

Challenging the Empirical Contribution of Thomas Piketty’s *Capital in the Twenty-First Century*

Phillip W. Magness

Institute for Humane Studies, George Mason University

Robert P. Murphy

Institute for Energy Research

Abstract

Thomas Piketty’s *Capital in the Twenty-First Century* has been widely debated on theoretical grounds, yet it continues to attract acclaim for its historically infused data analysis. In this study, we conduct a closer scrutiny of Piketty’s empirics than has appeared thus far, focusing upon his treatment of the United States. We find evidence of pervasive errors of historical fact, opaque methodological choices, and the cherry-picking of sources to construct favorable patterns from ambiguous data. Additional evidence suggests that Piketty used a highly distortive data assumption from the Soviet Union to accentuate one of his main historical claims about global “capitalism” in the twentieth century. Taken together, these problems suggest that Piketty’s highly praised and historically driven empirical work may actually be one of the book’s greatest weaknesses.

JEL Codes: N01, N10, D31, D63

Keywords: economic history, Thomas Piketty, capital, inequality

I. Introduction

Thomas Piketty’s *Capital in the Twenty-First Century* (2014b) begins with a bold claim. The ensuing work, he promises, is “based on much more extensive historical and comparative data than were available to previous researchers, data covering three centuries and more than twenty countries, as well as a new theoretical framework that affords a deeper understanding of the underlying mechanisms” of his subject matter (2014b, p. 1). Though he qualifies this assertion with an acknowledgment of an imperfect and incomplete data set, this concession should not be mistaken for modesty. As Piketty

repeatedly reminds his readers over the ensuing 700 pages, it is his unprecedented assemblage of data that supposedly sets his work apart from a literature on wealth inequality that—he contends—frequently suffers from “an abundance of prejudice and a paucity of fact” (2014b, p. 2).

When paired with an unconventional theoretical argument rooted in hypothesized “laws of capitalism,” and, perhaps more so, radical policy recommendations in the form of an 80 percent top marginal income tax rate and an *annual* 5 percent global wealth tax on the biggest fortunes (2014b, pp. 512, 530), Piketty’s claim to an empirically robust and data-heavy narrative has always been the strongest ecumenical feature of his work.¹ Empirics are also the root of much of the book’s claimed novelty, as well as its self-stated purpose of “patiently searching for facts and patterns and calmly analyzing economic, social, and political mechanisms that might explain them” (2014b, p. 3) in order to better inform the public discourse about the causes and consequences of global wealth inequality.

Data, and more specifically the story in those data—“as extensive as possible a set of historical data” (2014b, p. 16) as can be gathered—thus become the main evidentiary tool on which Piketty predicates his work. Indeed, he goes on to extol his own “novel historical sources” and lays claim to a patient, empirically driven search for “facts and patterns” within them twice more before the conclusion of the first chapter (2014b, pp. 20, 31–32). While Piketty’s product is part theoretical argument, part empirical exercise, and part policy recommendation, its unifying rationalization is an overarching historical narrative about the characteristics of human wealth accumulation, derived from and purportedly sustained in data.

Given these extensive claims, not to mention the heavy criticism directed toward certain other works in the wealth inequality genre, it might come as some surprise to learn that Piketty’s reported “three centuries” of empirics infrequently predate 1900 beyond a stray data point or two connected by a century’s worth of linear interpolation. His claimed global analysis only consistently examines three countries—France, Britain, and the United States—with more than passing rigor, with only occasional forays into Sweden and Germany

¹ Indeed, when the theoretical problems with Piketty’s book initially emerged, even his critics typically would concede that the empirical contribution was top-notch—at least until Chris Giles’s (2014) bombshell *FT* report, which was the first major piece to challenge this aspect of Piketty’s work.

beyond that. Even many of his twentieth-century figures, presumably constructed from better records and more readily available data sources, are often products of further interpolation and decennial averaging around multiyear and decade-long gaps. Taken alone, these circumstances might only attest to the inherent difficulties of amassing a large, continuous economic time series. A more serious problem emerges, though, when an author attempts to interpret highly specific historical events through data points that are substantially less thorough or conclusive than their initial presentation suggests.

Finally, the investigator may become downright alarmed when discovering the dubious foundation of some of Piketty's "novel" data sets, because Piketty's charts do not convey such weakness to the innocent reader. Furthermore, Piketty's narratives are occasionally peppered with wildly inaccurate historical "facts" that, coincidentally, seem to bolster his desired interpretation of the surrounding data. In this context, the various leaps and judgment calls that Piketty often makes in his historical reconstructions should raise alarm bells.

II. History: Misconstrued and Missing

At its most basic descriptive level, Piketty's presentation of major historical events at the center of his argument is laced with factual error. In addition to suggesting an inattentiveness to detail, a recurring problem of factual inaccuracy with historical events indicates that interpretative extrapolations from these errors, as well as more sophisticated data claims that appear throughout the book, may suffer from a basic fault in their underlying historical assumptions. While we will not endeavor to pick apart his most extensive historical recounting—the twentieth-century French economy—it is fair to note that he struggles, and struggles mightily at that, in many instances where he takes up the economic history of the United States. The book's favorable portrayal of FDR's New Deal policy initiatives, which function as a seminal event in Piketty's twentieth-century narrative as well as an important precedent for his prescription of confiscatory tax rates, is illustrative. Consider Piketty's descriptive retelling of Depression-era tax policy:

The Great Depression of the 1930s struck the United States with extreme force, and many people blamed the economic and financial elites for having enriched themselves while leading the country to ruin . . . Roosevelt came to power in

1933, when the crisis was already three years old and one-quarter of the country was unemployed. He immediately decided on a sharp increase in the top income tax rate, which had been decreased to 25 percent in the late 1920s and again under Hoover’s disastrous presidency. The top rate rose to 63 percent in 1933 and then to 79 percent in 1937. (2014b, pp. 506–07)

The problem with Piketty’s historical narrative in this instance is one of basic fact. Simply put, his dates are all wrong. As readily accessible tax records illustrate, the top marginal income tax rate was actually brought down to 25 percent by the year 1925, which is not “the late 1920s” and which was well within the presidency of Calvin Coolidge (with Hoover taking office on March 4, 1929).² More troubling still for Piketty’s narrative, it was under *Hoover* that the rate was raised to a decidedly punitive 63 percent under the Revenue Act of 1932. And just to round out Piketty’s tax-error trifecta, the top rate increased under FDR to 79 percent in 1936, not 1937 as Piketty claims.³

We see another example from this playbook—namely, inventing historical “facts” in order to support his narrative—a bit earlier in the book when Piketty informs his readers in a parenthetical remark, “Herbert Hoover, the US president in 1929, thought that limping businesses had to be ‘liquidated,’ and until Franklin Roosevelt

² See, for example, the Tax Policy Center’s chart, “Historical Income Tax Parameters.”

³ We call attention to the recurring complexity of Piketty’s error here as attesting to his ideological basis. If Piketty had only been wrong about 1932 versus 1933, and then 1936 versus 1937, we might attribute the errors to a misunderstanding of the tax year versus the filing year (or fiscal year versus calendar year, or even election year versus presidential inauguration year). But there is really no way to easily explain his mistake about the tax rate being reduced to 25 percent “in the late 1920s . . . under Hoover’s . . . presidency” when in fact it happened *four years* earlier under Coolidge. Whatever the explanation, notice that Piketty’s mistakes—which are so basic that they would discredit a high school history paper—serve to bolster his narrative of the low-tax, awful-economy Hoover versus the high-tax, economic-recovery FDR. The odd inclusion of the word “again” in the quote—when Piketty writes “and again under Hoover”—may refer back to page 473, where Piketty establishes that “the top rate under Hoover had been only 25 percent.” Since the book was translated into English from French, the extremely generous reader could acquit Piketty of intentionally misleading the reader here, but translation difficulties cannot explain why he claims that a 1925 tax-rate reduction occurred “in the late 1920s” and consistently mentions Hoover but not Coolidge.

replaced Hoover in 1933, they were” (2014b, p. 472). This claim is simply not true. Herbert Hoover (1952, pp. 30–31) in his memoirs quotes the (in)famous advice given to him by Treasury Secretary Andrew Mellon to “liquidate labor, liquidate stocks, liquidate the farmers, liquidate real estate.” But the rhetorical point of Hoover bringing up this advice was to assure his reader that he had *rejected* such tough love. Hoover was compassionate with his misguided subordinate, though, writing, “Secretary Mellon was not hard-hearted . . . He felt there would be less suffering if his course were pursued. The real trouble with him was that he insisted that this was just an ordinary boom-slump.” Piketty is not alone in attributing to Hoover the very view that Hoover explicitly renounced, but it is nonetheless one of many examples of demonstrably false statements in the book that conveniently align with Piketty’s historical worldview.

The common theme of these factual errors is that Piketty uses them to augment certain historical political events and figures that align with his own modern prescriptions. In this sense, a specific narrative construction of the past—even though factually erroneous and misconstrued—may be seen to lend favor to desired policies in the present day.⁴ We see a comparable episode when Piketty turns to more modern times and the U.S. federal minimum wage, when he writes, “From 1980 to 1990, under the presidents Ronald Reagan and George H. W. Bush, the federal minimum wage remained stuck at \$3.35, which led to a significant decrease in purchasing power when inflation is factored in. It then rose to \$5.25 under Bill Clinton in the 1990s and was frozen at that level under George W. Bush before being increased several times by Barack Obama after 2008” (2014b, p. 309).

Here again, this “history” is utterly wrong, as readily available federal sources reveal.⁵ Piketty’s description is so at odds with actual history that it is easiest if we present the correct information in a table.

⁴ Piketty’s historical narrative for the United States may be said to exhibit the pattern described by Hayek (1954, pp. 3–4) wherein “historical beliefs which guide us in the present are not always in accord with the facts; sometimes they are even the effects rather than the cause of political beliefs.”

⁵ The first person to our knowledge who caught Piketty’s dubious minimum wage discussion was Furchtgott-Roth (2014). Historical minimum wage rates are readily available from the U.S. Department of Labor, Wage and Hour Division, “History of Federal Minimum Wage Rates Under the Fair Labor Standards Act, 1938–2009.”

Table 1. U.S. Federal Minimum Wage, Select Periods

Date	Minimum Wage	President in Office
January 1, 1980	\$3.10	Jimmy Carter
January 1, 1981	\$3.35	Jimmy Carter
April 1, 1990	\$3.80	George H. W. Bush
April 1, 1991	\$4.25	George H. W. Bush
October 1, 1996	\$4.75	Bill Clinton
September 1, 1997	\$5.15	Bill Clinton
July 24, 2007	\$5.85	George W. Bush
July 24, 2008	\$6.55	George W. Bush
July 24, 2009	\$7.25	Barack Obama

Source: U.S. Department of Labor.

Piketty’s breezy discussion of the minimum wage is *almost* correct—though unbelievably misleading—if one were to look at his treatment up through Clinton. (Even here, he is wrong about 1980 versus 1981, and the minimum wage under Clinton was \$5.15, not \$5.25.) But for him to claim that the minimum wage was frozen under George W. Bush until being raised under Obama is utter nonsense. If we wanted to be pedantic, we could bring up the fact that the July 24, 2009 increase that occurred under Obama was due to legislation signed by George W. Bush, but that would detract from the more basic point that Piketty cannot even get his years, dollar amounts, and presidential administrations right. There are many problems with Piketty’s portrayal, given the ease with which a more conscientious researcher could have verified such basic information from U.S. Department of Labor tables. Indeed, Piketty’s own data files indicate his awareness of this source.⁶ Yet, Piketty’s bizarre errors aren’t completely without a pattern: they serve to paint ostensibly market-friendly Republican presidents as ogres, while liberal Democrats are the heroes of the working class and purveyors of policies that Piketty embraces. Keep these “easy” examples of Depression-era tax rates and minimum wages since 1980 in mind when we delve into harder areas of obscure data series on wealth distribution, where Piketty at many places asks the reader to trust him.

⁶ Piketty’s (2014b) data table TS9.1 actually shows he had access to U.S. federal minimum wage data by year from the same U.S. Department of Labor source we used, although he adjusts the timeline of each hike to the beginning of the nearest calendar year. While this adjustment certainly adds further confusion to his timeline, the errors in both wage rates and attributed presidential administrations persist even after we take it into account.

Before immersing ourselves in the wealth data, let's consider another error of omission in Piketty's figure 14.2 (2014b, p. 504), which is a comparative historical portrayal of the top estate-tax rate in the United States, Britain, and France. This chart shows the United States as a relative latecomer to the approvingly referenced practice of estate taxation, starting only after 1916. Piketty evinces no awareness that the United States actually began its modern experiments in estate taxation with the Spanish-American War Revenue Act of 1898.⁷

It is worth elaborating on this example because it typifies a problem of recurring shoddiness in Piketty's claims of historical "facts." At first glance, his figure 14.2 seems authoritative and authentic; the reader can see the different symbols representing the individual annual inheritance tax rates for each country. The issue is his careless omissions. What Piketty apparently did was research the origin of the *modern* federal inheritance tax in the United States, while simply assuming that the federal estate tax was nonexistent beforehand. The factual errors displayed in his chart thus reveal the absence of even cursory research. As an easily accessed Tax Foundation report notes, in contrast to Piketty's chart, "The federal government resorted once again to transfer taxes in the 1860's when the Civil War and subsequent reconstruction forced Congress to look for additional federal revenue. A series of Acts passed in 1862, 1864, and 1866 created and refined the first federal inheritance tax. In 1870 Congress repealed this tax as demands for federal revenue eased. When the Spanish-American War flared-up in 1898, Congress again relied on a transfer tax, this time an estate tax, to defray some of the costs of the conflict. This tax was repealed in 1902" (Fleenor 1994, pp. 3–4).

The 1898 act's estate tax was even upheld in a Supreme Court ruling in 1900 (*Knowlton v. Moore*, 178 U.S. 41), where part of the dispute concerned the *progressivity* of the tax, making Piketty's oversight all the more surprising. Yet, it also remains consistent with a historical narrative that casts the late nineteenth century as an "unenlightened" period of capitalism run amok.

While a superficial credulity for partisan talking points seems evident in Piketty's rendering of the Depression, historical tax rates,

⁷ The War Revenue Act of 1898 was repealed in 1902, although certain inheritance tax provisions from it did not expire until 1907, providing almost a decade of operations that go unnoticed by Piketty (2014b, p. 338), who traces the first American estate tax to 1916.

and the minimum wage, there is also something more elementary at play in his recounting of historical events. Piketty's approach to the economic history of the United States shows telltale signs of a scholar who is deeply unfamiliar with the historical particulars of his subject matter and frequently errs in recounting them, but who stakes strong interpretive claims upon his constructed "history" and chooses to enlist it for prescriptive purposes nonetheless.

III. Data Discrepancy or Data Discretion?

In the aforementioned cases, Piketty's facts are simply wrong or his data missing. As problematic as this may be, it is only the most easily detected of similar factual discrepancies and what appear to be wholly discretionary calls that are buried deep in his data files. While Piketty has received praise for following the increasingly common practice of making his source data available to other scholars by posting them online, this act of courtesy should not be mistaken for a license to impute validity to their contents. Consider Piketty's figure 13.1 (2014b, p. 474) portraying historical tax revenue to national income ratios across a series of countries since 1870. The chart itself suffers from another tax history imprecision that is virtually undetectable by the investigator who relies solely on the information in the book, and it only becomes apparent upon closely scrutinizing its source files. Two images of Piketty's root data file for the United States and its formulas appear in figure 1.

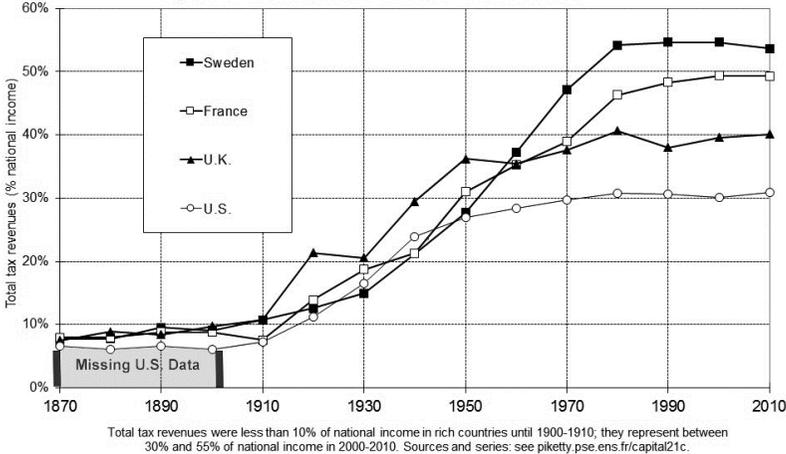
Note in particular that Piketty's figure for 1900 consists of a decennial average for that decade at 6 percent (expanded to 6.122 percent when rounding is removed). This figure comes from a U.S. Census Bureau-produced table of historical revenue data from 1902 to the present day. While his figure 13.1 suggests a complete accounting to the casual reader, Piketty actually has no data source for the thirty-two years prior to 1902, extending his series backward to the reported starting point of 1870. His claimed numbers for these three decades appear not to have been obtained from any actual source, but rather through the alternating addition or subtraction of 0.5 percentage points projected backward from the 1900 average. He evidently filled the gap in his own data collection—needed to bring the U.S. time series into line with his European records—by constructing it *ex nihilo*. Figure 2 shows Piketty's missing data points.

Figure 1. Manufactured Data Points in Piketty’s Accounting of Historical U.S. Tax Revenue

B6 =B7+0.005				B9 =AVERAGE(H538:H547)/0.9			
A		B		A		B	
Table S13.1. Tax revenues in rich countries, 1870-13.1)				Table S13.1. Tax revenues in rich countries, 1870-13.1)			
(tax revenues, % national income)	U.S.	U.K.		(tax revenues, % national income)	U.S.	U.K.	
1870	7%	8%		1870	6.622%	8%	
1880	6%	9%		1880	6.122%	9%	
1890	7%	8%		1890	6.622%	8%	
1900	6%	10%		1900	6.122%	10%	
1910	7%	11%		1910	7.210%	11%	
1920	11%	21%		1920	11.208%	21%	
1930	17%	21%		1930	16.559%	21%	
1940	24%	30%		1940	23.849%	30%	
1950	27%	36%		1950	26.910%	36%	
1960	28%	35%		1960	28.414%	35%	
1970	30%	38%		1970	29.717%	38%	
1980	31%	41%		1980	30.730%	41%	
1990	31%	38%		1990	30.653%	38%	
2000	30%	40%		2000	30.084%	40%	
2010	31%	40%		2010	30.850%	40%	

Source: Piketty (2014b) data appendix, table S13.1, <http://piketty.pse.ens.fr/en/capital21c2>.

Figure 2. Missing U.S. Data Points in Piketty’s Figure 13.1
Figure 13.1. Tax revenues in rich countries, 1870-2010



Source: Piketty (2014b), figure 13.1, with 1870–1901 U.S. data gap denoted.

It is not difficult to see how such an invention of numbers, even on a relatively small part of a comparative graph, could leave Piketty vulnerable to charges of data manipulation or contrivance. Most researchers, when faced with two or more historical series of different lengths, would simply start the plot from the shorter series later on the graph when the data became available. Piketty instead

decided to fill in his chart with numbers he pulled out of the air. Equally significant is that the missing data actually do exist, though it would have required more than superficial probing to incorporate them into his charts.⁸ It therefore becomes a legitimate cause for concern when other similarly questionable decisions pepper his data presentation in increasingly complex ways, and with claims that carry greater interpretive significance to his theory of capital accumulation.

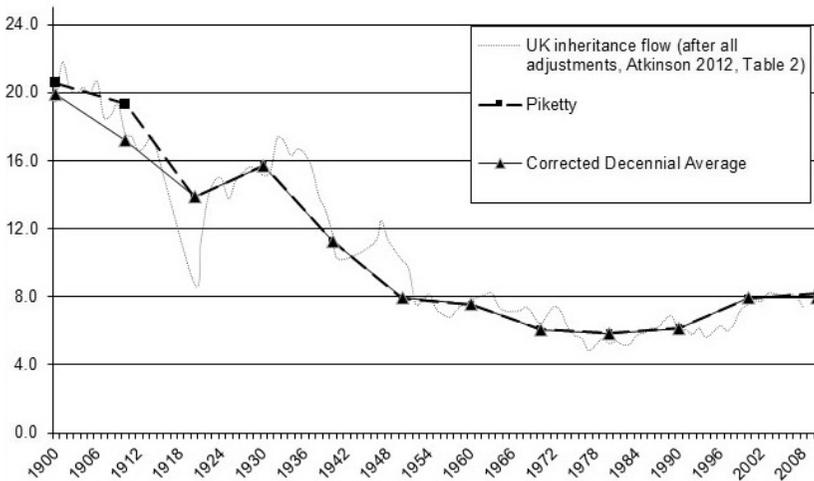
Another problem appears in Piketty's figure 11.12 (2014b, p. 423) and accompanying charts, which purport to illustrate wealth inheritance flows in the United Kingdom since the beginning of the twentieth century. In this instance Piketty uses a bizarrely creative averaging technique to obtain decennial averages for 1900 and 1910, deviating from the standard formula he uses for calculating the averages for the remainder of the twentieth century. In Piketty's presentation, these two data points show a relatively flat and stable trend for about a decade prior to the sudden drop in inheritance flows allegedly precipitated by the First World War, a seminal trigger event in the twentieth century's dissipation of concentrated capital, according to Piketty's main argument. Figure 3 compares Piketty's constructed trend line (in red) against the raw annual source data he uses, as well as our own "corrected decennial average" using the same source data.

Notice that our "corrected decennial average," included here for comparison against Piketty's alterations, both tracks the raw source numbers more carefully and shows a much sharper downward trend between 1900 and 1910. This outcome is not accidental. Whereas Piketty's other decennial averages are encompassing of the represented decade, as with our corrections (e.g., his 1920s figure averages the years 1920–1929), he appears to have intentionally back loaded his calculations for the 1900 and 1910 data points with numbers that are actually taken from their respective preceding decades. Deviating from the consistent decennial calculations of later decades, Piketty's resulting chart actually presents 1896–1902 figures in its average for the "1900s" and 1904–1910 figures for the "1910s," without any annotation or indication that he performed this backward shift. His intent is difficult to discern, but these

⁸ Annual U.S. federal tax revenue is easily attainable in a standardized annual form going back to 1821, and state revenues may be reliably ascertained or estimated throughout the nineteenth century. Several widely available estimates have been made for U.S. national income in the nineteenth century as well. Piketty simply did not incorporate them into his source material.

discretionary modifications do produce a picture that is more amenable to a grand narrative about a static, capital-hording aristocracy in the Belle Époque. Alternatively, a modest but already-declining inheritance flow pattern in the decade and a half before the war—as suggested in the raw data—might present a wrinkle that requires a more nuanced investigation than Piketty’s sweeping theory of capital accumulation permits.

Figure 3. UK Inheritance Flows, Piketty Chart vs. Corrected Decennial Average
UK Inheritance Flows as a % of National Income, 1900-present



Source: Original data and corrections made using Piketty (2014b), data appendix, table S11.3.

IV. From Theoretical Arguments to Data Illustrations

At this point, it is fair to query whether the aforementioned issues, even if acknowledged as problematic to Piketty’s historical narrative, pose a larger challenge to his theoretical argument. Briefly summarized, Piketty (p. 571) posits that capitalism is afflicted by an internal “contradiction” of sorts wherein an inherent propensity exists for the rate of return on capital—and thus the wealth of its owners—to outpace overall economic growth, and with it the employment income of workers. The implication of this theory is a long-term divergence between a “rentier” class of capital-sustained wealthy elite and those who rely on their labor to earn income, perhaps best epitomized by the wealth disparities of the late “Gilded Age.” Piketty appends this argument with a historical explanation for decreasing wealth inequality in the twentieth century, which he attributes to the ravages of two world wars and an intervening

depression upon the prior century's capital-hording aristocracies, as well as the "progressive" era rise of confiscatory taxation. His predictive and prescriptive arguments, in turn, contend that this dissipation is at an end because of policy changes and trends in the global economy. Rounding out this argument, he claims that the experience of the late twentieth century through the present day begins to resemble an oft-referenced "U-shaped" trend wherein capital-accumulation-fueled wealth disparities are presently on course to return to their Victorian levels, or something similar.

Piketty formalizes this narrative into a set of quasi-scientific "laws" of capitalism, each representing specific characteristics of capital accumulation and its return. These theoretical underpinnings have proven fertile ground for criticism, and while further inquiry of this vein is certainly warranted, it falls beyond the scope of our present examination.⁹ We accordingly view the theoretical line of scrutiny as a distinct issue with his work, although one that is complementary in its conclusions to our assessment of Piketty's empirics.

Returning to Piketty's data, their function is fundamentally evidentiary: that is, his data-driven historical analysis supplies the numbers that validate his theorizing about the nature of capital. He goes about this task through multiple examples drawn mostly from France, Britain, and the United States. Where examined as a century-long time series, his argument thus rests upon demonstrating the aforementioned "U-shaped" trend amidst metrics that attest to capital accumulation and wealth distributions in a country or region. A "U-shaped" resurgence in wealth disparity between the highest levels and the rest of society might accordingly indicate the long-term effects of returns on capital, "r," outpacing the rate of growth, "g." A "U-shaped" curve showing a divergent ratio between privately owned capital and national income might similarly validate Piketty's claimed causal source for concurrent inequality trends—that is, a reconstitution of the capital stock that he places at the center of his late-twentieth-century narrative and forecasts for the twenty-first century.

It would exceed the scope of a single article to dissect Piketty's data in their entirety, so our discussion will focus on two iterations of the "U shape" that cut to the core of his evidentiary claims. The first is Piketty's widely cited figure 10.5 (2014b, p. 347), in which he

⁹ See in particular McCloskey (forthcoming).

purports to represent 200 years of wealth inequality in the United States. The figure's title is itself a misnomer, as the entire first century of Piketty's graph is based on only two data points, 1810 and 1870, connected by linear interpolation. He attempts to weave a story of the nineteenth-century United States as something of a paradise lost whereby an initially tempered disparity of wealth grew to resemble the high-water mark of European aristocracy by the close of the century, but the quantitative evidence he offers is far too sporadic to sustain this case in any detail. The subsequent course is more subdued than its counterparts in Britain and France, but is one with both causal claims and predictive implications for the United States.

Turning to his post-1910 data, the familiar "U shape" begins to emerge, wherein a pre-World War I peak in wealth disparity dissipates across the twentieth century. Spurred by a "pioneering" progressive tax policy, this trend continues until reaching a stable bottom in the 1970s. It is then depicted as following a gradual yet sustained and certain uptick into the present day. Piketty explains this graph with a very specific historical claim: "In the United States . . . a (white) patrimonial middle class already existed in the nineteenth century. It suffered a setback during the Gilded Age, regained its health in the middle of the twentieth century, and then suffered another setback after 1980. This 'yo-yo' pattern is reflected in the history of US taxation" (2014b, p. 350).

After lauding the United States for revolutionizing the practice of "confiscatory taxation of excessive incomes" (2014b, p. 505)—a narrative he constructs from the aforementioned faulty recounting of Depression-era tax history—Piketty attempts to pin the observed uptick since the 1980s squarely on Reagan-era tax cuts, drawing it out as a claimed implication of an "explosion of executive salaries" (2014b, pp. 335, 508) that in turn validates his theories of capital accumulation.

But is Piketty's figure 10.5, the graph at the root of his American wealth disparity narrative, even accurate? Wealth inequality is a notoriously difficult concept to quantify, though two common estimation techniques do allow approximations. The first uses historical estate-tax records to estimate wealth shares annually. The second samples wealth distributions using the Federal Reserve's Survey of Consumer Finances (SCF), though only at the less frequent intervals permitted by the survey's collection. Piketty's central graph is really a composite of other studies, based on variations of these two techniques.

When Piketty’s book went to press, the most complete study of U.S. wealth inequality was from a 2004 article by Wojciech Kopczuk and Emmanuel Saez, who used estate-tax records to estimate wealth distributions from 1916 to 2000. Their time series, reprinted below as figure 4, shows a late-twentieth-century trend that is at odds with Piketty’s narrative.

Figure 4. Kopczuk and Saez Estate-Tax Estimate of U.S. Wealth Inequality

Figure 2. The Top 1% Wealth Share in the United States, 1916–2000



Source: Kopczuk and Saez (2004), figure 2.

Whereas Piketty’s figure 10.5 depicts a sharp uptick in inequality beginning in the 1980s—which fits his narrative that tax cuts for the wealthy are driving the new trend in inequality—the Kopczuk and Saez estate-tax study displays a trend that is at best ambiguous and gives the appearance of being flat through the end of the twentieth century.

A closer examination of the source data in Piketty’s figure 10.5 reveals that most of its root data are actually taken from the Kopczuk and Saez study, yet they are also “augmented” and extended through 2010 with other studies based on the SCF as well as a number of opaque adjustments that are simply hard-coded into Piketty’s source files. Unfortunately, neither Piketty’s annotation nor the supplemental document that he released in response to data criticisms contains an adequate or transparent explanation of how he performs these “augmentations” (Piketty 2014a, pp. 6–7). By reconstructing his graph from the provided data tables, however, it

becomes quickly apparent that he is not so much an aggregator of the existing literature as a cherry-picker.

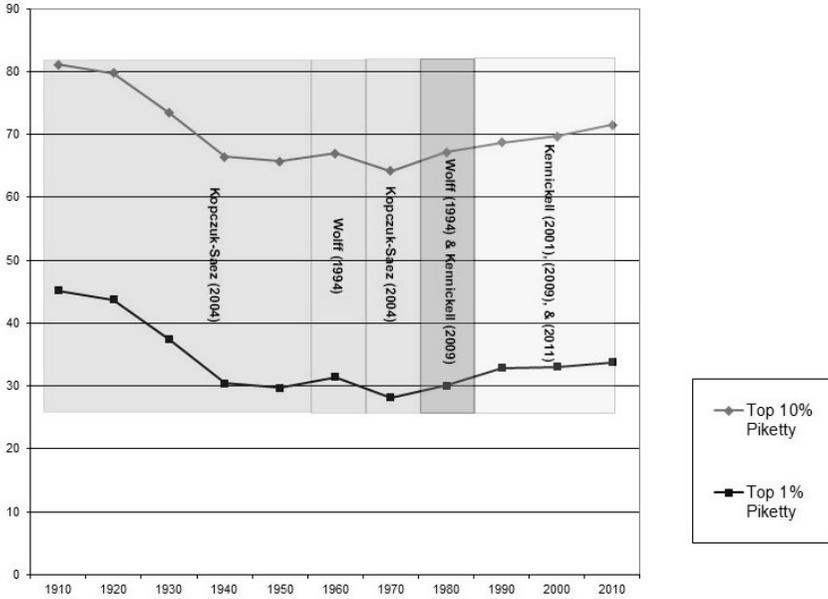
Beginning with the raw Kopczuk and Saez data set for the top 1 percent of the wealth distribution, Piketty first reconciles it upward through “corrective” adjustments to match other studies using the SCF methodology. Converting his output into decennial averages to account for gaps in the data, he retains Kopczuk and Saez from 1910 through 1950. He then merges their data with a single 1962 data point from an SCF-based study by Edward N. Wolff (1994), bringing the Kopczuk and Saez average for the 1960s into line with the SCF. Without citing this change in sources, and in fact suggesting otherwise in his limited annotation, he then migrates back to a figure derived from Kopczuk and Saez for the 1970s, obtained by weighting the SCF-reconciled 1960s data point with a marked drop-off shown in their estate tax series.¹⁰ This switch introduces a problematic dimension to his decennial averaging technique, as the 1970s contain the largest modern data gap in Kopczuk and Saez, containing only two years for the entire decade. Though the Kopczuk and Saez series improves for the 1980s and 1990s with eight and ten full years of data, respectively, this is also the point where their inequality trend line flattens out, thus becoming at odds with Piketty’s “U-shaped” expectations. Piketty therefore quietly abandons Kopczuk and Saez from 1980 onward, migrating back to Wolff (1994, 2010) and then to a separate SCF-based study by Arthur Kennickell (2009, 2011), as figure 5 shows.

The two sets of SCF-based studies do suggest a very modest increase in wealth inequality over the last two decades, but it is Piketty’s unconventional combination and rotation of them with the Kopczuk and Saez estate-tax estimate that removes a substantial amount of ambiguity to suggest a pronounced upward trend into the present day. The aforementioned pivot to the Kopczuk and Saez 1970s data point adds an accentuating effect to Piketty’s chart, as it produces a clear bottom point of the “U shape.” When merged with the later survey sources, this point gives the illusion of a steeper rebound than the post-1990 SCF studies show when taken alone. Piketty’s assembly process for his figure 10.5 continues from there as he extrapolates a second wealth distribution for the top 10 percent.

¹⁰ The SCF-based estimate in Wolff (1994) does not contain a data point for the 1970s.

Figure 5. Source Data Reconstruction of Piketty’s Figure 10.5

Wealth Inequality in the US, 1910-2010 (Piketty Figure 10.5)



Source: Piketty (2014b) figure 10.5 and data appendix, table TS10.1DetailsUS.

He begins this by simply appending a fixed 36 percentage points to his top 1 percent series from 1910 to 1950, and again in 1970, without alerting the reader that this (large) portion of his trend line is based *not* on underlying data but on his mere adding a fixed number to the lower trend line.¹¹ He then reconciles this new line with SCF data for the remaining decades. The end product is a Frankenstein graph, assembled from bits and pieces of the existing literature that seem to be added or dropped for the convenience of the trend line he wishes to see at the moment his preferred historical narrative expects it to appear.¹²

¹¹ Giles (2014) first drew attention to the hard-coding of this 36 percentage point adjustment into Piketty’s estimates for the top 10 percent from 1910 to 1950. We acknowledge that this number may derive from another unnamed source or method, but Piketty (2014a, pp.7–8) did not clarify its origin when specifically pressed by Giles to do so.

¹² Auerbach and Hassett (forthcoming, pp. 5–6) reach similar conclusions of Piketty’s figure 10.5, focusing on his use of the SCF-derived sources. When stripped of Piketty’s unconventional smoothing techniques, they note, the SCF data become “noisy” and the trend line “no longer rises without interruption in an apparently deterministic trend from 1970 onward.”

V. A Useful Digression: Estimating Wealth Distribution Using Estate versus Income Tax Data

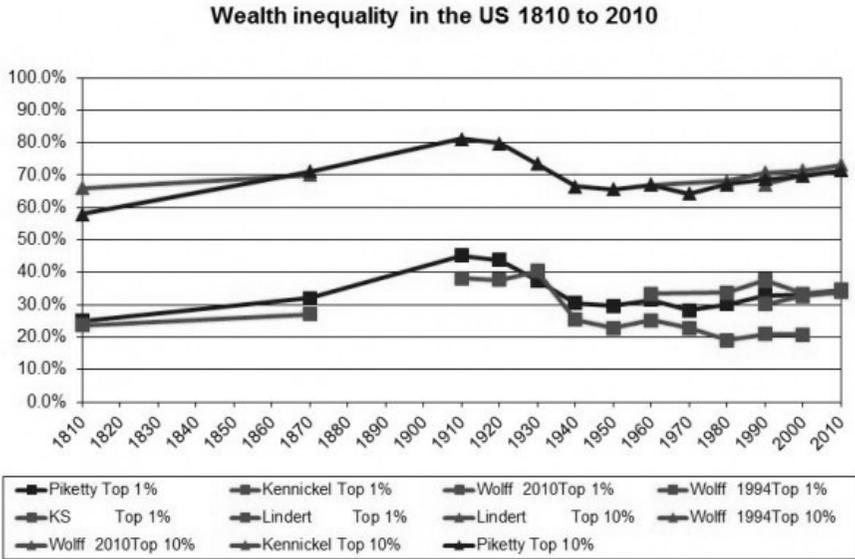
One of the most unfortunate outgrowths of Piketty's popular acclaim is that the public and much of the economics profession has accepted the settled "fact" of significantly increasing wealth inequality in the past three decades. In reality, the evidence for this conclusion is much more ambiguous than its loudest champions—including Piketty—would have us believe. In this section, we outline some of the controversies.

As Piketty's book zoomed to bestseller status soon after its release in April 2014, the academics and pundits who disagreed with its policy conclusions were caught somewhat flat-footed. It was a dense economics book running almost 700 pages, after all. Critics raised their objections, but they were drowned out by the constant drumbeat that Piketty's empirical work was of the highest caliber. People argued about the truth and significance of " $r > g$," but few doubted the accuracy of the historical charts that Piketty had assembled. These charts showed a rising concentration of wealth (in both the top 10 percent and the top 1 percent), which Piketty warned would continue if drastic policy countermeasures were not taken.

Then on May 23, 2014, Chris Giles of the *Financial Times* launched a major broadside against Piketty's empirical work, arguing (among other things) that Piketty's chart for U.S. wealth inequality conveniently sampled from various disparate data sets in order to give the appearance of an increase in wealth concentration that the original data sets lacked. As we showed in the preceding section, this selective sampling is largely responsible for the "U shape" of Piketty's figure 10.5. We reproduce Giles's U.S. chart in figure 6 to illustrate the ambiguity of its disparate components and other similar studies when taken apart.

In Giles's chart, the two continuous lines represent Piketty's estimates as presented in his book. They are based in turn on a "blending" of the various data sets depicted in the other lines (which are red in the original). The disconnected lines represent various source estimates. Note the lower source lines in particular, showing the concentration of wealth among the top 1 percent of wealth holders. Starting in 1910, Piketty's line tracks one of the source lines fairly closely (though overstating it).

Figure 6. Comparative U.S. Wealth Inequality Metrics from the *Financial Times* Analysis



Source: Giles (2014).

Yet, from the 1970s data point to the present, Piketty’s line and this original source diverge sharply. Here, Piketty performs his aforementioned pivot to the other data sets, which do not go back as far. But had Piketty’s line continued to track the original source line, Piketty would have shown that, as of 2000 at least, the wealth concentration in the upper 1 percent in the United States was just about the lowest and most stable it had been in recorded history, save for a slight dip in the 1980s.

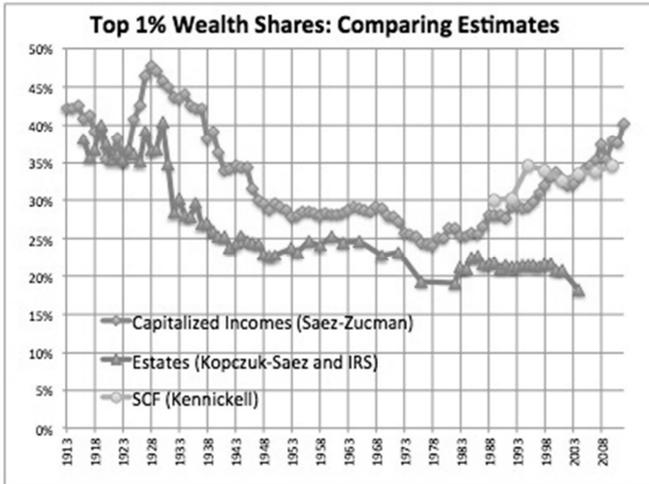
After Giles challenged Piketty, Piketty directed his critics to a PowerPoint presentation based on a forthcoming (though at the time not released) paper from economists Emmanuel Saez—the coauthor of the 2004 study—and Gabriel Zucman. We reproduce a key slide from the Saez and Zucman PowerPoint presentation in figure 7.

Far from accounting for his own questionable judgment calls on how to blend the various data sets in his figure 10.5, Piketty answered the *FT* critiques by claiming vindication in the new Saez and Zucman PowerPoint results and retorting, “If anything, my book underestimates the rise in [U.S.] wealth inequality” (2014a, p. 2). It is important to note the methodologies of these two series. The Kopczuk and Saez (2004) study uses estate tax data to estimate wealth concentration. The new Saez and Zucman (2014b) PowerPoint series, which was only released in full as part of an

NBER working paper in October, uses data on capital *income* and rates of return (by asset class) in order to generate a new estimate of the capital base that generated such income flows.

Figure 7. Saez and Zucman Slideshow Comparison on Inequality Estimates

Estate tax returns fail to capture rising top wealth shares



Source: March 2014 Saez and Zucman PowerPoint presentation, <http://gabriel-zucman.eu/files/SaezZucman2014Slides.pdf>.

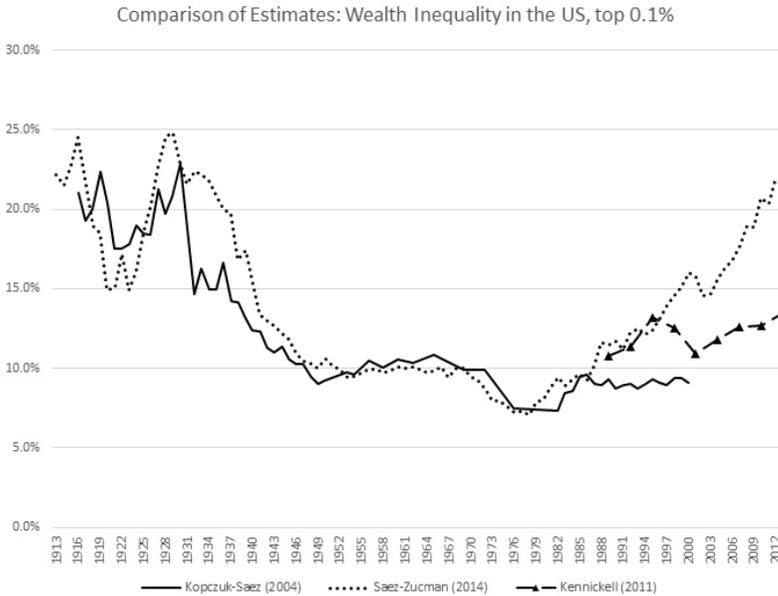
The amazing thing about the slide shown in figure 7 is that the two methods—directly assessing wealth from estate tax data, versus computing the wealth by dividing capital income by the relevant rate of return—tracked each other decently from the 1920s through the 1970s. The consistency of this tracking is even more pronounced when one looks at the top 0.1 percent of wealth ownership, as shown in the following comparison, figure 8, constructed from Kopczuk and Saez (2004), Saez and Zucman (2014b), and a separate SCF-based study by Kennickell et al. (2011).

But in 1985, the two studies diverge sharply, with the estate tax-based Kopczuk and Saez estimates remaining flat and even falling through the early 2000s, while the new Saez and Zucman estimate zooms upward about 10 percentage points in the same time period.¹³

¹³ We call the reader's attention to the stark explanatory reversal by coauthor Saez between these two papers. In 2004, Saez interpreted his findings by noting "we tentatively suggest (but do not prove) that steep progressive income and

A similar divergence follows in the SCF study in 1991. This major discrepancy between the new Saez and Zucman PowerPoint results and both of the older techniques for estimating wealth inequality received almost no acknowledgement from Piketty or his defenders in the rush to claim vindication for Piketty’s prior findings in his figure 10.5.

Figure 8. Saez and Zucman Divergence from Alternative Estimates



Source: Compiled from data appendix, Saez and Zucman (2014b)

estate taxation, by reducing the rate of wealth accumulation of the rich, may have been the most important factor preventing large fortunes to be reconstituted after the shocks of the 1929–1945 period.” The then-observed flat trend line was “consistent with the decreased importance of capital incomes at the top of the income distribution documented by Piketty and Saez (2003), and suggest[s] that the rentier class of the early century is not yet reconstituted.” In an accompanying policy brief for the 2014 paper, Saez, along with Zucman, asserts, “Wealth inequality, it turns out, has followed a spectacular U-shape[d] evolution over the past 100 years. From the Great Depression in the 1930s through the late 1970s there was a substantial democratization of wealth. The trend then inverted, with the share of total household wealth owned by the top 0.1 percent increasing to 22 percent in 2012 from 7 percent in the late 1970s.” Offering their own suggestions to avoid what they dub “this dystopian future,” they call for a resumption of “Progressive estate and income taxation” measures, which—they say—“were the key tools that reduced the concentration of wealth after the Great Depression.” Also note that the one constant in Saez’s work between these two otherwise divergent interpretations of conflicting data results *for the same time period* is a strong prescriptive endorsement of progressive income and estate taxation.

Although it is not essential for our present purpose, we can note several plausible explanations for the huge discrepancy between the “capitalization” method and the estate-tax data. One possibility is that the *income* inequality estimates—such as the (reputedly) definitive series produced by Piketty and Saez (2003)—are themselves distorted because of the significant changes to the U.S. federal income tax code introduced in the 1980s. The sharp reduction in the top marginal personal income-tax rate (from 70 percent in 1980 to 28 percent by 1988) gave an incentive to high-income earners to reorganize their businesses as S corporations and other pass-through entities, allowing what was previously considered “business” income (then taxed at the corporate rate) to be taxed under the new personal income tax rates. Thus, the surge in *income* inequality during the 1980s may have been partly an artifact of a mere rearrangement of income that was there all along.

On the other hand, much of the middle class’s capital income (such as capital gains, dividends, and interest earnings) disappeared from the tax data in the 1980s because of the rise of tax-deferred investment vehicles. Such capital income had been reported in the 1970s, meaning that the middle class’s capital income was artificially understated by comparison in the 1980s. After discussing all of these factors (and more), Alan Reynolds (2006, p. 80) concluded, “It is extremely deceptive to compare tax-based estimates of income distribution from 1970 to 1979 with any year after 1986.” To be clear, Reynolds was writing in the mid-2000s and was not addressing our present controversy. However, if his hypothesis about *income* inequality measures is right, then it also may be a major explanation for the divergence in *wealth* inequality measures. Specifically, Reynolds thinks the standard income inequality series became skewed in the mid-1980s and beyond, and this is precisely when the estate-tax method diverged sharply from the capitalization method for wealth inequality, whereas they had tracked each other tolerably well for the previous sixty years.

Saez and Zucman, of course, consider their new results to be correct. A glance back at their PowerPoint slide shows that it is titled, “Estate tax returns fail to capture rising top wealth shares.” In other words, they are *sure* that their results (at least the general trend) are correct, such that if the estate-tax method disagrees, then the estate-tax method must be wrong—it is missing the “rising top wealth shares” that Saez and Zucman are confident must really be there.

How do Saez and Zucman explain that the estate-tax method mimicked their preferred approach for decades, but then went wildly adrift starting in the 1980s? In their October 2014 paper, Saez and Zucman suggest that the older Kopczuk and Saez (2004) data set went awry by failing to account for a substantial mortality gradient between the super wealthy and the merely wealthy. If, in fact, the wealthiest 1 percent live significantly longer than the wealthiest 10 percent, then assuming a constant mortality rate would—using the estate-tax method—understate the true concentration of wealth among the living.

Kopczuk himself in November 2014 published a working paper addressing all of these issues, presenting the strengths and weaknesses of the main approaches to estimating wealth inequality (in addition to the capitalization, estate tax, and SCF methods, he also assesses a fourth method of ranking the wealthiest individuals, as put out in the *Forbes* 400 list). Kopczuk is balanced in his treatment, though he notes that both the SCF and estate-tax methods show at best a modest increase in inequality since the mid-1980s, whereas the capitalization method shows a dramatic increase. He brings up Saez and Zucman's (2014b) theory that different mortality rates explain the discrepancy in the results, then comments:

This explanation is conceptually plausible, but the estimated gap in mortality rates for the very wealthy is so large and unexplored elsewhere in the literature, that the subject clearly requires further research. For example, an alternative possible explanation for their finding of such a large mortality advantage at the very top of the wealth distribution rests on the observation that, by construction, they report mortality rates for individuals with high capital income (which they interpret as high wealth). If high capital income represents active rather than passive returns, because it is a form of compensation for actively running or managing a business, for example, then individuals with high capital income are partially selected on health—it is being healthy that allows them to be active beyond retirement. On the flip side, individuals who are sickly may instead have an incentive to engage in tax planning and not realize capital income; in particular, there is a strong tax incentive not to realize capital gains until death in order to benefit from the step up of the basis of capital gains at death. As I will argue in what follows,

it is likely that individuals at the top of the wealth distribution have become increasingly self-made, so that one might plausibly expect that this type of selection has become stronger over time. (Kopczuk 2014, p. 18)

Our purpose in walking through the debate in the inequality literature is not to single out one method or data series as the best. Rather, we are documenting that there is still lively debate among the top scholars in the field over the claimed “fact” of rapidly increasing wealth inequality since the 1980s. This debate stands in marked contrast to the portrayal of Piketty by his progressive champions, with Paul Krugman (2014a) recently crediting his “historical depth” for “demonstrating that we really are living in a new Gilded Age” and purporting the inability of his detractors to mount a “substantive” challenge to this claim. Despite further characterizations of Piketty’s critics as peddlers of “inequality denial” on the conspiratorial payroll of “powerful groups with a strong interest in rejecting the facts, or at least creating a fog of doubt,” Krugman (2014b) evinces little awareness of the widespread ambiguity that presently characterizes the academic literature outside of Piketty on this very same subject. As a final word on this topic, let us repeat that Piketty and his defenders retorted by endorsing *a PowerPoint presentation* from an unreleased and little-vetted study when challenged by critics. That presentation has only recently been appended by no more than a working paper, and one that creates more unanswered research questions about its own methods than it resolves for Piketty’s claimed trend line. When it comes to the allegation of a rapid increase in wealth inequality in recent decades, the mismatch between the actual strength of the empirical evidence and the rhetorical flourishes with which it is communicated to the public is breathtaking.

VI. Capital-Income Ratios, Soviet Distortions, and Intentional Mischief

Much of the debate around Piketty’s inequality figures is methodological in nature, as may be seen in the following passage he wrote in response to Giles and the *Financial Times*:

What is troubling about the FT methodological choices is that they use the estimates based upon estate-tax statistics for the older decades (until the 1980s), and then they shift to the survey based estimates for the more recent period. This is

problematic because we know that in every country wealth surveys tend to underestimate top wealth shares as compared to estimates based upon administrative fiscal data. (2014a, p. 8)

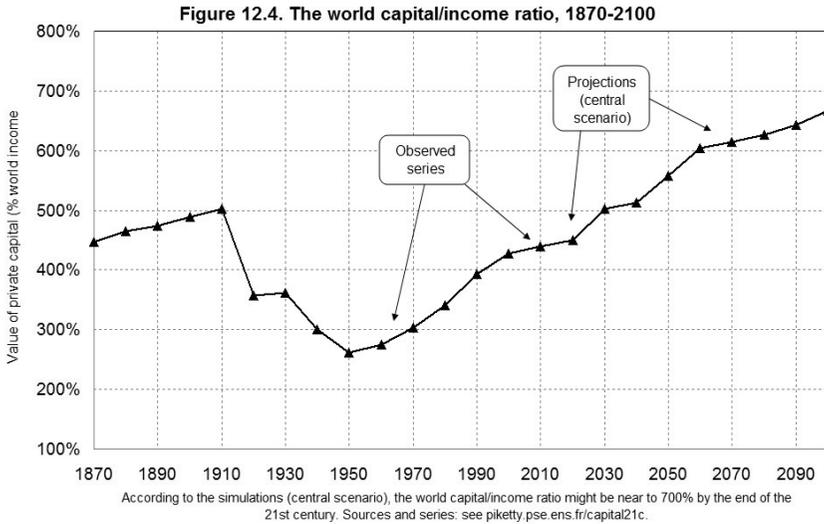
This excerpt refers to the *FT*'s own attempts to compare differing inequality measurement methods for the United Kingdom. Yet, also notice: Is Piketty not guilty of the very same methodological charge in his United States series, where he first enlisted the estate-tax figures of Kopczuk and Saez (2004) before shifting to the SCF-based survey estimate for the more recent period when selective use of the latter seemed to validate his narrative? Methodological concerns should weigh upon any evaluation of admittedly imprecise measurement tools, but such uneven and inconsistent applications suggest a severe confirmation bias at play throughout Piketty's book. He selectively pivots between methods that affirm his story while criticizing others for lesser indulgences in the same.

Indeed, most of the examples we have considered thus far display elements of confirmation bias, whether found in sloppy misstatements of simple tax rates and dates to augment Piketty's historical narrative or cherry-picked data points to construct a trend line that mirrors his predictions. Such errors, though sufficiently serious to call into question the claimed empirical soundness of his work, are quite distinct from acts of intentional mischief in the presentation of data. One final example warrants closer examination for reasons that may extend beyond simple biases, because it is qualitatively more dubious than the previous examples of questionable choices we have documented. We thus turn to the second "U-shaped" trend that Piketty enlists to support his central argument of " $r > g$ " and its implications.

As a multipronged attack on inequality and its theorized roots in the private capital stock, Piketty attempts to demonstrate his thesis by applying his argument globally and, with it, test his "second law" of capitalism, wherein he predicts an intrinsic tendency of the ratio of capital to national income to increase over time. Piketty presents a "world" depiction of this ratio for the years 1870 to 2100 in his figure 5.8, which also appears as figure 12.4 in a later chapter to illustrate a projected "international convergence" in the proportions of capital accumulation around the globe over the coming century (hence the title of Piketty's book). As the last ninety years of this graph are Piketty's predictions, we will focus only on the historical data.

Piketty’s historical interpretation of figure 5.8/12.4 takes on the same familiar narrative of the “U-shaped” trend he claims for the twentieth century, as figure 9 shows.

Figure 9. Global Capital-Income Ratio, As Estimated by Piketty



Source: Piketty (2014b), figure 12.4.

Here, he posits a global divergence of capital to national income at the peak of the Gilded Age, followed by a flattening of the capital stock from World War I and ensuing events including the imposition of heavy progressive taxation, and then a rebound from the 1970s to the present. The growth in capital to income ratios for the developing world is said to reflect this reconstitution of the capital stock as well as drive future divergence in the cumulative ratio as the global capital stock stabilizes around developed-world levels. Should this upswing continue on the course his forecast lays out, the “entire planet could look like Europe at the turn of the twentieth century, at least in terms of capital intensity” (Piketty 2014b, p. 196).

Piketty’s derivation of figure 5.8/12.4 poses a challenge, as it is poorly sourced and annotated in ways that are even less transparent than his U.S. wealth inequality example. His technical appendix annotates figure 5.8, the first iteration of this graph, as follows: “The series used to construct figure 5.8, replicated in the book on p. 196, are available in Table S12.4 (see appendix to chapter 12). All the details about the assumptions on which the series are build, especially for the period 2010–2100, are specified in the book, as well as in the corresponding excel file” (Piketty 2014c, p. 34).

When one turns to figure 12.4 and the corresponding Excel table S12.4, the accompanying note sends the reader on a circular return to the original note for figure 5.8: “These estimates are based on assumptions already presented in chapter 5 and in the appendix to chapter 5” (2014c, p. 71).

The product is nonetheless presented as clear evidence of the predicted “U-shaped” trend and thus validation of his theory of capital stock accumulation and, by extension, its theorized inequality link. In reconstructing this figure, it helps to examine Piketty’s derivation in steps, as questionable data decisions mar each. To begin, since Piketty is working with capital data at the national level (where such data exist), he needs to bring these figures together for a global metric. To do so, he aggregates national income data by region to determine its respective percentage of the “world output distribution” in a given year. He uses these percentages, in turn, to weight each region’s capital-to-income ratio, the cumulative result being the curve in the figure.

A closer examination of Piketty’s regional weights (found in table S12.4b of his data set) reveals the first of many problems. Piketty has few data points to perform this basic calibration and resorts to extreme dependence on linear interpolation to fill the gaps. Of fifteen decennial data points between 1870 and 2010, nine of them are actually interpolated. The source from the remaining six “decades” is referenced to another table, S1.1, which reveals them to be distributions for six individual years out of the 150 covered in the resulting graph, each presented as representative of its nearest decade mark. The following images of Piketty’s data tables (figure 10) illustrate the source years, along with his corresponding regional output weights. Interpolated years are highlighted, whereas weights are taken from individual years in the first table and transferred directly into the six noninterpolated rows.

While remaining fully cognizant that data limitations do indeed impinge upon our ability to assemble a comprehensive multicentury time series of this sort, to use only six individual years for a 150-year examination is almost as astounding as it is incomprehensible. Indeed, one of the largest gaps in Piketty’s output distribution places him without *any* data points with which to estimate his output distributions between 1913 and 1950 (note the distance between years in the first column of the first table shown in our figure 10). This deficiency leaves Piketty making the absurdly untenable assumption that the regional distribution of world output remained

static and fixed throughout both world wars and the Great Depression.

Figure 10. Piketty’s World Output Distribution Source Charts

	A	B	C	D	E	F
6						
7		World output	Europe	America	Africa	Asia
8	0	100%	18%	2%	8%	73%
9	1000	100%	14%	4%	11%	71%
10	1500	100%	25%	3%	8%	64%
11	1700	100%	30%	2%	7%	61%
12	1820	100%	33%	4%	5%	59%
13	1870	100%	46%	12%	4%	39%
14	1913	100%	47%	24%	3%	26%
15	1950	100%	39%	36%	4%	21%
16	1970	100%	40%	32%	4%	24%
17	1990	100%	34%	33%	4%	29%
18	2012	100%	25%	29%	4%	42%
19						

Table S12.4a. Private capital/national income ratio for the world, 1870-2100 (estimate) (series used fo figure 12.4)

private capital / national income	World	Europe	America	Africa	Asia
1880	100%	46%	18%	4%	32%
1890	100%	46%	18%	4%	32%
1900	100%	46%	18%	4%	32%
1910	100%	47%	24%	3%	26%
1920	100%	43%	30%	3%	23%
1930	100%	43%	30%	3%	23%
1940	100%	43%	30%	3%	23%
1950	100%	39%	36%	4%	21%
1960	100%	39%	34%	4%	23%
1970	100%	40%	32%	4%	24%
1980	100%	37%	33%	4%	27%
1990	100%	34%	33%	4%	29%
2000	100%	29%	31%	4%	36%

Source: Piketty (2014b), data appendix, tables S1.1 and S12.4.

Piketty turns next to the individual ratios of capital to national income by country, and again encounters a problem of nonexistent data. He at least partially acknowledges this instance in his text, as for most of the world there are “no truly reliable estimates” (2014b, p. 195) of private capital until the late twentieth century. Still, he is left with little more than figures for parts of Western Europe and North America across the 1870–2010 span, plus the addition of Japan, Australia, and New Zealand from 1970 onward. As with the regional output distributions he uses to weight these figures, the devil is in the missing details. Piketty simply “guesstimates” for the remaining countries by approximating where he expects them to be and assigning a value, as seen in the highlighted cells in our figure 11. And

naturally they all follow an assumed “U shape,” thereby reinforcing his expected trend when merged with the Western European and North American data.

More troubling, though, is a case of what appears to be some mischief afoot in Piketty’s approximations as he accounts for the communist regimes of the mid-twentieth century. Starting with the Soviet Union in 1920 and adding in Eastern Europe and China after 1950, Piketty arbitrarily reduces his estimates of their capital-to-income ratios to parity between the two indicators, thus implying a full reduction in the capital stock in these countries. This ratio reduction, as highlighted in dark gray in our figure 12, is represented as 100 percent, or 1:1 for parity, for the duration of each region’s rigid communist period.

While one may plausibly expect a scenario in which the onset of communism is highly disruptive to the capital stock, Piketty offers no explanation for this assumption in his annotation around the chart and no accounting for the unique ratio parity he assigns to the communist regions outside of a possible vague allusion appearing much later in the book.¹⁴ Nor does he attempt to justify the inclusion of an assumed data point from multiple communist economies with the premise of his “U-shaped” curve argument, which purports to illustrate the characteristics of global capitalism. In doing so, he evinces what may be a naïve *a priorism* toward the claimed effects of the collectivization of capital under communist ideological systems. The effect nonetheless registers prominently in Piketty’s figure 5.8/12.4. When evaluated for their effects upon the “global” index he purports to construct, his decisions with regard to the Soviet Union and other communist regimes exert a strong downward pull upon the “U-shaped” curve to coincide with its claimed mid-century trough.

¹⁴ In a later chapter of his book, Piketty describes the Soviet experiment thusly, possibly hinting at the basis of his underlying assumption: “By abolishing private ownership of the means of production, including land and buildings as well as industrial, financial, and business capital (other than a few individual plots of land and small cooperatives), the Soviet experiment simultaneously eliminated all private returns on capital” (2014b, p. 531). While this is perhaps an accurate expression of Soviet ideological claims, it is inexcusably naïve if not outright misleading to incorporate such claims into the construction of a historical time series purporting to illustrate operational characteristics of global capitalism. It is also negligent by its omission of the simultaneous economic effects of a state-driven collectivization of the capital stock, including pervasive economic inequality under the historical Soviet system. For a detailed discussion of this issue see Henderson, McNab, and Rózsás (2005).

Figure 11. Existent vs. Estimated Data for Piketty's Capital-Income Ratios

private capital / national income	Western Europe	Eastern Europe	Russia (+Ukraine/ Belarus/ Moldavia)	North America	Latin America	Northern Africa	Sub-Saharan Africa	China	India	Japan	Australia/NZ	Middle East (y.c. Turkey)	Central Asia	Other Asian countries
1870	680%	300%	300%	446%	300%	300%	200%	300%	300%	600%	300%	300%	300%	300%
1880	671%	325%	325%	437%	325%	325%	250%	325%	325%	600%	325%	325%	325%	325%
1890	643%	350%	350%	478%	350%	350%	300%	350%	350%	600%	350%	350%	350%	350%
1900	662%	375%	375%	448%	375%	375%	350%	375%	375%	600%	375%	375%	375%	375%
1910	659%	400%	400%	440%	400%	400%	400%	400%	400%	600%	400%	400%	400%	400%
1920	343%	200%	100%	407%	400%	400%	400%	400%	400%	600%	400%	400%	400%	400%
1930	386%	200%	100%	485%	300%	300%	300%	300%	300%	600%	400%	300%	300%	300%
1940	328%	200%	100%	328%	300%	300%	300%	300%	300%	500%	300%	300%	300%	300%
1950	232%	100%	100%	356%	300%	300%	300%	100%	300%	200%	300%	300%	300%	300%
1960	267%	100%	100%	361%	300%	300%	300%	100%	300%	300%	300%	300%	300%	300%
1970	285%	100%	100%	332%	400%	400%	400%	100%	400%	372%	343%	400%	400%	400%
1980	331%	100%	100%	357%	400%	400%	400%	100%	400%	531%	351%	400%	400%	400%
1990	403%	200%	200%	392%	400%	400%	400%	200%	400%	616%	407%	400%	400%	400%
2000	490%	300%	300%	447%	400%	400%	400%	300%	400%	586%	500%	400%	400%	400%
2010	545%	400%	400%	410%	400%	400%	400%	400%	400%	601%	518%	400%	400%	400%

Source: Piketty (2014b), data appendix, table TS12.4.

Figure 12. Assumption of Communist Capital-Income Parity in Piketty's Source Charts

private capital / national income	Western Europe	Eastern Europe	Russia (+Ukraine/ Belarus/ Moldavia)	North America	Latin America	Northern Africa	Sub-Saharan Africa	China	India	Japan	Australia/NZ	Middle East (y.c. Turkey)	Central Asia	Other Asian countries
1870	680%	300%	300%	446%	300%	300%	200%	300%	300%	600%	300%	300%	300%	300%
1880	671%	325%	325%	437%	325%	325%	250%	325%	325%	600%	325%	325%	325%	325%
1890	643%	350%	350%	478%	350%	350%	300%	350%	350%	600%	350%	350%	350%	350%
1900	662%	375%	375%	448%	375%	375%	350%	375%	375%	600%	375%	375%	375%	375%
1910	659%	400%	400%	440%	400%	400%	400%	400%	400%	600%	400%	400%	400%	400%
1920	343%	200%	100%	407%	400%	400%	400%	400%	400%	600%	400%	400%	400%	400%
1930	386%	200%	100%	485%	300%	300%	300%	300%	300%	600%	400%	300%	300%	300%
1940	328%	200%	100%	328%	300%	300%	300%	300%	300%	500%	300%	300%	300%	300%
1950	232%	100%	100%	356%	300%	300%	300%	100%	300%	200%	300%	300%	300%	300%
1960	267%	100%	100%	361%	300%	300%	300%	100%	300%	300%	300%	300%	300%	300%
1970	285%	100%	100%	332%	400%	400%	400%	100%	400%	372%	343%	400%	400%	400%
1980	331%	100%	100%	357%	400%	400%	400%	100%	400%	531%	351%	400%	400%	400%
1990	403%	200%	200%	392%	400%	400%	400%	200%	400%	616%	407%	400%	400%	400%
2000	490%	300%	300%	447%	400%	400%	400%	300%	400%	586%	500%	400%	400%	400%
2010	545%	400%	400%	410%	400%	400%	400%	400%	400%	601%	518%	400%	400%	400%

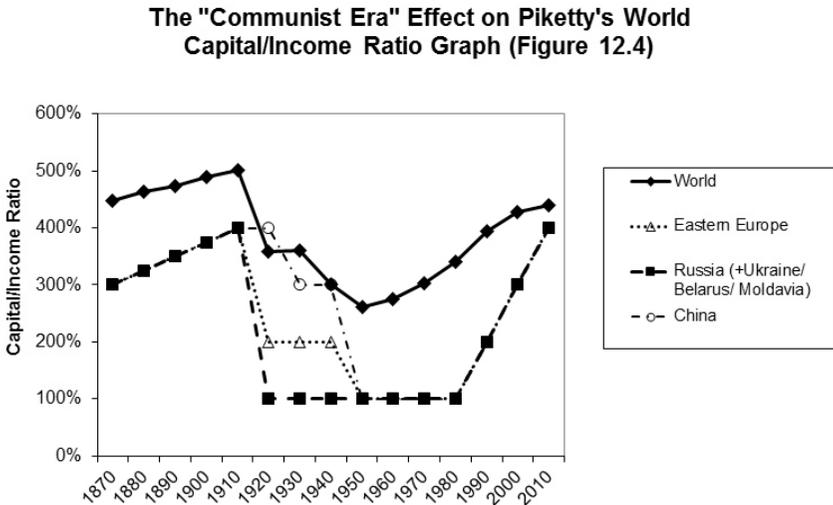
Note: External source data are depicted in white. Rough “guesstimation” figures as determined by Piketty are indicated in the lighter shade of gray. Piketty’s communist world data assumption of 1:1 C/I parity is indicated in the darker shade of gray for the applicable decades.

Source: Piketty (2014b), data appendix, table TS12.4.

The following graph, figure 13, displays the three communist regions overlaid with Piketty’s depiction of the global capital-to-income ratio. As the “world” index is, in effect, a time series constructed from weighted decennial averages, the inclusion of the communist bloc and the assumptions made about capital-to-income ratios therein end up being downwardly distortive to the constructed “world” trend line:

Nor is this distortion insignificant. At the 1950 trough, the weighting Piketty assigns to communist regimes is a sizable 17 percent of global output (compared to 27 percent for Western Europe and 29 percent for North America, making it the third largest bloc). With a 1:1 ratio falling far below either the North American ratio (3.6:1) or the Western European ratio (2.3:1) in the same year, the communist world’s inclusion in Piketty’s weighted average is biasing, severe, and unaccounted for in his book’s text or annotation. Rather, it is quietly hard-coded some three spreadsheets deep into his data file, well beyond the awareness of even the most diligent reader.

Figure 13. Effects of Communist Distortion on Piketty’s Global Capital-Income Ratio



Source: Chart created from Piketty (2014b), data appendix, table TS12.4.

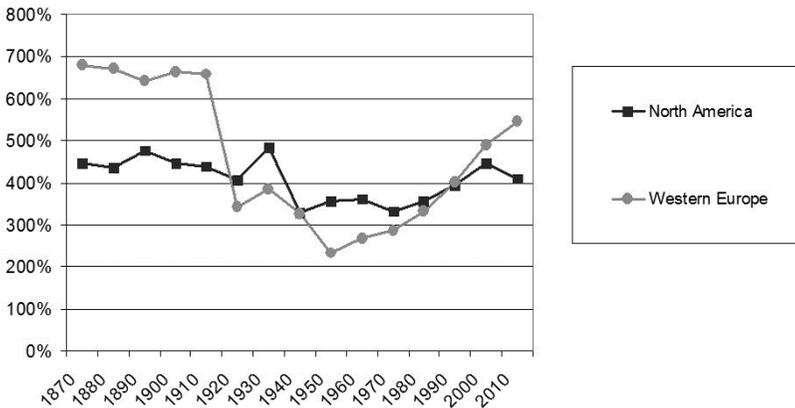
Similarly, the apparent rebound of the (formerly) communist countries after 1990 obviously pushes up the “world” capital-income ratio. Yet, most readers would be very surprised to see that this huge and identical shift among these countries (with a ratio of 1:1 rising to 4:1 over just three decades) does not derive from calculations

performed on underlying data sources, but rather occurs because Piketty typed them directly into his Excel sheet.

A point of contrast may also be seen in the disaggregated ratios for Western Europe and North America, which form the substance of his actual data sources in the cumulative “world” graph. While a “U shape” is indeed present for Western Europe when the two series are separated, a number of pertinent observations become clear. First, the bulk of the Western European “U shape” is attributable to a single decade’s movement coinciding with the First World War. Second and perhaps more important, the North American data actually defy the century long U-shaped prediction. As figure 14 shows, the North American ratio hovers around a comparatively stable and flat 4:1 for the entire twentieth century, with only mild fluctuations as happened in the Depression era. When considering these regional trends alone, Piketty applies a misleading descriptor to the North American case, portraying it as a U shape that is simply “smaller in amplitude in the United States than in Europe” (2014b, p. 154). In actuality, the century-long ratios for the two regions display two completely divergent trends, as figure 14 shows.

Figure 14. Comparative Capital-Income Ratios in Piketty’s Source Data

C/I ratio Western Europe vs. North America (Piketty data)



Source: Chart created from Piketty (2014b), data appendix, table TS12.4.

A true weighted average drawn only from these regions would show the effect of North America stabilizing and somewhat flattening the overall “U shape” found in the Western European data, particularly in conjunction with the former’s dramatically ascendant

share of total global economic output (and thus an increase in its weighting) at mid-century. Piketty's aggregate presentation shows the opposite effect, though, and precisely because it includes the distortive Soviet, Chinese, and Eastern European data points. The observed global "U shape" in Piketty's figure 5.8/12.4 is not a result of careful data analysis or even reasoned and articulated methodologies to account for its many data gaps, but rather the product of opaque and highly questionable assumptions Piketty has simply written into the file code his estimates for the rest of the world, with his unexplained communist region modifications exerting their strongest downward distortions precisely in sync with the 1950 bottom point of the claimed "U."

VII. Conclusions

In light of the foregoing analysis, it is increasingly evident that a sizable fissure exists between the bold data claims that Piketty makes at the outset of *Capital in the Twenty-First Century* and his actual empirics under closer scrutiny. We have highlighted a number of specific factual errors in the book as well as two of Piketty's larger evidentiary representations in the form of two claimed "U-shaped" patterns across the twentieth century, one for wealth inequality in the United States and another as it pertains to estimated capital-income ratios worldwide. Though these issues represent only a slice of Piketty's work, they exemplify the problems with the historical narrative that he uses to explain the twentieth century and to make predictive claims about the coming century. They severely undermine his justification for a heavy-handed global "correction" of prescriptive tax policy, and they invalidate several core pieces of empirical evidence behind his larger theoretical argument about the nature of the capital stock in a market economy.

Our study is by no means exhaustive, and the public prominence of Piketty's work has at minimum called attention to an important and ongoing line of research into economic history. We have focused on his claims about the United States in particular, and those most central to (1) his larger historical narrative and (2) United States tax policy, the latter being both a premise of the former and the primary implication of his policy prescriptions. While we have not exhaustively scrutinized Piketty's data claims in the other regions and countries he examines, we have noted that other investigations have found similar patterns of historical mistakes of fact and questionable methods in his data presentation (Sahlén and Furth 2014).

Whereas other data critiques, including that of the *Financial Times*, have raised important normative and methodological questions in Piketty's data presentations, the issues highlighted here suggest an even more fundamental problem of his reliance upon factually mistaken data claims, unsupported assertions of validity, and certain dubious chart constructions that only make sense in service to a preconceived narrative. The discrepancies we identify are pervasive in the book, beginning with misstatements of basic historical fact and extending to an abundance of political distortion and confirmation bias in his data selection and methodological choices. In the aforementioned use of communist data assumptions to accentuate the shape of a desired trend line, ostensibly explaining a hypothesized characteristic of capitalism, it is difficult to maintain a noble opinion of the scholarship involved. These problems afflict *Capital in the Twenty-First Century's* historical narrative in ways that are both large and small, although the frequency with which they occur is sufficient to warrant deep skepticism of the book as a whole and especially the many instances where Piketty substitutes an appeal to the reader's trust for annotation that is—at best—murky. It is therefore curious that the early reaction to *Capital* credited its data analysis despite other reservations with its contents and prescriptions. To the contrary, an abundance of questionable and problematic data claims may well mean that empirics are the book's weakest point.

References

- Auerbach, Alan, and Kevin Hassett. Forthcoming. "Capital Taxation in the 21st Century." *American Economic Review: Papers and Proceedings*.
- Fleener, Patrick. 1994. "A History and Overview of Estate Taxes in the United States." Washington, DC: The Tax Foundation, January.
- Furchtgott-Roth, Diana. 2014. "Piketty's Historic Minimum Wage Errors." *e21* (blog), Manhattan Institute, April 22.
- Giles, Chris. 2014. "Piketty Findings Undercut By Errors." *Financial Times*, May 23.
- Hayek, F. A. 1954. *Capitalism and the Historians*. Chicago: University of Chicago Press.
- Henderson, David R., Robert McNab, and Tamás Rózsás. 2005. "The Hidden Inequality in Socialism." *Independent Review*, 9(3): 389–412.
- Hoover, Herbert. 1952. *The Memoirs of Herbert Hoover: The Great Depression, 1929–1941*. New York: Macmillan.
- Kennickell, Arthur. 2009. "Ponds and Streams: Wealth and Income in the U.S., 1989 to 2007." Federal Reserve Board Working Paper.

- Kennickell, Arthur, Jesse Bricker, Brian K. Bucks, Traci L. Mach, and Kevin Moore. 2011. "Drowning or Weathering the Storm? Changes in Family Finances from 2007 to 2009." National Bureau of Economic Research Working Paper 16985.
- Kopczuk, Wojciech. 2014. "What Do We Know About Evolution of Top Wealth Shares in the United States?" NBER Working Paper No. 20734.
- Kopczuk, Wojciech, and Emmanuel Saez. 2004. "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate Tax Returns." *National Tax Journal*, 57(2, part 2): 445–88.
- Krugman, Paul. 2014a. "The Piketty Panic." *New York Times*, April 25.
- Krugman, Paul. 2014b. "On Inequality Denial." *New York Times*, June 1.
- McCloskey, Deirdre N. Forthcoming. "Measured, Unmeasured, Mismeasured, and Unjustified Pessimism: A Review Essay of Thomas Piketty's *Capital in the Twenty-First Century*." *Erasmus Journal of Philosophy and Ethics*.
- Piketty, Thomas. 2014a. "Addendum: Response to FT." May 28.
- Piketty, Thomas. 2014b. *Capital in the Twenty-First Century*. Cambridge, MA: Harvard University Press.
- Piketty, Thomas. 2014c. "Technical appendix of the book *Capital in the Twenty-First Century*." March.
- Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics*, 118(1): 1–39.
- Reynolds, Alan. 2006. *Income and Wealth*. London: Greenwood Press.
- Saez, Emmanuel, and Gabriel Zucman. 2014a. "Exploding Wealth Inequality in the United States." Washington, DC: Washington Center for Equitable Growth, October 20.
- Saez, Emmanuel, and Gabriel Zucman. 2014b. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." NBER Working Paper 20625.
- Sahlén, Malin, and Salim Furth. 2014. "Piketty Is Misleading about the Swedish Case." *Timbra*, November 7.
- Wolff, Edward N. 1994. "Trends in Household Wealth in the United States, 1962–83 and 1983–89." *Review of Income and Wealth*, 40(2): 143–74.
- Wolff, Edward N. 2010. "Recent Trends in Household Wealth in the United States: Rising Debt and the Middle-Class Squeeze—An Update to 2007." Levy Economics Institute Working Papers Series No. 159.