealth is 'Self made' in 1996, but it's not clear that this is an error when looking at
    other years, so we'll leave it be

for year in years:
  for name in names:
    wealth = data[name]['wealth'][year]
    if wealth != '':
      # change mixed to 'Both'
      type = data[name]['selfmade or inherited'][year]
      if type in ['Both', 'Built up inheritance', 'Inherited and growing']:
        data[name]['selfmade or inherited'][year] = 'Both'

## EXPORT PROCESSED DATA ##

# open file
filename = "./data/forbes_400_dataset_1982_to_" + str(last_year) + ".txt"
table = open(filename, 'w')
# write license
table.write("This work is licensed under a Creative Commons Attribution-ShareAlike 4.0
    International License. For more information on that license see
    http://creativecommons.org/licenses/by-sa/4.0/. Under the terms of the license, you are
    free to share and adapt the work for any purpose, but you must give appropriate credit and,
    if any changes are made to the original version, you must distribute your adapted version
    under the same license as the original. Preferred citation is: Capehart, Kevin W. Essays on
    the Wealthiest Americans. PhD dissertation, American University, Washington, DC,
    2014.\n\n")
# wite notes
table.write("Note:\n")
table.write("The following table is based on each year of Forbes Magazine's list of the 400
    wealthiest Americans between its inaugural year in 1982 and its most-recent year as of
    writing, which was 2013. Each row of the table corresponds to a unique individual who
    appeared on the magazine's list in at least one year between those years. Unique individuals
```

```
    were identified by the author in the manner discussed in the dissertation. For individuals
    who appeared on the magazine's list in a given year, the columns of the table show an
    individual's wealth (in millions of current dollars), his or her age, the specific source of his
    or her wealth, the general industry in which his or her wealth was made, and whether his or
    her wealth was self−made. If an individual did not appear on the list in a given year, then
    the corresponding entries in the table are blank. All of the entries in the table are based on,
    but not always the same as, the information reported by the magazine. As discussed in the
    dissertation, the magazine did not report a wealth estimate for Malcolm Stevenson Forbes
    during the years in which he appeared on its list, so we imputed his wealth in a given year
    as the median wealth of the 399 other people on the list in the given year. Some of the ages
    are also different than those reported by the magazine for the reasons discussed in the
    dissertation. Of note, information on whether someone's wealth was self−made, as well as
    information on the industry in which someone made his or her wealth, is only available for
    the years since 1996 when the magazine started making its list available online. \n\n")
# write first header for the table
header = '\t'
for info in infos:
  header += info + ' in ...' + '\t'
  for year in years[1:]:
    if (info == infos[−1]) and (year == years[−1]):
      header += '\n'
    else:
      header += '\t' # if last info and last year, then break line
table.write(header)
# write second header for the table
header = 'name/year\t'
for info in infos:
  for year in years:
    header += year
    if (info == infos[−1]) and (year == years[−1]):
      header += '\n'
    else:
      header += '\t'
table.write(header)
# write data
for name in names:
  table.write(name + '\t') # write name
  for info in infos:
    for year in years:
      table.write(data[name][info][year])
      if (info == infos[−1]) and (year == years[−1]):
        table.write('\n')
      else:
        table.write('\t')
# close file
table.close()
```

E.2 Code for Loading the Dataset

This section includes a program that loads the dataset used by each of the essays in the dissertation. Again, the data is based on *Forbes Magazine*'s annual list of the 400 wealthiest Americans. Programs specific to each essay are included in subsequent sections of this appendix. All of those programs execute the code in this section in order to load the dataset. The code is as follows.

```
'''
Filename: load_forbes_400_dataset.py
Python version: 2.7
Source: Capehart, Kevin W. Essays on the Wealthiest Americans. PhD dissertation, American
    University, Washington, DC, 2014.
Description: This file loads a dataset based on Forbes Magazine's annual list of the 400 wealthiest
    Americans. See the dissertation cited above for more details on the dataset. Some other
    relevant sets of data are also loaded, including data on the dates for which the magazine tried
    to take a snapshot of wealth, the people who dropped off the magazine's list because they
    died, and the people who dropped off the list because they renounced their American
    citizenship.
'''

from __future__ import division
import numpy
import datetime # import for dealing with dates

## LOAD FORBES 400 DATA ##

# define the most-recent year of the list
last_year = 2013
# define the years between the earliest and most-recent lists
years = [str(y) for y in range(1982, last_year+1)]
# define some other sets of years that will prove useful
years_less_initial = [str(y) for y in range(1983, last_year+1)]
years_less_last = [str(y) for y in range(1982, (last_year-1)+1)]
years_less_initial_and_last = [str(y) for y in range(1983, (last_year-1)+1)]
descending_years = [str(y) for y in range(last_year, 1982-1, -1)]

# load the Forbes 400 data from the data subdirectory
filename = "./data/forbes_400_dataset_1982_to_" + str(last_year) + ".txt"
header = ['name']
for info in ['wealth','age','source','industry','selfmade or inherited']:
  header.extend(info+year for year in years)
rawdata = numpy.loadtxt(filename, delimiter='\t', skiprows=7, dtype={'names':(header),
    'formats':(['<S99']*len(header))})
# create a dictionary of the data
data = dict(zip(rawdata['name'], [dict(zip(['wealth', 'age', 'source', 'industry', 'selfmade or
    inherited'], [dict(zip(years, [rawdata['wealth'+y][i] for y in years])), dict(zip(years,
```

```
    [rawdata['age'+y][i] for y in years])), dict(zip(years, [rawdata['source'+y][i] for y in years])),
    dict(zip(years, [rawdata['industry'+y][i] for y in years])), dict(zip(years, [rawdata['selfmade
    or inherited'+y][i] for y in years]))]))) for i in range(0, len(rawdata['name']))]))
# get list of names sorted alphabetically
names = sorted(data.keys())

# load data on Forbes 400 members who dropped out because they died
filename = "./data/exit_by_death_data.txt"
header = ['name', 'year']
rawdata = numpy.loadtxt(filename, delimiter='\t', skiprows=2, dtype={'names':(header),
    'formats':(['<S50' for _ in range(0, len(header))])})
# get dict of names of the dead by year
year_dead_in_for = dict(zip(rawdata['name'], rawdata['year']))
dead_in = dict(zip(years_less_initial, [[] for year in years_less_initial]))
for name in year_dead_in_for.keys():
  year = year_dead_in_for[name]
  if year in years:
    dead_in[year].append(name)

# load data on Forbes 400 members who dropped out because they renounced their citizenship
filename = "./data/exit_by_renunciation_of_citizenship_data.txt"
header = ['name', 'year']
rawdata = numpy.loadtxt(filename, delimiter='\t', skiprows=8, dtype={'names':(header),
    'formats':(['<S50' for _ in range(0, len(header))])})
# get dict of names of those who renounced their citizenship by year
year_renounced_citizenship_in_for = dict(zip(rawdata['name'], rawdata['year']))
renunciants_in = dict(zip(years_less_initial, [[] for year in years_less_initial]))
for name in year_renounced_citizenship_in_for.keys():
  year = year_renounced_citizenship_in_for[name]
  if year in years:
    renunciants_in[year].append(name)

## RECORD SOME SUMMARY STATISTICS ##

# Note that M.S. Forbes and his imputed wealth are included in the following.

# define dicts for storing infos
infos = ['wealths', 'len', 'sum', 'min', 'median', 'mean', 'max']
summary_stats = dict(zip(infos, [dict(zip(years, [[] for year in years])) for _ in infos]))

# get sorted wealths by year
for year in years:
  for name in names:
    wealth = data[name]['wealth'][year]
    if (wealth != ''):
      summary_stats['wealths'][year].append(float(wealth))
  summary_stats['wealths'][year].sort(reverse=True) # sort in descending order

# get stats on wealths
for year in years:
  summary_stats['len'][year] = len(summary_stats['wealths'][year])
  summary_stats['sum'][year] = sum(summary_stats['wealths'][year])
```

```
  summary_stats['min'][year] = min(summary_stats['wealths'][year])
  summary_stats['median'][year] = numpy.median(summary_stats['wealths'][year])
  summary_stats['mean'][year] = numpy.mean(summary_stats['wealths'][year])
  summary_stats['max'][year] = max(summary_stats['wealths'][year])

## DATES OF VALUATION ##

# load the dates (formatted as year, month, day) when Forbes priced publicly traded stocks
valued_on_in = dict(zip(years, [datetime.date(1982, 8, 25), # imputed (see comments below)
  datetime.date(1983, 8, 24),
  datetime.date(1984, 8, 17),
  datetime.date(1985, 9, 6),
  datetime.date(1986, 9, 12),
  datetime.date(1987, 9, 11),
  datetime.date(1988, 8, 30),
  datetime.date(1989, 9, 8),
  datetime.date(1990, 9, 5),
  datetime.date(1991, 9, 4),
  datetime.date(1992, 8, 27), # last Thursday of August
  datetime.date(1993, 8, 27), # last Friday of August *
  datetime.date(1994, 8, 31), # last day of August, a Wednesday
  datetime.date(1995, 8, 23), # second to last Wednesday of August
  datetime.date(1996, 8, 23), # second to last Friday of August *
  datetime.date(1997, 8, 22), # second to last Friday of August *
  datetime.date(1998, 9, 1), # first day of September, a Tuesday
  datetime.date(1999, 8, 27), # imputed
  datetime.date(2000, 8, 25), # imputed
  datetime.date(2001, 8, 27), # last Monday of August
  datetime.date(2002, 8, 16), # mid-August, a Friday *
  datetime.date(2003, 8, 29), # imputed
  datetime.date(2004, 8, 27), # imputed
  datetime.date(2005, 8, 26), # imputed
  datetime.date(2006, 8, 31), # last day of August, a Thursday
  datetime.date(2007, 8, 31), # last day of August, a Friday *
  datetime.date(2008, 8, 29), # last Friday of August *
  datetime.date(2009, 9, 10), # second Thursday of September
  datetime.date(2010, 8, 25), # last Wednesday of August
  datetime.date(2011, 8, 26), # last Friday of August *
  datetime.date(2012, 8, 24), # second to last Friday of August *
  datetime.date(2013, 8, 23)])) # second to last Friday of August *

# In the 1982 Forbes 400, the valuation date was reported as "mid-August," 1982. We'll
    impute the valuation date as Wednesday, August 25, 1982, based on the fact that the
    valuation date for the 1983 Forbes 400 was Wednesday, August 24, 1983. To our knowledge,
    neither Forbes Magazine nor Forbes.com reported the valuation dates in 1999, 2000, or 2003
    to 2005. We'll impute the valuation dates as the last Friday of August of the respective
    years, given that, in recent years (i.e., since 1992), most of the valuation dates (9 out of 15)
    have been on a Friday towards the end of August (the last or second to last Friday of the
    month).
years_w_imputed_valuation_dates = ['1982', '1999', '2000', '2003', '2004', '2005']
```

E.3 Code for the First Essay

The code for generating the results reported in the first essay is as follows.

```
'''
Filename: essay_on_estimation_of_wealth_code.py
Python version: 2.7
Source: Capehart, Kevin W. Essays on the Wealthiest Americans. PhD dissertation, American
    University, Washington, DC, 2014.
Description: This file generates results reported by the essay on the estimation of the wealth of
    the wealthiest Americans.
'''

from __future__ import division
from load_forbes_400_dataset import * # load our Forbes 400 dataset
import decimal # import for trying to avoid any severe floating point errors
import scipy.stats # import for calculating quantiles
from os.path import isfile # import for checking whether a file exists
import pickle # import for dumping and loading computationally intensive results

## DEFINE A FUNCTION FOR CONVERTING FLOATS ##

# The estate-multiplier method often involves taking the reciprocal of small mortality rates
    and multiplying them by large wealths, so we will use Python's module for decimal floating
    point arithmetic in order to try to avoid any severe floating point errors

def dec(floating_point):
  ''' Returns a decimal instead of a float '''
  return decimal.Decimal(str(floating_point))

## LOAD SOME DATA ##

# load the names of the women who dropped off the Forbes 400 because they died
filepath = "./data/data_specific_to_first essay/"
filename = filepath + "women_who_died_data.txt"
raw_data = numpy.loadtxt(filename, delimiter='\t', skiprows=2, dtype={'names':['name'],
    'formats':['<S50']})
names_of_dead_who_were_women = raw_data['name']

# load mortality rates used by Kopczuk and Saez (2004a,b)
filename = filepath + "mortality_rates_data.txt"
header = ['year', 'age', 'male', 'female', 'total']
raw_mortality_rates = numpy.loadtxt(filename, delimiter='\t', skiprows=9,
    dtype={'names':header, 'formats':['<S50']*len(header)})
filename = filepath + "relative_mortality_rates_data.txt"
header = ['age', 'male', 'female']
raw_relative_mortality_rates = numpy.loadtxt(filename, delimiter='\t', skiprows=7,
    dtype={'names':header, 'formats':['<S50']*len(header)})

# define dicts for mortality rates
genders = ['male', 'female']
```

```
genders_and_total = [‘male’, ‘female’, ‘total’]
ages = raw_relative_mortality_rates[‘age’]
years_btw_1982_and_2012 = [str(y) for y in range(1982, 2012+1)]
years_btw_1982_to_2012_but_also_1918 = [‘1918’]
years_btw_1982_to_2012_but_also_1918.extend([year for year in years_btw_1982_and_2012])
mortality_rate_of = dict(zip(genders_and_total, [dict(zip(years_btw_1982_to_2012_but_also_1918,
    [dict(zip(raw_mortality_rates[‘age’][i:i+119+1],
    raw_mortality_rates[gender_or_total][i:i+119+1])) for i in range(0, (2012−1982+1+1)∗119,
    119+1)])) for gender_or_total in genders_and_total]))
relative_mortality_rate_of = dict(zip(genders, [dict(zip(raw_relative_mortality_rates[‘age’],
    raw_relative_mortality_rates[gender])) for gender in genders]))

# load Kopczuk and Saez (2004a,b) data on wealthiest 0.01%
filename = filepath + “ks_wealthiest_0_01_percent_data.txt”
header = [‘year’, ‘population’, ‘top 0.01% average wealth’, ‘top 0.01% wealth threshold’]
years_btw_1982_and_2000 = [str(y) for y in range(1982, 2000+1)] # only have estimates up to
    the year 2000
raw_ks_wealthiest_0_01_percent_data = numpy.loadtxt(filename, delimiter=‘\t’, skiprows=4,
    dtype={‘names’:header, ‘formats’:[‘<S50’]∗len(header)})
ks_wealthiest_0_01_percent_data = dict(zip([‘population’, ‘top 0.01% average wealth’, ‘top 0.01%
    wealth threshold’], [dict(zip(years_btw_1982_and_2000,
    [raw_ks_wealthiest_0_01_percent_data[_][i] for i in range(0, 2000−1982+1)])) for _ in
    [‘population’, ‘top 0.01% average wealth’, ‘top 0.01% wealth threshold’]]))

# load the price index used by Kopczuk and Saez (2004a,b), which is an earlier version of the
    CPI−U−RS with 2000 as the base year
filename = filepath + “ks_price_index_data.txt”
raw_data = numpy.loadtxt(filename, delimiter=‘\t’, skiprows=3, dtype={‘names’:[‘year’,
    ‘index’], ‘formats’:[‘<S50’,‘f’]})
price_index_used_by_ks = dict(zip(raw_data[‘year’], raw_data[‘index’]))

# load the latest version of the CPI−U−RS
filename = “./data/cpi_u_rs_data.txt”
raw_cpi_data = numpy.loadtxt(filename, delimiter=‘\t’, skiprows=4, dtype={‘names’:[‘year’,
    ‘index’], ‘formats’:[‘<S50’,‘f’]})
cpi_u_rs_w_1977_as_base_year = dict(zip(raw_cpi_data[‘year’], raw_cpi_data[‘index’]))
cpi_u_rs_w_2000_as_base_year = dict(zip(years, [dec(100.0) ∗
    (dec(cpi_u_rs_w_1977_as_base_year[year]) / dec(cpi_u_rs_w_1977_as_base_year[‘2000’])) for year
    in years]))

## DEFINE FUNCTIONS FOR CONVERTING DOLLAR AMOUNTS ##

# Kopczuk and Saez’s (2004a,b) dollar figures are in CPI−U−RS dollars, but the version of the
    CPI−U−RS that they used is an older version, so we will re−inflate their constant dollar
    figures into current dollars in order to then deflate them using the latest version of the index

def not_in_2000_dollars(constant_dollar_amount, year):
  ‘‘‘ Returns a 2000 dollar amount in current dollars using the old version of the CPI-U-RS used
      by Kopczuk and Saez (2004a,b) ’’’
  return constant_dollar_amount ∗ (dec(price_index_used_by_ks[year]) / dec(100.0))

def in_2000_dollars(current_dollar_amount, year):
```

```
    ''' Returns a current dollar amount in 2000 dollars using the latest version of the CPI-U-RS '''
    return current_dollar_amount / (dec(cpi_u_rs_w_2000_as_base_year[year]) / dec(100.0))

## SEC. A.4: K&S VS. SCF ESTIMATES OF THE WEALTH OF WEALTHIEST ONE
    PERCENT ##

# The following code compares Kopczuk and Saez's (2004a,b) estate-multiplier estimates of the
    total wealth of the wealthiest one percent, on the one hand, to survey estimates of that
    group's wealth.

# load SCF data
years_of_scf = ['1989', '1992', '1995', '1998', '2001', '2004', '2007', '2010']
scf_mean_wealth_in_thousands_of_2009_dollars_in = dict(zip(years_of_scf, [315.3, 279.5, 296.1,
    373.2, 483.0, 509.6, 575.9, 493.1])) # Source: Table 2 of Kennickell's (2012) "The Other,
    Other Half: Changes in the Finances of the Least Wealthy 50%, 2007-2009."
number_of_households_in_millions = dict(zip(years_of_scf, [93.1, 95.9, 99.0, 102.6, 106.5, 112.1,
    116.1, 117.6])) # Sources: Kennickell and Starr (1994, p. 880, n. 22); Kennickell et al. (2000,
    p. 27, n. 35); Bricker et al. (2012, p. 78, table A.3)
perct_of_wealth_held_by_wealthiest_one_perct_in = dict(zip(years_of_scf, [30.1, 30.1, 34.6, 33.9,
    32.6, 33.4, 33.8, 34.5])) # Source: Ibid.

# define a function for converting Kennickell's 2009 constant dollars to current dollars
cpi_u_rs_w_2009_as_base_year = dict(zip(years, [dec(100.0) *
    (dec(cpi_u_rs_w_1977_as_base_year[year]) / dec(cpi_u_rs_w_1977_as_base_year['2009'])) for year
    in years]))
def not_in_2009_dollars(constant_dollar_amount, year):
    ''' Returns a 2009 dollar amount in current dollars using the latest version of the CPI-U-RS '''
    return constant_dollar_amount * (dec(cpi_u_rs_w_2009_as_base_year[year]) / dec(100.))

# calculate wealth of the wealthiest one percent based on the SCF data
scf_estimate_of_wealth_of_wealthiest_one_percent_in = dict(zip(years_less_last,
    ['TBD']*len(years_of_scf)))
for year in years_of_scf:
    mean_wealth = in_2000_dollars( not_in_2009_dollars(
        dec(scf_mean_wealth_in_thousands_of_2009_dollars_in[year]) * dec(1000.), year=year),
        year=year) # in 2000 dollars
    number_of_households = dec(number_of_households_in_millions[year]) * dec(1000000.) # in
        households
    total_wealth = number_of_households * mean_wealth # in 2000 dollars
    scf_estimate_of_wealth_of_wealthiest_one_percent_in[year] =
        ((dec(perct_of_wealth_held_by_wealthiest_one_perct_in[year]) / dec(100.)) * total_wealth) /
        dec(1000000000.) # in billions of 2000 dollars

# load Kopczuk and Saez's (2004a,b) estimates of the wealth of the wealthiest one percent
filename = filepath + "ks_wealthiest_one_percent_data.txt"
header = ['year', 'total wealth', 'top 1% wealth share']
raw_ks_wealthiest_one_percent_data = numpy.loadtxt(filename, delimiter='\t', skiprows=4,
    dtype={'names':header, 'formats':['<S50', '<S50']*len(header)})
ks_wealthiest_one_percent_data = dict(zip(['total wealth', 'top 1% wealth share'],
    [dict(zip(years_btw_1982_and_2000, [raw_ks_wealthiest_one_percent_data[_][i] for i in range(0,
    2000-1982+1)])) for _ in ['total wealth', 'top 1% wealth share']]))
```

```
ks_estimate_of_wealth_of_wealthiest_one_percent_in = dict(zip(years_less_last,
      ['TBD']*len(years_btw_1982_and_2000)))
for year in years_btw_1982_and_2000:
  total_wealth = in_2000_dollars( not_in_2000_dollars(
      constant_dollar_amount=dec(ks_wealthiest_one_percent_data['total wealth'][year]),
      year=year), year=year)
  ks_estimate_of_wealth_of_wealthiest_one_percent_in[year] = total_wealth *
      (dec(ks_wealthiest_one_percent_data['top 1% wealth share'][year]) / dec(100.)) # billions of
      2000 dollars

# print estimates of the wealth of the wealthiest one percent
for year in years:
  if year == years[0]: print "\nestimates of the wealth of the wealthiest one percent \nyear
      \tSCF estimate \tK&S estimate"
  print year, "\t",
  if year in years_of_scf:
    print round(scf_estimate_of_wealth_of_wealthiest_one_percent_in[year]), "\t",
  else:
    print "\t",
  if year in years_btw_1982_and_2000:
    print round(ks_estimate_of_wealth_of_wealthiest_one_percent_in[year]), "\n",
  else:
    print "\n",

## SEC. 2.2 TO 2.3: DIRECT VS. ESTATE-MULTIPLIER ESTIMATES OF THE WEALTH
    OF THE 400 WEALTHIEST ##

# The following code uses Kopczuk and Saez's (2004a,b) estate-multiplier estimates for the
    smallest group they consider (the wealthiest 0.01%) and also uses their Pareto extrapolation
    to method in order to get estate-multiplier estimates of the wealth of the 400 wealthiest.
    Those estate-multiplier estimates are then compared to Forbes Magazine's direct estimates.

ks_estimated_wealth_of_400_wealthiest_in = dict(zip(years_btw_1982_and_2000, ['TBD' for year in
    years_btw_1982_and_2000]))
difference_in = dict(zip(years_less_last, ['TBD' for year in years_less_last]))
for year in years_btw_1982_and_2000:
  avg_wealth_at_top = dec(ks_wealthiest_0_01_percent_data['top 0.01% average wealth'][year]) #
      in their 2000 dollars
  number_in_top = (dec(0.01) / dec(100.)) *
      dec(ks_wealthiest_0_01_percent_data['population'][year])
  total_wealth_at_top = avg_wealth_at_top * number_in_top
  wealth_threshold_for_top = dec(ks_wealthiest_0_01_percent_data['top 0.01% wealth
      threshold'][year]) # in their 2000 dollars
  a = total_wealth_at_top / (total_wealth_at_top - wealth_threshold_for_top * number_in_top)
  ks_estimated_wealth_of_400_wealthiest_in_year_in_their_2000_dollars = total_wealth_at_top *
      ((dec(400.) / number_in_top) ** (dec(1) - (dec(1) / a)))
  ks_estimated_wealth_of_400_wealthiest_in[year] = in_2000_dollars( not_in_2000_dollars(
      constant_dollar_amount =
      ks_estimated_wealth_of_400_wealthiest_in_year_in_their_2000_dollars, year=year), year=year)
  direct_estimate = in_2000_dollars(dec(summary_stats['sum'][year]), year)
  difference_in[year] = ks_estimated_wealth_of_400_wealthiest_in[year] - direct_estimate
  # print results
```

```
if year == years_btw_1982_and_2000[0]: print "\ndifference btw direct and estate−multiplier
    estimates of the wealth of the 400 wealthiest Americans \nyear \testate−multiplier
    estimate \tdirect estimate \t difference in dollars \t difference as a percentage of the
    direct estimate"
print year, "\t",
print round(ks_estimated_wealth_of_400_wealthiest_in[year]), "\t",
print round(direct_estimate), "\t",
print round(difference_in[year]), "\t",
print round(dec(100.0) * difference_in[year] / direct_estimate, 2)

## SOME CALCULATIONS ##

print "\n400 wealthiest's wealth (millions of 2000 dollars) in 1990, according to K&S:\t",
print round(ks_estimated_wealth_of_400_wealthiest_in['1990'])
print "400 wealthiest's wealth in 2000, according to K&S:\t",
print round(ks_estimated_wealth_of_400_wealthiest_in['2000'])
print "400 wealthiest's share of wealth in 1990, according to K&S:\t",
print round(dec(100.) * ks_estimated_wealth_of_400_wealthiest_in['1990'] / (in_2000_dollars(
    not_in_2000_dollars( constant_dollar_amount=dec(21588.2728070478), year='1990'),
    year='1990') * dec(1000.)), 2)
print "400 wealthiest's share of wealth in 2000, according to K&S:\t",
aggregate_wealth_in_billions_according_to_ks = dec(32936.4508843353) # in billions of dollars
print round(dec(100.) * ks_estimated_wealth_of_400_wealthiest_in['2000'] /
    (aggregate_wealth_in_billions_according_to_ks * dec(1000)), 2)
print "difference in 2000 (millions of dollars)\t",
print difference_in['2000']
print "difference in 2000 as a percent of 400 wealthiest's wealth, according to K&S:\t",
print round(dec(100.) * difference_in['2000'] / ks_estimated_wealth_of_400_wealthiest_in['2000'])
print "difference in 2000 as a percent of 400 wealthiest's wealth, according to Forbes:\t",
print round(dec(100.) * difference_in['2000'] / in_2000_dollars(dec(summary_stats['sum']['2000']),
    '2000'))
print "difference in 2000 as a percent of top 0.01%'s wealth, according to K&S:\t",
top_0_01_perct_wealth = (dec(3.89579841645598) / dec(100.)) * (in_2000_dollars(
    not_in_2000_dollars( constant_dollar_amount=aggregate_wealth_in_billions_according_to_ks,
    year='2000'), year='2000') * dec(1000)) # in millions
print round(dec(100.) * difference_in['2000'] / top_0_01_perct_wealth)
print "difference in 2000 as a percent of top 2%'s wealth, according to K&S:\t",
top_2_perct_wealth = (dec(26.4262930798121) / dec(100.)) * (in_2000_dollars(
    not_in_2000_dollars( constant_dollar_amount=aggregate_wealth_in_billions_according_to_ks,
    year='2000'), year='2000') * dec(1000)) # in millions
print round(dec(100.) * difference_in['2000'] / top_2_perct_wealth)

## SEC. 2.4.4: ESTIMATE MORTALITY RATES AMONG THE FORBES 400 ##

# The probability that someone on the Forbes 400 list in a given year will die by the next year
    can be estimated by assuming that members of the Forbes 400 live and die according to the
    Gompertz−Makeham law of mortality, as discussed in the dissertation. The parameters
    associated with that law must be estimated by using numerical methods. We estimated
    them using the econometric program gretl 1.9.14. The following Python code exports the
    data that we used, can print the gretl code that we used, and imports estimates made using
    gretl and our gretl code.
```

```
# open file for all years
filename = filepath + “exported_gompertz_data_for_all_years.txt”
file_for_all_years = open(filename, ‘w’)
# write header
file_for_all_years.write(“t,r,d,y\n”)
for year in years_less_last:
  next_year = str(int(year) + 1)
  # open file for given year
  filename = filepath + “exported_gompertz_data_for_” + year + “.txt”
  file_for_one_year = open(filename, ‘w’)
  # write header
  file_for_one_year.write(“t,r,d\n”)
  # write data
  for name in names:
    if (data[name][‘wealth’][year] != ‘’):
      age = data[name][‘age’][year]
      s = age + ‘,’
      wealth = float(data[name][‘wealth’][year])
      rank = str(summary_stats[‘wealths’][year].index(wealth) + 1)
      s += rank + ‘,’
      if (name in dead_in[next_year]):
        s += str(1)
      else:
        s += str(0)
      file_for_all_years.write(s + ‘,’ + str(years_less_last.index(year)) + ‘\n’)
      file_for_one_year.write(s + ‘\n’)
  # close file
  file_for_one_year.close()
file_for_all_years.close()

# de−comment to print gretl code
‘‘‘
for year in years_less_last:
  print “\nset seed 250624”
  print ‘open “ABSOLUTE_PATH/data/data_specific_to_first
      essay/exported_gompertz_data_for_’ + year + ‘.txt”’ # note that the open command is
      not available in loop mode in gretl
  print “scalar a = 0.01”
  print “scalar b = 0.10”
  print “scalar c = 0”
  print “mle loglik = check ? (1 − d) ∗ ((a/b) ∗ (exp(b∗t) − exp(b∗(t+1))) − c) + d ∗ log(1 −
      exp(((a/b) ∗ (exp(b∗t) − exp(b∗(t+1))) − c))) : NA”
  print “\tscalar check = (a>=0) && (b>=0) && (c>=0)”
  print “\tparams a b”
  print “end mle −−quiet”
  print “scalar a = a”
  print “scalar b = b”
  print “scalar c = a”
  print “mle loglik = check ? (1 − d) ∗ ((a/b) ∗ (exp(b∗t) − exp(b∗(t+1))) − c) + d ∗ log(1 −
      exp(((a/b) ∗ (exp(b∗t) − exp(b∗(t+1))) − c))) : NA”
  print “\tscalar check = (a>=0) && (b>=0) && (c>=0)”
  print “\tparams a b c” # note that you cannot export results to a file in gretl
```

```
  print "end mle --quiet --lbfgs"
  print "print a b c"
  print "restrict"
  print "\tb[3] = 0"
  print "end restrict"
'''

# import estimates that were made using gretl and the gretl code given above
filename = filepath + "gompertz_estimates.txt"
header = ['year', 'a', 'b', 'c']
raw_data = numpy.loadtxt(filename, delimiter='\t', skiprows=3, dtype={'names':header,
    'formats':['f']*len(header)})
a_parameter_in = dict(zip(years_less_last, raw_data['a']))
b_parameter_in = dict(zip(years_less_last, raw_data['b']))
c_parameter_in = dict(zip(years_less_last, raw_data['c']))

# define a function that returns the mortality rates
def estimated_mortality_rate_in(year, age):
  ''' Returns the probability that someone on the Forbes 400 list in one year would die by the
      next. The estimates were made by assuming that members of the Forbes 400 live and die
      according to the Gompertz-Makeham law of mortality '''
  return 1. - numpy.exp((a_parameter_in[year] / b_parameter_in[year]) *
      (numpy.exp(b_parameter_in[year] * float(age)) - numpy.exp(b_parameter_in[year] *
      (float(age) + 1.))) - c_parameter_in[year])

## SEC. 2.5: APPLY THE ESTATE-MULTIPLIER METHOD TO THE FORBES 400 LIST
    ##

class eme_of_wealth_of_400_or_other_wealthiest_in(object):
  '''
 Class for applying the estate-multiplier method to the Forbes 400 list in the manner discussed in
    the initial exercise in the essay

 Parameters
 ==========
 year : string of the year of one of the Forbes 400 list
 number_of_wealthiest : int of the number of people whose wealth will be estimated
 use_estimated_mortality_rates : bool for whether to use our estimated mortality rates or Kopczuk
    and Saez's (2004a,b)
 cut_wealths_in_half : bool for whether to undervalue a person's wealth by half
  '''

  def __init__(self, year, number_of_wealthiest=400, use_estimated_mortality_rates=False,
      cut_wealths_in_half=False):
    self.year = year
    self.number_of_wealthiest = dec(number_of_wealthiest)
    self.use_estimated_mortality_rates = use_estimated_mortality_rates
    self.cut_wealths_in_half = cut_wealths_in_half
    self.names_of_dead = dead_in[str(int(self.year) + 1)] # names of people on the Forbes 400
        who died by next year
    self.estimated_wealth_of_number_of_wealthiest = 'TBD'
    self.bias = 'TBD'
```

```
def calculate_raw_estimates(self):
  ''' Apply the estate-multiplier method by assuming that we observe the wealths, ages, and
      genders of people who were on the Forbes 400, but died by the next year. Note that, in
      general, the estate-multiplier estimate of the number of people is not going to be equal to
      the number of people whose wealth we might wish to estimate, so we must interpolate or
      extrapolate the raw estate-multiplier estimates. Returns None '''
  self.W = dec(0) # estimate of total wealth
  self.N = dec(0) # estimate of number of people
  self.wealths_and_implied_numbers = list() # store info
  self.var_of_raw_total = dec(0) # store variance
  for name in self.names_of_dead:
    age, wealth = data[name]['age'][self.year],
        in_2000_dollars(dec(data[name]['wealth'][self.year]), self.year) # wealth is in millions of
        2000 dollars
    if self.use_estimated_mortality_rates:
      mortality_rate = dec(estimated_mortality_rate_in(year=self.year, age=age))
    else:
      if name in names_of_dead_who_were_women:
        gender = 'female'
      else:
        gender = 'male'
      mortality_rate = dec(mortality_rate_of[gender][self.year][age]) *
          dec(relative_mortality_rate_of[gender][age]) # adjust mortality rates for social
          differential
    if self.cut_wealths_in_half:
      wealth *= dec(0.5) # cut wealth in half
    self.W += wealth / mortality_rate
    self.N += dec(1.) / mortality_rate
    self.wealths_and_implied_numbers.append((wealth, dec(1) / mortality_rate)) # note: list of
        tuples of wealths and number of people (which is typically not a whole number) with
        those wealths
    self.var_of_raw_total += ((dec(1) - mortality_rate) / (mortality_rate ** dec(2.))) * (wealth
        ** dec(2.))

def get_interpolated_wealth_of_the_number_of_wealthiest(self):
  ''' Returns the interpolated wealth of the 400 or other wealthiest. The interpolation method
      follows Kopczuk and Saez (2004b, appendix D.2) '''
  # if the raw estate-multiplier estimate of the total number of people is already
      number_of_wealthiest, return the raw estate-multiplier estimate of their wealth
  if self.N == self.number_of_wealthiest:
    estimated_wealth_of_number_of_wealthiest = self.W
  else:
    # make sure that the 400 or other wealthiest can be interpolated
    assert self.N > self.number_of_wealthiest, "Wealth must be extrapolated"
    # if the raw estate-multipier estimate is different, then go through a list of sorted wealths
        and the number of people with those wealths until we get to number_of_wealthiest
    estimated_wealth_of_number_of_wealthiest = dec(0)
    people_so_far = dec(0)
    self.wealths_and_implied_numbers.sort(reverse=True) # sorted in descending order
    for wealth, number in self.wealths_and_implied_numbers:
      if self.number_of_wealthiest < (people_so_far + number):
```

```
            estimated_wealth_of_number_of_wealthiest += (self.number_of_wealthiest −
                people_so_far) ∗ wealth
            people_so_far += (self.number_of_wealthiest − people_so_far)
        else:
            estimated_wealth_of_number_of_wealthiest += number ∗ wealth
            people_so_far += number
    return estimated_wealth_of_number_of_wealthiest

def get_extrapolated_wealth_of_the_number_of_wealthiest(self):
    ''' Returns the extrapolated wealth of the 400 or other wealthiest. Assumes that wealth
        follows a Pareto distribution, following Kopczuk and Saez (2004b, p. 7, note 10) '''
    # assert that the sample can be extrapolated
    assert ((self.w ∗ self.N == self.W) != True), "Can't apply the Pareto extrapolation method"
    # apply the Pareto extrapolation method
    w, W, N = self.w, self.W, self.N
    a = W / (W − N ∗ w)
    estimated_min_wealth_of_number_of_wealthiest = W / (self.number_of_wealthiest ∗ (a / (a −
        dec(1.))) ∗ (N / self.number_of_wealthiest) ∗∗ (dec(1.) − (dec(1.) / a)))
    estimated_wealth_of_number_of_wealthiest = self.number_of_wealthiest ∗ (a / (a − dec(1.))) ∗
        estimated_min_wealth_of_number_of_wealthiest
    return estimated_wealth_of_number_of_wealthiest

def get_estimate(self):
    ''' Returns the estate-multiplier estimate of the wealth of the 400 or other wealthiest '''
    # calculate raw estimates
    self.calculate_raw_estimates()
    # assert that at least somebody died
    assert ((not self.wealths_and_implied_numbers) != True), "Nobody died"
    # if implied population is less than 400, then extrapolate, if extrapolation is possible
    if self.N < self.number_of_wealthiest:
        # get minimum wealth
        self.w = min(self.wealths_and_implied_numbers)[0]
        # extrapolate the sample
        self.estimated_wealth_of_number_of_wealthiest =
            self.get_extrapolated_wealth_of_the_number_of_wealthiest()
    # if implied population is greater than or equal to 400, then interpolate
    else:
        self.estimated_wealth_of_number_of_wealthiest =
            self.get_interpolated_wealth_of_the_number_of_wealthiest()
    return self.estimated_wealth_of_number_of_wealthiest # in millions of current dollars

def get_bias(self):
    ''' Returns the bias of the estate-multiplier estimator '''
    if self.estimated_wealth_of_number_of_wealthiest == 'TBD':
        self.get_estimate()
    self.bias = self.estimated_wealth_of_number_of_wealthiest − in_2000_dollars( dec( sum(
        summary_stats['wealths'][self.year][0:int(self.number_of_wealthiest)])), self.year)
    return self.bias # in millions of current dollars

def get_comparison_of_bias_w_combined_wealth_of_the_wealthiest(self):
    ''' Returns an interperation of any bias in the estate-multiplier method's estimate '''
    # get bias if not already calculated
```

```
    if self.bias == 'TBD':
      self.get_bias()
    # get whether bias is negative or positive
    if self.bias < 0:
      over_or_under_estimates = "under-estimates"
    elif self.bias > 0:
      over_or_under_estimates = "over-estimates"
    else:
      return "The estate-multiplier method as it was applied estimates the total wealth of the "
          + str(number_of_wealthiest) + " exactly"
    # compare the absolute value of the bias to the combined wealth of some of the wealthiest
    i = 1
    wealth_of_some_of_the_wealthiest = dec(0)
    while (wealth_of_some_of_the_wealthiest <= abs(self.bias)) and (i <
        int(self.number_of_wealthiest)):
      wealth_of_some_of_the_wealthiest +=
          in_2000_dollars(dec(summary_stats['wealths'][self.year][int(i - 1)]), self.year)
      i += 1
    if i == int(self.number_of_wealthiest):
      more_than_or_about = "more than"
    else:
      more_than_or_about = "about"
      wealth_of_some_of_the_wealthiest -=
          in_2000_dollars(dec(summary_stats['wealths'][self.year][int(i - 1)]), self.year)
      i -= 1
    # return result of comparison
    return "The estate-multiplier method as it was applied " + over_or_under_estimates + " the
        total wealth of the " + str(number_of_wealthiest) + " by about " +
        str(format(round(abs(self.bias) / dec(1000000), 2), '.2f')) + " trillion constant (2000)
        dollars, which was " + more_than_or_about + " the total wealth of the " + str(i) + "
        wealthiest Americans on the magazine's list"

# use the mortality rates assumed by Kopczuk and Saez (2004a,b)
eme_of_wealth_of_400_wealthiest_in_w_ks_mortality_rates = dict(zip(years_less_last,
    ['TBD']*len(years_less_last)))
for year in years_less_last:
  estimator = eme_of_wealth_of_400_or_other_wealthiest_in(year)
  eme_of_wealth_of_400_wealthiest_in_w_ks_mortality_rates[year] = estimator.get_estimate()
  # print results
  if year == years_less_last[0]: print "\nwealth of the Forbes 400 using K&S mortality
      rates\nyear \testimated \tactual \tdifference"
  estimated = eme_of_wealth_of_400_wealthiest_in_w_ks_mortality_rates[year]
  actual = in_2000_dollars(dec(summary_stats['sum'][year]), year)
  difference = estimated - actual
  print year, "\t", round(estimated), "\t", round(actual), "\t", round(difference)

# use the mortality rates assumed by Kopczuk and Saez (2004a,b) and also under-value
    wealths by half
eme_of_wealth_of_400_wealthiest_in_w_ks_mortality_rates_and_halved_wealths =
    dict(zip(years_less_last, ['TBD']*len(years_less_last)))
for year in years_less_last:
  estimator = eme_of_wealth_of_400_or_other_wealthiest_in(year, cut_wealths_in_half=True)
```

```
    eme_of_wealth_of_400_wealthiest_in_w_ks_mortality_rates_and_halved_wealths[year] =
        estimator.get_estimate()
    # print results
    if year == years_less_last[0]: print “\nwealth of the Forbes 400 using K&S mortality rates and
        under−valuing wealths by half\nyear \testimated \tactual\t difference”
    estimated = eme_of_wealth_of_400_wealthiest_in_w_ks_mortality_rates_and_halved_wealths[year]
    actual = in_2000_dollars(dec(summary_stats[‘sum’][year]), year)
    difference = estimated − actual
    print year, “\t”, round(estimated), “\t”, round(actual), “\t”, round(difference)

## SEC. 2.6.1: PROBABILITY OF UNDERESTIMATION ##

class sample_and_estimate_total_wealth_of_the_400_or_other_wealthiest_from(object):
  ‘‘‘
 Class for sampling from a given group of people and estimating the wealth of the group in the
     manner discussed in the extended exercise in the essay

 Parameters
 ==========
 wealth_age_pairs : pairs of wealths and ages to sample from
 year_for_mortality_rates : string of the year to use for mortality rates
 number_of_wealthiest : int of the number of people whose wealth will be estimated. Must be less
     than or equal to 400
 use_estimated_mortality_rates_to_sample : bool for whether to use our estimated mortality rates
     or Kopczuk and Saez’s (2004a,b) to sample people
 use_estimated_mortality_rates_to_inflate : bool for whether to use our estimated mortality rates or
     Kopczuk and Saez’s (2004a,b) to inflate wealths
 cut_wealths_in_half : bool for whether to undervalue a person’s wealth by half
  ’’’

  def __init__(self, wealth_age_pairs, year_for_mortality_rates, number_of_wealthiest=400,
        cut_wealths_in_half=False, use_estimated_mortality_rates_to_sample=True,
        use_estimated_mortality_rates_to_inflate=True):
    self.wealth_age_pairs = wealth_age_pairs
    self.year_for_mortality_rates = year_for_mortality_rates
    self.number_of_wealthiest = dec(number_of_wealthiest) # number of people whose wealth
        will be estimated
    self.cut_wealths_in_half = cut_wealths_in_half
    self.use_estimated_mortality_rates_to_sample = use_estimated_mortality_rates_to_sample
    self.use_estimated_mortality_rates_to_inflate = use_estimated_mortality_rates_to_inflate
    self.min_of_wealth_for_wealth_age_pairs = min(wealth for wealth, age in wealth_age_pairs)
    self.empty_sample = False

  def draw_sample(self):
    ‘‘‘ Draws a sample from the wealth and age pairs. Returns None ’’’
    self.W = dec(0)
    self.N = dec(0)
    self.wealths_and_implied_numbers = list()
    for pair in self.wealth_age_pairs:
      wealth, age = pair
      # get mortality rate for sampling
      if self.use_estimated_mortality_rates_to_sample:
```

```
                true_mortality_rate = dec(estimated_mortality_rate_in(year=self.year_for_mortality_rates, age=age))
            else:
                true_mortality_rate = dec(mortality_rate_of[‘male’][self.year_for_mortality_rates][age]) * dec(relative_mortality_rate_of[‘male’][age]) # use male mortality rates
            # sample wealth age pair?
            if numpy.random.uniform(0, 1) < float(true_mortality_rate):
                # misestimate mortality rates?
                if self.use_estimated_mortality_rates_to_inflate:
                    assumed_mortality_rate =
                        dec(estimated_mortality_rate_in(year=self.year_for_mortality_rates, age=age))
                else:
                    assumed_mortality_rate =
                        dec(mortality_rate_of[‘male’][self.year_for_mortality_rates][age]) *
                        dec(relative_mortality_rate_of[‘male’][age])
                # cut wealth in half?
                if self.cut_wealths_in_half:
                    wealth *= dec(0.50)
                self.W += dec(wealth) / assumed_mortality_rate
                self.N += dec(1) / assumed_mortality_rate
                self.wealths_and_implied_numbers.append((wealth, dec(1) / assumed_mortality_rate))
        # if sample is empty, then assume that number_of_wealthiest have the minimum wealth for the reasons discussed in the dissertation
        if (not self.wealths_and_implied_numbers):
            known_min_wealth = dec(self.min_of_wealth_for_wealth_age_pairs)
            if self.cut_wealths_in_half:
                known_min_wealth *= dec(0.50)
            self.W = known_min_wealth * self.number_of_wealthiest
            self.N = self.number_of_wealthiest
            self.wealths_and_implied_numbers = [(known_min_wealth, self.number_of_wealthiest)]
            self.empty_sample = True

    def get_estimate(self):
        ‘‘‘ Returns the estate-multiplier estimate of the wealth of the 400 or other wealthiest ’’’
        self.draw_sample()
        W, N, wealths_and_implied_numbers = self.W, self.N, self.wealths_and_implied_numbers
        # extrapolate
        if N < self.number_of_wealthiest:
            # if each of the N < number_of_wealthiest have the same wealth, extrapolate by assuming that number_of_wealthiest have the same wealth, the shape parameter of the Pareto distribution can be thought of as infinite
            w = min(wealths_and_implied_numbers)[0]
            if (w * N == W):
                estimated_wealth_of_number_of_wealthiest = w * self.number_of_wealthiest
            # otherwise, extrapolate by assuming a pareto distribution
            else:
                a = W / (W - N * w)
                estimated_min_wealth_of_number_of_wealthiest = W / ( self.number_of_wealthiest * (a / (a - dec(1))) * (N / self.number_of_wealthiest) ** (dec(1) - (dec(1) / a)) )
                estimated_wealth_of_number_of_wealthiest = self.number_of_wealthiest * (a / (a - dec(1))) * estimated_min_wealth_of_number_of_wealthiest
        # interpolate
```

```
    else:
      estimated_wealth_of_number_of_wealthiest = dec(0)
      people_so_far = dec(0)
      wealths_and_implied_numbers.sort(reverse=True)
      for wealth, number in wealths_and_implied_numbers:
        if self.number_of_wealthiest < (people_so_far + number):
          estimated_wealth_of_number_of_wealthiest += (self.number_of_wealthiest −
              people_so_far) ∗ wealth
          people_so_far += (self.number_of_wealthiest − people_so_far)
        else:
          estimated_wealth_of_number_of_wealthiest += number ∗ wealth
          people_so_far += number
    # return estimate
    return estimated_wealth_of_number_of_wealthiest

# differences that were observed between the direct and estate−multiplier estimates
difference_in_2000_as_a_fraction_of_total_wealth_of_forbes_400 =
    (dec(ks_estimated_wealth_of_400_wealthiest_in['2000']) −
    dec(in_2000_dollars(dec(summary_stats['sum']['2000']), '2000'))) /
    dec(in_2000_dollars(dec(summary_stats['sum']['2000']), '2000'))
difference_in_year_as_a_fraction_of_total_wealth_of_forbes_400 = dict()
for year in years_btw_1982_and_2000:
  difference_in_year_as_a_fraction_of_total_wealth_of_forbes_400[year] =
      (dec(ks_estimated_wealth_of_400_wealthiest_in[year]) −
      dec(in_2000_dollars(dec(summary_stats['sum'][year]), year))) /
      dec(in_2000_dollars(dec(summary_stats['sum'][year]), year))

# define dicts for storing results
set_ups = ['right wealths and right rates', 'wrong wealths but right rates', 'right wealths but
    wrong rates', 'wrong wealths and wrong rates']
median_bias_for = dict(zip(set_ups, [dict(zip(years_less_last, ['TBD' for year in years_less_last]))
    for set_up in set_ups]))
prob_under_estimate_by_2000_amount_for = dict(zip(set_ups, [dict(zip(years_less_last, ['TBD' for
    year in years_less_last])) for set_up in set_ups]))
prob_under_estimate_by_same_amount_for = dict(zip(set_ups, [dict(zip(years_btw_1982_and_2000,
    ['TBD' for year in years_btw_1982_and_2000])) for set_up in set_ups]))
cumulative_prob_1996_to_2000_for = dict(zip(set_ups, [dec(1.) for set_up in set_ups]))
# set the following to True to re−estimate if already estimated
re_estimate = False
# check if already been estimated
filenames = [filepath + _ + ".pkl" for _ in ["median_bias_for",
    "prob_under_estimate_by_2000_amount_for", "prob_under_estimate_by_same_amount_for",
    "cumulative_prob_1996_to_2000_for"]]
already_estimated = False not in [isfile(filename) for filename in filenames]
# estimate or re−estimate
if not already_estimated or re_estimate:
  # set up simulations
  for set_up in set_ups:
    # number of resamples
    number_of_resamples = 10000
    # cut wealths in half?
    if 'wrong wealths' in set_up:
```

```
  cut_wealths_in_half = True
else:
  cut_wealths_in_half = False
# misestimate mortality rates?
if ‘wrong rates’ in set_up:
  use_estimated_mortality_rates_to_sample = True
  use_estimated_mortality_rates_to_inflate = False
else:
  use_estimated_mortality_rates_to_sample = True
  use_estimated_mortality_rates_to_inflate = True
print set_up
# run simulations for each year
for year in years_less_last:
  numpy.random.seed(250624)
  biases_for_year = list()
  median_bias_for_year = list()
  prob_under_estimate_by_2000_amount_for_year = 0
  prob_under_estimate_by_same_amount_for_year = 0
  wealth_age_pairs = [(in_2000_dollars(dec(data[name][‘wealth’][year]), year),
      data[name][‘age’][year]) for name in names if (data[name][‘wealth’][year] != ‘’)]
  unknown_total_wealth_of_400_wealthiest = sum(zip(∗wealth_age_pairs)[0])
  year_for_mortality_rates = year
  # for each resample
  for _ in range(1, number_of_resamples+1):
    # draw a sample
    estimate = sample_and_estimate_total_wealth_of_the_400_or_other_wealthiest_from(
        wealth_age_pairs = wealth_age_pairs, year_for_mortality_rates =
        year_for_mortality_rates, cut_wealths_in_half = cut_wealths_in_half,
        use_estimated_mortality_rates_to_sample = use_estimated_mortality_rates_to_sample,
        use_estimated_mortality_rates_to_inflate =
        use_estimated_mortality_rates_to_inflate).get_estimate()
    # get bias
    bias = (estimate − unknown_total_wealth_of_400_wealthiest)
    biases_for_year.append(float(bias))
    # bias bigger than given amount?
    bias_as_a_fraction_of_total_wealth_of_forbes_400 = bias /
        (unknown_total_wealth_of_400_wealthiest) # bias as a percent
    if not bias_as_a_fraction_of_total_wealth_of_forbes_400 >
        difference_in_2000_as_a_fraction_of_total_wealth_of_forbes_400:
      prob_under_estimate_by_2000_amount_for_year += 1
    if year in years_btw_1982_and_2000:
      if not bias_as_a_fraction_of_total_wealth_of_forbes_400 >
          difference_in_year_as_a_fraction_of_total_wealth_of_forbes_400[year]:
        prob_under_estimate_by_same_amount_for_year += 1
  # record results
  median_bias_for[set_up][year] = str(round(scipy.stats.mstats.mquantiles( biases_for_year,
      prob=[50./100.], alphap=1/3., betap=1/3.)[0], 2))
  prob_under_estimate_by_2000_amount_for[set_up][year] = str(dec(100.) ∗
      dec(prob_under_estimate_by_2000_amount_for_year) / dec(number_of_resamples))
  if year in years_btw_1982_and_2000:
    prob_under_estimate_by_same_amount_for[set_up][year] = str(dec(100.) ∗
        dec(prob_under_estimate_by_same_amount_for_year) / dec(number_of_resamples))
```

```
            if int(1996) <= int(year) <= int(2000):
                cumulative_prob_1996_to_2000_for[set_up] *=
                        dec(prob_under_estimate_by_same_amount_for_year) / dec(number_of_resamples)
        print year
    # dump results
    filename = filepath + “median_bias_for.pkl”
    pickle.dump(median_bias_for, open(filename, ‘wb’))
    filename = filepath + “prob_under_estimate_by_2000_amount_for.pkl”
    pickle.dump(prob_under_estimate_by_2000_amount_for, open(filename, ‘wb’))
    filename = filepath + “prob_under_estimate_by_same_amount_for.pkl”
    pickle.dump(prob_under_estimate_by_same_amount_for, open(filename, ‘wb’))
    filename = filepath + “cumulative_prob_1996_to_2000_for.pkl”
    pickle.dump(cumulative_prob_1996_to_2000_for, open(filename, ‘wb’))
else:
    # load results that were dumped using pickle
    median_bias_for = pickle.load(open(filepath + “median_bias_for.pkl”, ‘rb’))
    prob_under_estimate_by_2000_amount_for = pickle.load(open(filepath +
          “prob_under_estimate_by_2000_amount_for.pkl”, ‘rb’))
    prob_under_estimate_by_same_amount_for = pickle.load(open(filepath +
          “prob_under_estimate_by_same_amount_for.pkl”, ‘rb’))
    cumulative_prob_1996_to_2000_for = pickle.load(open(filepath +
          “cumulative_prob_1996_to_2000_for.pkl”, ‘rb’))

# print results
for set_up in set_ups:
    for year in years_btw_1982_and_2000:
        if year == years_btw_1982_and_2000[0]: print “\nyear\t probability of underestimating by at
              least about 76 percent if ” + set_up + “ \tmedian misestmation”
        print year, “\t”,
        print round(float(prob_under_estimate_by_2000_amount_for[set_up][year]), 2), “\t”,
        print round(float(median_bias_for[set_up][year]))
set_up = ‘wrong wealths and wrong rates’
for year in years_btw_1982_and_2000:
    if year == years_btw_1982_and_2000[0]: print “\nyear\t observed difference\t probability of
          observing that difference if ” + set_up
    print year, “\t”,
    print round(float(dec(100.) *
          difference_in_year_as_a_fraction_of_total_wealth_of_forbes_400[year])), “\t”,
    print round(float(prob_under_estimate_by_same_amount_for[set_up][year]))
print “\nprobability of underestimating by observed differences in each year between 1996 and
      2000:\t”, cumulative_prob_1996_to_2000_for[set_up]

# Of note, medians were calculated by using one of the empirical quantile functions recommend
      by Hyndman and Yanan’s (1996) “Sample Quantiles in Statistical Packackages”
      (specifically, the eighth function they consider). According to the authors, it gives
      approximately median−unbiased estimates of a quartile regardless of the underlying
      distribution. Numpy’s median function yields similar results.

## SEC. 2.6.2, ESP. FIG. 2.9: EFFECT OF INEQUALITY ##

# The following code looks at the probability of underestimating the total wealth of the Forbes
      400 if the distribution of their wealth is changed and all other things are held equal.
```

```
def get_pareto_index_if_wealthiest_perct_owns_one_minus_perct(perct):
  ''' Returns the Pareto index that corresponds to a percentage of the population owning one
      minus their percentage of the population '''
  perct_of_pop = perct
  pareto_index = numpy.log(perct_of_pop) / numpy.log(perct_of_pop / (1. - perct_of_pop))
  return pareto_index

def wealths_w_pareto_index_of(pareto_index, total_wealth=dec(summary_stats['sum']['2000']),
    number_of_people=400):
  ''' Returns wealths that precisely follow a Pareto distribution '''
  wealths = list()
  for i in range(1, number_of_people+1):
    p = dec(i) / dec(number_of_people) # perct of pop corresponding to person
    p_1 = dec(i - 1) / dec(number_of_people) # perct of pop corresponding to next person
    wealth = total_wealth * (p ** (dec(1.) - (dec(1.) / dec(pareto_index))) - p_1 ** (dec(1.) -
        (dec(1.) / dec(pareto_index))))
    wealths.append(wealth)
  return wealths

# define dicts for storing results
percts_of_wealth = [_ for _ in (100. * numpy.array(range(201, 399+1, 2)) / 400.)]
biases_for_w_different_inequality = dict(zip(percts_of_wealth, [list() for perct_of_wealth in
    percts_of_wealth]))
lower_bias_for_w_different_inequality = dict(zip(percts_of_wealth, ['TBD' for perct_of_wealth in
    percts_of_wealth]))
median_bias_for_w_different_inequality = dict(zip(percts_of_wealth, ['TBD' for perct_of_wealth in
    percts_of_wealth]))
upper_bias_for_w_different_inequality = dict(zip(percts_of_wealth, ['TBD' for perct_of_wealth in
    percts_of_wealth]))
prob_under_estimate_by_fixed_amount_for_w_different_inequality = dict(zip(percts_of_wealth, [0
    for perct_of_wealth in percts_of_wealth]))
prob_over_estimate_by_fixed_amount_for_w_different_inequality = dict(zip(percts_of_wealth, [0 for
    perct_of_wealth in percts_of_wealth]))
# set the following to True to re-estimate if already estimated
re_estimate = False
# check if already been estimated
filenames = [filepath + _ + ".pkl" for _ in ["lower_bias_for_w_different_inequality",
    "median_bias_for_w_different_inequality", "upper_bias_for_w_different_inequality",
    "prob_under_estimate_by_fixed_amount_for_w_different_inequality",
    "prob_over_estimate_by_fixed_amount_for_w_different_inequality"]]
already_estimated = False not in [isfile(filename) for filename in filenames]
# estimate or re-estimate
if not already_estimated or re_estimate:
  # set seed
  numpy.random.seed(250624)
  # number of resamples
  number_of_resamples = 10000
  # cut wealths in half
  cut_wealths_in_half = True
  # misestimate mortality rate
  use_estimated_mortality_rates_to_sample = True
```

```
use_estimated_mortality_rates_to_inflate = False
# get ages, sorted by descending wealth with order of names breaking ties
ranks = range(0, 400)
ages_of = dict(zip(ranks, [list() for rank in ranks]))
for name in names:
  if (data[name]['wealth']['2000'] != ''):
    rank = summary_stats['wealths']['2000'].index(float(data[name]['wealth']['2000']))
    age = str(data[name]['age']['2000'])
    ages_of[rank].append(age)
ages_to_pair_w_wealths = list()
for rank in ranks:
  ages_to_pair_w_wealths.extend(ages_of[rank])
# get fixed amount
fixed_amount = abs(difference_in['2000'])
# for each perct of wealth
for perct_of_wealth in percts_of_wealth:
  # get wealths to pair with ages
  wealths_to_pair_w_ages = wealths_w_pareto_index_of( pareto_index =
      get_pareto_index_if_wealthiest_perct_owns_one_minus_perct( perct = (100. −
      perct_of_wealth) / 100.)) # different distributions of wealth
  unknown_total_wealth_of_400_wealthiest = sum(wealths_to_pair_w_ages)
  # for each resample
  for _ in range(1, number_of_resamples+1):
    # pair wealths and ages
    wealth_age_pairs = zip(wealths_to_pair_w_ages, ages_to_pair_w_wealths)
    # draw a sample
    estimate = sample_and_estimate_total_wealth_of_the_400_or_other_wealthiest_from(
        wealth_age_pairs = wealth_age_pairs, year_for_mortality_rates='2000',
        cut_wealths_in_half = cut_wealths_in_half, use_estimated_mortality_rates_to_sample =
        use_estimated_mortality_rates_to_sample, use_estimated_mortality_rates_to_inflate =
        use_estimated_mortality_rates_to_inflate).get_estimate()
    # get bias
    biases_for_w_different_inequality[perct_of_wealth].append(float(estimate −
        unknown_total_wealth_of_400_wealthiest))
    if abs(estimate − unknown_total_wealth_of_400_wealthiest) > fixed_amount:
      if estimate − unknown_total_wealth_of_400_wealthiest < 0:
        prob_under_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth] += 1
      if estimate − unknown_total_wealth_of_400_wealthiest > 0:
        prob_over_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth] += 1
  # record results
  median_bias_for_w_different_inequality[perct_of_wealth] =
      str(round(scipy.stats.mstats.mquantiles(
      biases_for_w_different_inequality[perct_of_wealth], prob=[50./100.], alphap=1/3.,
      betap=1/3.)[0]))
  lower_bias_for_w_different_inequality[perct_of_wealth] =
      str(round(scipy.stats.mstats.mquantiles(
      biases_for_w_different_inequality[perct_of_wealth], prob=[2.5/100.], alphap=1/3.,
      betap=1/3.)[0], 1))
  upper_bias_for_w_different_inequality[perct_of_wealth] =
      str(round(scipy.stats.mstats.mquantiles(
      biases_for_w_different_inequality[perct_of_wealth], prob=[97.5/100.], alphap=1/3.,
      betap=1/3.)[0], 1))
```

```
        prob_under_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth] =
            str(dec(100.) *
            dec(prob_under_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth]) /
            dec(number_of_resamples))
        prob_over_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth] =
            str(dec(100.) *
            dec(prob_over_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth]) /
            dec(number_of_resamples))
        print perct_of_wealth
    # dump results
    filename = filepath + "lower_bias_for_w_different_inequality.pkl"
    pickle.dump(lower_bias_for_w_different_inequality, open(filename, 'wb'))
    filename = filepath + "median_bias_for_w_different_inequality.pkl"
    pickle.dump(median_bias_for_w_different_inequality, open(filename, 'wb'))
    filename = filepath + "upper_bias_for_w_different_inequality.pkl"
    pickle.dump(upper_bias_for_w_different_inequality, open(filename, 'wb'))
    filename = filepath + "prob_under_estimate_by_fixed_amount_for_w_different_inequality.pkl"
    pickle.dump(prob_under_estimate_by_fixed_amount_for_w_different_inequality, open(filename,
        'wb'))
    filename = filepath + "prob_over_estimate_by_fixed_amount_for_w_different_inequality.pkl"
    pickle.dump(prob_over_estimate_by_fixed_amount_for_w_different_inequality, open(filename,
        'wb'))
else:
    # load results that were dumped using pickle
    lower_bias_for_w_different_inequality = pickle.load(open(filepath +
        "lower_bias_for_w_different_inequality.pkl", 'rb'))
    median_bias_for_w_different_inequality = pickle.load(open(filepath +
        "median_bias_for_w_different_inequality.pkl", 'rb'))
    upper_bias_for_w_different_inequality = pickle.load(open(filepath +
        "upper_bias_for_w_different_inequality.pkl", 'rb'))
    prob_under_estimate_by_fixed_amount_for_w_different_inequality = pickle.load(open(filepath +
        "prob_under_estimate_by_fixed_amount_for_w_different_inequality.pkl", 'rb'))
    prob_over_estimate_by_fixed_amount_for_w_different_inequality = pickle.load(open(filepath +
        "prob_over_estimate_by_fixed_amount_for_w_different_inequality.pkl", 'rb'))

print "\nmisestimation with different levels of inequality \nperct of wealth \tlower bias
    \tmedian bias \tupper bias \tprob underestimate \tprob overestimate"
for perct_of_wealth in percts_of_wealth:
    print perct_of_wealth, "\t", lower_bias_for_w_different_inequality[perct_of_wealth], "\t",
        median_bias_for_w_different_inequality[perct_of_wealth], "\t",
        upper_bias_for_w_different_inequality[perct_of_wealth], "\t",
        prob_under_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth], "\t",
        prob_over_estimate_by_fixed_amount_for_w_different_inequality[perct_of_wealth]

## SEC. 2.6.2, ESP. FIG. 2.10: ESTIMATES WITH EXTREME INEQUALITY ##

# define a function for rounding wealth estimates
def get_rounded(dec_estimate_in_billions):
    ''' Takes a dollar amount in billions and rounds it the nearest billion, if the nearest one billion
        dollars, if it was greater than five billion, and the nearest 100 million, otherwise. Dollar
        amounts are decimals, not floats '''
    # less than five billion?
```

```
  if dec_estimate_in_billions < dec(5.0):
    round_to_nearest = dec(0.100) # hundred million
  else:
    round_to_nearest = dec(1.0) # billion
  return dec(round(float(dec_estimate_in_billions / round_to_nearest), 0)) * round_to_nearest

# define dict for storing results
prob_of_estimating_w_extreme_inequality = dict()
# set the following to True to re-estimate if already estimated
re_estimate = False
# check if already been estimated
filename = filepath + "prob_of_estimating_w_extreme_inequality.pkl"
already_estimated = isfile(filename)
# estimate or re-estimate
if not already_estimated or re_estimate:
  # set seed
  numpy.random.seed(250624)
  # number of resamples
  number_of_resamples = 10000
  # cut wealths in half
  cut_wealths_in_half = True
  # misestimate mortality rate
  use_estimated_mortality_rates_to_sample = True
  use_estimated_mortality_rates_to_inflate = False
  # get ages, sorted by descending wealth with order of names breaking ties
  ranks = range(0, 400)
  ages_of = dict(zip(ranks, [list() for rank in ranks]))
  for name in names:
    if (data[name]['wealth']['2000'] != ''):
      rank = summary_stats['wealths']['2000'].index(float(data[name]['wealth']['2000']))
      age = str(data[name]['age']['2000'])
      ages_of[rank].append(age)
  ages_to_pair_w_wealths = list()
  for rank in ranks:
    ages_to_pair_w_wealths.extend(ages_of[rank])
  # estimate probabilities
  biases = list()
  # for the most-extreme level of inequality
  for perct_of_wealth in [99.75]:
    # get wealths to pair with ages
    wealths_to_pair_w_ages = wealths_w_pareto_index_of( pareto_index =
        get_pareto_index_if_wealthiest_perct_owns_one_minus_perct( perct = (100. -
        perct_of_wealth) / 100.))
    unknown_total_wealth_of_400_wealthiest = sum(wealths_to_pair_w_ages)
    # for each resample
    for resample in range(1, number_of_resamples+1):
      # pair wealths and ages
      wealth_age_pairs = zip(wealths_to_pair_w_ages, ages_to_pair_w_wealths)
      # draw a sample
      estimate = sample_and_estimate_total_wealth_of_the_400_or_other_wealthiest_from(
          wealth_age_pairs = wealth_age_pairs, year_for_mortality_rates='2000',
          cut_wealths_in_half = cut_wealths_in_half, use_estimated_mortality_rates_to_sample =
```

```
            use_estimated_mortality_rates_to_sample, use_estimated_mortality_rates_to_inflate =
            use_estimated_mortality_rates_to_inflate).get_estimate()
        estimate = str(get_rounded(estimate / dec(1000.))) # string of rounded estimate in billions
        if estimate in prob_of_estimating_w_extreme_inequality.keys():
            prob_of_estimating_w_extreme_inequality[estimate] += 1
        else:
            prob_of_estimating_w_extreme_inequality[estimate] = 1
        print resample
    for estimate in prob_of_estimating_w_extreme_inequality.keys():
        prob_of_estimating_w_extreme_inequality[estimate] = round(100 *
            (prob_of_estimating_w_extreme_inequality[estimate] / number_of_resamples), 2)
    # dump results
    filename = filepath + "prob_of_estimating_w_extreme_inequality.pkl"
    pickle.dump(prob_of_estimating_w_extreme_inequality, open(filename, 'wb'))
else:
    # load results that were dumped using pickle
    prob_of_estimating_w_extreme_inequality = pickle.load(open(filepath +
        "prob_of_estimating_w_extreme_inequality.pkl", 'rb'))

## SEC. 2.6.2, ESP. TABLE 2.1: EFFECT OF CHANGES RELATIVE TO BASE YEAR ##

# The following code looks at the effect of changes between base_year and 2000 on the
    probability of underestimating the total wealth of the Forbes 400.

base_year = '1990'
changes = ['No changes relative to base year', 'Ages', 'Sampling rates', 'Ages and sampling
    rates', 'Wealths', 'All changes']
# define dict for storing results
prob_under_estimate_by_2000_amount_for_w_changes = dict(zip(set_ups, [dict(zip(changes, [0 for
    change in changes])) for set_up in set_ups]))
# set the following to True to re−estimate if already estimated
re_estimate = False
# check if already been estimated
filename = filepath + "prob_under_estimate_by_2000_amount_for_w_changes.pkl"
already_estimated = isfile(filename)
# estimate or re−estimate
if not already_estimated or re_estimate:
    # set up simulations
    for set_up in set_ups:
        # number of resamples
        number_of_resamples = 10000
        # cut wealths in half?
        if 'wrong wealths' in set_up:
            cut_wealths_in_half = True
        else:
            cut_wealths_in_half = False
        # misestimate mortality rates?
        if 'wrong rates' in set_up:
            use_estimated_mortality_rates_to_sample = True
            use_estimated_mortality_rates_to_inflate = False
        else:
            use_estimated_mortality_rates_to_sample = True
```

```
  use_estimated_mortality_rates_to_inflate = True
print set_up
# for each change
for change in changes:
  # set seed
  numpy.random.seed(250624)
  # get wealths in descending order by year
  # also get ages by year, sorted by descending wealth with order of names breaking ties
  ranks = range(0, 400)
  wealths_in_base_year = dict(zip(ranks, [list() for rank in ranks]))
  wealths_in_2000 = dict(zip(ranks, [list() for rank in ranks]))
  ages_in_base_year = dict(zip(ranks, [list() for rank in ranks]))
  ages_in_2000 = dict(zip(ranks, [list() for rank in ranks]))
  for name in names:
    if (data[name]['wealth'][base_year] != ''):
      rank = summary_stats['wealths'][base_year].index(float(data[name]['wealth'][base_year]))
      wealth = in_2000_dollars(dec(data[name]['wealth'][base_year]), base_year)
      age = str(data[name]['age'][base_year])
      wealths_in_base_year[rank].append(wealth)
      ages_in_base_year[rank].append(age)
    if (data[name]['wealth']['2000'] != ''):
      rank = summary_stats['wealths']['2000'].index(float(data[name]['wealth']['2000']))
      wealth = in_2000_dollars(dec(data[name]['wealth']['2000']), '2000')
      age = str(data[name]['age']['2000'])
      wealths_in_2000[rank].append(wealth)
      ages_in_2000[rank].append(age)
  wealths_in_base_year_to_pair_w_ages = list()
  wealths_in_2000_to_pair_w_ages = list()
  ages_in_base_year_to_pair_w_wealths = list()
  ages_in_2000_to_pair_w_wealths = list()
  for rank in ranks:
    wealths_in_base_year_to_pair_w_ages.extend(wealths_in_base_year[rank])
    wealths_in_2000_to_pair_w_ages.extend(wealths_in_2000[rank])
    ages_in_base_year_to_pair_w_wealths.extend(ages_in_base_year[rank])
    ages_in_2000_to_pair_w_wealths.extend(ages_in_2000[rank])
  # changes wealth, ages, or year of mortality rates
  if change == 'No changes relative to base year':
    # base year wealths, ages, and mortality rates
    wealth_age_pairs = [(in_2000_dollars(dec(data[name]['wealth'][base_year]), base_year),
        data[name]['age'][base_year]) for name in names if (data[name]['wealth'][base_year]
        != '')] # same order, so that we get the same result with the same seed
    unknown_total_wealth_of_400_wealthiest = sum(zip(*wealth_age_pairs)[0])
    year_for_mortality_rates = base_year
  elif change == 'All changes':
    # 2000 wealths, ages, and mortality rates
    wealth_age_pairs = [(in_2000_dollars(dec(data[name]['wealth']['2000']), '2000'),
        data[name]['age']['2000']) for name in names if (data[name]['wealth']['2000'] != '')] #
        same order, so that we get the same result with the same seed
    unknown_total_wealth_of_400_wealthiest = sum(zip(*wealth_age_pairs)[0])
    year_for_mortality_rates = '2000'
  else:
    # change wealth?
```

```
          if change == 'Wealths':
             # 2000 wealths, but 1990 ages
             wealth_age_pairs = zip(wealths_in_2000_to_pair_w_ages,
                   ages_in_base_year_to_pair_w_wealths)
          # change ages?
          if 'Ages' in change:
             # 1990 wealths, but 2000 ages
             wealth_age_pairs = zip(wealths_in_base_year_to_pair_w_ages,
                   ages_in_2000_to_pair_w_wealths)
          # total wealth
          unknown_total_wealth_of_400_wealthiest = sum(zip(*wealth_age_pairs)[0])
          # change mortality rates?
          if 'sampling rates' in change.lower():
             year_for_mortality_rates = '2000'
          else:
             year_for_mortality_rates = base_year
       # for each resample
       for _ in range(1, number_of_resamples+1):
          # draw a sample
          estimate = sample_and_estimate_total_wealth_of_the_400_or_other_wealthiest_from(
                wealth_age_pairs = wealth_age_pairs, year_for_mortality_rates =
                year_for_mortality_rates, cut_wealths_in_half = cut_wealths_in_half,
                use_estimated_mortality_rates_to_sample = use_estimated_mortality_rates_to_sample,
                use_estimated_mortality_rates_to_inflate =
                use_estimated_mortality_rates_to_inflate).get_estimate()
          # get bias
          bias = (estimate - unknown_total_wealth_of_400_wealthiest) /
                (unknown_total_wealth_of_400_wealthiest) # bias as a percent
          if not bias > difference_in_2000_as_a_fraction_of_total_wealth_of_forbes_400:
             prob_under_estimate_by_2000_amount_for_w_changes[set_up][change] += 1
       prob_under_estimate_by_2000_amount_for_w_changes[set_up][change] = str(dec(100.) *
             dec(prob_under_estimate_by_2000_amount_for_w_changes[set_up][change]) /
             dec(number_of_resamples))
       print change
   # dump results
   filename = filepath + "prob_under_estimate_by_2000_amount_for_w_changes.pkl"
   pickle.dump(prob_under_estimate_by_2000_amount_for_w_changes, open(filename, 'wb'))
else:
   # load results that were dumped using pickle
   prob_under_estimate_by_2000_amount_for_w_changes = pickle.load(open(filepath +
         "prob_under_estimate_by_2000_amount_for_w_changes.pkl", 'rb'))

# print results
print "\neffect of changes between " + base_year + " and 2000"
for set_up in set_ups:
   print "\nset up: " + set_up + " \nchange \tprob of underestimating by 2000 amount given
         change"
   for change in changes:
      print change, "\t", prob_under_estimate_by_2000_amount_for_w_changes[set_up][change]

## SEC. A.5: EFFECT OF CHANGES BETWEEN 1918 AND 2000 ##
```

```
# get data on thirty wealthiest Americans in 1918
filename = filepath + "thirty_wealthiest_in_1918_data.txt"
header = ['Rank', 'Name', 'Wealth (in millions)', 'Born', 'Died', 'Gender', 'Age in 1918
    (assuming birthday already occurred)', 'Prob of dying in 1918']
raw_thirty_wealthiest_in_1918_data = numpy.loadtxt(filename, delimiter='\t', skiprows=5,
    dtype={'names':header, 'formats':['<S50']*len(header)})
wealths_in_1918, ages_in_1918 = list(), list()
for i in range(0, len(raw_thirty_wealthiest_in_1918_data['Rank'])):
  wealth = dec(raw_thirty_wealthiest_in_1918_data['Wealth (in millions)'][i])
  age = raw_thirty_wealthiest_in_1918_data['Age in 1918 (assuming birthday already occurred)'][i]
  gender = raw_thirty_wealthiest_in_1918_data['Gender'][i].lower() # continue to ignore gender
      for simplicity
  wealths_in_1918.append(wealth)
  ages_in_1918.append(age)
wealth_age_pairs_for_1918 = zip(wealths_in_1918, ages_in_1918)

# get data on thirty wealthiest Americans in 2000
wealths_in_2000, ages_in_2000 = list(), list()
for year in ['2000']:
  for name in names:
    if data[name]['wealth'][year] != '':
      wealth = data[name]['wealth'][year]
      rank = summary_stats['wealths'][year].index(float(wealth)) # note: no ties for 30th
          wealthiest in 2000
      if rank < 30:
        wealths_in_2000.append(dec(wealth))
        age = data[name]['age'][year]
        if name in ['Walton, Alice L', 'Walton, Helen R', 'Anthony, Barbara Cox', 'Chambers,
            Anne Cox', 'Johnson, Abigail']:
          gender = 'female'
        else:
          gender = 'male' # continue to ignore gender for simplicity
        ages_in_2000.append(age)
wealth_age_pairs_for_2000 = zip(wealths_in_2000, ages_in_2000)

base_year = '1918'
# define dicts for storing results
biases_for = dict(zip(changes, [list() for change in changes]))
median_bias_as_a_perct_of_wealth_for_w_changes = dict(zip(changes, ['TBD' for change in
    changes]))
# set the following to True to re-estimate if already estimated
re_estimate = False
# check if already been estimated
filename = filepath + "median_bias_as_a_perct_of_wealth_for_w_changes.pkl"
already_estimated = isfile(filename)
# estimate or re-estimate
if not already_estimated or re_estimate:
  # number of resamples
  number_of_resamples = 10000
  # fully value wealths
  cut_wealths_in_half = False
  # use K&S rates to sample and inflate
```

```
use_estimated_mortality_rates_to_sample = False
use_estimated_mortality_rates_to_inflate = False
# for each change
for change in changes:
  # set seed
  numpy.random.seed(250624)
  # changes wealth, ages, or year of sampling rates
  if change == 'No changes relative to base year':
    # base year wealths, ages, and sampling rates
    wealth_age_pairs = wealth_age_pairs_for_1918
    unknown_total_wealth_of_30_wealthiest = sum(zip(*wealth_age_pairs)[0])
    year_for_mortality_rates = base_year
    wrong_rates_in_year = False
  elif change == 'All changes':
    # 2000 wealths, ages, and sampling rates
    wealth_age_pairs = wealth_age_pairs_for_2000
    unknown_total_wealth_of_30_wealthiest = sum(zip(*wealth_age_pairs)[0])
    year_for_mortality_rates = '2000'
    wrong_rates_in_year = False
  else:
    # change wealth?
    if change == 'Wealths':
      # 2000 wealths, but 1918 ages
      wealth_age_pairs = zip(wealths_in_2000, ages_in_1918)
    # change ages?
    if 'Ages' in change:
      # 1918 wealths, but 2000 ages
      wealth_age_pairs = zip(wealths_in_1918, ages_in_2000)
    # total wealth
    unknown_total_wealth_of_30_wealthiest = sum(zip(*wealth_age_pairs)[0])
    # change sampling rates?
    if 'sampling rates' in change.lower():
      year_for_mortality_rates = '2000'
      wrong_rates_in_year = False
    else:
      year_for_mortality_rates = base_year
      wrong_rates_in_year = False
  # for each resample
  for _ in range(1, number_of_resamples+1):
    # draw a sample and get estimate
    estimate = sample_and_estimate_total_wealth_of_the_400_or_other_wealthiest_from(
        wealth_age_pairs = wealth_age_pairs, year_for_mortality_rates =
        year_for_mortality_rates, number_of_wealthiest=30, cut_wealths_in_half =
        cut_wealths_in_half, use_estimated_mortality_rates_to_sample =
        use_estimated_mortality_rates_to_sample, use_estimated_mortality_rates_to_inflate =
        use_estimated_mortality_rates_to_inflate).get_estimate()
    # get bias
    bias = (estimate - unknown_total_wealth_of_30_wealthiest) /
        (unknown_total_wealth_of_30_wealthiest) # bias as a percent
    biases_for[change].append(100. * float(bias))
  # record results
```

```
    median_bias_as_a_perct_of_wealth_for_w_changes[change] =
        str(round(scipy.stats.mstats.mquantiles( biases_for[change], prob=[50./100.],
        alphap=1/3., betap=1/3.)[0], 2))
    print change
  # dump results
  filename = filepath + "median_bias_as_a_perct_of_wealth_for_w_changes.pkl"
  pickle.dump(median_bias_as_a_perct_of_wealth_for_w_changes, open(filename, 'wb'))
else:
  # load results that were dumped using pickle
  median_bias_as_a_perct_of_wealth_for_w_changes = pickle.load(open(filepath +
      "median_bias_as_a_perct_of_wealth_for_w_changes.pkl", 'rb'))

# print results
print "\neffect of changes between 1918 and 2000 \nset up: right wealths and right rates
    \nchange \tmedian misestimation as a perct of wealth given change"
for change in changes:
  print change, "\t", median_bias_as_a_perct_of_wealth_for_w_changes[change]
```

E.4 Code for the Second Essay

The code for generating the results reported in the second essay is as follows.

```
'''
Filename: essay_on_distribution_of_wealth_code.py
Python version: 2.7
Source: Capehart, Kevin W. Essays on the Wealthiest Americans. PhD dissertation, American
    University, Washington, DC, 2014.
Description: This file generates results reported by the essay on the distribution of the wealth of
    the wealthiest Americans.
'''

from __future__ import division
import numpy
import scipy.stats # import for confidence intervals
import scipy.optimize # import for maximum likelihood estimation
import scipy.special # import for gamma distribution
from scipy.special import erf # import for error function erf(z) = 2 / sqrt(pi) * integral(exp(-x
    ** 2), x=0...z)
from os.path import isfile # import for checking whether a file exists
import pickle # import for dumping and loading computationally intensive results

## SEC. 3.2: REPLICATION OF KLASS ET AL. (2006)

# The following code estimates the shape parameters of Pareto distributions using Klass et al.'s
    (2006) data and methods

def get_pareto_index_using_klass_et_al_method(wealths, shift_term_for_ranks=0., ignore_ties=True,
    drop_wealthiest_and_poorest=True, regress_wealth_on_rank_instead_of_rank_on_wealth=False):
  '''
```

```
Returns OLS estimate of shape parameter of pareto distribution from a regression of log-ranks on
    log-wealths

Parameters
==========
shift_term_for_ranks : float for shift term to subtract from ranks as in the rank-1/2 OLS method
ignore_ties : bool for whether to ignore or account for ties in wealth when ranking wealths
drop_wealthiest_and_poorest : bool for whether to drop wealths that are strictly poorer than the
    400th and strictly wealthier than the 10th
regress_wealth_on_rank_instead_of_rank_on_wealth : bool for whether to run a log-log weak-rank
    rather than a log-log rank-wealth regression
 '''
 # sort wealths in descending order
 wealths.sort(reverse=True)
 # create ranks
 # ignore ties?
 if ignore_ties:
   ranks = list(rank - shift_term_for_ranks for rank in range(1, len(wealths)+1))
 else:
   ranks = [1]
   for i in range(1, len(wealths)):
     if wealths[i-1] == wealths[i]:
       ranks.append(ranks[-1])
     else:
       ranks.append(i+1)
 # drop wealthiest and poorest?
 if drop_wealthiest_and_poorest:
   # drop wealths that are strictly poorer than the 400th
   while ranks[-1] > 400:
     ranks.pop()
     wealths.pop()
   # drop wealths that are strictly wealthier than the 10th
   i = 0
   rank = ranks[i]
   while rank < 10:
     i += 1
     rank = ranks[i]
   index_of_10th_wealthiest = i
   wealths, ranks = wealths[index_of_10th_wealthiest:], ranks[index_of_10th_wealthiest:] # note:
       re-defining wealths and ranks
 # log wealths and ranks
 log_wealths = numpy.log(wealths)
 log_ranks = numpy.log(ranks)
 # get means of log wealth and ranks
 mean_of_log_wealths = numpy.mean(log_wealths)
 mean_of_log_ranks = numpy.mean(log_ranks)
 # estimate shape parameter
 # regress log wealths on log ranks or vice versa?
 if regress_wealth_on_rank_instead_of_rank_on_wealth:
   cov_btw_log_ranks_and_log_wealths = (1. / len(ranks)) * sum([(log_ranks[i] -
       mean_of_log_ranks) * (log_wealths[i] - mean_of_log_wealths) for i in range(0, len(ranks))])
```

```
    var_of_log_ranks = (1. / len(ranks)) * sum([(log_ranks[i] − mean_of_log_ranks) ** 2. for i in
        range(0, len(ranks))])
    shape_parameter = − 1. / (cov_btw_log_ranks_and_log_wealths / var_of_log_ranks)
  else:
    cov_btw_log_wealths_and_log_ranks = (1. / len(wealths)) * sum([(log_wealths[i] −
        mean_of_log_wealths) * (log_ranks[i] − mean_of_log_ranks) for i in range(0, len(wealths))])
    var_of_log_wealths = (1. / len(wealths)) * sum([(log_wealths[i] − mean_of_log_wealths) ** 2.
        for i in range(0, len(wealths))])
    shape_parameter = − 1. * (cov_btw_log_wealths_and_log_ranks / var_of_log_wealths)
  # return estimated shape parameter
  return shape_parameter

# load Klass et al.'s (2006) data
filepath = "./data/data_specific_to_second_essay/"
filename = filepath + "data_for_klass_et_al_2006.txt"
select_years = list(str(y) for y in range(1988, 2003+1))
header = ['Name']
header.extend(select_years)
rawdata = numpy.loadtxt(filename, delimiter=',', skiprows=5, dtype={'names':(header),
    'formats':(['<S50' for _ in range(0, len(header))])})

# define a dict for storing their data with its errors while preparing to correct errors
wealths_for_klass_et_al_data_in = dict(zip(select_years, [dict(zip(['with errors', 'without errors'],
    ['TBD' for _ in ['with errors', 'without errors']])) for year in select_years]))
# store raw data, removing zeros
for year in select_years:
  wealths_for_klass_et_al_data_in[year]['with errors'] = [float(rawdata[year][i]) for i in range(0,
      len(rawdata['Name']))]
  while 0. in wealths_for_klass_et_al_data_in[year]['with errors']:
    wealths_for_klass_et_al_data_in[year]['with errors'].remove(0.)

# estimate Pareto indexes using klass et al.'s (2006) data while also preparing to estimate them
    after correcting errors in their data
estimated_pareto_index_using_klass_et_al_methods_in = dict(zip(select_years, list(dict(zip(['with
    errors', 'without errors'], list(dict(zip(['point', 'se'], ['TBD', 'TBD'])) for _ in ['with errors',
    'without errors']))) for year in select_years)))
w_or_wo = 'with errors'
for year in select_years:
  wealths = wealths_for_klass_et_al_data_in[year][w_or_wo]
  point = get_pareto_index_using_klass_et_al_method(wealths, shift_term_for_ranks=0.,
      ignore_ties=True, drop_wealthiest_and_poorest=True,
      regress_wealth_on_rank_instead_of_rank_on_wealth=False)
  se = point * (2. / 391.) ** 0.5 # use the standard error suggested by Gabaix and Ibragimov
      (2011)
  estimated_pareto_index_using_klass_et_al_methods_in[year][w_or_wo]['point'] = point
  estimated_pareto_index_using_klass_et_al_methods_in[year][w_or_wo]['se'] = se
  # print results
  if year == select_years[0]: print "\npareto indexes estimated using Klass et al.'s (2006) data
      and methods \nyear \tpoint estimate"
  print year, "\t",
      round(estimated_pareto_index_using_klass_et_al_methods_in[year][w_or_wo]['point'], 2)
```

```
## 3.3.1: EFFECT OF ERRORS IN KLASS ET AL.’s (2006) DATA ##

# The following code shows the effect of errors in Klass et al.’s (2006) data on estimates of the
    Pareto indexes. Their estimation method is still used.

# correct errors related to the wealths of individuals/families who appear in Klass et al.’s (2006)
    data
# Richard Alexander Manoogian’s wealth was 625 million dollars in 1988 (24 October 1988, p.
    188, 340), not 885 million dollars, as in Klass et al.’s (2006) data.
rawdata[‘1988’][list(rawdata[‘Name’]).index(‘ManoogianRichardAlexander’)] = ‘625’
# Henry Ross Perot Sr.’s wealth was 2500 million dollars in 1989 (23 October 1989, p. 156, 352),
    not 500 million dollars, as in Klass et al.’s (2006) data.
rawdata[‘1989’][list(rawdata[‘Name’]).index(‘PerotHenryRoss’)] = ‘2500’
# Klass et al.’s (2006) data includes Roy Michael Huffington with a worth of 400 million dollars
    in the year 1994, but he dropped off the 1994 Forbes Four Hundred (17 October 1994, p.
    331, 310).
rawdata[‘1994’][list(rawdata[‘Name’]).index(‘HuffingtonRoyMichael’)] = ‘0’
# Leonard Samuel Skaggs Jr.’s wealth was 950 million million dollars in 1996 (14 October 1996,
    p. 204, 354)
rawdata[‘1996’][list(rawdata[‘Name’]).index(‘SkaggsLeonardSamuelJr’)] = ‘950’

# store corrected data, removing zeros
for year in select_years:
  wealths_for_klass_et_al_data_in[year][‘without errors’] = [float(rawdata[year][i]) for i in range(0,
      len(rawdata[‘Name’]))]
  while 0. in wealths_for_klass_et_al_data_in[year][‘without errors’]:
    wealths_for_klass_et_al_data_in[year][‘without errors’].remove(0.)

# correct errors related to individuals/families who do not appear in Klass et al.’s (2006) data
# Klass et al.’s (2006) data includes the family of Charles E. Smith with a worth of 290 million
    dollars in the year 1988, but the 1988 Forbes Four Hundred included Charles E. Smith and
    Robert H. Smith each with worths of 290 million dollars (24 October 1988, p. 258,
    344−−345).
wealths_for_klass_et_al_data_in[‘1988’][‘without errors’].append(290.)
# Klass et al.’s (2006) data does not include Shelby Cullom Davis for the year 1993, but he was
    on the 1993 Forbes Four Hundred with a worth of 800 million dollars (18 October 1993, p.
    180).
wealths_for_klass_et_al_data_in[‘1993’][‘without errors’].append(800.)
# Klass et al.’s (2006) data includes a Frank Batten with a worth 1.6 billion dollars in the year
    1999, but there was a Frank Batten Sr. who was worth twice that much (i.e., 2.1 billion
    dollars) and a Frank Batten Jr. who was worth half that much (i.e., 1.1 billion dollars) on
    the 1999 Forbes Four Hundred (11 October 1999, p. 242).
wealths_for_klass_et_al_data_in[‘1999’][‘without errors’].remove(1600.)
wealths_for_klass_et_al_data_in[‘1999’][‘without errors’].append(2100.)
wealths_for_klass_et_al_data_in[‘1999’][‘without errors’].append(1100.)

# estimate Pareto indexes without errors in the data
w_or_wo = ‘without errors’
for year in select_years:
  wealths = wealths_for_klass_et_al_data_in[year][w_or_wo]
  point = get_pareto_index_using_klass_et_al_method(wealths, shift_term_for_ranks=0.,
      ignore_ties=True, drop_wealthiest_and_poorest=True,
```

```
      regress_wealth_on_rank_instead_of_rank_on_wealth=False)
  se = point * (2. / 391.) ** 0.5 # use the standard error suggested by Gabaix and Ibragimov
      (2011)
  estimated_pareto_index_using_klass_et_al_methods_in[year][w_or_wo]['point'] = point
  estimated_pareto_index_using_klass_et_al_methods_in[year][w_or_wo]['se'] = se
  # print results
  if year == select_years[0]: print "\npareto indexes with and without errors \nyear \twith
      \twithout \tdiff"
  with_errors, without_errors = estimated_pareto_index_using_klass_et_al_methods_in[year]['with
      errors']['point'], estimated_pareto_index_using_klass_et_al_methods_in[year]['without
      errors']['point']
  print year, "\t", format(round(with_errors, 2), '.2f'), "\t",
  print format(round(without_errors, 2), '.2f'), "\t",
  print format(round(with_errors - without_errors, 2), '.2f')

## LOAD OUR FORBES 400 DATASET ##

from load_forbes_400_dataset import *

## SEC. 3.4.1: VARIATION OVER TIME IN PARETO INDEX ##

# The following code looks at variation over time in the maximum-likelihood estimate of the
    shape parameter of a Pareto distribution fit the wealths of the Forbes 400

class fit_pareto_to(object):
  ''' Class for getting the (unbiased) maximum-likelihood estimates of the parameters of a
      two-parameter Pareto distribution '''

  def __init__(self, wealths):
    self.wealths = wealths # observations
    self.n = len(wealths) # number of observations
    self.min_wealth = min(self.wealths) # minimum observation

  def estimate_biased_parameters(self):
    ''' Calculates biased MLEs of lower bound and shape parameters. Returns None '''
    self.biased_lower_bound = self.min_wealth
    self.biased_shape = self.n / sum(list(numpy.log(wealth / self.min_wealth) for wealth in
        self.wealths))

  def get_estimated_parameters(self):
    ''' Returns unbiased MLEs of lower bound and shape parameters '''
    self.estimate_biased_parameters()
    self.lower_bound = (1. - (1. / ((self.n - 1) * self.biased_shape))) * self.biased_lower_bound
    self.shape = ((self.n - 2.) / self.n) * self.biased_shape
    return self.lower_bound, self.shape

  def get_standard_errors_of_estimated_parameters(self):
    ''' Returns standard errors of estimated lower bound and shape parameters '''
    se_of_lower_bound = numpy.sqrt((self.lower_bound ** 2.) * (self.shape ** -1.) * ((self.n -
        1.) ** -1.) * ((self.shape * self.n - 2.) ** -1.))
    se_of_shape = numpy.sqrt((self.shape ** 2.) * ((self.n - 3.) ** -1.))
    return se_of_lower_bound, se_of_shape
```

```
# define dict for storing estimates
estimated_pareto_index_in = dict(zip(years, [dict(zip([‘point’, ‘se’], [‘TBD’ for _ in [‘point’, ‘se’]]))
    for year in years]))
# get estimates
for year in years:
  wealths = summary_stats[‘wealths’][year]
  fit = fit_pareto_to(wealths=wealths)
  _, estimated_pareto_index_in[year][‘point’] = fit.get_estimated_parameters()
  _, estimated_pareto_index_in[year][‘se’] = fit.get_standard_errors_of_estimated_parameters()
  # print results
  if year == years[0]: print “\nmles of pareto indexes \nyear \tpoint \tse”
  print year, “\t”, estimated_pareto_index_in[year][‘point’], “\t”,
      estimated_pareto_index_in[year][‘se’]

## SEC. 3.4.2: PARTS RELATED TO FIG. 3.1 ##

# The following code generates the empirical CCDF, fitted CCDF, and largest absolute
    difference between those two distributions shown in Fig. 3.1 of the essay

class k_s_like_test_against_unknown_pareto_distribution(object):
  ‘‘‘ Class for a Kolmogorov−Smirnov−like test against a two−parameter Pareto distribution.
      The test is exactly like a K−S−like test, except it is on the CCDF rather than the CDF,
      but that has no effect on the conclusions as long as the empirical CCDF and CDF are the
      complement of one another.

 Parameters
 ==========
 wealths : observed wealths
 a_priori_parameters_of_null_distribution : None for unknown parameters of Pareto distribution or
    tuple for specifying parameters a priori
  ’’’

  def __init__(self, wealths, a_priori_parameters_of_null_distribution=None):
    self.wealths = wealths
    self.wealths.sort() # sort wealths in ascending order
    self.number_of_observations = len(self.wealths)
    self.distinct_wealths = list(set(self.wealths))
    self.distinct_wealths.sort() # sort distinct wealths in ascending order
    self.number_of_distinct_observations = len(self.distinct_wealths)
    if a_priori_parameters_of_null_distribution == None:
      self.lower_bound_parameter, self.shape_parameter = self.get_estimated_parameters()
    else:
      self.lower_bound_parameter, self.shape_parameter = a_priori_parameters_of_null_distribution
    self.max_discrepancy = None # maximum discrepancy

  def get_estimated_parameters(self):
    fit = fit_pareto_to(wealths=self.wealths)
    return fit.get_estimated_parameters()

  def get_predicted_ccdf_at(self, given_wealth):
    if given_wealth < self.lower_bound_parameter:
```

```
        return 1.
    else:
        return (self.lower_bound_parameter / given_wealth) ** self.shape_parameter

def get_predicted_wealth_for(self, prob_wealth_greater_than_or_equal_to):
    return self.lower_bound_parameter * (prob_wealth_greater_than_or_equal_to ** (−
        self.shape_parameter))

def get_empirical_ccdf_at(self, given_wealth):
    ''' Returns proportion of wealths that are strictly greater than the given wealth '''
    return sum(list(int(wealth > given_wealth) for wealth in self.wealths)) /
        float(self.number_of_observations)

def get_discrepancy_statistic(self):
    ''' Returns the largest absolute difference between an empirical and fitted CCDF '''
    # get discrepancies
    discrepancies = list()
    # observations, except left most
    for i in range(1, self.number_of_distinct_observations):
        discrepancy = max(abs(self.get_empirical_ccdf_at(self.distinct_wealths[i−1]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[i])),
            abs(self.get_empirical_ccdf_at(self.distinct_wealths[i]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[i])))
        discrepancies.append(discrepancy)
    # left−most observation
    discrepancy = max(abs(1. − self.get_predicted_ccdf_at(self.distinct_wealths[0])),
        abs(self.get_empirical_ccdf_at(self.distinct_wealths[0]) −
        self.get_predicted_ccdf_at(self.distinct_wealths[0])))
    discrepancies.append(discrepancy)
    # get maximum discrepancy
    max_discrepancy = max(discrepancies)
    self.max_discrepancy = max_discrepancy
    return max_discrepancy

def get_wealths_and_probs_at_max_discrepancy(self):
    ''' Returns wealth(s) and prob(s) at the maximum discrepancy between an empirical and
        fitted CCDF '''
    if self.max_discrepancy == None:
        self.get_discrepancy_statistic()
    else:
        pass
    # get wealths and probs at max discrepancy
    wealths_at_max_discrepancy = list()
    probs_at_max_discrepancy = list()
    # left−most observation
    discrepancy = max(abs(1. − self.get_predicted_ccdf_at(self.distinct_wealths[0])),
        abs(self.get_empirical_ccdf_at(self.distinct_wealths[0]) −
        self.get_predicted_ccdf_at(self.distinct_wealths[0])))
    if discrepancy < self.max_discrepancy:
        pass
    else:
        wealths_at_max_discrepancy.append(self.distinct_wealths[0])
```

```
        # just before step?
        if abs(1. − self.get_predicted_ccdf_at(self.distinct_wealths[0])) >
                abs(self.get_empirical_ccdf_at(self.distinct_wealths[0]) −
                self.get_predicted_ccdf_at(self.distinct_wealths[0])):
            probs_at_max_discrepancy.append([1., self.get_predicted_ccdf_at(self.distinct_wealths[0])])
        # just after step?
        else:
            probs_at_max_discrepancy.append([self.get_empirical_ccdf_at(self.distinct_wealths[0]),
                    self.get_predicted_ccdf_at(self.distinct_wealths[0])])
    # others observations
    for i in range(1, self.number_of_distinct_observations):
        discrepancy = max(abs(self.get_empirical_ccdf_at(self.distinct_wealths[i−1]) −
                self.get_predicted_ccdf_at(self.distinct_wealths[i])),
                abs(self.get_empirical_ccdf_at(self.distinct_wealths[i]) −
                self.get_predicted_ccdf_at(self.distinct_wealths[i])))
        if discrepancy < self.max_discrepancy:
            pass
        else:
            wealths_at_max_discrepancy.append(self.distinct_wealths[i])
            # just before step?
            if abs(self.get_empirical_ccdf_at(self.distinct_wealths[i−1]) −
                    self.get_predicted_ccdf_at(self.distinct_wealths[i])) >
                    abs(self.get_empirical_ccdf_at(self.distinct_wealths[i]) −
                    self.get_predicted_ccdf_at(self.distinct_wealths[i])):
                probs_at_max_discrepancy.append([self.get_empirical_ccdf_at(self.distinct_wealths[i−1]),
                        self.get_predicted_ccdf_at(self.distinct_wealths[i])])
            # just after step?
            else:
                probs_at_max_discrepancy.append([self.get_empirical_ccdf_at(self.distinct_wealths[i]),
                        self.get_predicted_ccdf_at(self.distinct_wealths[i])])
    return wealths_at_max_discrepancy, probs_at_max_discrepancy

# get empirical CCDF for wealth of the 2003 Forbes 400
wealths_in_2003 = summary_stats['wealths']['2003']
k_s_like_test = k_s_like_test_against_unknown_pareto_distribution(wealths_in_2003) # use this
    classes empirical CCDF function
empirical_ccdf_at_for_2003 = [k_s_like_test.get_empirical_ccdf_at(wealth) for wealth in
    wealths_in_2003]
# get fitted CCDF
fitted_ccdf_for_2003 = [k_s_like_test.get_predicted_ccdf_at(wealth) for wealth in wealths_in_2003]
# get largest discrepancy and the wealth(s) and prob(s) associated with that discrepancy
discrepancy_statistic_in_2003 = k_s_like_test.get_discrepancy_statistic()
wealths_at_max_discrepancy_for_2003, probs_at_max_discrepancy_for_2003 =
    k_s_like_test.get_wealths_and_probs_at_max_discrepancy()
# largest discrepancy only occurs once, so only need first element of lists
assert len(wealths_at_max_discrepancy_for_2003) == 1, "Largest discrepancy occurs more than
    once"
wealth_at_max_discrepancy_for_2003 = wealths_at_max_discrepancy_for_2003[0]
probs_at_max_discrepancy_for_2003 = probs_at_max_discrepancy_for_2003[0]

## FUNCTION AND CLASS FOR CRITICAL VALUES OF K−S−LIKE TEST AGAINST
    UNBOUNDED PARETO ##
```

```
def random_pareto(lower_bound, shape_parameter, size):
  ''' Returns random draws from a two-parameter Pareto distribution '''
  uniforms = 1. - numpy.random.random(size=size) # uniform draws from (0.0, 1.0]
  paretos = lower_bound / (uniforms ** (1. / shape_parameter))
  return list(paretos)

class resample_d_stats_for_unknown_pareto_distribution(object):
  ''' Class for resampling the discrepancy statistic for a K-S-like test against an unknown
      Pareto distribution

 Parameters
 ==========
 wealths : observed wealths
 round_to_nearest_for : dict of instructions for way in which the wealths of billionaires and
    sub-billionaires were rounded
 number_of_resamples : int for number of resamples to draw
 fitting : True for estimating parameters or tuple for specifying parameters a priori
 rounding : bool for whether to round resamples using instructions from round_to_nearest_for
 diffs_to_draw_from : False for no perturbation of resamples or set of kernel density estimates to
    draw from
  '''

  def __init__(self, wealths=list(), round_to_nearest_for={}, number_of_resamples=1, fitting=True,
      rounding=True, diffs_to_draw_from=None):
    self.wealths = wealths
    self.number_of_observations = len(self.wealths)
    self.round_to_nearest_for = round_to_nearest_for
    k_s_like_test = k_s_like_test_against_unknown_pareto_distribution(self.wealths)
    self.lower_bound_parameter, self.shape_parameter = k_s_like_test.get_estimated_parameters()
    self.number_of_resamples = number_of_resamples
    if fitting:
      self.a_priori_parameters_of_null_distribution = None
    else:
      self.a_priori_parameters_of_null_distribution = (self.lower_bound_parameter,
          self.shape_parameter)
    self.rounding = rounding
    self.diffs_to_draw_from = diffs_to_draw_from
    if self.diffs_to_draw_from != None:
      self.kde = kernel_density_estimator(x=diffs_to_draw_from)
    self.resampled_discrepancy_statistics = list()

  def get_resample_discrepancy_statistics(self):
    ''' Returns discrepancy statistics from applying the same K-S-like test to sets of random
        draws from a bounded Pareto distribution '''
    for resample in range(1, self.number_of_resamples+1):
      resampled_wealths = self.get_resampled_wealths()
      k_s_like_test_on_resampled_wealths =
          k_s_like_test_against_unknown_pareto_distribution(resampled_wealths,
          a_priori_parameters_of_null_distribution=self.a_priori_parameters_of_null_distribution)
      discrepancy_statistic = k_s_like_test_on_resampled_wealths.get_discrepancy_statistic()
      self.resampled_discrepancy_statistics.append(discrepancy_statistic)
```

```
    self.resampled_discrepancy_statistics.sort(reverse=True) # sort in descending order
    return self.resampled_discrepancy_statistics

def get_p_value(self, discrepancy_statistic):
    ''' Returns estimated p-value associated with given a discrepancy statistic """'''
    # resample discrepancy statistic, if haven't done so already
    if len(self.resampled_discrepancy_statistics) == 0:
        self.get_resample_discrepancy_statistics()
    else:
        pass
    # find the proportion of resampled discrepancy statistics that are larger than the given
        discrepancy_statistic
    i = 0
    while (discrepancy_statistic < self.resampled_discrepancy_statistics[i]) and (i <
        self.number_of_resamples-1):
        i += 1
    return i / self.number_of_resamples

def get_perct_critical_value(self, perct=5):
    ''' Returns the critical value associated with a given p-value '''
    # resample discrepancy statistic, if haven't done so already
    if len(self.resampled_discrepancy_statistics) == 0:
        self.get_resample_discrepancy_statistics()
    else:
        pass
    # find the proportion of resampled discrepancy statistics that are larger than the given
        discrepancy_statistic
    return self.resampled_discrepancy_statistics[int(int(round((perct / 100.) *
        self.number_of_resamples, 0)) - 1)]

def get_resampled_wealths(self):
    ''' Returns a set of random draws from a Pareto distribution '''
    if self.diffs_to_draw_from == None:
        wealths = random_pareto(lower_bound=self.lower_bound_parameter,
            shape_parameter=self.shape_parameter, size=self.number_of_observations)
    else:
        uniform_randoms = 1. - numpy.random.uniform(low=0.0, high=1.0,
            size=self.number_of_observations * 10.) # order of magnitude more observations
        wealths = list((self.lower_bound_parameter / 10.) / (uniform_randoms ** (1. /
            self.shape_parameter))) # order of magnitude lower lower-bound parameter
        wealths = list(wealth * self.kde.get_random_sample_from_estimated_kernel_density() for
            wealth in wealths) # perturb
        wealths.sort(reverse=True)
        wealths = wealths[0:self.number_of_observations] # only use the number_of_observations
            largest samples
    # round wealths
    if self.rounding:
        wealths = list(self.get_rounded(wealth) for wealth in wealths)
    # return resampled and rounded wealth
    return wealths

def get_rounded(self, wealth):
```

```
        ''' Returns a wealth rounded according to the rounding instructions '''
        if wealth < 1000.:
            round_to_nearest = self.round_to_nearest_for['sub-billionaires']
        else:
            round_to_nearest = self.round_to_nearest_for['billionaires']
        wealth = round(wealth / round_to_nearest, 0) * round_to_nearest
        return wealth

## CLASSES FOR FITTING TO A BOUNDED PARETO AND TESTING THE GOODNESS
    OF THE FIT ##

class fit_bounded_pareto_to(object):
    ''' Class for getting the maximum-likelihood estimates of the three parameters of a truncated or
        tappered or bounded Pareto distribution '''

    def __init__(self, wealths):
        self.wealths = wealths # the data
        self.n = len(wealths) # number of observations
        self.min_wealth = min(self.wealths) # minimum observation
        self.max_wealth = max(self.wealths) # maximum observation
        self.min_over_max = (self.min_wealth / self.max_wealth)
        self.found_root = False

    def get_estimated_parameters(self):
        ''' Returns unbiased MLEs of parameters '''
        # get MLE of shape parameter
        tries = 0
        initial_guess_for_shape = self.get_initial_guess_for_shape_parameter()
        while (self.found_root == False) and (tries < 1000+1):
            guess_for_shape = numpy.random.uniform(initial_guess_for_shape / 2.,
                    initial_guess_for_shape * 2.)
            biased_shape = scipy.optimize.newton(func=self.func, x0=guess_for_shape, fprime=None)
            self.found_root = (not (self.func(biased_shape) > 1.48e-08))
            tries += 1
        # adjust for bias by following the suggestion of Maschberger and Kroupa 2009
        bias_adjusted_shape = ((self.n - 3.) / self.n) * biased_shape
        # get bias-adjusted MLEs of upper- and lower-bound parameters by following the
            suggestions of Zhang (2012)
        lower_bound = self.min_wealth * (1. + (((self.min_over_max ** biased_shape) - 1.) / (self.n
            * biased_shape)))
        upper_bound = self.max_wealth * (1. + (numpy.log(self.max_wealth / self.min_wealth) /
            self.n))
        return lower_bound, upper_bound, bias_adjusted_shape

    def get_initial_guess_for_shape_parameter(self):
        ''' Returns an initial guess for the shape parameter based on the MLE of the shape parameter
            for an unbounded Pareto '''
        return self.n / sum(list(numpy.log(wealth / self.min_wealth) for wealth in self.wealths))

    def func(self, shape):
        ''' Returns the shape parameter that solves func = 0 '''
```

```
        return (self.n / shape) + ((self.n * ((self.min_over_max ** shape) *
            numpy.log(self.min_over_max))) / (1. - (self.min_over_max ** shape))) -
            sum(list(numpy.log(wealth / self.min_wealth) for wealth in self.wealths))

    def get_found_root(self):
        ''' Returns a bool for whether a root was found or not '''
        return self.found_root

class k_s_like_test_against_unknown_bounded_pareto_distribution(object):
    ''' Class for a K-S-like test against an unknown bounded pareto distribution '''

    def __init__(self, wealths):
        self.wealths = wealths
        self.wealths.sort() # sort wealths in ascending order
        self.number_of_observations = len(self.wealths)
        self.distinct_wealths = list(set(self.wealths))
        self.distinct_wealths.sort() # sort distinct wealths in ascending order
        self.number_of_distinct_observations = len(self.distinct_wealths)
        self.lower_bound_parameter, self.upper_bound_parameter, self.shape_parameter =
            self.get_estimated_parameters()
        self.max_discrepancy = None # maximum discrepancy

    def get_estimated_parameters(self):
        fit = fit_bounded_pareto_to(wealths=self.wealths)
        estimated_parameters = fit.get_estimated_parameters()
        self.found_root = fit.get_found_root()
        return estimated_parameters

    def get_found_root(self):
        return self.found_root

    def get_predicted_ccdf_at(self, given_wealth):
        if given_wealth < self.lower_bound_parameter:
            return 1.
        elif given_wealth > self.upper_bound_parameter:
            return 0.
        else:
            return ((self.lower_bound_parameter ** self.shape_parameter) * ((given_wealth ** (-
                self.shape_parameter)) - (self.upper_bound_parameter ** (- self.shape_parameter))))
                / (1. - ((self.lower_bound_parameter / self.upper_bound_parameter) **
                self.shape_parameter))

    def get_empirical_ccdf_at(self, given_wealth):
        return sum(list(int(wealth > given_wealth) for wealth in self.wealths)) /
            float(self.number_of_observations)

    def get_discrepancy_statistic(self):
        # record discrepancies
        discrepancies = list()
        # observations, except left most
        for i in range(1, self.number_of_distinct_observations):
```

```
      discrepancy = max(abs(self.get_empirical_ccdf_at(self.distinct_wealths[i−1]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[i])),
            abs(self.get_empirical_ccdf_at(self.distinct_wealths[i]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[i])))
      discrepancies.append(discrepancy)
    # left−most observation
    discrepancy = max(abs(1. − self.get_predicted_ccdf_at(self.distinct_wealths[0])),
          abs(self.get_empirical_ccdf_at(self.distinct_wealths[0]) −
          self.get_predicted_ccdf_at(self.distinct_wealths[0])))
    discrepancies.append(discrepancy)
    # maximum discrepancy
    max_discrepancy = max(discrepancies)
    self.max_discrepancy = max_discrepancy
    return max_discrepancy

  def get_wealths_and_probs_at_max_discrepancy(self):
    if self.max_discrepancy == None:
      self.get_discrepancy_statistic()
    else:
      pass
    wealths_at_max_discrepancy = list()
    probs_at_max_discrepancy = list()
    # left−most observation
    discrepancy = max(abs(1. − self.get_predicted_ccdf_at(self.distinct_wealths[0])),
          abs(self.get_empirical_ccdf_at(self.distinct_wealths[0]) −
          self.get_predicted_ccdf_at(self.distinct_wealths[0])))
    if discrepancy < self.max_discrepancy:
      pass
    else:
      wealths_at_max_discrepancy.append(self.distinct_wealths[0])
      # just before step?
      if abs(1. − self.get_predicted_ccdf_at(self.distinct_wealths[0])) >
            abs(self.get_empirical_ccdf_at(self.distinct_wealths[0]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[0])):
        probs_at_max_discrepancy.append([1., self.get_predicted_ccdf_at(self.distinct_wealths[0])])
      # just after step?
      else:
        probs_at_max_discrepancy.append([self.get_empirical_ccdf_at(self.distinct_wealths[0]),
              self.get_predicted_ccdf_at(self.distinct_wealths[0])])
    # others observations
    for i in range(1, self.number_of_distinct_observations):
      discrepancy = max(abs(self.get_empirical_ccdf_at(self.distinct_wealths[i−1]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[i])),
            abs(self.get_empirical_ccdf_at(self.distinct_wealths[i]) −
            self.get_predicted_ccdf_at(self.distinct_wealths[i])))
      if discrepancy < self.max_discrepancy:
        pass
      else:
        wealths_at_max_discrepancy.append(self.distinct_wealths[i])
        # just before step?
        if abs(self.get_empirical_ccdf_at(self.distinct_wealths[i−1]) −
              self.get_predicted_ccdf_at(self.distinct_wealths[i])) >
```

```
                abs(self.get_empirical_ccdf_at(self.distinct_wealths[i]) −
                self.get_predicted_ccdf_at(self.distinct_wealths[i]))):
              probs_at_max_discrepancy.append([self.get_empirical_ccdf_at(self.distinct_wealths[i−1]),
                  self.get_predicted_ccdf_at(self.distinct_wealths[i])])
          # just after step?
          else:
            probs_at_max_discrepancy.append([self.get_empirical_ccdf_at(self.distinct_wealths[i]),
                  self.get_predicted_ccdf_at(self.distinct_wealths[i])])
      return wealths_at_max_discrepancy, probs_at_max_discrepancy

def random_bounded_pareto(lower_bound, shape_parameter, upper_bound, size):
  ''' Returns random draws from a bounded Pareto distribution '''
  uniform_randoms = 1. − numpy.random.uniform(low=0.0, high=1.0, size=size)
  bounded_paretos = list((− (uniform_randoms ∗ (upper_bound ∗∗ shape_parameter) −
      uniform_randoms ∗ (lower_bound ∗∗ shape_parameter) − (upper_bound ∗∗
      shape_parameter)) / ((lower_bound ∗∗ shape_parameter) ∗ (upper_bound ∗∗
      shape_parameter))) ∗∗ (− 1. / shape_parameter))
  return list(bounded_paretos)

class resample_d_stats_for_unknown_bounded_pareto_distribution(object):
  ''' Class for resampling the discrepancy statistic for a K-S-like test against an unknown bounded
      Pareto distribution '''

  def __init__(self, wealths=list(), round_to_nearest_for={}, number_of_resamples=10000,
      diffs_to_draw_from=None):
    self.wealths = wealths
    self.number_of_observations = len(self.wealths)
    self.round_to_nearest_for = round_to_nearest_for
    k_s_like_test = k_s_like_test_against_unknown_bounded_pareto_distribution(self.wealths)
    self.lower_bound_parameter, self.upper_bound_parameter, self.shape_parameter =
        k_s_like_test.get_estimated_parameters()
    self.number_of_resamples = number_of_resamples
    self.diffs_to_draw_from = diffs_to_draw_from
    if self.diffs_to_draw_from != None:
      self.kde = kernel_density_estimator(x=diffs_to_draw_from)
    self.resampled_discrepancy_statistics = list()

  def get_resample_discrepancy_statistics(self):
    for resample in range(1, self.number_of_resamples+1):
      #tries = 0
      found_root = False
      while found_root == False:
        resampled_wealths = self.get_resampled_wealths()
        k_s_like_test_on_resampled_wealths =
            k_s_like_test_against_unknown_bounded_pareto_distribution(resampled_wealths)
        found_root = k_s_like_test_on_resampled_wealths.get_found_root()
      discrepancy_statistic = k_s_like_test_on_resampled_wealths.get_discrepancy_statistic()
      self.resampled_discrepancy_statistics.append(discrepancy_statistic)
    self.resampled_discrepancy_statistics.sort(reverse=True) # sort in descending order
    return self.resampled_discrepancy_statistics

  def get_p_value(self, discrepancy_statistic):
```

```
    # resample discrepancy statistic, if haven't done so already
    if len(self.resampled_discrepancy_statistics) == 0:
      self.get_resample_discrepancy_statistics()
    else:
      pass
    # find the proportion of resampled discrepancy statistics that are larger than the given
        discrepancy_statistic
    i = 0
    while (discrepancy_statistic < self.resampled_discrepancy_statistics[i]) and (i <
        self.number_of_resamples-1):
      i += 1
    return i / self.number_of_resamples

  def get_resampled_wealths(self):
    if self.diffs_to_draw_from == None:
      wealths = random_bounded_pareto(lower_bound=self.lower_bound_parameter,
          shape_parameter=self.shape_parameter, upper_bound=self.upper_bound_parameter,
          size=self.number_of_observations)
    else:
      uniform_randoms = 1. - numpy.random.uniform(low=0.0, high=1.0,
          size=self.number_of_observations * 10.) # order of magnitude more observations
      wealths = list((- (uniform_randoms * (self.upper_bound_parameter **
          self.shape_parameter) - uniform_randoms * ((self.lower_bound_parameter / 10.) **
          self.shape_parameter) - (self.upper_bound_parameter ** self.shape_parameter)) /
          (((self.lower_bound_parameter / 10.) ** self.shape_parameter) *
          (self.upper_bound_parameter ** self.shape_parameter))) ** (- 1. /
          self.shape_parameter)) # order of magnitude lower lower-bound parameter
      wealths = list(wealth * self.kde.get_random_sample_from_estimated_kernel_density() for
          wealth in wealths) # mismeasure
      wealths.sort(reverse=True)
      wealths = wealths[0:self.number_of_observations] # only use the number_of_observations
          largest samples
    # round wealths, following the methodology of forbes
    wealths = list(self.get_rounded(wealth) for wealth in wealths)
    # return resampled and rounded wealth
    return wealths

  def get_rounded(self, wealth):
    ''' Returns a wealth rounded according to the rounding instructions '''
    if wealth < 1000.:
      round_to_nearest = self.round_to_nearest_for['sub-billionaires']
    else:
      round_to_nearest = self.round_to_nearest_for['billionaires']
    wealth = round(wealth / round_to_nearest, 0) * round_to_nearest
    return wealth

## K-S-LIKE TEST WITH ROUNDING ERRORS ##

# The following code runs the K-S-like tests

# define dict for storing results for K-S-like test with rounding errors against unbounded and
    bounded Pareto distributions
```

```
unbounded_and_bounded_pareto = ['pareto', 'bounded pareto']
k_s_like_test_against_w_rounding_errors = dict(zip(unbounded_and_bounded_pareto,
    [dict(zip(years, [dict(zip(['D stat', 'p-value', 'D stats'], ['TBD' for _ in ['D stat', 'p-value',
    'D stats']])) for year in years])) for distribution in unbounded_and_bounded_pareto]))
# set the following to True to re-estimate if already estimated
re_estimate = False
# check if already been estimated
filename = filepath + "k_s_like_test_against_w_rounding_errors.pkl"
already_estimated = isfile(filename)
# estimate or re-estimate
if not already_estimated or re_estimate:
  # set number of resamples
  number_of_resamples = 10000
  # for each year
  for year in years:
    # get wealths
    wealths = summary_stats['wealths'][year]
    # get rounding instructions
    round_to_nearest_for = dict()
    unique_wealths = list(set(wealths))
    unique_wealths.sort(reverse=True)
    # if wealthiest person isn't a billionaire
    if summary_stats['max'][year] < 1000.0:
      min_difference = min(list(unique_wealths[i] - unique_wealths[i+1] for i in range(0,
          len(unique_wealths)-1) if unique_wealths[i] < 1000.0))
      round_to_nearest_for['billionaires'] = min_difference # if a billion dollar wealth is drawn,
          use the same rounding
      round_to_nearest_for['sub-billionaires'] = min_difference
    # if poorest person is a billionaire
    elif summary_stats['min'][year] >= 1000.0:
      min_difference = min(list(unique_wealths[i] - unique_wealths[i+1] for i in range(0,
          len(unique_wealths)-1) if unique_wealths[i+1] >= 1000.0))
      round_to_nearest_for['billionaires'] = min_difference # if a sub-billion dollar wealth is
          drawn, use the same rounding
      round_to_nearest_for['sub-billionaires'] = min_difference
    else:
      round_to_nearest_for['billionaires'] = min(list(unique_wealths[i] - unique_wealths[i+1] for i
          in range(0, len(unique_wealths)-1) if unique_wealths[i+1] >= 1000.0))
      round_to_nearest_for['sub-billionaires'] = min(list(unique_wealths[i] - unique_wealths[i+1]
          for i in range(0, len(unique_wealths)-1) if unique_wealths[i] < 1000.0))
    ## k-s test with errors against unbounded pareto ##
    # get discrepancy statistic
    numpy.random.seed(250624)
    k_s_like_test = k_s_like_test_against_unknown_pareto_distribution(wealths)
    discrepancy_statistic = k_s_like_test.get_discrepancy_statistic()
    k_s_like_test_against_w_rounding_errors['pareto'][year]['D stat'] = discrepancy_statistic
    # resample discrepancy statistic
    numpy.random.seed(250624)
    resample_d_stats = resample_d_stats_for_unknown_pareto_distribution(wealths=wealths,
        round_to_nearest_for=round_to_nearest_for, number_of_resamples=number_of_resamples,
        fitting=True, rounding=True)
```

```
        k_s_like_test_against_w_rounding_errors['pareto'][year]['p-value'] =
            resample_d_stats.get_p_value(discrepancy_statistic)
        #k_s_like_test_against_w_rounding_errors['pareto'][year]['D stats'] =
            resample_d_stats.resampled_discrepancy_statistics # de-comment to store D-stats
        ## k-s test with errors against bounded pareto ##
        # get discrepancy statistic against bounded pareto distribution
        numpy.random.seed(250624)
        k_s_like_test = k_s_like_test_against_unknown_bounded_pareto_distribution(wealths=wealths)
        discrepancy_statistic = k_s_like_test.get_discrepancy_statistic()
        k_s_like_test_against_w_rounding_errors['bounded pareto'][year]['D stat'] =
            discrepancy_statistic
        # resample discrepancy statistic
        numpy.random.seed(250624)
        resample_d_stats =
            resample_d_stats_for_unknown_bounded_pareto_distribution(wealths=wealths,
            round_to_nearest_for=round_to_nearest_for, number_of_resamples=number_of_resamples)
            # always fitting and rounding
        k_s_like_test_against_w_rounding_errors['bounded pareto'][year]['p-value'] =
            resample_d_stats.get_p_value(discrepancy_statistic)
        #k_s_like_test_against_w_rounding_errors['bounded pareto'][year]['D stats'] =
            resample_d_stats.resampled_discrepancy_statistics # de-comment to store D-stats
        print year
    # dump results
    pickle.dump(k_s_like_test_against_w_rounding_errors, open(filename, 'wb'))
else:
    # load result
    k_s_like_test_against_w_rounding_errors = pickle.load(open(filename, 'rb'))

## BLOOMBERG VS. FORBES DIFFERENCES ##

class kernel_density_estimator(object):

    def __init__(self, x):
        self.x = x
        self.N = len(self.x)
        self.global_bandwidth = self.get_silverman_plug_in_estimate()

    def get_sample_inter_quartile_range(self):
        ''' Computes the sample interquartile range using one of the functions recommended by
            Hyndman and Fan (1996), specifically, the eighth fuction they consider. According to the
            authors, it gives approximately median-unbiased estimates of a quartile regardless of the
            underlying distribution '''
        upper_quartile = scipy.stats.mstats.mquantiles(self.x, prob=[75/100.], alphap=1/3.,
            betap=1/3.)[0]
        lower_quartile = scipy.stats.mstats.mquantiles(self.x, prob=[25/100.], alphap=1/3.,
            betap=1/3.)[0]
        return upper_quartile - lower_quartile

    def get_silverman_plug_in_estimate(self):
        ''' Return's Silverman's plug-in estimate of the optimal band width '''
        constant = (3 / (8 * numpy.sqrt(numpy.pi))) ** -0.2 # = about 1.3643
        delta = (1 / (2 * numpy.sqrt(numpy.pi))) ** 0.2 # = about 0.7764
```

```
    sample_standard_deviation = numpy.std(self.x, ddof=1) # standard deviation with (N-1)
        adjustment
    iqr = self.get_sample_inter_quartile_range()
    normalized_iqr = iqr / (scipy.stats.norm.ppf(0.75) - scipy.stats.norm.ppf(0.25)) # if the
        distribution is normal, then the standard deviation should be the inter-quartile range
        deflated by about 1.349
    return constant * delta * ( self.N ** (-0.2) ) * min(sample_standard_deviation,
        normalized_iqr)

  def guassian_kernel(self, z):
    ''' Returns Guassian kernel '''
    return ((2. * numpy.pi) ** (-0.5)) * numpy.exp(-0.5 * (z ** 2.))

  def get_kernel_density_estimates(self, xrange=None):
    ''' Returns a dictionary of the Guassian kernel density estimates at each data point.
        Bandwidth is equal to Silverman's plug-in estimate '''
    density_estimate_at = dict()
    if xrange == None:
      self.xrange = self.x
    else:
      self.xrange = xrange
    bandwidth_at = dict(zip(self.x, [self.global_bandwidth] * self.N))
    for x0 in self.xrange:
      density_estimate_at[x0] = (1. / self.N) * sum([(1. / bandwidth_at[xi]) *
          self.guassian_kernel((x0 - xi) / bandwidth_at[xi]) for xi in self.x])
    return density_estimate_at

  def get_random_sample_from_estimated_kernel_density(self):
    ''' Returns random draws from the kernel density estimate '''
    return numpy.random.normal(self.x[numpy.random.randint(low=int(0),
        high=int(self.N-1))], self.global_bandwidth)

# load forbes vs. bloomberg data
filename = filepath + "bloomberg_vs_forbes_data.txt"
header = ['name', 'forbes wealth', 'bloomberg wealth', 'year']
raw_data = numpy.loadtxt(filename, delimiter='\t', skiprows=8, dtype={'names':header,
    'formats':['<S50']*len(header)})

# define dicts for storing results
sources = ['forbes 2012', 'bloomberg 2012', 'forbes 2013', 'bloomberg 2013']
wealths_from = dict(zip(sources, [list() for source in sources]))
n_on = dict(zip(sources, [0 for source in sources]))
comparisions = ['forbes and bloomberg 2012', 'forbes and bloomberg 2013']
n_on_both = dict(zip(comparisions, [0 for comparison in comparisions])) # ignoring Rick Cohen
diffs_btw = dict(zip(comparisions, [list() for comparison in comparisions]))
for i in range(len(raw_data)):
  name = raw_data['name'][i]
  forbes_wealth = raw_data['forbes wealth'][i]
  bloomberg_wealth = raw_data['bloomberg wealth'][i]
  year = raw_data['year'][i]
  if (forbes_wealth != ''):
    n_on['forbes' + ' ' + year] += 1
```

```
      wealths_from['forbes' + ' ' + year].append(float(forbes_wealth))
      if (bloomberg_wealth != ''):
        n_on_both['forbes and bloomberg' + ' ' + year] += 1
        diff = (float(bloomberg_wealth) / float(forbes_wealth)) # bloomberg as perct of forbes
        diffs_btw['forbes and bloomberg' + ' ' + year].append(diff)
  if (bloomberg_wealth != ''):
    n_on['bloomberg' + ' ' + year] += 1
    wealths_from['bloomberg' + ' ' + year].append(float(bloomberg_wealth))
    # note: diff already recorded
for comparison in comparisions:
  diffs_btw[comparison].sort() # sort in ascending order

print "\nnumber on forbes and bloomberg lists, respectively"
for year in ['2012', '2013']:
  print year, "\t", str(n_on['forbes' + ' ' + year]), "\t", str(n_on['bloomberg' + ' ' + year])

print "\nnumber on both lists",
for year in ['2012', '2013']:
  print year, "\t", str(n_on_both['forbes and bloomberg' + ' ' + year])

print "\nrange of differences between the forbes and bloomberg estimates of someone's wealth
    (as a percentage of the Forbes estimate)"
for year in ['2012', '2013']:
  print "min/max in " + str(year) + ":\t", str(100. * diffs_btw['forbes and bloomberg' + ' ' +
      year][0]), "\t", str(round(100. * diffs_btw['forbes and bloomberg' + ' ' + year][-1], 0))

# kde for 2013 diffs
xrange_for_bloomberg_forbes_diffs = list(set(diffs_btw['forbes and bloomberg
    2012']).union(set(diffs_btw['forbes and bloomberg 2013']))) # use diffs from both years
xrange_for_bloomberg_forbes_diffs.append(0) # include zero to pick up any left tail
xrange_for_bloomberg_forbes_diffs.sort()
kde = kernel_density_estimator(x=diffs_btw['forbes and bloomberg 2013'])
kde_for_bloomberg_and_forbes_2013_diffs_at =
    kde.get_kernel_density_estimates(xrange=xrange_for_bloomberg_forbes_diffs)
# kde for 2012 diffs
kde = kernel_density_estimator(x=diffs_btw['forbes and bloomberg 2012'])
kde_for_bloomberg_and_forbes_2012_diffs_at =
    kde.get_kernel_density_estimates(xrange=xrange_for_bloomberg_forbes_diffs)

## K-S-LIKE TEST WITH ROUNDING AND OTHER MEASUREMENT ERRORS ##

# define dict for storing results for K-S-like test with rounding and other measurement errors
    against unbounded and bounded Pareto distributions
k_s_like_test_against_w_rounding_and_other_measurement_errors =
    dict(zip(unbounded_and_bounded_pareto, [dict(zip(years, [dict(zip(['D stat', 'p-value', 'D
    stats'], ['TBD' for _ in ['D stat', 'p-value', 'D stats']])) for year in years])) for distribution in
    unbounded_and_bounded_pareto]))
# set the following to True to re-estimate if already estimate
re_estimate = False
# check if already been estimated
filename = filepath + "k_s_like_test_against_w_rounding_and_other_measurement_errors.pkl"
already_estimated = isfile(filename)
```

```
# estimate or re−estimate
if not already_estimated or re_estimate:
  # set number of resamples
  number_of_resamples = 10000
  # for each year
  for year in years:
    # get wealths
    wealths = summary_stats['wealths'][year]
    # get rounding instructions
    round_to_nearest_for = dict()
    unique_wealths = list(set(wealths))
    unique_wealths.sort(reverse=True)
    # if wealthiest person isn't a billionaire
    if summary_stats['max'][year] < 1000.0:
      min_difference = min(list(unique_wealths[i] − unique_wealths[i+1] for i in range(0,
           len(unique_wealths)−1) if unique_wealths[i] < 1000.0))
      round_to_nearest_for['billionaires'] = min_difference # if a billion dollar wealth is drawn,
           use the same rounding
      round_to_nearest_for['sub−billionaires'] = min_difference
    # if poorest person is a billionaire
    elif summary_stats['min'][year] >= 1000.0:
      min_difference = min(list(unique_wealths[i] − unique_wealths[i+1] for i in range(0,
           len(unique_wealths)−1) if unique_wealths[i+1] >= 1000.0))
      round_to_nearest_for['billionaires'] = min_difference # if a sub−billion dollar wealth is
           drawn, use the same rounding
      round_to_nearest_for['sub−billionaires'] = min_difference
    else:
      round_to_nearest_for['billionaires'] = min(list(unique_wealths[i] − unique_wealths[i+1] for i
           in range(0, len(unique_wealths)−1) if unique_wealths[i+1] >= 1000.0))
      round_to_nearest_for['sub−billionaires'] = min(list(unique_wealths[i] − unique_wealths[i+1]
           for i in range(0, len(unique_wealths)−1) if unique_wealths[i] < 1000.0))
    ## k−s like test with rounding and other measurement errors against unbounded pareto
         ##
    # get discrepancy statistic
    numpy.random.seed(250624)
    k_s_like_test = k_s_like_test_against_unknown_pareto_distribution(wealths)
    discrepancy_statistic = k_s_like_test.get_discrepancy_statistic()
    k_s_like_test_against_w_rounding_and_other_measurement_errors['pareto'][year]['D stat'] =
         discrepancy_statistic
    # resample discrepancy statistic
    numpy.random.seed(250624)
    resample_d_stats = resample_d_stats_for_unknown_pareto_distribution(wealths=wealths,
         round_to_nearest_for=round_to_nearest_for, number_of_resamples=number_of_resamples,
         fitting=True, rounding=True, diffs_to_draw_from=diffs_btw['forbes and bloomberg
         2013'])
    k_s_like_test_against_w_rounding_and_other_measurement_errors['pareto'][year]['p−value'] =
         resample_d_stats.get_p_value(discrepancy_statistic)
    #k_s_like_test_against_w_rounding_and_other_measurement_errors['pareto'][year]['D stats'] =
         resample_d_stats.resampled_discrepancy_statistics # de−comment to store D−stats
    ## k−s like test with rounding and other measurement errors against bounded pareto ##
    # get discrepancy statistic against bounded pareto distribution
    numpy.random.seed(250624)
```

```
    k_s_like_test = k_s_like_test_against_unknown_bounded_pareto_distribution(wealths=wealths)
    discrepancy_statistic = k_s_like_test.get_discrepancy_statistic()
    k_s_like_test_against_w_rounding_and_other_measurement_errors['bounded pareto'][year]['D
        stat'] = discrepancy_statistic
    # resample discrepancy statistic
    numpy.random.seed(250624)
    resample_d_stats =
        resample_d_stats_for_unknown_bounded_pareto_distribution(wealths=wealths,
        round_to_nearest_for=round_to_nearest_for, number_of_resamples=number_of_resamples,
        diffs_to_draw_from=diffs_btw['forbes and bloomberg 2013']) # always rounding and
        fitting
    k_s_like_test_against_w_rounding_and_other_measurement_errors['bounded
        pareto'][year]['p-value'] = resample_d_stats.get_p_value(discrepancy_statistic)
    #k_s_like_test_against_w_rounding_and_other_measurement_errors['bounded pareto'][year]['D
        stats'] = resample_d_stats.resampled_discrepancy_statistics # de-comment to store
        D-stats
    print year
  # dump results
  pickle.dump(k_s_like_test_against_w_rounding_and_other_measurement_errors, open(filename,
      'wb'))
else:
  # load result
  k_s_like_test_against_w_rounding_and_other_measurement_errors = pickle.load(open(filename,
      'rb'))

# print results
distribution = 'pareto'
print "\nresults of k-s test for " + distribution + "\nt-stat \tp-value accounting for rounding
    \tp-value accounting for rounding and other measurement errors"
for year in years:
  _ = k_s_like_test_against_w_rounding_errors['pareto'][year]
  print year, "\t", _['D stat'], "\t", _['p-value'], "\t",
  print k_s_like_test_against_w_rounding_and_other_measurement_errors['pareto'][year]['p-value']

## CLASSES FOR FITTING A TRUNCATED LOG-NORMAL AND COMPARING ITS
    FIT TO THAT OF A PARETO ##

class fit_truncated_log_normal_distribution_to(object):
  ''' Class for getting the maximum likelihood estimates of the parameters of a two-parameter
      log-normal distribution truncated from below at a known truncation point '''

  def __init__(self, x, truncated_from_below_at):
    self.x = x # the data
    self.truncated_from_below_at = truncated_from_below_at # the truncation point
    self.converged = False # optimization method converged

  def pdf_at(self, xi, parameters):
    ''' Returns the PDF at a given xi of a log-normal with given mean and stdev parameters '''
    # unpack the parameters
    mean, stdev = parameters
    # return the pdf evaluated at xi
```

```
        return (1. / (xi * numpy.sqrt(2. * numpy.pi * (stdev ** 2.)))) * numpy.exp((-
            (numpy.log(xi) - mean) ** 2.) / (2. * (stdev ** 2.)))

    def cdf_at(self, xi, parameters):
        ''' Returns the CDF at a given xi of a log-normal with given mean and stdev parameters '''
        # unpack the parameters
        mean, stdev = parameters
        # return the cdf evaluated at xi
        return 0.5 * (1. + erf((numpy.log(xi) - mean) / numpy.sqrt(2. * (stdev ** 2.))))

    def negative_log_likelihood_function(self, parameters):
        ''' Returns the negative log-likelihood function for given data and parameters '''
        # create a variable for the sum of contributions to the negative log likelihood function
        sum_of_contributions_to_log_likelihood_function = 0.
        # create scaling factor
        ln_of_1_minus_cdf_at_truncation_point = numpy.log((1. -
            self.cdf_at(self.truncated_from_below_at, parameters)))
        # sum the contributions
        for xi in self.x:
            sum_of_contributions_to_log_likelihood_function += numpy.log(self.pdf_at(xi, parameters))
                - ln_of_1_minus_cdf_at_truncation_point
        # return the (negative of the) sum of contributions
        return - sum_of_contributions_to_log_likelihood_function

    def partial_derivative_of_ln_of_pdf_wrt_mean_at(self, xi, parameters):
        mean, stdev = parameters
        return (numpy.log(xi) - mean) / (stdev ** 2.)

    def partial_derivative_of_ln_of_pdf_wrt_stdev_at(self, xi, parameters):
        mean, stdev = parameters
        return (- 1. / stdev) + (((numpy.log(xi) - mean) ** 2.) / (stdev ** 2.))

    def partial_derivative_of_cdf_wrt_mean_at(self, xi, parameters):
        mean, stdev = parameters
        return (- 1. / numpy.sqrt(2 * numpy.pi * (stdev ** 2.))) * numpy.exp((- (numpy.log(xi) -
            mean) ** 2.) / (2. * (stdev ** 2.)))

    def partial_derivative_of_cdf_wrt_stdev_at(self, xi, parameters):
        mean, stdev = parameters
        return (- (numpy.log(xi) - mean) / ((stdev ** 2.) * numpy.sqrt(2. * numpy.pi))) *
            numpy.exp((- (numpy.log(xi) - mean) ** 2.) / (2. * (stdev ** 2.)))

    def gradient_of_negative_log_likelihood_function(self, parameters):
        ''' Returns the gradient of the negative log-likelihood function '''
        # create variables for the components of the gradient
        partial_derivative_wrt_mean = 0.
        partial_derivative_wrt_stdev = 0.
        # create scaling factors
        cdf_at_truncation_point = self.cdf_at(self.truncated_from_below_at, parameters)
        partial_derivative_of_cdf_wrt_mean_at_truncation_point =
            self.partial_derivative_of_cdf_wrt_mean_at(self.truncated_from_below_at, parameters)
```

```
        partial_derivative_of_cdf_wrt_stdev_at_truncation_point =
            self.partial_derivative_of_cdf_wrt_stdev_at(self.truncated_from_below_at, parameters)
        # sum the contributions
        for xi in self.x:
            partial_derivative_wrt_mean += self.partial_derivative_of_ln_of_pdf_wrt_mean_at(xi,
                parameters) + (1. / cdf_at_truncation_point) *
                partial_derivative_of_cdf_wrt_mean_at_truncation_point
            partial_derivative_wrt_stdev += self.partial_derivative_of_ln_of_pdf_wrt_stdev_at(xi,
                parameters) + (1. / cdf_at_truncation_point) *
                partial_derivative_of_cdf_wrt_stdev_at_truncation_point
        # return the (negative of the) components of the gradient
        return numpy.array([- partial_derivative_wrt_mean, - partial_derivative_wrt_stdev])

    def get_initial_guess(self):
        ''' Returns an initial guess for the parameters based on the MLEs of an untruncated
            log-normal '''
        mean0 = (1. / len(self.x)) * sum([numpy.log(xi) for xi in self.x])
        stdev0 = numpy.sqrt((1. / len(self.x)) * sum([(numpy.log(xi) - mean0) ** 2. for xi in
            self.x]))
        return mean0, stdev0

    def get_estimated_parameters(self):
        ''' Returns the MLEs of the parameters '''
        # if optimization fails to converge, try a new guess
        tries = 0
        fmins = list()
        initial_guess = self.get_initial_guess()
        while (self.converged == False) and (tries < 100+1):
            guess = numpy.random.uniform(initial_guess[0] / 2., initial_guess[0] * 2.),
                numpy.random.uniform(initial_guess[1] / 2., initial_guess[1] * 2.) # take anywhere
                between half and twice the initial guess
            fmin = scipy.optimize.fmin_tnc(func=self.negative_log_likelihood_function,
                fprime=self.gradient_of_negative_log_likelihood_function, bounds=[(None, None),
                (0.+10**-100, None)], x0=guess, maxfun=100, messages=0) # 8 is "Exit reasons"
            parameters = (fmin[0][0], fmin[0][1])
            negative_log_likelihood = self.negative_log_likelihood_function(parameters)
            converged = (fmin[2] == 1) # function converged
            if numpy.isnan(negative_log_likelihood) or numpy.isinf(negative_log_likelihood):
                pass
            else:
                fmins.append((negative_log_likelihood, parameters, converged))
                fmins.sort()
                self.estimated_parameters = fmins[0][1]
                self.converged = fmins[0][2]
            tries += 1
            if len(fmins) == 0:
                tries -= 1 # try again, if only found log-likelihood of NaN or +/Infinity
            if (self.converged == True) and (tries < 10+1):
                self.converged = False # try at least ten starting points
        return self.estimated_parameters

    def get_converged(self):
```

```
        return self.converged

    def get_std_errors_of_estimated_parameters(self):
        ''' Returns std errors of the estimated mean and stdev parameters '''
        # following gretl, standard errors are based on the outer product of the gradient
        G = list()
        # create scaling factors
        cdf_at_truncation_point = self.cdf_at(self.truncated_from_below_at, self.estimated_parameters)
        partial_derivative_of_cdf_wrt_mean_at_truncation_point =
            self.partial_derivative_of_cdf_wrt_mean_at(self.truncated_from_below_at,
            self.estimated_parameters)
        partial_derivative_of_cdf_wrt_stdev_at_truncation_point =
            self.partial_derivative_of_cdf_wrt_stdev_at(self.truncated_from_below_at,
            self.estimated_parameters)
        # sum the contributions
        for xi in self.x:
            partial_derivative_of_ln_of_pdf_wrt_mean_at_xi =
                self.partial_derivative_of_ln_of_pdf_wrt_mean_at(xi, self.estimated_parameters) + (1. /
                cdf_at_truncation_point) * partial_derivative_of_cdf_wrt_mean_at_truncation_point
            partial_derivative_of_ln_of_pdf_wrt_stdev_at_xi =
                self.partial_derivative_of_ln_of_pdf_wrt_stdev_at(xi, self.estimated_parameters) + (1. /
                cdf_at_truncation_point) * partial_derivative_of_cdf_wrt_stdev_at_truncation_point
            G.append([- partial_derivative_of_ln_of_pdf_wrt_mean_at_xi, -
                partial_derivative_of_ln_of_pdf_wrt_stdev_at_xi])
        G = numpy.matrix(G)
        VCV = numpy.array((G.T * G).I)
        return numpy.sqrt(VCV[0][0]), numpy.sqrt(VCV[1][1])

class vuong_test_for_pareto_against_truncated_log_normal_distribution(object):
    ''' Class for a Vuong (1989) test of a two-parameter Pareto distribution against a truncated
        log-normal distribution. Follows Cameron and Trivedi (2006, sec. 8.5.3–5) '''

    def __init__(self, wealths, truncated_from_below_at):
        self.wealths = wealths
        self.truncated_from_below_at = truncated_from_below_at # for truncated log-normal
        self.fit_distributions()
        self.n = len(wealths)

    def fit_distributions(self):
        # fit parteo
        fit_to_pareto = fit_pareto_to(wealths=self.wealths)
        self.estimated_parameters_for_pareto = fit_to_pareto.get_estimated_parameters()
        # fit log normal
        self.fit_to_truncated_log_normal = fit_truncated_log_normal_distribution_to(x=self.wealths,
            truncated_from_below_at=self.truncated_from_below_at)
        self.estimated_parameters_for_truncated_log_normal =
            self.fit_to_truncated_log_normal.get_estimated_parameters()

    def get_log_likelihood_at_for_fitted_pareto(self, wealth):
        lower_bound, shape = self.estimated_parameters_for_pareto # unpack parameters
        return numpy.log(shape * (lower_bound ** shape) * (wealth ** (- 1. - shape)))
```

```
    def get_log_likelihood_at_for_fitted_truncated_log_normal(self, wealth):
        ln_of_1_minus_cdf_at_truncation_point = numpy.log((1. - self.fit_to_truncated_log_normal.cdf_at(self.truncated_from_below_at, self.estimated_parameters_for_truncated_log_normal)))
        return numpy.log(self.fit_to_truncated_log_normal.pdf_at(xi=wealth, parameters=self.estimated_parameters_for_truncated_log_normal)) - ln_of_1_minus_cdf_at_truncation_point

    def get_test_statistic(self):
        ''' Returns the test statistic '''
        likelihood_ratios = list(self.get_log_likelihood_at_for_fitted_pareto(wealth=wealth) - self.get_log_likelihood_at_for_fitted_truncated_log_normal(wealth=wealth) for wealth in self.wealths)
        LR = sum(likelihood_ratios)
        omega_squared = (1. / self.n) * sum(list(likelihood_ratio ** 2. for likelihood_ratio in likelihood_ratios)) - ((1. / self.n) * LR) ** 2.
        return LR / numpy.sqrt(self.n * omega_squared)

## CLASSES FOR FITTING A TRUNCATED GAMMA AND COMPARING ITS FIT TO THAT OF A PARETO ##

class fit_truncated_gamma_distribution_to(object):
    ''' Class for getting the maximum likelihood estimates of the parameters of a two-parameter gamma distribution truncated from below at a known truncation point '''

    def __init__(self, x, truncated_from_below_at):
        self.x = x # the data
        self.truncated_from_below_at = truncated_from_below_at # the truncation point
        self.converged = False # optimization method converged

    def pdf_at(self, xi, parameters):
        # unpack the parameters
        shape, scale = parameters
        # return the pdf evaluated at xi
        return (1. / scipy.special.gamma(shape)) * (scale ** (- shape)) * (xi ** (shape - 1)) * numpy.exp(- xi / scale)

    def cdf_at(self, xi, parameters):
        # unpack the parameters
        shape, scale = parameters
        # return the cdf evaluated at xi
        return scipy.special.gammainc(shape, xi / scale)
        # note that scipy.special.gammainc is the regularized version of the lower incomplete gamma function

    def negative_log_likelihood_function(self, parameters):
        # create a variable for the sum of contributions to the negative log likelihood function
        sum_of_contributions_to_log_likelihood_function = 0.
        # create scaling factor, note that the scipy.special.gammaincc could be used in calculating the complement
        ln_of_1_minus_cdf_at_truncation_point = numpy.log((1. - self.cdf_at(self.truncated_from_below_at, parameters)))
```

```
    # sum the contributions
    for xi in self.x:
        sum_of_contributions_to_log_likelihood_function += numpy.log(self.pdf_at(xi, parameters))
            − ln_of_1_minus_cdf_at_truncation_point
    # return the (negative of the) sum of contributions
    return − sum_of_contributions_to_log_likelihood_function

def partial_derivative_of_ln_of_pdf_wrt_shape_at(self, xi, parameters):
    shape, scale = parameters
    return numpy.log(xi) − numpy.log(scale) − scipy.special.psi(shape)

def partial_derivative_of_ln_of_pdf_wrt_scale_at(self, xi, parameters):
    shape, scale = parameters
    return − (shape / scale) − (xi / scale)

def partial_derivative_of_cdf_wrt_shape_at(self, xi, parameters):
    shape, scale = parameters
    return "TBD"

def partial_derivative_of_cdf_wrt_scale_at(self, xi, parameters):
    shape, scale = parameters
    return − (numpy.exp(− x / scale) * ((x / scale) ** shape)) / (scale *
        scipy.special.gamma(shape))

def gradient_of_negative_log_likelihood_function(self, parameters):
    # create variables for the components of the gradient
    partial_derivative_wrt_shape = 0.
    partial_derivative_wrt_scale = 0.
    # create scaling factors
    cdf_at_truncation_point = self.cdf_at(self.truncated_from_below_at, parameters)
    partial_derivative_of_cdf_wrt_shape_at_truncation_point =
        self.partial_derivative_of_cdf_wrt_shape_at(self.truncated_from_below_at, parameters)
    partial_derivative_of_cdf_wrt_scale_at_truncation_point =
        self.partial_derivative_of_cdf_wrt_scale_at(self.truncated_from_below_at, parameters)
    # sum the contributions
    for xi in self.x:
        partial_derivative_wrt_shape += self.partial_derivative_of_ln_of_pdf_wrt_shape_at(xi,
            parameters) + (1. / cdf_at_truncation_point) *
            partial_derivative_of_cdf_wrt_shape_at_truncation_point
        partial_derivative_wrt_scale += self.partial_derivative_of_ln_of_pdf_wrt_scale_at(xi,
            parameters) + (1. / cdf_at_truncation_point) *
            partial_derivative_of_cdf_wrt_scale_at_truncation_point
    # return the (negative of the) components of the gradient
    return numpy.array([− partial_derivative_wrt_shape, − partial_derivative_wrt_scale])

def get_initial_guess(self):
    ''' Returns an initial guess for the parameters based on Thom's (1958) approximation to the
        MLEs for an untruncated gamma '''
    arithmetic_mean = (1. / len(self.x)) * sum(self.x)
    d = numpy.log(arithmetic_mean) − (1. / len(self.x)) * sum(list(numpy.log(xi) for xi in
        self.x))
    shape0 = (1. + numpy.sqrt(1. + ((4. * d) / 3.))) / (4. * d)
```

```
        scale0 = arithmetic_mean / shape0
        return shape0, scale0

    def get_estimated_parameters(self):
        ''' Returns the MLEs '''
        tries = 0
        fmins = list()
        initial_guess = self.get_initial_guess()
        while (self.converged == False) and (tries < 100+1):
            guess = numpy.random.uniform(initial_guess[0] / 2., initial_guess[0] * 2.),
                numpy.random.uniform(initial_guess[1] / 2., initial_guess[1] * 2.)
            fmin = scipy.optimize.fmin_l_bfgs_b(func=self.negative_log_likelihood_function, x0=guess,
                fprime=None, approx_grad=True, bounds=[(0.+10**-100, None), (0.+10**-100,
                None)], maxfun=100)
            parameters = (fmin[0][0], fmin[0][1])
            negative_log_likelihood = float(fmin[1])
            converged = (fmin[2]['warnflag'] == 0)
            if numpy.isnan(negative_log_likelihood) or numpy.isinf(negative_log_likelihood):
                pass
            else:
                fmins.append((negative_log_likelihood, parameters, converged))
                fmins.sort()
                self.estimated_parameters = fmins[0][1]
                self.converged = fmins[0][2]
            tries += 1
            if len(fmins) == 0:
                tries -= 1 # try again, if only found log-likelihood of NaN or +/Inf
            if (self.converged == True) and (tries < 10+1):
                self.converged = False # try at least ten starting points
        return self.estimated_parameters

    def get_converged(self):
        return self.converged

class vuong_test_for_pareto_against_truncated_gamma_distribution(object):
    ''' Class for a Vuong (1989) test of of a two-parameter Pareto distribution against a
        left-truncated gamma distribution '''

    def __init__(self, wealths, truncated_from_below_at):
        self.wealths = wealths
        self.truncated_from_below_at = truncated_from_below_at # for truncated gamma
        self.fit_distributions()
        self.n = len(wealths)

    def fit_distributions(self):
        # fit parteo
        fit_to_pareto = fit_pareto_to(wealths=self.wealths)
        self.estimated_parameters_for_pareto = fit_to_pareto.get_estimated_parameters()
        # fit truncated gamma
        self.fit_to_truncated_gamma = fit_truncated_gamma_distribution_to(x=self.wealths,
            truncated_from_below_at=self.truncated_from_below_at)
```

```
    self.estimated_parameters_for_truncated_gamma =
        self.fit_to_truncated_gamma.get_estimated_parameters()

  def get_log_likelihood_at_for_fitted_pareto(self, wealth):
    lower_bound, shape = self.estimated_parameters_for_pareto # unpack parameters
    return numpy.log(shape * (lower_bound ** shape) * (wealth ** (- 1. - shape)))

  def get_log_likelihood_at_for_fitted_truncated_gamma(self, wealth):
    ln_of_1_minus_cdf_at_truncation_point = numpy.log((1. -
        self.fit_to_truncated_gamma.cdf_at(xi=self.truncated_from_below_at,
        parameters=self.estimated_parameters_for_truncated_gamma)))
    return numpy.log(self.fit_to_truncated_gamma.pdf_at(xi=wealth,
        parameters=self.estimated_parameters_for_truncated_gamma)) -
        ln_of_1_minus_cdf_at_truncation_point

  def get_test_statistic(self):
    likelihood_ratios = list(self.get_log_likelihood_at_for_fitted_pareto(wealth=wealth) -
        self.get_log_likelihood_at_for_fitted_truncated_gamma(wealth=wealth) for wealth in
        self.wealths)
    LR = sum(likelihood_ratios)
    omega_squared = (1. / self.n) * sum(list(likelihood_ratio ** 2. for likelihood_ratio in
        likelihood_ratios)) - ((1. / self.n) * LR) ** 2.
    return LR / numpy.sqrt(self.n * omega_squared)

## CLASS FOR COMPARING FIT OF BOUNDED AND UNBOUNDED PARETOS ##

class lr_test_for_pareto_against_bounded_pareto_distribution(object):
  ''' Class for a likelihood ratio test of an unbounded Pareto distribution against a bounded one.
      A standard likelihood ratio test can be applied because the former is nested in the latter.
      The two distributions are the same as the bounding parameter goes to infinity '''

  def __init__(self, wealths):
    self.wealths = wealths
    self.fit_distributions()

  def fit_distributions(self):
    # fit parteo
    fit_to_pareto = fit_pareto_to(wealths=self.wealths)
    self.estimated_parameters_for_pareto = fit_to_pareto.get_estimated_parameters()
    # fit bounded pareto
    fit_to_bounded_pareto = fit_bounded_pareto_to(wealths=self.wealths)
    self.estimated_parameters_for_bounded_pareto =
        fit_to_bounded_pareto.get_estimated_parameters()

  def get_log_likelihood_at_for_fitted_pareto(self, wealth):
    lower_bound, shape = self.estimated_parameters_for_pareto # unpack parameters
    return numpy.log(shape * (lower_bound ** shape) * (wealth ** (- 1. - shape)))

  def get_log_likelihood_at_for_fitted_bounded_pareto(self, wealth):
    lower_bound, upper_bound, shape = self.estimated_parameters_for_bounded_pareto # unpack
        parameters
```

```
    return numpy.log((shape * (lower_bound ** shape) * (wealth ** (- 1. - shape))) / (1. -
        (lower_bound / upper_bound) ** shape))

  def get_test_statistic(self):
    ''' Returns the test statistic '''
    likelihood_for_pareto = sum(list(self.get_log_likelihood_at_for_fitted_pareto(wealth=wealth)
        for wealth in self.wealths))
    likelihood_for_bounded_pareto =
        sum(list(self.get_log_likelihood_at_for_fitted_bounded_pareto(wealth=wealth) for wealth
        in self.wealths))
    return -2. * (likelihood_for_pareto - likelihood_for_bounded_pareto)

## LIKELIHOOD-RATIO TESTS OF PARETO AGAINST OTHER DISTRIBUTIONS ##

def get_p_value(z_stat):
  ''' Returns p-value for z-test '''
  return scipy.stats.norm.cdf(-abs(z_stat), loc=0, scale=1) + scipy.stats.norm.sf(abs(z_stat),
      loc=0, scale=1)

# define dicts for storing results of tests
other_distributions = ['gamma', 'log-normal', 'bounded pareto']
lr_test_of_pareto_against = dict(zip(other_distributions, [dict(zip(years, [dict(zip(['t-stat',
    'p-value'], ['TBD' for _ in ['t-stat', 'p-value']])) for year in years])) for distribution in
    other_distributions]))
# set the following to True to re-estimate if already estimate
re_estimate = False
# check if already been estimated
filename = filepath + "lr_test_of_pareto_against.pkl"
already_estimated = isfile(filename)
# estimate or re-estimate
if not already_estimated or re_estimate:
  # for each year
  for year in years:
    # get wealths
    wealths = summary_stats['wealths'][year]
    # get rounding
    round_to_nearest_for = dict()
    unique_wealths = list(set(wealths))
    unique_wealths.sort(reverse=True)
    # if wealthiest person isn't a billionaire
    if summary_stats['max'][year] < 1000.0:
      min_difference = min(list(unique_wealths[i] - unique_wealths[i+1] for i in range(0,
          len(unique_wealths)-1) if unique_wealths[i] < 1000.0))
      round_to_nearest_for['billionaires'] = min_difference # if a billion dollar wealth is drawn,
          use the same rounding
      round_to_nearest_for['sub-billionaires'] = min_difference
    # if poorest person is a billionaire
    elif summary_stats['min'][year] >= 1000.0:
      min_difference = min(list(unique_wealths[i] - unique_wealths[i+1] for i in range(0,
          len(unique_wealths)-1) if unique_wealths[i+1] >= 1000.0))
      round_to_nearest_for['billionaires'] = min_difference # if a sub-billion dollar wealth is
          drawn, use the same rounding
```

```
        round_to_nearest_for['sub-billionaires'] = min_difference
      else:
        round_to_nearest_for['billionaires'] = min(list(unique_wealths[i] - unique_wealths[i+1] for i
            in range(0, len(unique_wealths)-1) if unique_wealths[i+1] >= 1000.0))
        round_to_nearest_for['sub-billionaires'] = min(list(unique_wealths[i] - unique_wealths[i+1]
            for i in range(0, len(unique_wealths)-1) if unique_wealths[i] < 1000.0))
      # get truncation point
      truncated_from_below_at = summary_stats['min'][year] -
          (round_to_nearest_for['sub-billionaires'] / 2.)
      ## test pareto against gamma ##
      numpy.random.seed(250624)
      vuong_test = vuong_test_for_pareto_against_truncated_gamma_distribution(wealths=wealths,
          truncated_from_below_at=truncated_from_below_at)
      test_statistic = vuong_test.get_test_statistic()
      lr_test_of_pareto_against['gamma'][year]['t-stat'] = test_statistic
      lr_test_of_pareto_against['gamma'][year]['p-value'] = get_p_value(z_stat=test_statistic)
      ## test pareto against truncated log-normal ##
      numpy.random.seed(250624)
      vuong_test =
          vuong_test_for_pareto_against_truncated_log_normal_distribution(wealths=wealths,
          truncated_from_below_at=truncated_from_below_at)
      test_statistic = vuong_test.get_test_statistic()
      lr_test_of_pareto_against['log-normal'][year]['t-stat'] = test_statistic
      lr_test_of_pareto_against['log-normal'][year]['p-value'] = get_p_value(z_stat=test_statistic)
      ## test pareto against bounded pareto ##
      # note that the unbounded pareto is nested in the bounded pareto
      numpy.random.seed(250624)
      lr_test = lr_test_for_pareto_against_bounded_pareto_distribution(wealths=wealths)
      test_statistic = lr_test.get_test_statistic()
      lr_test_of_pareto_against['bounded pareto'][year]['t-stat'] = test_statistic
      lr_test_of_pareto_against['bounded pareto'][year]['p-value'] = 1. -
          scipy.stats.chi2.cdf(test_statistic, 1) # chi-squared with one degree of freedom
      print year
    # dump results
    pickle.dump(lr_test_of_pareto_against, open(filename, 'wb'))
else:
    # load result
    lr_test_of_pareto_against = pickle.load(open(filename, 'rb'))

# print results
print "\nresults of LR tests"
for distribution in other_distributions:
    # intepretation of Vuong (1989) tests
    if distribution not in ['bounded pareto']:
      n_years_favors_dist = int(sum([lr_test_of_pareto_against[distribution][year]['t-stat'] > 0. and
          lr_test_of_pareto_against[distribution][year]['p-value'] < 0.10 for year in years]))
      n_years_favors_pareto = int(sum([lr_test_of_pareto_against[distribution][year]['t-stat'] < 0.
          and lr_test_of_pareto_against[distribution][year]['p-value'] < 0.10 for year in years]))
      neither_favored = len(years) - (n_years_favors_dist + n_years_favors_pareto)
    # intepretation of standard LR tests
    else:
```

```
    n_years_favors_dist = int(sum([lr_test_of_pareto_against[distribution][year]['p-value'] < 0.10
        for year in years]))
    n_years_favors_pareto = len(years)
  print "A LR test favors a " + distribution + " over an unbounded pareto in " +
      str(n_years_favors_dist) + " years.",
  if distribution not in ['bounded pareto']:
    print "The same test favors the pareto over the " + distribution + " in " +
        str(n_years_favors_pareto) + " years.",
    print "Neither is favored over the other in " + str(neither_favored) + " years."
```

E.5 Code for the Third Essay

The code for generating the results reported in the third essay is as follows.

```
'''
Filename: essay_on_duration_of_wealth_code.py
Python version: 2.7
Source: Capehart, Kevin W. Essays on the Wealthiest Americans. PhD dissertation, American
    University, Washington, DC, 2014.
Description: This file generates results reported by the essay on the duration of the wealth of the
    wealthiest Americans.
'''

from __future__ import division
from load_forbes_400_dataset import * # load our Forbes 400 dataset
import scipy.stats # import for confidence intervals
from scipy.special import btdtri # import for quantiles of beta distribution
import statsmodels.api as sm # import for multionomial logit model
from os.path import isfile # import for checking whether a file exists
import pickle # import for dumping and loading computationally intensive results

## SEC. 4.3 AND C.1: NUMBER APPEARING AGAIN OR DROPPING OFF ##

# The following code looks at the year-over-year movement of unique individuals onto, off of,
    and throughout the Forbes 400.

dyears = [year + '--' + str(int(year) + 1) for year in years_less_last]
dyear_types = ['remain wealthy', 'remain wealthy and move up', 'remain wealthy but move
    down', 'remain wealthy and stay the same', 'exit', 'exit by decline', 'exit by death', 'exit by
    renunciation of citizenship', 'entrants', 'first-timers', 'returnees']
n_by_dyear_types_in = dict(zip(dyears, [dict(zip(dyear_types, [0 for type in dyear_types])) for
    dyear in dyears]))

for dyear in dyears:
  year, next_year = dyear.split('--')
  for name in names:
    if data[name]['wealth'][year] != '':
      # remained wealthy?
      if (data[name]['wealth'][next_year] != ''):
        n_by_dyear_types_in[dyear]['remain wealthy'] += 1
```

```
            rank_in_year = summary_stats['wealths'][year].index(float(data[name]['wealth'][year]))
            rank_in_next_year =
                summary_stats['wealths'][next_year].index(float(data[name]['wealth'][next_year]))
            # up, down, or same?
            if rank_in_next_year > rank_in_year:
              n_by_dyear_types_in[dyear]['remain wealthy and move up'] += 1
            elif rank_in_next_year < rank_in_year:
              n_by_dyear_types_in[dyear]['remain wealthy but move down'] += 1
            else:
              assert (rank_in_next_year == rank_in_year), "If a person didn't move up or down in
                  the rankings, their rank should be the same"
              n_by_dyear_types_in[dyear]['remain wealthy and stay the same'] += 1
          # exited?
          else:
            n_by_dyear_types_in[dyear]['exit'] += 1
            # exited by death, decline, or renunciation of citizenship?
            if (name in dead_in[next_year]):
              n_by_dyear_types_in[dyear]['exit by death'] += 1
            elif (name in renunciants_in[next_year]):
              n_by_dyear_types_in[dyear]['exit by renunciation of citizenship'] += 1
            else:
              n_by_dyear_types_in[dyear]['exit by decline'] += 1
        else:
          # not on list in first year, but comes on by next?
          if data[name]['wealth'][next_year] != '':
            n_by_dyear_types_in[dyear]['entrants'] += 1
            # first-timer or returnee?
            if True in [data[name]['wealth'][str(y)] != '' for y in range(1982, int(year))]:
              n_by_dyear_types_in[dyear]['returnees'] += 1
              assert (True in [data[name]['wealth'][str(y)] != '' for y in range(1982, int(year))]) ==
                  (True in [data[name]['wealth'][str(y)] != '' for y in range(1982, int(next_year))])
            else:
              n_by_dyear_types_in[dyear]['first-timers'] += 1

print "\nyear \tremain wealthy \texit by decline \texit by death \texit by renunciation of
    citizenship"
for dyear in dyears: print year, "\t", n_by_dyear_types_in[dyear]['remain wealthy'], "\t",
    n_by_dyear_types_in[dyear]['exit by decline'], "\t", n_by_dyear_types_in[dyear]['exit by
    death'], "\t", n_by_dyear_types_in[dyear]['exit by renunciation of citizenship']

print "\ntype \tavg number across years"
for type in dyear_types: print type, "\t", round(numpy.average([n_by_dyear_types_in[dyear][type]
    for dyear in dyears]), 0)

# quick calculation to compare to highest income 400
print "\nnumber of unique individuals on Forbes 400 btw 1992 and 2009:\t", len(list(set([name
    for name in names for year in [str(y) for y in range(1992, 2009+1)] if
    (data[name]['wealth'][year] != '')])))

## SEC. 4.3: PATTERNS OF APPEARANCES ###
```

```
# The following code uses Ashworth et al.'s (1994) classification scheme in order to look at the
    patterns with which unique individuals appeared on the Forbes 400 over the years.

patterns = ['transient', 'occasional', 'recurrent', 'persistent', 'chronic', 'permanent']
persons_by_pattern = dict(zip(patterns, [list() for pattern in patterns]))
longest_spell_for_multiple_spells = 0
# for each name
for name in names:
  # get spells
  spells = []
  spell = []
  for year in years:
    wealth = data[name]['wealth'][year]
    if wealth != '':
      spell.append(year)
    else:
      spells.append(spell)
      spell = []
  spells.append(spell)
  # remove empty spells
  while [] in spells:
    spells.remove([])
  # record single-spell patterns
  if len(spells) == 1:
    # 'transient': one 'short' (i.e., one year) spell
    if len(spells[0]) == 1:
      persons_by_pattern['transient'].append(name)
    # 'permanent': one continuous spell
    elif len(spells[0]) == len(years):
      persons_by_pattern['permanent'].append(name)
    # 'persistent': one 'non-short' spell (i.e., longer than one year), but not continuous
    else:
      persons_by_pattern['persistent'].append(name)
  # record multiple-spell patterns
  else:
    # get shortest spell
    shortest_spell = min([len(spell) for spell in spells])
    # get longest spell, too
    longest_spell = max([len(spell) for spell in spells])
    if longest_spell > longest_spell_for_multiple_spells:
      longest_spell_for_multiple_spells = longest_spell
    # get longest time between observed spells
    longest_time_between_spells = []
    for i in range(0, len(spells)-1):
      longest_time_between_spells.append(int(spells[i+1][0]) - int(spells[i][-1]) - 1)
    longest_time_between_spells = max(longest_time_between_spells)
    # 'occasional': multiple 'short' spells (all lasting one year)
    if (shortest_spell == 1):
      persons_by_pattern['occasional'].append((name, [len(spell) for spell in spells]))
    # 'recurrent': multiple 'non-short' spells (some lasting over one year) and 'non-short'
        out-of-poverty spells (some lasting over one year)
    elif (shortest_spell > 1) and (longest_time_between_spells > 1):
```

```
        persons_by_pattern['recurrent'].append((longest_time_between_spells, name, [len(spell) for
            spell in spells]))
    # 'chronic': multiple 'non-short' poverty spells (some lasting over one year) and 'short'
        out-of-poverty spells (lasting one year)
    else:
        persons_by_pattern['chronic'].append((name, [len(spell) for spell in spells]))
    #if len(spells) == 4:
    # print name

## SOME CALCULATIONS ##

# persons for all patterns
persons_by_pattern['all'] = list()
for pattern in patterns:
    persons_by_pattern['all'].extend(persons_by_pattern[pattern])
total_number_of_persons = len(persons_by_pattern['all'])

# single spell patterns
print "\nBetween 1982 and 2013,",
print "out of the " + str(format(total_number_of_persons, ',.0f')) + " people who appeared at
    least once on the magazine's list of the 400 wealthiest Americans,",
print "over three quarters of them (",
number_with_single_spell = len(persons_by_pattern['transient']) +
    len(persons_by_pattern['persistent']) + len(persons_by_pattern['permanent'])
print str(format(number_with_single_spell, ',.0f')) + " or about " + str(int(round(100. *
    number_with_single_spell / total_number_of_persons))),
print "percent of them) came onto the list, appeared in consecutive years, and then dropped off
    without reappearing again.",
# transient, persistent, permanent
print "Out of these people who had only one spell on the magazine's list,",
print "a minority (",
print str(len(persons_by_pattern['transient'])) + " or about " + str(int(round(100. *
    len(persons_by_pattern['transient']) / number_with_single_spell))),
print "percent) only appeared in one year, a majority (",
print str(len(persons_by_pattern['persistent']) + len(persons_by_pattern['permanent'])) + " or
    about " + str(int(round(100. * (len(persons_by_pattern['persistent']) +
    len(persons_by_pattern['permanent'])) / number_with_single_spell))),
print "percent) appeared in more than one year, and a small fraction",
print "(" + str(len(persons_by_pattern['permanent'])),
print "people or about",
print str(int(round(100. * len(persons_by_pattern['permanent']) / number_with_single_spell))),
print "percent) appeared in every year.",
number_with_multiple_spells = len(persons_by_pattern['occasional']) +
    len(persons_by_pattern['recurrent']) + len(persons_by_pattern['chronic'])
# occassional
print "\nOut of the " + str(number_with_multiple_spells) + " people who had more than one
    spell, about half of them (",
print str(len(persons_by_pattern['occasional'])) + " or about " + str(int(round(100. *
    len(persons_by_pattern['occasional']) / number_with_multiple_spells))),
print "percent of them) had a pattern wherein all of their spells (of which the median number
    was",
```

```
print str(int(round(numpy.median([len(spells) for name, spells in
      persons_by_pattern['occasional']])))),
print "and the maximum number was",
print str(max([len(spells) for name, spells in persons_by_pattern['occasional']])),
print "spells) only lasted one year.",
# recurrent
print str(len(persons_by_pattern['recurrent'])) + " or about " + str(int(round(100. *
      len(persons_by_pattern['recurrent']) / number_with_multiple_spells))),
print "percent of them had a pattern wherein at least one of their spells (of which the median
      number was",
print str(int(round(numpy.median([len(spells) for name, spells, longest_time_between_spells in
      persons_by_pattern['recurrent']])))),
print "and the maximum number was",
print str(max([len(spells) for name, spells, longest_time_between_spells in
      persons_by_pattern['recurrent']])),
print "spells) lasted longer than one year, but at least one of the times in between their spells
      was longer than one year.",
# chronic
print str(len(persons_by_pattern['chronic'])) + " or about " + str(int(round(100. *
      len(persons_by_pattern['chronic']) / number_with_multiple_spells))),
print "percent of them had a pattern wherein at least one of their spells (of which the median
      number was",
print str(int(round(numpy.median([len(spells) for name, spells in
      persons_by_pattern['chronic']])))),
print "and the maximum number was",
print str(max([len(spells) for name, spells in persons_by_pattern['chronic']])),
print "spells) lasted longer than one year, and none of the times in between their spells were
      longer than one year.",
# longest and shortest
print "The longest spell for anyone with more than one spell was",
print str(longest_spell_for_multiple_spells),
print "years (the heiress Phoebe Hearst Cooke and the media mogul Sumner Murray Redstone
      both had spells of that length),",
print "while the longest time between any two spells was",
persons_by_pattern['recurrent'].sort(reverse=True)
assert persons_by_pattern['recurrent'][0][0] != persons_by_pattern['recurrent'][1][0] # no ties for
      the longest time between spells
print str(persons_by_pattern['recurrent'][0][0]),
print "years (Kenneth Stanley ''Bud'' Adams Jr. was off the list for that long between his two
      spells)."

## DEFINE FUNCTIONS FOR CI'S AND P-VALUES ##

def get_normal_ci(point, se, confidence_level=0.95):
  ''' Returns lower and upper confidence intervals based on a Normal distribution '''
  z_stat = scipy.stats.norm.ppf(1. - ((1. - confidence_level) / 2.))
  lower = point - z_stat * se
  upper = point + z_stat * se
  return lower, upper

def get_p_value(z_stat):
  ''' Returns p-value for a z-test '''
```

```
  return scipy.stats.norm.cdf(−abs(z_stat), loc=0, scale=1) + scipy.stats.norm.sf(abs(z_stat),
      loc=0, scale=1)

def get_cp_ci(trials, failures, confidence_level=0.95):
  ''' Returns lower and upper Clopper-Pearson confidence intervals, which are calculated based on
      the quantiles of a beta distribution '''
  lower = btdtri(failures, trials − failures + 1., (1. − confidence_level) / 2.)
  upper = btdtri(failures + 1., trials − failures, 1. − ((1. − confidence_level) / 2.))
  return lower, upper

## SEC. 4.3, ESP. FIG. 4.3: SURVIVAL FUNCTION ##

# The following code looks at consecutive appearances of unique invidiauls on the Forbes 400
    and calculates a survival function for appearing again and again for at least a given number
    of consecutive years.

durations = range(1, len(years), 1)
n_at_risk_at = dict(zip(durations, [0 for duration in durations]))
n_exits_by_decline_in_wealth_at = dict(zip(durations, [0 for duration in durations]))
n_exits_by_death_at = dict(zip(durations, [0 for duration in durations]))
n_exits_by_renunciation_of_citizenship_at = dict(zip(durations, [0 for duration in durations]))

# omit left−censored spells?
omit_left_censored_spells = False
if omit_left_censored_spells:
  durations = range(1, len(years)−1, 1) # re−defines durations
# for each name
for name in names:
  # get spells
  spells = []
  spell = []
  for year in years:
    wealth = data[name]['wealth'][year]
    if wealth != '':
      spell.append(year)
    else:
      spells.append(spell)
      spell = []
  spells.append(spell)
  # change left−censored spells to empty spells, so they'll be removed when empty spells are
      removed
  if omit_left_censored_spells:
    for i in range(0, len(spells)):
      if '1982' in spells[i]:
        spells[i] = []
  # remove empty spells
  while [] in spells:
    spells.remove([])
  # for each spell, record data
  while len(spells) > 0:
    spell = spells.pop()
    # as long as the spell didn't start in the last year
```

```
        if spell == [years[-1]]:
            pass
        else:
            # for each year of the spell
            for i in range(0, len(spell)):
                # year
                year = spell[i]
                next_year = str(int(year)+1)
                # if year is not the last year
                if year != years[-1]:
                    # record at risk
                    duration = (i + 1)
                    n_at_risk_at[duration] += 1
                    # remains wealthy?
                    remains_wealthy = int(i != len(spell)-1)
                    # exits by decline in wealth?
                    exits_by_decline_in_wealth = int(data[name]['wealth'][next_year] == '')
                    # exits by death?
                    exits_by_death = int(name in dead_in[next_year])
                    # exits by renunciation of citizenship?
                    exits_by_renunciation_of_citizenship = int(name in renunciants_in[next_year])
                    # record any incidence
                    if exits_by_renunciation_of_citizenship:
                        n_exits_by_renunciation_of_citizenship_at[duration] += 1
                    elif exits_by_death:
                        n_exits_by_death_at[duration] += 1
                    elif exits_by_decline_in_wealth:
                        n_exits_by_decline_in_wealth_at[duration] += 1
                    else: # remains wealthy
                        pass

n_exits_at = dict(zip(durations, [n_exits_by_decline_in_wealth_at[duration] +
        n_exits_by_death_at[duration] + n_exits_by_renunciation_of_citizenship_at[duration] for
        duration in durations]))
point_estimate_for_survivor_function = dict()
lower_ci_for_survivor_function, upper_ci_for_survivor_function = dict(), dict()
duration = durations[0]
lower_ci_for_survivor_function[duration], point_estimate_for_survivor_function[duration],
        upper_ci_for_survivor_function[duration] = 1., 1., 1.

print "\nduration \tpoint estimate for survivor function \t std error"
print duration, "\t", point_estimate_for_survivor_function[duration], "\t", 0.0
for duration in durations[1:]:
    shorter_durations = range(1, duration, 1)
    point_estimate_for_survivor_function[duration] = numpy.prod([1. -
            (n_exits_at[shorter_duration] / n_at_risk_at[shorter_duration]) for shorter_duration in
            shorter_durations])
    var = (point_estimate_for_survivor_function[duration] ** 2.) * sum([n_exits_at[shorter_duration]
            / (n_at_risk_at[shorter_duration] * (n_at_risk_at[shorter_duration] -
            n_exits_at[shorter_duration])) for shorter_duration in shorter_durations])
    lower_ci_for_survivor_function[duration], upper_ci_for_survivor_function[duration] =
            get_normal_ci(point=point_estimate_for_survivor_function[duration], se=numpy.sqrt(var))
```

```
  print duration, "\t", point_estimate_for_survivor_function[duration], "\t", 100. *
      numpy.sqrt(var)
print "\nOf note, if survival curves with and without left−censored observations are computed,
    then the largest difference between them is 0.027692595 and that occurs once at three years,
    but that difference is not quite statistically significant at the 10 percent level (with a
    p−value of about " + str(get_p_value(z_stat=(0.657313437 − 0.629620841) /
    numpy.sqrt(0.000124073 + 0.000164852295641))) + ")."

## SEC. 4.3, ESP. FIG. 4.2: HAZARD FUNCTION ##

# The following code calculates a hazard function for appearing again after a given number of
    consecutive appearances. Normal CI's contain impossible values less than zero, so we use
    Clopper−Pearson CI's instead.

durations = range(1, len(years), 1)
point_estimate_for_hazard_function = dict()
lower_ci_for_hazard_function, upper_ci_for_hazard_function = dict(), dict()
for duration in durations:
  # Normal CI's
  point_estimate_for_hazard_function[duration] = n_exits_at[duration] / n_at_risk_at[duration]
  se = numpy.sqrt((point_estimate_for_hazard_function[duration] * (1. −
      point_estimate_for_hazard_function[duration])) / n_at_risk_at[duration])
  lower, upper = get_normal_ci(point=point_estimate_for_hazard_function[duration], se=se)
  # print results
  if duration == durations[0]: print "\nduration \tpoint estimate for hazard function duration
      \tlower Normal CI \tupper Normal CI \tlower CP CI \tupper CP CI"
  print duration, "\t", point_estimate_for_hazard_function[duration], "\t", lower, "\t", upper,
      "\t",
  # Clopper−Pearson CI's
  lower, upper = get_cp_ci(trials=n_at_risk_at[duration], failures=n_exits_at[duration])
  lower_ci_for_hazard_function[duration] = lower
  upper_ci_for_hazard_function[duration] = upper
  # print results
  print lower, "\t", upper

## SEC. 4.4: GET DATA FOR BASELINE MODEL ##

# keep track of names associated with each spell
names_associated_w_spells = list()
# keep track of ages and ranks
ages_as_dependent_vars, ranks_as_dependent_vars = list(), list()
# list of data for spells
data_for_spells = list()
# for each name
for name in names:
  # get spells
  spells = []
  spell = []
  for year in years:
    wealth = data[name]['wealth'][year]
    if wealth != '':
      spell.append(year)
```

```
  else:
    spells.append(spell)
    spell = []
spells.append(spell)
# remove empty spells
while [] in spells:
  spells.remove([])
# for each spell, record data
while len(spells) > 0:
  spell = spells.pop()
  # as long as the spell didn't start in the last year
  if spell == [years[−1]]:
    pass
  else:
    # list of data for spell
    data_for_spell = list()
    # for each year of the spell
    for i in range(0, len(spell)):
      # year
      year = spell[i]
      next_year = str(int(year)+1)
      # if year is not the last year
      if year != years[−1]:
        # data for given year of the spell
        data_for_year_of_spell = list()
        # remains wealthy?
        remains_wealthy = int(i != len(spell)−1)
        # exits by decline in wealth?
        exits_by_decline_in_wealth = int(data[name]['wealth'][next_year] == '')
        # exits by death? superseding exit by decline
        exits_by_death = int(name in dead_in[next_year])
        # exits by renunciation of citizenship? superseding exit by decline
        exits_by_renunciation_of_citizenship = int(name in renunciants_in[next_year])
        # record dependent variable
        if exits_by_renunciation_of_citizenship:
          dependent_var = 3 # where 3 = exits by renunciation of citizenship:
        elif exits_by_death:
          dependent_var = 2 # where 2 = exits by death
        elif exits_by_decline_in_wealth:
          dependent_var = 1 # where 1 = exits by decline in wealth
        else: # remains wealthy
          dependent_var = 0 # where 0 = remains wealthy
        data_for_year_of_spell.append(dependent_var)
        # independent variables
        # age
        age = int(data[name]['age'][year])
        age_squared = age ∗ age
        # rank
        rank = summary_stats['wealths'][year].index(float(data[name]['wealth'][year])) + 1
        # left−censored dummy
        left_censored = int('1982' in spell)
```

```
            data_for_year_of_spell.extend([1., age, age_squared, rank, left_censored]) # include a
                constant
            # dummies for years, but drop the dummy for 1982
            data_for_year_of_spell.extend([int(year == _) for _ in years_less_initial_and_last])
            # test:
            if 1 in data_for_year_of_spell[6:6+len(years_less_initial_and_last)]:
              assert year ==
                    years[data_for_year_of_spell[6:6+len(years_less_initial_and_last)].index(1)+1]
            else:
              assert year == '1982'
            # dummies for durations of given number of years or 28 or more years (see below),
                but drop the dummy for a duration of one year
            duration = (i + 1)
            for _ in range(2, 27+1):
              data_for_year_of_spell.append(int(duration == _))
            data_for_year_of_spell.append(int(duration > 27))
            # add to list
            if (data_for_year_of_spell[0] != 3): # ignore spells that end with exit by renunciation of
                citizenship for reasons discussed in dissertation
              data_for_spell.append(data_for_year_of_spell)
              ages_as_dependent_vars.append(age)
              ranks_as_dependent_vars.append(rank)
        # add to list
        if (data_for_year_of_spell[0] != 3): # again, ignorning spells that end with exit by
            renunciation of citizenship
          data_for_spells.append(data_for_spell)
          names_associated_w_spells.append(name)

# Each of the 25 people who were on the list for 30 years by 2011 remained on the list into 2012
    and 2013, so there will be perfect prediction problem if we include a dummy variable for a
    duration of 30 or more years. Moreover, none of the people who who were on the list for 29
    years died, they all either remained wealthy or exited by a decline in wealth, so there would
    again be a perfect prediction problem if we included a dummy variable for a duration of 29
    years. In order to avoid these problems while still allowing for a non-parameteric approach
    to duration dependence, we use a dummy for a duration of 28 or more years, instead of
    dummies for durations of 28, 29, and 30 years.

print "\nThere are " + str(sum(len(_) for _ in data_for_spells)) + " person-year observations
    associated with " + str(len(data_for_spells)) + " spells of " +
    str(len(list(set(names_associated_w_spells)))) + " unique individuals."

# Observations of 400 people in each year for 31 years should yield 400 * 31 = 12,400
    person-year observations. But if we ignore the observations associated with spells for the
    four people who renounced their citizenship, then we get 12,379 person-year observations
    associated with the 1,854 spells of 1,451 unique individuals. (One person who would
    eventually exit by renunciation of citizenship, John Thompson Dorrance III, had an earlier
    spell.)

## ORGANIZE DATA AND EXPORT IT ##

# organize data for estimation
dependent_vars = list()
```

```
independent_vars = list()
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    dependent_vars.append(data_for_year_of_spell[0])
    independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
# export to file
filepath = "./data/data_specific_to_third_essay/"
filename = filepath + "exported_survival_data.csv"
f = open(filename, 'w')
# write header
s = "dependent_var,constant,age,age_squared,rank,left_censored,"
for year in years_less_initial_and_last:
  s += "d" + year + ","
for duration in range(2, 27+1):
  s += "d" + str(duration) + ","
s += "d28_or_more" + "\n"
f.write(s)
header = s.replace('\n', '')
header = header.split(',')
n_of_independent_vars = len(header[1:])
# write data
for i in range(0, len(dependent_vars)):
  s = str(dependent_vars[i])
  for var in independent_vars[i]:
    s += ','
    s += str(var)
  s += '\n'
  f.write(s)
# close file
f.close()

## SOME SUMMARY STATISTICS ##

# import the data that was just exported
raw_data = numpy.loadtxt(filename, delimiter=',', skiprows=1, dtype={'names':header,
    'formats':['f']*len(header)})
# calculate some summary statistics
for var in ['age','rank']:
  print "\nThe minimum, median, and maximum " + var + " are " + str(min(raw_data[var])) +
      ', ' + str(numpy.median(raw_data[var])) + ", and " + str(max(raw_data[var])) + ",
      respectively."
for var in header[5:]:
  if var == header[5:][0]: print "\nThe total number of observations where each dummy variable
      is true are as follows."
  print var, "\t", sum(raw_data[var])

## ESTIMATE BASELINE MODEL ##

model = sm.MNLogit(dependent_vars, independent_vars)
estimate_using_gretl = False
```

```
if estimate_using_gretl:
  # use the optimizers used by the MNLogit function in StatsModel to estimate the model
  converged = False
  while converged == False:
    fitted_for_baseline = model.fit(disp=False)
    converged = fitted_for_baseline.mle_retvals['converged']
else:
  # use estimated parameters from gretl and just use the MNLogit function for var-covar
      estimates
  start_params = []
  filename = filepath + "start_params_for_baseline_model.txt"
  f = open(filename, "r")
  lines = f.readlines()
  for line in lines[len(lines) - n_of_independent_vars * 2:]:
    start_params.append(float(line.replace('\n', '')))
  f.close()
  # fit model
  fitted_for_baseline = model.fit(method='bfgs', maxiter=0, disp=False,
      start_params=numpy.array(start_params))
#print fitted_for_baseline.summary(yname=header[0], xname=header[1:]) # de-comment to
    print summary of estimated model
# As a check, it can be shown that the econometric program gretl 1.9.14 generates the same
    estimates (ignoring floating point errors) as the MNLogit function in the StatsModel, so we
    will simply use the estimates generated by the latter.

## SEC. 4.4.3., ESP. FIGS. 4.4 TO 4.7: PREDICTED PROBABILITIES ##

exits = ['remain wealthy', 'exit by decline', 'exit by death'] # ignoring exit by renunciation of
    citizenship for reasons discussed in the essay
ages = range(20, 100+1) # select ages
ranks = [1]
ranks.extend(range(10, 400+1, 10)) # select ranks
durations = range(1, 28+1) # select durations

# set the following to True to re-estimate confidence intervals if they've already been estimated
re_estimate = False
# check if confidence intervals have already been estimated
filenames = [filepath + _ + ".pkl" for _ in ["prob_of_exit_by_age", "prob_of_exit_by_rank",
    "prob_of_exit_by_duration", "prob_of_exit_by_year"]]
already_estimated = False not in [isfile(filename) for filename in filenames]
# estimate or re-estimate
if not already_estimated or re_estimate:

  # construct Krinsky-Robb confidence intervals by taking a large number of draws from a
      multivariate normal with means equal to the estimated parameters and var-covar matrix
      equal to the estimated var-covar matrix
  numpy.random.seed(250624)
  redrawn_parameters =
      numpy.random.multivariate_normal(mean=fitted_for_baseline.params.flatten('F'),
      cov=fitted_for_baseline.cov_params(), size=10000)

  # probs of exits by ages
```

```
prob_of_exit_by_age = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'], [dict(zip(ages, [[] for
    age in ages])) for _ in ['lower', 'point', 'upper']])) for exit in exits]))
# for each age
for age in ages:
  X = [1, age, age * age, 194, 0] # rank of 194, no left censoring
  X.extend([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]) #
      1996
  X.extend([0 for _ in range(2, 28+1)]) # duration of one year
  n = len(X)
  # point estimates for probs, given estimated parameters
  eXb1 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[0:n]))
  eXb2 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[n:]))
  denominator = 1. + eXb1 + eXb2
  prob_of_exit_by_age['remain wealthy']['point'][age] = 1. / denominator
  prob_of_exit_by_age['exit by decline']['point'][age] = eXb1 / denominator
  prob_of_exit_by_age['exit by death']['point'][age] = eXb2 / denominator
  # point estimates for probs, given redrawn parameters
  estimates_for = dict(zip(exits, [[] for exit in exits]))
  for b in redrawn_parameters:
    eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, b[n:]))
    denominator = 1. + eXb1 + eXb2
    estimates_for['remain wealthy'].append(1. / denominator)
    estimates_for['exit by decline'].append(eXb1 / denominator)
    estimates_for['exit by death'].append(eXb2 / denominator)
  # get 95% CI
  for exit in exits:
    prob_of_exit_by_age[exit]['lower'][age] = scipy.stats.mstats.mquantiles(estimates_for[exit],
        prob=[2.5/100.], alphap=1/3., betap=1/3.)[0]
    prob_of_exit_by_age[exit]['upper'][age] = scipy.stats.mstats.mquantiles(estimates_for[exit],
        prob=[97.5/100.], alphap=1/3., betap=1/3.)[0]
# dump results
filename = filepath + "prob_of_exit_by_age.pkl"
pickle.dump(prob_of_exit_by_age, open(filename, 'wb'))

# probs of exits by rank
prob_of_exit_by_rank = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'], [dict(zip(ranks, [None
    for rank in ranks])) for _ in ['lower', 'point', 'upper']])) for exit in exits]))
# for each rank
for rank in ranks:
  X = [1, 64, 4096, rank, 0] # 64 years of age, no left censoring
  X.extend([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]) #
      1996
  X.extend([0 for _ in range(2, 28+1)]) # duration of one year
  n = len(X)
  # point estimates for probs, given estimated parameters
  eXb1 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[0:n]))
  eXb2 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[n:]))
  denominator = 1. + eXb1 + eXb2
  prob_of_exit_by_rank['remain wealthy']['point'][rank] = 1. / denominator
  prob_of_exit_by_rank['exit by decline']['point'][rank] = eXb1 / denominator
  prob_of_exit_by_rank['exit by death']['point'][rank] = eXb2 / denominator
```

```
  # point estimates for probs, given redrawn parameters
  estimates_for = dict(zip(exits, [[] for exit in exits]))
  for b in redrawn_parameters:
    eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, b[n:]))
    denominator = 1. + eXb1 + eXb2
    estimates_for['remain wealthy'].append(1. / denominator)
    estimates_for['exit by decline'].append(eXb1 / denominator)
    estimates_for['exit by death'].append(eXb2 / denominator)
  # get 95% CI
  for exit in exits:
    prob_of_exit_by_rank[exit]['lower'][rank] = scipy.stats.mstats.mquantiles(estimates_for[exit],
        prob=[2.5/100.], alphap=1/3., betap=1/3.)[0]
    prob_of_exit_by_rank[exit]['upper'][rank] = scipy.stats.mstats.mquantiles(estimates_for[exit],
        prob=[97.5/100.], alphap=1/3., betap=1/3.)[0]
# dump results
filename = filepath + "prob_of_exit_by_rank.pkl"
pickle.dump(prob_of_exit_by_rank, open(filename, 'wb'))

# probs of exits by duration
prob_of_exit_by_duration = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(years_less_last, [None for duration in durations])) for _ in ['lower', 'point',
    'upper']])) for exit in exits]))
# for each duration
for duration in durations:
  X = [1, 64, 4096, 194, 0] # 64 years of age, rank of 194, no left censoring
  X.extend([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]) #
      1996
  duration_dummies = [0 for _ in range(2, 28+1)]
  if duration == 1:
    pass
  else:
    duration_dummies[range(2, 28+1).index(duration)] = 1
  X.extend(duration_dummies)
  n = len(X)
  # point estimates for probs, given estimated parameters
  eXb1 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[0:n]))
  eXb2 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[n:]))
  denominator = 1. + eXb1 + eXb2
  prob_of_exit_by_duration['remain wealthy']['point'][duration] = 1. / denominator
  prob_of_exit_by_duration['exit by decline']['point'][duration] = eXb1 / denominator
  prob_of_exit_by_duration['exit by death']['point'][duration] = eXb2 / denominator
  # point estimates for probs, given redrawn parameters
  estimates_for = dict(zip(exits, [[] for exit in exits]))
  for b in redrawn_parameters:
    eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, b[n:]))
    denominator = 1. + eXb1 + eXb2
    estimates_for['remain wealthy'].append(1. / denominator)
    estimates_for['exit by decline'].append(eXb1 / denominator)
    estimates_for['exit by death'].append(eXb2 / denominator)
  # get 95% CI
```

```
    for exit in exits:
        prob_of_exit_by_duration[exit]['lower'][duration] =
            scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[2.5/100.], alphap=1/3.,
            betap=1/3.)[0]
        prob_of_exit_by_duration[exit]['upper'][duration] =
            scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[97.5/100.], alphap=1/3.,
            betap=1/3.)[0]
# dump results
filename = filepath + "prob_of_exit_by_duration.pkl"
pickle.dump(prob_of_exit_by_duration, open(filename, 'wb'))

# probs of exits by year
prob_of_exit_by_year = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(years_less_last, [None for year in years_less_last])) for _ in ['lower', 'point',
    'upper']])) for exit in exits]))
# for each year
for year in years_less_last:
    X = [1, 64, 4096, 194, 0] # 64 years of age, rank of 194, no left censoring
    year_dummies = [0 for _ in years_less_initial_and_last]
    if year == '1982':
        pass
    else:
        year_dummies[years_less_initial_and_last.index(year)] = 1
    X.extend(year_dummies)
    X.extend([0 for _ in range(2, 28+1)]) # duration of one year
    n = len(X)
    # point estimates for probs, given estimated parameters
    eXb1 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[n:]))
    denominator = 1. + eXb1 + eXb2
    prob_of_exit_by_year['remain wealthy']['point'][year] = 1. / denominator
    prob_of_exit_by_year['exit by decline']['point'][year] = eXb1 / denominator
    prob_of_exit_by_year['exit by death']['point'][year] = eXb2 / denominator
    # point estimates for probs, given redrawn parameters
    estimates_for = dict(zip(exits, [[] for exit in exits]))
    for b in redrawn_parameters:
        eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
        eXb2 = numpy.exp(numpy.dot(X, b[n:]))
        denominator = 1. + eXb1 + eXb2
        estimates_for['remain wealthy'].append(1. / denominator)
        estimates_for['exit by decline'].append(eXb1 / denominator)
        estimates_for['exit by death'].append(eXb2 / denominator)
    # get 95% CI
    for exit in exits:
        prob_of_exit_by_year[exit]['lower'][year] = scipy.stats.mstats.mquantiles(estimates_for[exit],
            prob=[2.5/100.], alphap=1/3., betap=1/3.)[0]
        prob_of_exit_by_year[exit]['upper'][year] = scipy.stats.mstats.mquantiles(estimates_for[exit],
            prob=[97.5/100.], alphap=1/3., betap=1/3.)[0]

# dump results
filename = filepath + "prob_of_exit_by_year.pkl"
pickle.dump(prob_of_exit_by_year, open(filename, 'wb'))
```

```
else:

  # load results that were dumped using pickle
  prob_of_exit_by_age = pickle.load(open(filepath + "prob_of_exit_by_age.pkl", 'rb'))
  prob_of_exit_by_rank = pickle.load(open(filepath + "prob_of_exit_by_rank.pkl", 'rb'))
  prob_of_exit_by_duration = pickle.load(open(filepath + "prob_of_exit_by_duration.pkl", 'rb'))
  prob_of_exit_by_year = pickle.load(open(filepath + "prob_of_exit_by_year.pkl", 'rb'))

## SEC. 4.4.3: PREDICTION ERRORS ##

prediction_error_for = dict()
translation_for_exit_number = dict(zip([0, 1, 2], ['Appeared again', 'Exited by decline', 'Exited by
    death']))
for i in range(0, len(data_for_spells)):
  data_for_spell = data_for_spells[i]
  name = names_associated_w_spells[i]
  for data_for_year_of_spell in data_for_spell:
    # point estimates for probs, given estimated parameters
    prob_of = dict()
    X = data_for_year_of_spell[1:]
    n = len(X)
    eXb1 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, fitted_for_baseline.params.flatten('F')[n:]))
    denominator = 1. + eXb1 + eXb2
    prob_of['Appeared again'] = 1. / denominator
    prob_of['Exited by decline'] = eXb1 / denominator
    prob_of['Exited by death'] = eXb2 / denominator
    # actual outcome
    actual_outcome = translation_for_exit_number[data_for_year_of_spell[0]]
    # actual outcome was most likely outcome(s)?
    max_prob = max([prob_of[outcome] for outcome in ['Appeared again', 'Exited by decline',
        'Exited by death']])
    most_prob_outcome = list()
    for outcome in ['Appeared again', 'Exited by decline', 'Exited by death']:
      if prob_of[outcome] == max_prob:
        most_prob_outcome.append(outcome)
    # if actual outcome was not the most probable outcome, get the prediction error
    if actual_outcome in most_prob_outcome:
      pass
    else:
      if 1 in data_for_year_of_spell[6:6+len(years_less_initial_and_last)]:
        year = years[data_for_year_of_spell[6:6+len(years_less_initial_and_last)].index(1)+1]
      else:
        year = '1982'
      next_year = str(int(year) + 1)
      dyear = year + '--' + next_year
      # prediction error
      for outcome in ['Appeared again', 'Exited by decline', 'Exited by death']:
        if outcome == actual_outcome:
          assert len(most_prob_outcome) == 1, "Ties in most-probable outcome"
```

```
        if i==0: print "\nname \tdyear \tactual outcome \tmost probable outcome
            \tprediction error"
        # only print especially egregious prediction errors
        if (1. - prob_of[outcome]) > 0.995:
          print name, "\t", dyear, "\t", actual_outcome, "\t", most_prob_outcome[0], "\t", (1.
              - prob_of[outcome])

## SEC. C.3: TEST OF THE IIA ASSUMPTION FOR BASELINE DURATION MODEL ##

# define dict for storing results
results_of_iaa_test = dict()
# set the following to True to re-estimate if already estimated
re_estimate = False
# check if already been estimated
filename = filepath + _ + "results_of_iaa_test.pkl"
already_estimated = isfile(filename)
# estimate or re-estimate
if not already_estimated or re_estimate:

  # get point estimates, var-covar matrix, and relative risks for full model
  b_full = fitted_for_baseline.params.flatten('F')
  cov_full = fitted_for_baseline.cov_params()
  X = [1, 64, 4096, 194, 0] # 64 years of age, rank of 194, no left censoring
  X.extend([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]) # 1996
  X.extend([0 for _ in range(2, 28+1)]) # duration of one year
  n = len(X)
  eXb1 = numpy.exp(numpy.dot(X, b_full[0:n]))
  eXb2 = numpy.exp(numpy.dot(X, b_full[n:]))
  denominator = 1. + eXb1 + eXb2
  prob_of_remain_wealthy = 1. / denominator
  prob_of_exit_by_decline = eXb1 / denominator
  prob_of_exit_by_death = eXb2 / denominator
  relative_risk_of_exit_by_decline_to_remain_wealthy = prob_of_exit_by_decline /
      prob_of_remain_wealthy
  relative_risk_of_exit_by_death_to_remain_wealthy = prob_of_exit_by_death /
      prob_of_remain_wealthy
  relative_risk_of_exit_by_decline_to_exit_by_death = prob_of_exit_by_decline /
      prob_of_exit_by_death

  # re-estimate model ignoring exit by decline
  dependent_vars = list()
  independent_vars = list()
  for data_for_spell in data_for_spells:
    for data_for_year_of_spell in data_for_spell:
      dependent_var = data_for_year_of_spell[0]
      if (data_for_year_of_spell[0] == 0) or (data_for_year_of_spell[0] == 2): # ignore spells that
          end with exit by decline
        dependent_vars.append(data_for_year_of_spell[0])
        independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
  dependent_vars = numpy.array(dependent_vars)
  independent_vars = numpy.array(independent_vars)
  model_subset = sm.MNLogit(dependent_vars, independent_vars)
```

```
numpy.random.seed(250624)
fitted_subset = model_subset.fit(disp=False)
# get t-stat and p-value
b_subset = fitted_subset.params.flatten('F')
cov_subset = fitted_subset.cov_params()
n = len(fitted_subset.params)
stat = numpy.matrix(b_subset - b_full[n:]) * numpy.matrix(cov_subset - cov_full[n:,n:]).I *
    numpy.matrix(b_subset - b_full[n:]).T
dof = numpy.linalg.matrix_rank(cov_subset - cov_full[n:,n:])
# store results
_ = "t-stat and p-value for ignoring exit by decline"
results_of_iaa_test[_] = (stat[(0,0)], 1.0 - scipy.stats.chi2.cdf(stat[(0,0)], dof))
# get relative risks
eXb2 = numpy.exp(numpy.dot(X, b_subset))
denominator = 1. + eXb2
prob_of_remain_wealthy = 1. / denominator
prob_of_exit_by_death = eXb2 / denominator
relative_risk_of_exit_by_death_to_remain_wealthy_subset = prob_of_exit_by_death /
    prob_of_remain_wealthy
# store results
_ = "relative risk of exit by death to remain wealthy, full model"
results_of_iaa_test[_] = relative_risk_of_exit_by_death_to_remain_wealthy
_ = "relative risk of exit by death to remain wealthy, ignoring exit by decline"
results_of_iaa_test[_] = relative_risk_of_exit_by_death_to_remain_wealthy_subset

# re-estimate model ignoring exit by death
dependent_vars = list()
independent_vars = list()
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    dependent_var = data_for_year_of_spell[0]
    if (data_for_year_of_spell[0] == 0) or (data_for_year_of_spell[0] == 1): # ignore spells that
        end with exit by death
      dependent_vars.append(data_for_year_of_spell[0])
      independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
model_subset = sm.MNLogit(dependent_vars, independent_vars)
# estimate model on subset
numpy.random.seed(250624)
fitted_subset = model_subset.fit(disp=False)
# get t-stat and p-value
b_subset = fitted_subset.params.flatten('F')
cov_subset = fitted_subset.cov_params()
n = len(fitted_subset.params)
stat = numpy.matrix(b_subset - b_full[0:n]) * numpy.matrix(cov_subset - cov_full[0:n,0:n]).I *
    numpy.matrix(b_subset - b_full[0:n]).T
dof = numpy.linalg.matrix_rank(cov_subset - cov_full[0:n,0:n])
# store results
_ = "t-stat and p-value for ignoring exit by death"
results_of_iaa_test[_] = (stat[(0,0)], 1.0 - scipy.stats.chi2.cdf(stat[(0,0)], dof))
# get relative risk for model estimated on subset
```

```
eXb1 = numpy.exp(numpy.dot(X, b_subset))
denominator = 1. + eXb1
prob_of_remain_wealthy = 1. / denominator
prob_of_exit_by_decline = eXb1 / denominator
relative_risk_of_exit_by_decline_to_remain_wealthy_subset = prob_of_exit_by_decline /
    prob_of_remain_wealthy
# store results
_ = "relative risk of exit by decline to remain wealthy, full model"
results_of_iaa_test[_] = relative_risk_of_exit_by_decline_to_remain_wealthy
_ = "relative risk of exit by decline to remain wealthy, ignoring exit by death"
results_of_iaa_test[_] = relative_risk_of_exit_by_decline_to_remain_wealthy_subset

# re-estimate model ignoring remain wealthy
dependent_vars = list()
independent_vars = list()
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    dependent_var = data_for_year_of_spell[0]
    if (data_for_year_of_spell[0] == 1) or (data_for_year_of_spell[0] == 2): # ignore remain
        wealthy
      dependent_vars.append(data_for_year_of_spell[0])
      independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
model_subset = sm.MNLogit(dependent_vars, independent_vars)
# estimate model on subset
numpy.random.seed(250624)
fitted_subset = model_subset.fit(disp=False)
b_subset = fitted_subset.params.flatten('F')
cov_subset = fitted_subset.cov_params()
n = len(fitted_subset.params)
# re-estimate full model with exit by decline as base so that the estimates are comparable
dependent_vars = list()
independent_vars = list()
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    # re-define dependent variable so that exit by decline is the base
    if data_for_year_of_spell[0] == 0:
      dependent_vars.append(1) # 0=remain wealthy becomes 1
    if data_for_year_of_spell[0] == 1:
      dependent_vars.append(0) # 1=exit by decline becomes 0
    if data_for_year_of_spell[0] == 2:
      dependent_vars.append(2) # 2=exit by death stays as 2
    independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
model = sm.MNLogit(dependent_vars, independent_vars)
numpy.random.seed(250624)
fitted_w_different_base = model.fit(disp=False)
# get t-stat and p-value
b_full = fitted_w_different_base.params.flatten('F')
cov_full = fitted_w_different_base.cov_params()
```

```
    stat = numpy.matrix(b_subset − b_full[n:]) ∗ numpy.matrix(cov_subset − cov_full[n:,n:]).I ∗
        numpy.matrix(b_subset − b_full[n:]).T
    dof = numpy.linalg.matrix_rank(cov_subset − cov_full[n:,n:])
    # store results
    _ = "t−stat and p−value for ignoring remaining wealthy"
    results_of_iaa_test[_] = (stat[(0,0)], 1.0 − scipy.stats.chi2.cdf(stat[(0,0)], dof))
    # get relative risk for full model with different base
    eXb2 = numpy.exp(numpy.dot(X, b_subset))
    denominator = 1. + eXb2
    prob_of_exit_by_decline = 1. / denominator
    prob_of_exit_by_death = eXb2 / denominator
    relative_risk_of_exit_by_decline_to_exit_by_death_subset = prob_of_exit_by_decline /
        prob_of_exit_by_death
    # store results
    _ = "relative risk of exit by decline to exit by death, full model"
    results_of_iaa_test[_] = relative_risk_of_exit_by_decline_to_exit_by_death
    _ = "relative risk of exit by decline to exit by death, ignoring remaining wealthy"
    results_of_iaa_test[_] = relative_risk_of_exit_by_decline_to_exit_by_death_subset

    # dump all results
    filename = filepath + "results_of_iaa_test.pkl"
    pickle.dump(results_of_iaa_test, open(filename, 'wb'))

else:

    # load results that were dumped using pickle
    results_of_iaa_test = pickle.load(open(filepath + "results_of_iaa_test.pkl", 'rb'))

organized_keys = ['t−stat and p−value for ignoring exit by decline',
    't−stat and p−value for ignoring exit by death',
    't−stat and p−value for ignoring remaining wealthy',
    'relative risk of exit by death to remain wealthy, full model',
    'relative risk of exit by death to remain wealthy, ignoring exit by decline',
    'relative risk of exit by decline to remain wealthy, full model',
    'relative risk of exit by decline to remain wealthy, ignoring exit by death',
    'relative risk of exit by decline to exit by death, full model',
    'relative risk of exit by decline to exit by death, ignoring remaining wealthy']

# print results
print "\nresults related to test of IIA assumption in baseline duration model"
for key in organized_keys:
    print key, "\t", results_of_iaa_test[key]

## GIVING IT AWAY? ##

# The following code relates to the extension to the baseline duration model that includes a
    philanthropic dummy

## GET SURVIVAL DATA FOR PHILANTHROPIC MODEL ##

# load data on Forbes 400 members on the Philanthropy 50 or Slate 60 (note that we only have
    that data since 1996)
```

```
filename = filepath + “philanthropy_50_and_slate_60_data.txt”
header = [‘year’, ‘philanthropy 50 or slate 60 donor name(s)’, ‘forbes 400 name(s)’, ‘total
    amount committed’, ‘total amount committed in millions’]
rawdata = numpy.loadtxt(filename, delimiter=‘\t’, skiprows=19, dtype={‘names’:(header),
    ‘formats’:([‘<S100’ for _ in range(0, len(header))])})
# get dict of names on Philanthropy 50 and Slate 60 data
years_since_1996 = [str(y) for y in range(1996, 2013+1)]
philanthropists_in = dict(zip(years_since_1996, [[] for year in years_since_1996]))
for i in range(0, len(rawdata[‘year’])):
  year = rawdata[‘year’][i]
  if year in years:
    donor_or_donors = rawdata[‘forbes 400 name(s)’][i]
    # in the data, different names are separated by ‘ and ’
    if ‘ and ’ in donor_or_donors:
      donors = donor_or_donors.split(‘ and ’)
    else:
      donors = [donor_or_donors]
    for name in donors:
      if name in names:
        if data[name][‘wealth’][year] != ‘’:
          philanthropists_in[year].append(name)

# philantropy dummy for just the same year or any earlier year?
just_the_same_year = False
# select years
years_since_1996 = [str(y) for y in range(1996, 2013+1)]
years_since_1996_less_initial_and_last = years_since_1996[1:−1]
# keep track of names associated with each spell
names_associated_w_spells = list()
# keep track of ages and ranks
ages_as_dependent_vars, ranks_as_dependent_vars = list(), list()
# list of data for spells
data_for_spells = list()
# for each name
for name in names:
  # get spells
  spells = []
  spell = []
  for year in years:
    wealth = data[name][‘wealth’][year]
    if wealth != ‘’:
      spell.append(year)
    else:
      spells.append(spell)
      spell = []
  spells.append(spell)
  # remove empty spells
  while [] in spells:
    spells.remove([])
  # for each spell, record data
  while len(spells) > 0:
    spell = spells.pop()
```

```
# as long as the spell didn't start in the last year
if spell == [years[−1]]:
  pass
else:
  # list of data for spell
  data_for_spell = list()
  # for each year of the spell
  for i in range(0, len(spell)):
    # year
    year = spell[i]
    next_year = str(int(year)+1)
    # if year is not the last year
    if year != years[−1]:
      # data for given year of the spell
      data_for_year_of_spell = list()
      # remains wealthy?
      remains_wealthy = int(i != len(spell)−1)
      # exits by decline in wealth?
      exits_by_decline_in_wealth = int(data[name]['wealth'][next_year] == '')
      # exits by death?
      exits_by_death = int(name in dead_in[next_year])
      # exits by renunciation of citizenship?
      exits_by_renunciation_of_citizenship = int(name in renunciants_in[next_year])
      # record dependent variable
      if exits_by_renunciation_of_citizenship:
        dependent_var = 3 # where 3 = exits by renunciation of citizenship:
      elif exits_by_death:
        dependent_var = 2 # where 2 = exits by death
      elif exits_by_decline_in_wealth:
        dependent_var = 1 # where 1 = exits by decline in wealth
      else: # remains wealthy
        dependent_var = 0 # where 0 = remains wealthy
      data_for_year_of_spell.append(dependent_var)
      # independent variables
      # age
      age = int(data[name]['age'][year])
      age_squared = age ∗ age
      # rank
      rank = summary_stats['wealths'][year].index(float(data[name]['wealth'][year])) + 1
      # left−censored dummy
      left_censored = int('1982' in spell)
      data_for_year_of_spell.extend([1., age, age_squared, rank, left_censored]) # include a
          constant
      # philanthropists?
      if int(year) > 1995:
        # same year?
        if just_the_same_year:
          philanthropic = int(name in philanthropists_in[year])
        # any earlier year or same year?
        else:
          philanthropic = int(True in [name in philanthropists_in[str(earlier_year)] for
              earlier_year in range(1996, int(year)+1)])
```

```
            # append to data set
            data_for_year_of_spell.append(philanthropic)
          else:
            data_for_year_of_spell.append('−999')
          # dummies for years, but drop 1996
          data_for_year_of_spell.extend([int(year == _) for _ in
               years_since_1996_less_initial_and_last])
          # dummies for durations of given number of years, except when there is a perfect
               prediction problem
          duration = (i + 1)
          for _ in range(2, 27+1):
            if 13 <= _ <= 15:
              pass
            else:
              data_for_year_of_spell.append(int(duration == _))
          data_for_year_of_spell.append(int(13 <= duration <= 15)) # 13, 14, or 15 years
          data_for_year_of_spell.append(int(duration > 27)) # 28 or more years
          # add to list
          # ignore spells that end with exit by renunciation of citizenship
          # also, ignore years of spells that are before 1996
          if (data_for_year_of_spell[0] != 3) and (int(year) > 1995):
            data_for_spell.append(data_for_year_of_spell)
            ages_as_dependent_vars.append(age)
            ranks_as_dependent_vars.append(rank)
      # add to list
      if (data_for_year_of_spell[0] != 3) and (int(year) > 1995):
        data_for_spells.append(data_for_spell)
        names_associated_w_spells.append(name)

# No one died after their 14th consecutive year on the list, so we'll make a dummy for 13, 14, or
    15 years.

## ORGANIZE PHILANTHROPIC DATA AND EXPORT IT ##

dependent_vars = list()
independent_vars = list()
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    dependent_vars.append(data_for_year_of_spell[0])
    independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
# export to file
f = open(filepath + "exported_survival_data_w_philanthropic_just_the_same_year_" +
    str(just_the_same_year) + ".csv", 'w')
# write header
s = "dependent_var,constant,age,age_squared,rank,left_censored,philanthropic,"
for year in years_since_1996_less_initial_and_last:
  s += "d" + year + ","
for duration in range(2, 27+1):
  if 13 <= duration <=15:
    pass
```

```
  else:
    s += "d" + str(duration) + ","
s += "d13_to_15" + ","
s += "d28_or_more" + "\n"
f.write(s)
# write data
for i in range(0, len(dependent_vars)):
  s = str(dependent_vars[i])
  for var in independent_vars[i]:
    s += ','
    s += str(var)
  s += '\n'
  f.write(s)
# close file
f.close()

## ESTIMATE PHILANTHROPIC MODEL ##

model = sm.MNLogit(dependent_vars, independent_vars)
fitted_for_philanthropic = model.fit(disp=False)
#print fitted_for_philanthropic.summary()

# define dict for storing results on probs of exits for a typical person
philanthropic_types = ['on philanthropy 50 list', 'not on philanthropy 50 list']
prob_of_exit_by_philanthropic_type = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(philanthropic_types, [None for type in philanthropic_types])) for _ in ['lower',
    'point', 'upper']])) for exit in exits]))

# set the following to True to re-estimate if already estimated
re_estimate = False
# check if already been estimated
filename = filepath + _ + "prob_of_exit_by_philanthropic_type_just_the_same_year_" +
    str(just_the_same_year) + ".pkl"
already_estimated = isfile(filename)
# estimate or re-estimate
if not already_estimated or re_estimate:
  # construct Krinsky-Robb confidence intervals by taking a large number of draws from a
      multivariate normal with means equal to the estimated parameters and var-covar matrix
      equal to the estimated var-covar matrix
  numpy.random.seed(250624)
  redrawn_parameters =
      numpy.random.multivariate_normal(mean=fitted_for_philanthropic.params.flatten('F'),
      cov=fitted_for_philanthropic.cov_params(), size=10000)
  # for each type
  for type in philanthropic_types:
    X = [1, 64, 4096, 194, 0] # 64 years of age, rank of 194, no left censoring
    if type == 'on philanthropy 50 list':
      X.append(1)
    else:
      X.append(0)
    X.extend([0 for _ in years_since_1996_less_initial_and_last]) # 1996
    X.extend([0 for _ in range(1, 25+1)]) # duration of one year
```

```
    n = len(X)
    # point estimates for probs, given estimated parameters
    eXb1 = numpy.exp(numpy.dot(X, fitted_for_philanthropic.params.flatten('F')[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, fitted_for_philanthropic.params.flatten('F')[n:]))
    denominator = 1. + eXb1 + eXb2
    prob_of_exit_by_philanthropic_type['remain wealthy']['point'][type] = 1. / denominator
    prob_of_exit_by_philanthropic_type['exit by decline']['point'][type] = eXb1 / denominator
    prob_of_exit_by_philanthropic_type['exit by death']['point'][type] = eXb2 / denominator
    # point estimates for probs, given redrawn parameters
    estimates_for = dict(zip(exits, [[] for exit in exits]))
    for b in redrawn_parameters:
      eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
      eXb2 = numpy.exp(numpy.dot(X, b[n:]))
      denominator = 1. + eXb1 + eXb2
      estimates_for['remain wealthy'].append(1. / denominator)
      estimates_for['exit by decline'].append(eXb1 / denominator)
      estimates_for['exit by death'].append(eXb2 / denominator)
    # get 95% CI
    for exit in exits:
      prob_of_exit_by_philanthropic_type[exit]['lower'][type] =
          scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[2.5/100.], alphap=1/3.,
          betap=1/3.)[0]
      prob_of_exit_by_philanthropic_type[exit]['upper'][type] =
          scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[97.5/100.], alphap=1/3.,
          betap=1/3.)[0]
  # dump results
  filename = filepath + "prob_of_exit_by_philanthropic_type_just_the_same_year_" +
      str(just_the_same_year) + ".pkl"
  pickle.dump(prob_of_exit_by_philanthropic_type, open(filename, 'wb'))
else:
  # load results that were dumped using pickle
  prob_of_exit_by_philanthropic_type = pickle.load(open(filepath +
      "prob_of_exit_by_philanthropic_type_just_the_same_year_" + str(just_the_same_year) +
      ".pkl", 'rb'))

print "\non philanthropy 50 list vs. not on, typical person"
for exit in exits:
  for type in philanthropic_types:
    print exit, "\t", type, "\t", prob_of_exit_by_philanthropic_type[exit]['lower'][type], "\t",
        prob_of_exit_by_philanthropic_type[exit]['point'][type], "\t",
        prob_of_exit_by_philanthropic_type[exit]['upper'][type]

## PROFLIGATE HEIRS? ##

# The following code relates to the extension to the baseline duration model that includes a
    self-made dummy

## GET SURVIVAL DATA FOR SELF-MADE MODEL ##

# keep track of names associated with each spell
names_associated_w_spells = list()
# keep track of ages and ranks
```

```
ages_as_dependent_vars, ranks_as_dependent_vars = list(), list()
# list of data for spells
data_for_spells = list()
# for each name
for name in names:
  # get spells
  spells = []
  spell = []
  for year in years:
    wealth = data[name]['wealth'][year]
    if wealth != '':
      spell.append(year)
    else:
      spells.append(spell)
      spell = []
  spells.append(spell)
  # remove empty spells
  while [] in spells:
    spells.remove([])
  # for each spell, record data
  while len(spells) > 0:
    spell = spells.pop()
    # as long as the spell didn't start in the last year
    if spell == [years[-1]]:
      pass
    else:
      # list of data for spell
      data_for_spell = list()
      # for each year of the spell
      for i in range(0, len(spell)):
        # year
        year = spell[i]
        next_year = str(int(year)+1)
        # if year is not the last year
        if year != years[-1]:
          # data for given year of the spell
          data_for_year_of_spell = list()
          # remains wealthy?
          remains_wealthy = int(i != len(spell)-1)
          # exits by decline in wealth?
          exits_by_decline_in_wealth = int(data[name]['wealth'][next_year] == '')
          # exits by death?
          exits_by_death = int(name in dead_in[next_year])
          # exits by renunciation of citizenship?
          exits_by_renunciation_of_citizenship = int(name in renunciants_in[next_year])
          # record dependent variable
          if exits_by_renunciation_of_citizenship:
            dependent_var = 3 # where 3 = exits by renunciation of citizenship:
          elif exits_by_death:
            dependent_var = 2 # where 2 = exits by death
          elif exits_by_decline_in_wealth:
            dependent_var = 1 # where 1 = exits by decline in wealth
```

```
            else: # remains wealthy
                dependent_var = 0 # where 0 = remains wealthy
            data_for_year_of_spell.append(dependent_var)
            # independent variables
            # age
            age = int(data[name]['age'][year])
            age_squared = age * age
            # rank
            rank = summary_stats['wealths'][year].index(float(data[name]['wealth'][year])) + 1
            # left-censored dummy
            left_censored = int('1982' in spell)
            data_for_year_of_spell.extend([1., age, age_squared, rank, left_censored]) # include a
                constant
            # self-made?
            if int(year) > 1995:
                data_for_year_of_spell.append(int(data[name]['selfmade or inherited'][year] == 'Self
                    made'))
            else:
                data_for_year_of_spell.append('-999')
            # dummies for years, but drop 1996
            data_for_year_of_spell.extend([int(year == _) for _ in
                years_since_1996_less_initial_and_last])
            # dummies for durations of given number of years, except when there is a perfect
                prediction problem
            duration = (i + 1)
            for _ in range(2, 27+1):
                if 13 <= _ <= 15:
                    pass
                else:
                    data_for_year_of_spell.append(int(duration == _))
            data_for_year_of_spell.append(int(13 <= duration <= 15)) # 13, 14, or 15 years
            data_for_year_of_spell.append(int(duration > 27)) # 28 or more years
            # add to list
            # ignore spells that end with exit by renunciation of citizenship
            # also, ignore years of spells that are before 1996
            if (data_for_year_of_spell[0] != 3) and (int(year) > 1995):
                data_for_spell.append(data_for_year_of_spell)
                ages_as_dependent_vars.append(age)
                ranks_as_dependent_vars.append(rank)
        # add to list
        if (data_for_year_of_spell[0] != 3) and (int(year) > 1995):
            data_for_spells.append(data_for_spell)
            names_associated_w_spells.append(name)

# No one died after their 14th consecutive year on the list, so we'll make a dummy for 13, 14, or
    15 years

## ORGANIZE SELF-MADE DATA AND EXPORT IT ##

# organize data for estimation
dependent_vars = list()
independent_vars = list()
```

```
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    dependent_vars.append(data_for_year_of_spell[0])
    independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
# export to file
f = open(filepath + "exported_survival_data_w_self_made.csv", 'w')
# write header
s = "dependent_var,constant,age,age_squared,rank,left_censored,self_made,"
for year in years_since_1996_less_initial_and_last:
  s += "d" + year + ","
for duration in range(2, 27+1):
  if 13 <= duration <=15:
    pass
  else:
    s += "d" + str(duration) + ","
s += "d13_to_15" + ","
s += "d28_or_more" + "\n"
f.write(s)
# write data
for i in range(0, len(dependent_vars)):
  s = str(dependent_vars[i])
  for var in independent_vars[i]:
    s += ','
    s += str(var)
  s += '\n'
  f.write(s)
# close file
f.close()

## ESTIMATE SELF−MADE MODEL ##

model = sm.MNLogit(dependent_vars, independent_vars)
fitted_for_self_made = model.fit(disp=False)
#print fitted_for_self_made.summary()

# probs of exits by self made or not
self_made_types = ['self made', 'not self made']
prob_of_exit_by_self_made_type = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(self_made_types, [None for type in self_made_types])) for _ in ['lower', 'point',
    'upper']])) for exit in exits]))
prob_of_exit_diff_btw_self_made_and_not = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(ranks, [None for rank in ranks])) for _ in ['lower', 'point', 'upper']])) for exit in
    exits]))

# set the following to True to re−estimate confidence intervals if they've already been estimated
re_estimate = False
# check if confidence intervals have already been estimated
filenames = [filepath + _ + ".pkl" for _ in ["prob_of_exit_by_self_made_type",
    "prob_of_exit_diff_btw_self_made_and_not"]]
already_estimated = False not in [isfile(filename) for filename in filenames]
```

```
# estimate or re−estimate
if not already_estimated or re_estimate:

  # construct Krinsky−Robb confidence intervals by taking a large number of draws from a
      multivariate normal with means equal to the estimated parameters and var−covar matrix
      equal to the estimated var−covar matrix
  numpy.random.seed(250624)
  redrawn_parameters =
      numpy.random.multivariate_normal(mean=fitted_for_self_made.params.flatten('F'),
      cov=fitted_for_self_made.cov_params(), size=10000)

  # probs of exits for typical person who is or is not philanthropic
  for type in self_made_types:
    X = [1, 64, 4096, 194, 0] # 64 years of age, rank of 194, no left censoring
    if type == 'self made': # self made or not
      X.append(1)
    else:
      X.append(0)
    X.extend([0 for _ in years_since_1996_less_initial_and_last]) # 1996
    X.extend([0 for _ in range(1, 25+1)]) # duration of one year
    n = len(X)
    # point estimates for probs, given estimated parameters
    eXb1 = numpy.exp(numpy.dot(X, fitted_for_self_made.params.flatten('F')[0:n]))
    eXb2 = numpy.exp(numpy.dot(X, fitted_for_self_made.params.flatten('F')[n:]))
    denominator = 1. + eXb1 + eXb2
    prob_of_exit_by_self_made_type['remain wealthy']['point'][type] = 1. / denominator
    prob_of_exit_by_self_made_type['exit by decline']['point'][type] = eXb1 / denominator
    prob_of_exit_by_self_made_type['exit by death']['point'][type] = eXb2 / denominator
    # point estimates for probs, given redrawn parameters
    estimates_for = dict(zip(exits, [[] for exit in exits]))
    for b in redrawn_parameters:
      eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
      eXb2 = numpy.exp(numpy.dot(X, b[n:]))
      denominator = 1. + eXb1 + eXb2
      estimates_for['remain wealthy'].append(1. / denominator)
      estimates_for['exit by decline'].append(eXb1 / denominator)
      estimates_for['exit by death'].append(eXb2 / denominator)
    # get 95% CI
    for exit in exits:
      prob_of_exit_by_self_made_type[exit]['lower'][type] =
          scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[2.5/100.], alphap=1/3.,
          betap=1/3.)[0]
      prob_of_exit_by_self_made_type[exit]['upper'][type] =
          scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[97.5/100.], alphap=1/3.,
          betap=1/3.)[0]

  # 95% CI for differences by rank
  for rank in ranks:
    X = dict(zip(self_made_types, [[1, 64, 4096, rank, 0] for type in self_made_types])) # age of
        64, no left censoring
    # self made dummy
    for type in self_made_types:
```

```
    if type == 'self made':
      X[type].append(1)
    else:
      X[type].append(0)
  # year is 1996
  year_dummies = [0 for _ in years_since_1996_less_initial_and_last]
  for type in self_made_types:
    X[type].extend(year_dummies)
  # duration dummies
  for type in self_made_types:
    X[type].extend([0 for _ in range(1, 25+1)]) # duration of one year
  n = len(X[self_made_types[0]])
  # point estimates for probs, given estimated parameters
  eXb1 = dict(zip(self_made_types, [None for type in self_made_types]))
  eXb2 = dict(zip(self_made_types, [None for type in self_made_types]))
  denominator = dict(zip(self_made_types, [None for type in self_made_types]))
  for type in self_made_types:
    eXb1[type] = numpy.exp(numpy.dot(X[type], fitted_for_self_made.params.flatten('F')[0:n]))
    eXb2[type] = numpy.exp(numpy.dot(X[type], fitted_for_self_made.params.flatten('F')[n:]))
    denominator[type] = 1. + eXb1[type] + eXb2[type]
  prob_of_exit_diff_btw_self_made_and_not['remain wealthy']['point'][rank] = (1. /
      denominator['self made']) - (1. / denominator['not self made'])
  prob_of_exit_diff_btw_self_made_and_not['exit by decline']['point'][rank] = (eXb1['self made'] /
      denominator['self made']) - (eXb1['not self made'] / denominator['not self made'])
  prob_of_exit_diff_btw_self_made_and_not['exit by death']['point'][rank] = (eXb2['self made'] /
      denominator['self made']) - (eXb2['not self made'] / denominator['not self made'])
  # point estimates for probs, given redrawn parameters
  estimates_for = dict(zip(exits, [[] for exit in exits]))
  for b in redrawn_parameters:
    eXb1 = dict(zip(self_made_types, [None for type in self_made_types]))
    eXb2 = dict(zip(self_made_types, [None for type in self_made_types]))
    denominator = dict(zip(self_made_types, [None for type in self_made_types]))
    for type in self_made_types:
      eXb1[type] = numpy.exp(numpy.dot(X[type], b[0:n]))
      eXb2[type] = numpy.exp(numpy.dot(X[type], b[n:]))
      denominator[type] = 1. + eXb1[type] + eXb2[type]
    estimates_for['remain wealthy'].append((1. / denominator['self made']) - (1. /
        denominator['not self made']))
    estimates_for['exit by decline'].append((eXb1['self made'] / denominator['self made']) -
        (eXb1['not self made'] / denominator['not self made']))
    estimates_for['exit by death'].append((eXb2['self made'] / denominator['self made']) -
        (eXb2['not self made'] / denominator['not self made']))
  # get 95% CI
  for exit in exits:
    prob_of_exit_diff_btw_self_made_and_not[exit]['lower'][rank] =
        scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[2.5/100.], alphap=1/3.,
        betap=1/3.)[0]
    prob_of_exit_diff_btw_self_made_and_not[exit]['upper'][rank] =
        scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[97.5/100.], alphap=1/3.,
        betap=1/3.)[0]

# dump results
```

```
    filename = filepath + "prob_of_exit_by_self_made_type.pkl"
    pickle.dump(prob_of_exit_by_self_made_type, open(filename, 'wb'))
    filename = filepath + "prob_of_exit_diff_btw_self_made_and_not.pkl"
    pickle.dump(prob_of_exit_diff_btw_self_made_and_not, open(filename, 'wb'))

else:

    # load results that were dumped using pickle
    prob_of_exit_by_self_made_type = pickle.load(open(filepath +
        "prob_of_exit_by_self_made_type.pkl", 'rb'))
    # load results that were dumped using pickle
    prob_of_exit_diff_btw_self_made_and_not = pickle.load(open(filepath +
        "prob_of_exit_diff_btw_self_made_and_not.pkl", 'rb'))

print "\nself made vs. not, typical person"
for exit in exits:
    for type in self_made_types:
        print exit, "\t", type, "\t", prob_of_exit_by_self_made_type[exit]['lower'][type], "\t",
            prob_of_exit_by_self_made_type[exit]['point'][type], "\t",
            prob_of_exit_by_self_made_type[exit]['upper'][type]

print "\ndifference by rank"
for exit in exits:
    for rank in ranks:
        print exit, "\t", rank, "\t",
        print prob_of_exit_diff_btw_self_made_and_not[exit]['lower'][rank], "\t",
            prob_of_exit_diff_btw_self_made_and_not[exit]['point'][rank], "\t",
            prob_of_exit_diff_btw_self_made_and_not[exit]['upper'][rank]

## A GREAT RECESSION? ##

# The following code relates to the extension to the baseline duration model that includes a
    FIRE industry dummy and interaction terms between that dummy and the year dummies

## GET SURVIVAL DATA FOR FIRE MODEL ##

# no one in the FIRE industry died in either 1999, 2000, or 2012, so we'll just focus on the years
    between 2001 and 2011
years_from_2001_to_2011 = [str(y) for y in range(2001, 2011+1)]
# keep track of names associated with each spell
names_associated_w_spells = list()
# keep track of ages and ranks
ages_as_dependent_vars, ranks_as_dependent_vars = list(), list()
# list of data for spells
data_for_spells = list()
# for each name
for name in names:
    # get spells
    spells = []
    spell = []
    for year in years:
        wealth = data[name]['wealth'][year]
```

```
  if wealth != '':
    spell.append(year)
  else:
    spells.append(spell)
    spell = []
spells.append(spell)
# remove empty spells
while [] in spells:
  spells.remove([])
# for each spell, record data
while len(spells) > 0:
  spell = spells.pop()
  # as long as the spell didn't start in the last year
  if spell == [years[-1]]:
    pass
  else:
    # list of data for spell
    data_for_spell = list()
    # for each year of the spell
    for i in range(0, len(spell)):
      # year
      year = spell[i]
      next_year = str(int(year)+1)
      # if year is not the last year
      if year != years[-1]:
        # data for given year of the spell
        data_for_year_of_spell = list()
        # remains wealthy?
        remains_wealthy = int(i != len(spell)-1)
        # exits by decline in wealth?
        exits_by_decline_in_wealth = int(data[name]['wealth'][next_year] == '')
        # exits by death?
        exits_by_death = int(name in dead_in[next_year])
        # exits by renunciation of citizenship?
        exits_by_renunciation_of_citizenship = int(name in renunciants_in[next_year])
        # record dependent variable
        if exits_by_renunciation_of_citizenship:
          dependent_var = 3 # where 3 = exits by renunciation of citizenship:
        elif exits_by_death:
          dependent_var = 2 # where 2 = exits by death
        elif exits_by_decline_in_wealth:
          dependent_var = 1 # where 1 = exits by decline in wealth
          # would he or she been on the list with the same wealth as last year?
          #if not float(data[name]['wealth'][year]) >= summary_stats['min'][next_year]:
          # print name, "\t", year, "\t", data[name]['wealth'][year], "\t",
              summary_stats['min'][next_year]
        else: # remains wealthy
          dependent_var = 0 # where 0 = remains wealthy
        data_for_year_of_spell.append(dependent_var)
        # independent variables
        # age
        age = int(data[name]['age'][year])
```

```
            age_squared = age * age
            # rank
            rank = summary_stats['wealths'][year].index(float(data[name]['wealth'][year])) + 1
            # left-censored dummy
            left_censored = int('1982' in spell)
            data_for_year_of_spell.extend([1., age, age_squared, rank, left_censored]) # include a
                constant
            # ' dummy and fire-year interaction dummies
            if int(year) > 1995:
                # fire dummy
                industry = data[name]['industry'][year]
                fire_industry = int(industry in ['Finance', 'Investments', 'Real estate'])
                data_for_year_of_spell.append(fire_industry)
                # fire-year interaction dummies
                for _ in years_from_2001_to_2011:
                    data_for_year_of_spell.append(int(fire_industry * int(year == _)))
            else:
                data_for_year_of_spell.append('-999')
                for _ in years_from_2001_to_2011:
                    data_for_year_of_spell.append('-999')
            # dummies for years, but drop 1996
            data_for_year_of_spell.extend([int(year == _) for _ in
                years_since_1996_less_initial_and_last])
            # dummies for durations of given number of years, except when there is a perfect
                prediction problem
            duration = (i + 1)
            for _ in range(2, 27+1):
                if 13 <= _ <= 15:
                    pass
                else:
                    data_for_year_of_spell.append(int(duration == _))
            data_for_year_of_spell.append(int(13 <= duration <= 15)) # 13, 14, or 15 years
            data_for_year_of_spell.append(int(duration > 27)) # 28 or more years
            # add to list
            # ignore spells that end with exit by renunciation of citizenship
            # also, ignore years of spells that are before 1996
            if (data_for_year_of_spell[0] != 3) and (int(year) > 1995):
                data_for_spell.append(data_for_year_of_spell)
                ages_as_dependent_vars.append(age)
                ranks_as_dependent_vars.append(rank)
        # add to list
        if (data_for_year_of_spell[0] != 3) and (int(year) > 1995):
            data_for_spells.append(data_for_spell)
            names_associated_w_spells.append(name)

# No one died after their 14th consecutive year on the list, so we'll make a dummy for 13, 14, or
    15 years

## ORGANIZE FIRE DATA AND EXPORT IT ##

# organize data for estimation
dependent_vars = list()
```

```
independent_vars = list()
for data_for_spell in data_for_spells:
  for data_for_year_of_spell in data_for_spell:
    dependent_vars.append(data_for_year_of_spell[0])
    independent_vars.append(numpy.array(data_for_year_of_spell[1:]))
dependent_vars = numpy.array(dependent_vars)
independent_vars = numpy.array(independent_vars)
# export to file
f = open(filepath + "exported_survival_data_w_fire.csv", 'w')
# write header
s = "dependent_var,constant,age,age_squared,rank,left_censored,fire,"
for year in years_from_2001_to_2011:
  s += "fire" + "d" + year + ","
for year in years_since_1996_less_initial_and_last:
  s += "d" + year + ","
for duration in range(2, 27+1):
  if 13 <= duration <=15:
    pass
  else:
    s += "d" + str(duration) + ","
s += "d13_to_15" + ","
s += "d28_or_more" + "\n"
f.write(s)
# write data
for i in range(0, len(dependent_vars)):
  s = str(dependent_vars[i])
  for var in independent_vars[i]:
    s += ','
    s += str(var)
  s += '\n'
  f.write(s)
# close file
f.close()

## ESTIMATE FIRE MODEL ##

model = sm.MNLogit(dependent_vars, independent_vars)
fitted_for_fire = model.fit(disp=False)
#print fitted_for_fire.summary()

# probs of exits by fire industry or not
industries = ['fire', 'other']
prob_of_exit_by_industry = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(years_from_2001_to_2011, [dict(zip(industries, [None for industry in industries])) for
    year in years_from_2001_to_2011])) for _ in ['lower', 'point', 'upper']])) for exit in exits]))
prob_of_exit_diff_btw_fire_and_other = dict(zip(exits, [dict(zip(['lower', 'point', 'upper'],
    [dict(zip(years_from_2001_to_2011, [None for year in years_from_2001_to_2011])) for _ in
    ['lower', 'point', 'upper']])) for exit in exits]))
# set the following to True to re-estimate confidence intervals if they've already been estimated
re_estimate = False
# check if confidence intervals have already been estimated
```

```
filenames = [filepath + _ + ".pkl" for _ in ["prob_of_exit_by_industry",
    "prob_of_exit_diff_btw_fire_and_other"]]
already_estimated = False not in [isfile(filename) for filename in filenames]
# estimate or re-estimate
if not already_estimated or re_estimate:

  # construct Krinsky-Robb confidence intervals by taking a large number of draws from a
      multivariate normal with means equal to the estimated parameters and var-covar matrix
      equal to the estimated var-covar matrix
  numpy.random.seed(250624)
  redrawn_parameters =
      numpy.random.multivariate_normal(mean=fitted_for_fire.params.flatten('F'),
      cov=fitted_for_fire.cov_params(), size=10000)
  import scipy.stats # for percentile confidence intervals

  # probs of exits for each industry and each year
  for industry in industries:
    for year in years_from_2001_to_2011:
      X = [1, 64, 4096, 194, 0] # 64 years of age, rank of 194, no left censoring
      # fire dummy and fire-year interaction terms
      if industry == 'fire':
        X.append(1)
        fire_year_interaction_terms = [0 for _ in years_from_2001_to_2011]
        if year == '1996':
          pass
        else:
          fire_year_interaction_terms[years_from_2001_to_2011.index(year)] = 1
      else:
        X.append(0)
        fire_year_interaction_terms = [0 for _ in years_from_2001_to_2011]
      X.extend(fire_year_interaction_terms)
      # year dummies
      year_dummies = [0 for _ in years_since_1996_less_initial_and_last]
      if year == '1996':
        pass
      else:
        year_dummies[years_since_1996_less_initial_and_last.index(year)] = 1
      X.extend(year_dummies)
      # duration dummies
      X.extend([0 for _ in range(1, 25+1)]) # duration of one year
      n = len(X)
      # point estimates for probs, given estimated parameters
      eXb1 = numpy.exp(numpy.dot(X, fitted_for_fire.params.flatten('F')[0:n]))
      eXb2 = numpy.exp(numpy.dot(X, fitted_for_fire.params.flatten('F')[n:]))
      denominator = 1. + eXb1 + eXb2
      prob_of_exit_by_industry['remain wealthy']['point'][year][industry] = 1. / denominator
      prob_of_exit_by_industry['exit by decline']['point'][year][industry] = eXb1 / denominator
      prob_of_exit_by_industry['exit by death']['point'][year][industry] = eXb2 / denominator
      # point estimates for probs, given redrawn parameters
      estimates_for = dict(zip(exits, [[] for exit in exits]))
      for b in redrawn_parameters:
        eXb1 = numpy.exp(numpy.dot(X, b[0:n]))
```

```
            eXb2 = numpy.exp(numpy.dot(X, b[n:]))
            denominator = 1. + eXb1 + eXb2
            estimates_for['remain wealthy'].append(1. / denominator)
            estimates_for['exit by decline'].append(eXb1 / denominator)
            estimates_for['exit by death'].append(eXb2 / denominator)
        # get 95% CI
        for exit in exits:
            prob_of_exit_by_industry[exit]['lower'][year][industry] =
                scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[2.5/100.], alphap=1/3.,
                betap=1/3.)[0]
            prob_of_exit_by_industry[exit]['upper'][year][industry] =
                scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[97.5/100.], alphap=1/3.,
                betap=1/3.)[0]

# 95% CI for differences by year
for year in years_from_2001_to_2011:
    X = dict(zip(industries, [[1, 64, 4096, 194, 0] for industry in industries])) # 64 years of age,
        rank of 194, no left censoring
    # fire dummy and fire-year interaction terms
    for industry in industries:
        if industry == 'fire':
            X[industry].append(1)
            fire_year_interaction_terms = [0 for _ in years_from_2001_to_2011]
            if year == '1996':
                pass
            else:
                fire_year_interaction_terms[years_from_2001_to_2011.index(year)] = 1
        else:
            X[industry].append(0)
            fire_year_interaction_terms = [0 for _ in years_from_2001_to_2011]
        X[industry].extend(fire_year_interaction_terms)
    # year dummies
    year_dummies = [0 for _ in years_since_1996_less_initial_and_last]
    if year == '1996':
        pass
    else:
        year_dummies[years_since_1996_less_initial_and_last.index(year)] = 1
    for industry in industries:
        X[industry].extend(year_dummies)
    # duration dummies
    for industry in industries:
        X[industry].extend([0 for _ in range(1, 25+1)]) # duration of one year
    n = len(X[industries[0]])
    # point estimates for probs, given estimated parameters
    eXb1 = dict(zip(industries, [None for industry in industries]))
    eXb2 = dict(zip(industries, [None for industry in industries]))
    denominator = dict(zip(industries, [None for industry in industries]))
    for industry in industries:
        eXb1[industry] = numpy.exp(numpy.dot(X[industry],
            fitted_for_fire.params.flatten('F')[0:n]))
        eXb2[industry] = numpy.exp(numpy.dot(X[industry], fitted_for_fire.params.flatten('F')[n:]))
        denominator[industry] = 1. + eXb1[industry] + eXb2[industry]
```

```
        prob_of_exit_diff_btw_fire_and_other[‘remain wealthy’][‘point’][year] = (1. /
              denominator[‘fire’]) − (1. / denominator[‘other’])
        prob_of_exit_diff_btw_fire_and_other[‘exit by decline’][‘point’][year] = (eXb1[‘fire’] /
              denominator[‘fire’]) − (eXb1[‘other’] / denominator[‘other’])
        prob_of_exit_diff_btw_fire_and_other[‘exit by death’][‘point’][year] = (eXb2[‘fire’] /
              denominator[‘fire’]) − (eXb2[‘other’] / denominator[‘other’])
        # point estimates for probs, given redrawn parameters
        estimates_for = dict(zip(exits, [[] for exit in exits]))
        for b in redrawn_parameters:
            eXb1 = dict(zip(industries, [None for industry in industries]))
            eXb2 = dict(zip(industries, [None for industry in industries]))
            denominator = dict(zip(industries, [None for industry in industries]))
            for industry in industries:
                eXb1[industry] = numpy.exp(numpy.dot(X[industry], b[0:n]))
                eXb2[industry] = numpy.exp(numpy.dot(X[industry], b[n:]))
                denominator[industry] = 1. + eXb1[industry] + eXb2[industry]
            estimates_for[‘remain wealthy’].append((1. / denominator[‘fire’]) − (1. /
                  denominator[‘other’]))
            estimates_for[‘exit by decline’].append((eXb1[‘fire’] / denominator[‘fire’]) − (eXb1[‘other’]
                  / denominator[‘other’]))
            estimates_for[‘exit by death’].append((eXb2[‘fire’] / denominator[‘fire’]) − (eXb2[‘other’] /
                  denominator[‘other’]))
        # get 95% CI
        for exit in exits:
            prob_of_exit_diff_btw_fire_and_other[exit][‘lower’][year] =
                  scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[2.5/100.], alphap=1/3.,
                  betap=1/3.)[0]
            prob_of_exit_diff_btw_fire_and_other[exit][‘upper’][year] =
                  scipy.stats.mstats.mquantiles(estimates_for[exit], prob=[97.5/100.], alphap=1/3.,
                  betap=1/3.)[0]

    # dump results
    filename = filepath + “prob_of_exit_by_industry.pkl”
    pickle.dump(prob_of_exit_by_industry, open(filename, ‘wb’))
    filename = filepath + “prob_of_exit_diff_btw_fire_and_other.pkl”
    pickle.dump(prob_of_exit_diff_btw_fire_and_other, open(filename, ‘wb’))

else:

    # load results that were dumped using pickle
    prob_of_exit_by_industry = pickle.load(open(filepath + “prob_of_exit_by_industry.pkl”, ‘rb’))
    # load results that were dumped using pickle
    prob_of_exit_diff_btw_fire_and_other = pickle.load(open(filepath +
          “prob_of_exit_diff_btw_fire_and_other.pkl”, ‘rb’))

# print results
print “\n\t\t”, “fire”, “\t\t\t”, “other”, “\t\t\t”, “difference”
for exit in exits:
    for year in years_from_2001_to_2011:
        print exit, “\t”, year, “\t”, prob_of_exit_by_industry[exit][‘lower’][year][‘fire’], “\t”,
              prob_of_exit_by_industry[exit][‘point’][year][‘fire’], “\t”,
              prob_of_exit_by_industry[exit][‘upper’][year][‘fire’], “\t”,
```

```
print prob_of_exit_by_industry[exit]['lower'][year]['other'], "\t",
      prob_of_exit_by_industry[exit]['point'][year]['other'], "\t",
      prob_of_exit_by_industry[exit]['upper'][year]['other'], "\t",
print prob_of_exit_diff_btw_fire_and_other[exit]['lower'][year], "\t",
      prob_of_exit_diff_btw_fire_and_other[exit]['point'][year], "\t",
      prob_of_exit_diff_btw_fire_and_other[exit]['upper'][year]
```

REFERENCES

Aban, Inmaculada, Mark Meerschaert, and Anna Panorska (2006). "Parameter Estimation for the Truncated Pareto Distribution." *Journal of the American Statistical Association* 101(473), 270–77.

Adams, Susan (2014). "The Americans Who Gave Away the Most Money in 2013." *Forbes Magazine*, forbes.com/sites/susanadams/2014/02/10/the-americans-who-gave-away-the-most-money-in-2013/ (accessed May 11, 2014).

Alchian, Armen and Benjamin Klein (1973). "On a Correct Measure of Inflation." *Journal of Money, Credit, and Banking* 5(1), 173–91.

Allegretto, Sylvia (2011). "The Few, the Proud, the Very Rich." *The Berkeley Blog: Topical Questions, Campus Experts, and Public Opinion from UC Berkeley*. blogs.berkeley.edu/2011/12/05/the-few-the-proud-the-very-rich/ (accessed May 11, 2014).

Allen, Jodie (1996). "The 1996 Slate 60: Introduction." *Slate*, slate.com/articles/life/the_slate60/1996/12/the_1996_slate_60_2.html (accessed May 11, 2014).

Allison, Paul (1982). "Discrete-time Methods for the Analysis of Event Histories." *Sociological Methodology* 13(1), 61–98.

Alvaredo, Facundo (2011). "A Note on the Relationship Between Top Income Shares and the Gini Coefficient." *Economics Letters* 110(3), 274–77.

Andreoni, James (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100(401), 464–77.

Arnold, Barry (1983). *Pareto Distributions.* Fairland, MD: International Cooperative Publishing House.

Ashworth, Karl, Martha Hill, and Robert Walker (1994). "Patterns of Childhood Poverty: New Challenges for Policy." *Journal of Policy Analysis and Management* 13(4), 658–80.

Associated Press (1986). "Randal Tyson, Executive in Family Food Company, Dies at 34." September 19.

Associated Press (2005). "John T. Walton, 58, Donor and Heir to Wal-Mart Fortune, Dies." June 29.

Atkinson, Anthony (2008). "Concentration Among the Rich." In James Davies, ed., *Personal Wealth from a Global Perspective*, pp. 64–90. Oxford, UK: Oxford University Press.

Atkinson, Anthony and Allan Harrison (1978). *Distribution of Personal Wealth in Britain*. Cambridge, UK: Cambridge University Press.

Backhouse, Roger and Steven Medema (2009). "On the Definition of Economics." *The Journal of Economic Perspectives* 23(1), 221–34.

Barsky, Neil (1990). "Negative Net Worth of $291 Million is Shown for Trump." *The Wall Street Journal*. August 16.

Basu, Debabrata ([1971] 2011). "An Essay on the Logical Foundations of Survey Sampling: Part One." In Anirban DasGupta, ed., *Selected Works of Debabrata Basu*, pp. 167–206. New York, NY: Springer.

Bernstein, Peter and Annalyn Swan (2007). *All the Money in the World: How the Forbes Four Hundred Make and Spend Their Fortunes*. New York, NY: Alfred K. Knopf.

Bertaut, Carol and Martha Starr (2002). "Household Portfolios in the United States." In Luigi Guiso, Michael Haliassos, and Tullio Jappelli, eds., *Household Portfolios*, pp. 181–218. Cambridge, MA: MIT Press.

Bivens, Josh (2012). "Inequality, Exhibit A: Walmart and the Wealth of American Families." *Working Economics: The Economic Policy Institute Blog*. epi.org/blog/inequality-exhibit-wal-mart-wealth-american/ (accessed May 11, 2014).

Blitz, Rudolph C. and John J. Siegfried (1992). "How Did the Wealthiest Americans Get So Rich?" *Quarterly Review of Economics and Finance* 32(1), 5–26.

Brazauskas, Vytaras and Robert Serfling (2000). "Robust Estimation of Tail Parameters for Two-Parameter Pareto and Exponential Models via Generalized Quantile Statistics." *Extremes* 3(3), 231–49.

Bricker, Jesse, Brian Bucks, Arthur Kennickell, Traci Mach, and Kevin Moore (2011). "Drowning or Weathering the Storm? Changes in Family Finances from 2007 to 2009." Working Paper 16985, National Bureau of Economic Research.

Bricker, Jesse, Arthur Kennickell, Kevin Moore, and John Sabe (2012). "Changes in US Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances." *Federal Reserve Bulletin* 98(2), 1–80.

Broom, Leonard and William Shay (2000). "Discontinuities in the Distribution of Great Wealth: Sectoral Forces Old and New." Working Paper 308, Levy Economics Institute of Bard College, Annandale-on-Hudson, NY.

Brown, Heidi (2001). "Dynamism Hurts." *Forbes Magazine*, July 9. forbes.com/global/2001/0709/030tab1.html (accessed May 11, 2014).

Brown, Jeffrey R., Jeffrey B. Liebman, and Joshua Pollet (2002). "Estimating Life Tables That Reflect Socioeconomic Differences in Mortality." In Martin Feldstein and Jeffrey Liebman, eds., *The Distributional Aspects of Social Security and Social Security Reform*, pp. 447–57. Chicago, IL: University of Chicago Press.

Brzezinski, Michal (2012). "Wealth Distributions and Power Laws: Evidence from 'Rich Lists.'" Paper presented at the Conference of the International Association for Research in Income and Wealth, Boston, MA, August 5–11. iariw.org/papers/2012/BrzezinskiPaper.pdf (accessed May 11, 2014).

Bureau of Labor Statistics (2013). "Updated CPI-U-RS, All items, 1978–2013." bls.gov/cpi/cpirsai1978-2013.pdf (accessed May 11, 2014).

Burris, Val (2000). "The Myth of Old Money Liberalism: The Politics of the Forbes Four Hundred Richest Americans." *Social Problems* 47(3), 360–78.

Cagetti, Marco and Mariacristina De Nardi (2006). "Entrepreneurship, Frictions, and Wealth." *Journal of Political Economy* 114(5), 835–70.

Cagetti, Marco and Mariacristina De Nardi (2008). "Wealth Inequality: Data and Models." *Macroeconomic Dynamics* 12(S2), 285–313.

Cameron, Colin and Pravin Trivedi (2005). *Microeconometrics: Methods And Applications*. Cambridge, UK: Cambridge University Press.

Canterbery, E. Ray and E. Joe Nosari (1985). "The Forbes Four Hundred: The Determinants of Super-Wealth." *Southern Economic Journal* 51(4), 1073–83.

Capehart, Kevin W. (2014). "Is the Wealth of the Worlds Billionaires Not Paretian?" *Physica A: Statistical Mechanics and Its Applications* 395, 255–60.

Capehart, Kevin W. (forthcoming). "Hyman Minsky's Interpretation of Donald Trump." *Journal of Post Keynesian Economics*.

Carroll, Christopher (2000). "Why Do the Rich Save So Much?" In Joel Slemrod, ed., *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, pp. 465–84. Boston, MA: Harvard University Press.

Castaldi, Carolina and Mishael Milakovic (2007). "Turnover Activity in Wealth Portfolios." *Journal of Economic Behavior and Organization* 63(3), 537–52.

Cheng, Simon and J. Scott Long (2007). "Testing for IIA in the Multinomial Logit Model." *Sociological Methods and Research* 35(4), 583–600.

Choi, Young Back (1999). "On the Rich Getting Richer and the Poor Getting Poorer." *Kyklos* 52(2), 239–58.

Choi, Young Back (2002). "Misunderstanding Distribution." *Social Philosophy and Policy* 19(1), 110–39.

Clark, Charles (2002). "Wealth and Poverty: On the Social Creation of Scarcity." *Journal of Economic Issues* 36(2), 415–21.

Clauset, Aaron, Cosma Rohilla Shalizi, and M.E.J. Newman (2009). "Power-Law Distributions in Empirical Data." *SIAM Review* 51(4), 661–703.

Coffey, Brendan (2011). "CEOs Who Make One Dollar (Or Less) A Year." *Forbes Magazine*, forbes.com/sites/brendancoffey/2011/05/16/ceos-who-make-one-dollar-or-less-a-year/ (accessed May 11, 2014).

Coffey, Brendan and Zohair Siraj (2013). "Hidden Billionaire Cohen Hauls Fortune in Unmarked Trucks." *Bloomberg*, bloomberg.com/news/2013-08-05/hidden-billionaire-cohen-hauls-fortune-in-unmarked-trucks.html (accessed May 11, 2014).

Cohen, A. Clifford (1991). *Truncated and Censored Samples: Theory and Applications*. New York, NY: Marcel Dekker.

Cowell, Frank (2000). "Measurement of Inequality." In Anthony Atkinson and Francois Bourguignon, eds., *Handbook of Income Distribution*, Volume 1, pp. 87–166. Amsterdam, Netherlands: Elsevier.

Davies, James (2009). "Wealth and Economic Inequality." In Wiemer Salverda, Brian Nolan, and Timothy Smeeding, eds., *Oxford Handbook of Economic Inequality*, pp. 127–49. Oxford, UK: Oxford University Press.

Davies, James, Susanna Sandstrom, Anthony Shorrocks, and Edward Wolff (2008). "The World Distribution of Household Wealth." In James Davies, ed., *Personal Wealth from a Global Perspective*, pp. 395–418. Oxford, UK: Oxford University Press.

Davies, James and Anthony Shorrocks (2000). "The Distribution of Wealth." In Anthony Atkinson and Francois Bourguignon, eds., *Handbook of Income Distribution*, Volume 1, pp. 605–75. Amsterdam, Netherlands: Elsevier.

de Jong, David (2012). "America's Fourth-Richest Woman Unveiled With Koch Stake." *Bloomberg*, bloomberg.com/news/2012-09-16/america-s-fourth-richest-woman-unveiled-with-koch-stake.html (accessed May 11, 2014).

Di Mento, Maria (2014). "How The Chronicle Compiled the Philanthropy 50 List." *The Chronicle of Philanthropy*, philanthropy.com/article/How-The-Chronicle-Compiled-the/144563/ (accessed May 11, 2014).

Diaz-Gimenez, Javier, Andy Glover, and Jose-Vctor Rios-Rull (2011). "Facts on the Distributions of Earnings, Income, and Wealth in the United States: 2007 Update." *Federal Reserve Bank of Minneapolis Quarterly Review* 34(1), 2–31.

Diener, Ed, Jeff Horwitz, and Robert Emmons (1985). "Happiness of the Very Wealthy." *Social Indicators Research* 16(3), 263–74.

Edlund, Lena and Wojciech Kopczuk (2009). "Women, Wealth, and Mobility." *American Economic Review* 99(1), 146–78.

Einstein, David (2000). "What Bill Gates Is Really Worth." *Forbes Magazine*, forbes.com/2000/09/28/0929forbes400.html (accessed May 11, 2014).

Essock-Vitale, Susan (1984). "The Reproductive Success of Wealthy Americans." *Ethology and Sociobiology* 5(1), 45–49.

Esteve-Perez, Silviano, Francisco Requena-Silvente, and Vicente Pallardo-Lopez (2013). "The Duration of Firm-Destination Export Relationships: Evidence from Spain, 1997–2006." *Economic Inquiry* 51(1), 159–80.

Fine, Ben (2007). "Eleven Hypotheses on the Conceptual History of Social Capital: A Response to James Farr." *Political Theory* 35(1), 47–53.

Fine, Ben (2013). "Economics: Unfit for Purpose; the Director's Cut." Department of Economics Working Paper 176, School of Oriental and African Studies.

Fitch, Stephane (2006). "What Is Trump Worth?" *Forbes Magazine*, forbes.com/2006/09/20/trump-wealth-biz_06rich400_0921trump.html (accessed May 11, 2014).

Flying Magazine (2001). "Atlas Air CEO and Reporter Killed in Crash." April.

Forbes, Steve and Elizabeth Ames ([2009] 2011). *How Capitalism Will Save Us: Why Free People and Free Markets Are the Best Answer in Today's Economy; Updated and Revised Edition*. New York, NY: Crown Business.

Forbes Magazine (1918). "America's 30 Richest: Who They Are and How Much They Own." March 2.

Forbes Magazine (1982). Forbes Four Hundred. September 13.

Forbes Magazine (1983). Forbes Four Hundred. Fall.

Forbes Magazine (1984). Forbes Four Hundred. October 1.

Forbes Magazine (1985). Forbes Four Hundred. October 28.

Forbes Magazine (1986). Forbes Four Hundred. October 27.

Forbes Magazine (1987). Forbes Four Hundred. October 26.

Forbes Magazine (1988). Forbes Four Hundred. October 24.

Forbes Magazine (1989). Forbes Four Hundred. October 23.

Forbes Magazine (1990). Forbes Four Hundred. October 22.

Forbes Magazine (1991). Forbes Four Hundred. October 21.

Forbes Magazine (1992). Forbes Four Hundred. October 19.

Forbes Magazine (1993). Forbes Four Hundred. October 18.

Forbes Magazine (1994). Forbes Four Hundred. October 17.

Forbes Magazine (1995). Forbes Four Hundred. October 16.

Forbes Magazine (1996). Forbes Four Hundred. October 14.

Forbes Magazine (1997). Forbes Four Hundred. October 13.

Forbes Magazine (1998). Forbes Four Hundred. October 12.

Forbes Magazine (1999). Forbes Four Hundred. October 11.

Forbes Magazine (2000). Forbes Four Hundred. October 9.

Forbes Magazine (2001). Forbes Four Hundred. October 8.

Forbes Magazine (2002). Forbes Four Hundred. September 30.

Forbes Magazine (2003). Forbes Four Hundred. October 6.

Forbes Magazine (2004). Forbes Four Hundred. October 11.

Forbes Magazine (2005). Forbes Four Hundred. October 10.

Forbes Magazine (2006). Forbes Four Hundred. October 9.

Forbes Magazine (2007). Forbes Four Hundred. October 8.

Forbes Magazine (2008). Forbes Four Hundred. October 6.

Forbes Magazine (2009). Forbes Four Hundred. October 19.

Forbes Magazine (2010). Forbes Four Hundred. October 11.

Forbes Magazine (2011). Forbes Four Hundred. October 10.

Forbes Magazine (2012). Forbes Four Hundred. October 8.

Forbes Magazine (2013). Forbes Four Hundred. October 7.

Fortune Magazine (1957). "The Fifty-Million-Dollar Man." November.

Fortune Magazine (1968). "America's Centimillionaires." May.

Foster, John Bellamy and Hannah Holleman (2010). "The Financial Power Elite." *Monthly Review* 61(1), 1–20.

Francis, Johanna (2007). "Saving, Investment, and the Entrepreneurship Decisions of the Wealthy." PhD diss., Johns Hopkins University.

Gabaix, Xavier and Rustam Ibragimov (2011). "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents." *Journal of Business and Economic Statistics* 29(1), 24–39.

Galbraith, John Kenneth (1977). *The Age of Uncertainity.* London, UK: British Broadcasting Corporation.

Gelman, Andrew and Hal Stern (2006). "The Difference Between 'Significant' and 'Not Significant' Is Not Itself Statistically Significant." *The American Statistician* 60(4), 328–331.

Goldstein, Michel, Steven Morris, and Gary Yen (2004). "Problems with Fitting to the Power-law Distribution." *European Physical Journal B: Condensed Matter* 41(2), 255–8.

Gompertz, Benjamin (1825). "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies." *Philosophical Transactions of the Royal Society of London* 115, 513–83.

Graetz, Michael (2011). "The Lloyd Leva Plaine Distinguished Lecture: The Politics and Policy of the Estate Tax; Past, Present, and Future." Paper 3810, Yale Law School Faculty Scholarship Series.

Grimm, Michael (2007). "Removing the Anonymity Axiom in Assessing Pro-Poor Growth." *Journal of Economic Inequality* 5(2), 179–97.

Hardy, Michael (2010). "Pareto's Law." *Mathematical Intelligencer* 32(3), 38–43.

Hausman, Jerry and Daniel McFadden (1984). "Specification Tests for the Multinomial Logit Model." *Econometrica* 52(5), 1219–40.

Heligman, Larry and John Pollard (1980). "The Age Pattern of Mortality." *Journal of the Institute of Actuaries* 107(1), 49–80.

Hertz, Thomas (2008). "Mobility Among the Hyper-Rich: Results from the Forbes Four Hundred." Presentation at the Century Foundation Conference on Billionaires and Their Impact. old.tcf.org/events/pdfs/ev227/HERTZMobilityAmongtheHyper-Rich.pdf (accessed May 11, 2014).

Hollander, Myles and Douglas Wolfe (1999). *Nonparametric Statistical Methods*. New York, NY: John Wiley & Sons.

Horvitz, D.G. and D.J. Thompson (1952). "A Generalization of Sampling Without Replacement From a Finite Universe." *Journal of the American Statistical Association* 47(260), 663–85.

Hurst, Erik, Ming Ching Luoh, Frank Stafford, and William Gale (1998). "The Wealth Dynamics of American Families, 1984–1994." *Brookings Papers on Economic Activity* 1998(1), 267–337.

Internal Revenue Service (2003). "The 400 Individual Income Tax Returns Reporting the Highest Adjusted Gross Incomes Each Year, 1992–2000." *Statistics of Income Bulletin* Spring, 7–9.

Internal Revenue Service (2012). "The 400 Individual Income Tax Returns Reporting the Highest Adjusted Gross Incomes Each Year, 1992–2009. irs.gov/file_source/pub/irs-soi/09intop400.pdf (accessed November 11, 2012).

Jenkins, Stephen (1995). "Easy Estimation Methods for Discrete-Time Duration Models." *Oxford Bulletin of Economics and Statistics* 57(1), 129–36.

Juster, F. Thomas, James Smith, and Frank Stafford (1999). "The Measurement and Structure of Household Wealth." *Labour Economics* 6(2), 253–75.

Kaplan, Steven and Joshua Rauh (2013). "It's the Market: The Broad-Based Rise in the Return to Top Talent." *Journal of Economic Perspectives* 27(3), 35–55.

Keister, Lisa (2005). *Getting Rich: America's New Rich and How They Got that Way*. Cambridge, UK: Cambridge University Press.

Kennickell, Arthur (1999). "Using Income Data to Predict Wealth." Federal Reserve Board, Washington, DC. federalreserve.gov/econresdata/scf/files/wealthincome7web.pdf (accessed May 11, 2014).

Kennickell, Arthur (2001). "Modeling Wealth with Multiple Observations of Income: Redesign of the Sample for the 2001 Survey of Consumer Finances." Federal Reserve Board, Washington, DC. federalreserve.gov/econresdata/scf/files/scf2001list.sampleredesign9.pdf (accessed May 11, 2014).

Kennickell, Arthur (2003). "A Rolling Tide: Changes in the Distribution of Wealth in the US, 1989–2001. Paper presented at the Conference on International Perspectives on Household Wealth, Levy Economics Institute of Bard College, Annandale-on-Hudson, NY, October 17–18. Audio recording available at levyinstitute.org/publications/?docid=677 (accessed May 11, 2014).

Kennickell, Arthur (2006). "A Rolling Tide: Changes in the Distribution of Wealth in the US, 1989–2001." In Edward Wolff, ed., *International Perspectives on Household Wealth*, pp. 19–88. Northampton, MA: Edward Elgar.

Kennickell, Arthur (2007). "The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances." Federal Reserve Board, Washington, DC. federalreserve.gov/econresdata/scf/files/isi2007.pdf (accessed May 11, 2014).

Kennickell, Arthur (2011). "Tossed and Turned: Wealth Dynamics of US Households, 2007–2009." *Finance and Economics Discussion Series*, Working Paper 2011-51, Revised May 29, 2012, Federal Reserve Board, Washington, DC.

Kennickell, Arthur (2012). "The Other, Other Half: Changes in the Finances of the Least Wealthy 50 Percent, 2007–2009." *Finance and Economics Discussion Series*, Working Paper 2012-40, Federal Reserve Board, Washington, DC.

Kennickell, Arthur and Martha Starr (1994). "Changes in Family Finances from 1989 to 1992: Evidence from the Survey of Consumer Finances." *Federal Reserve Bulletin* 80(October), 861–82.

Kennickell, Arthur and Martha Starr (1997). "Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel." *Review of Income and Wealth* 43(4), 381–99.

Kennickell, Arthur, Martha Starr, and Brian Surette (2000). "Recent Changes in US Family Finances: Results from the 1998 Survey of Consumer Finances." *Federal Reserve Bulletin* 86(January), 1–29.

Kilachand, Sean (2012). "The Forbes Four Hundred Hall Of Fame: 36 Members of Our Debut Issue Still in Ranks." *Forbes Magazine*, forbes.com/sites/seankilachand/2012/09/20/the-forbes-400-hall-of-fame-36-members-of-our-debut-issue-still-in-ranks/ (accessed May 11, 2014).

Klass, Oren, Ofer Biham, Moshe Levy, Ofer Malcai, and Sorin Solomon (2006). "The Forbes Four Hundred and the Pareto Wealth Distribution." *Economics Letters* 90(2), 290–95.

Kondlo, Lwando Orbet (2010). "Estimation of Pareto Distribution Functions from Samples Contaminated by Measurement Errors." PhD diss., University of the Western Cape.

Kopczuk, Wojciech and Emmanuel Saez (2004a). "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate-Tax Returns." *National Tax Journal* 57(2), 445–87.

Kopczuk, Wojciech and Emmanuel Saez (2004b). "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate-Tax Returns." Working Paper 10399, National Bureau of Economic Research.

Krinsky, Itzhak and A. Leslie Robb (1986). "On Approximating the Statistical Properties of Elasticities." *Review of Economics and Statistics* 68(4), 715–19.

Kroll, Luisa (2012). "The Forbes Four Hundred: The Richest People in America." *Forbes Magazine*, forbes.com/sites/luisakroll/2012/09/19/the-forbes-400-the-richest-people-in-america/ (accessed May 11, 2014).

Kroll, Luisa (2014). "Inside The 2014 Billionaires List: Facts and Figures." *Forbes Magazine*, forbes.com/sites/luisakroll/2014/03/03/inside-the-2014-forbes-billionaires-list-facts-and-figures/ (accessed May 11, 2014).

Krugman, Paul (2010). "Inequality and Crises: Coincidence or Causality? Keynote address at conference on Inequality and the Status of the Middle Class, University of Luxembourg, June 28–30.

Lampman, Robert (1962). *The Share of Top Wealth-Holders in National Wealth.* Princeton, NJ: Princeton University Press.

Lauer, Charlotte (2005). *Education and Labor Market Outcomes: A French-German Comparison.* Heidelberg, Germany: Physica-Verlag.

Levy, Moshe and Sorin Solomon (1997). "New Evidence for the Power-Law Distribution of Wealth." *Physica A: Statistical Mechanics and Its Applications* 242(1–2), 90–94.

Lewis, Mark (2002). "The Founder's Club." *Forbes Magazine*, forbes.com/2002/09/13/400redistribution.html (accessed May 11, 2014).

Lewis, Nicole (2014). "Giving Pledge Signers Gave Big in 2013 but Not Much for Today's Needs." *The Chronicle of Philanthropy*, philanthropy.com/article/Signers-of-the-Giving-Pledge/144567/ (accessed May 11, 2014).

Lilliefors, Hubert (1969). "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown." *Journal of the American Statistical Association* 64(325), 387–9.

Los Angeles Times (1993). "Sharon Lund; Daughter of Walt Disney." February 17.

Makeham, William (1860). "On the Law of Mortality and the Construction of Annuity Tables." *Assurance Magazine: The Journal of the Institute of Actuaries* 8(6), 301–10.

Maschberger, Thomas and Pavel Kroupa (2009). "Estimators for the Exponent and Upper limit, and Goodness-of-Fit Tests for (Truncated) Power-law Distributions." *Monthly Notices of the Royal Astronomical Society* 395(2), 931–42.

McCubbin, Janet (1994). "Improving Wealth Estimates Derived From Estate Tax Data." In Barry Johnson, ed., *Compendium of Federal Estate Tax and Personal Wealth Studies*, pp. 363–69. Washington, DC: International Revenue Service.

Miller, Leslie (1956). "Table of Percentage Points of Kolmogorov Statistics." *Journal of the American Statistical Association* 51(273), 111–21.

Miller, Matthew and Peter Newcomb (2012). "Slim Beats Gates in Global Daily Ranking of Billionaires." *Bloomberg*, bloomberg.com/news/2012-03-05/slim-beats-gates-in-global-daily-ranking-of-billionaires.html (accessed May 11, 2014).

Mishel, Lawrence, Josh Bivens, Elise Gould, and Heidi Shierholz (2012). *The State of Working America* (12th ed.). Washington, DC: Economic Policy Institute.

Navarro, Andrea (2013). "American Dreams: Foreign-Born Members of the Forbes 400." *Forbes Magazine*, forbes.com/sites/andreanavarro/2013/09/18/american-dreams-foreign-born-members-of-the-forbes-400/ (accessed May 11, 2014).

Neumayer, Eric (2004). "The Super-Rich in Global Perspective: A Quantitative Analysis of the *Forbes* List of Billionaires." *Applied Economics Letters* 11(13), 793–6.

New York Times (1990a). "Robert H. Lurie: Teams' Owner, 48." June 23.

New York Times (1990b). "Malcolm Forbes, Publisher, Dies at 70." February 26.

New York Times (1991). "John Heinz, 52, Heir to a Fortune And Senator From Pennsylvania." April 5.

New York Times (1994). "Robert Miner, 52, Software Executive." November 17.

New York Times (2002). "Harry V. Quadracci, 66, Chief Of Magazine Printing Company." July 31.

New York Times (2006). "William B. Ziff Jr., 76, Builder of Magazine Empire, Dies." September 12.

New York Times (2011). "Apple's Visionary Redefined Digital Age." October 5.

Officer, Lawrence and Samuel Williamson (2006). "Better Measurements of Worth." *Challenge: The Magazine of Economic Affairs* 49(4), 86–110.

Ogwang, Tomson (2013). "Is the Wealth of the World's Billionaires Paretian?" *Physica A: Statistical Mechanics and Its Applications* 392(4), 757–62.

Olshansky, S. Jay and Bruce Carnes (1997). "Ever Since Gompertz." *Demography* 34(1), 1–15.

Panel Study of Income Dynamics (2007). Supplemental Wealth File. Produced and distributed by the Institute for Social Research, Survey Research Center, University of Michigan, Ann Arbor, MI. simba.isr.umich.edu/Zips/zipSupp.aspx (accessed May 11, 2014).

Pareto, Vilfredo ([1896] 2001). "The Curve of the Distribution of Wealth." In Mario Baldassarri and Pierluigi Ciocca, eds., *Roots of the Italian School of Economics and Finance: From Ferrara (1857) to Einaudi (1944); Volume 2*, pp. 231–46. Houndmills, UK: Palgrave. Translation of an extract from Volume 2 of Pareto's *Cours d'Economie Politique*.

Peichl, Andreas, Thilo Schaefer, and Christoph Scheicher (2010). "Measuring Richness and Poverty: A Micro Data Application to Europe and Germany." *Review of Income and Wealth* 56(3), 597–619.

Pence, Karen (2006). "The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving." *B.E. Journals in Economic Analysis and Policy: Contributions to Economic Analysis and Policy* 5(1), 1–24.

Persky, Joseph (1992). "Retrospectives: Pareto's Law." *Journal of Economic Perspectives* 6(2), 181–92.

Petras, James and Christian Davenport (1990). "The Changing Wealth of the US Ruling Class." *Monthly Review* 42(7), 33–38.

Piketty, Thomas (2014). *Capital in the Twenty-First Century*. Trans. Arthur Goldhammer. Cambridge, MA: Belknap Press.

Piketty, Thomas and Emmanuel Saez (2003). "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics* 118(1), 1–39.

Pires, Ana and Conceicao Amado (2008). "Interval Estimators for a Binomial Proportion: Comparison of 20 Methods." *REVSTAT: Statistical Journal* 6(2), 165–97.

Plotz, David (2006). "Competitive Philanthropy: The History of the Slate 60." *Slate*, slate.com/articles/life/the_slate60/2006/02/competitive_philanthropy.html (accessed May 11, 2014).

Polanyi, George and John Wood (1974). "How Unevenly is Wealth Spread Today?" In William Letwin, ed., *Against Equality: Readings on Economic and Social Policy*, pp. 229–60. London, UK: Macmillan Press.

Potts, Jason (2006). "How Creative Are the Super-Rich?" *Agenda: A Journal of Policy Analysis and Reform* 13(4), 339–50.

Rae, Matthew and Robert DeHaan (2012). "'Listening to Other People': Arthur Kennickell; Leadership on the Survey of Consumer Finances." *Profiles in Statistical Courage.* statisticalcourage.info/downloads%5CArthur%20Kennickel.pdf (accessed May 11, 2014).

Rahman, Mezbahur and Larry Pearson (2003). "A Note on Estimating Parameters in Two-Parameter Pareto Distributions." *International Journal of Mathematical Education in Science and Technology* 34(2), 298–306.

Raskall, P. and G. Matheson (1992). "Understanding the Gini Coefficient." *Social Policy Research Centre Newsletter* 46, 11.

Raub, Brian, Barry Johnson, and Joe Newcomb (2010). "A Comparison of Wealth Estimates for America's Wealthiest Decedents Using Tax Data and Data from the Forbes Four Hundred." Paper presented at the Annual Meeting of the National Tax Association, Chicago, IL, November 20. piketty.pse.ens.fr/files/RaubJohnsonNewcomb2010.pdf (accessed May 11, 2014).

Sanandaji, Tino (2012). "The International Mobility of the Super-Rich." Working Paper 904, Research Institute of Industrial Economics (IFN).

Sanandaji, Tino and Peter Leeson (2013). "Billionaires." *Industrial and Corporate Change* 22(1), 313–37.

Scheuren, Fritz and Janet McCubbin (1987). "Piecing Together Personal Wealth Distributions." In *Statistics of Income and Related Administrative Record Research: 1986–1987*, pp. 25–44. Washington, DC: International Revenue Service.

Schumpeter, Joseph (1949). "Vilfredo Pareto (1848–1923)." *Quarterly Journal of Economics* 63(2), 147–73.

Seneker, Harold (1987). "The World's Billionaires." *Forbes Magazine*, October 5.

Smeeding, Timothy (2003). Discussion of session on "Wealth Changes in the United States Over the 1990s," Conference on International Perspectives on Household Wealth, Levy Economics Institute of Bard College, Annandale-on-Hudson, NY, October 17–18. Audio recording available at levyinstitute.org/publications/?docid=677 (accessed May 11, 2014).

Stehman, Stephen and Scott Overton (1994). "Comparison of Variance Estimators of the Horvitz-Thompson Estimator for Randomized Variable Probability Systematic Sampling." *Journal of the American Statistical Association* 89(425), 30–43.

Stiglitz, Joseph (2012). *The Price of Inequality: How Today's Divided Economy Endangers Our Future*. New York, NY: W.W. Norton.

Survey of Consumer Finances (2007). "Codebook for 2007 SCF." federalreserve.gov/econresdata/scf/files/codebk2007.txt (accessed May 11, 2014).

Survey of Consumer Finances (2010a). "Codebook for 2010 SCF." federalreserve.gov/econresdata/scf/files/codebk2010.txt (accessed May 11, 2014).

Survey of Consumer Finances (2010b). "Questionnaire Outline for 2010 SCF." federalreserve.gov/econresdata/scf/files/2010_scfoutline.pdf (accessed May 11, 2014).

Torgler, Benno and Marco Piatti (2013). "Extraordinary Wealth, Globalization, and Corruption." *Review of Income and Wealth* 59(2), 341–59.

Vijverberg, Wim (2011). "Testing for IIA with the Hausman-McFadden Test." Discussion Paper 5826, Institute for the Study of Labor, Bonn, Germany.

Vos, Paul and Suzanne Hudson (2005). "Evaluation Criteria for Discrete Confidence Intervals: Beyond Coverage and Length." *The American Statistician* 59(2), 137–42.

Vuong, Quang (1989). "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica* 57(2), 307–33.

Welsh, Alan (2011). "Basu on Survey Sampling." In Anirban DasGupta, ed., *Selected Works of Debabrata Basu*, pp. 45–50. New York, NY: Springer.

Wilmoth, John (1997). *Berkeley Mortality Database*. demog.berkeley.edu/∼bmd (accessed May 11, 2014).

Wingfield, Brian (2007). "America's Youngest And Oldest Billionaires." *Forbes Magazine*, forbes.com/2007/11/06/billionaires-wealth-forbes-400-biz-cx_bw_1031400ages.html (accessed May 11, 2014).

Wolff, Edward (2012). "The Asset Price Meltdown and the Wealth of the Middle Class." Working Paper 18559, National Bureau of Economic Research.

Zhang, Jin (2013). "Reducing Bias of the Maximum-Likelihood Estimation for the Truncated Pareto Distribution." *Statistics* 47(4), 792–99.