WORKING PAPERS ON

CAPITAL AS POWER

No. 2019/02

How the Rich are Different

Hierarchical Power as the Basis of Income Size and Class

Blair Fix

v.1 April 2019
v.2 October 2019

http://www.capitalaspower.com/?p=2652

How the Rich Are Different:

Hierarchical Power as the Basis of Income Size and Class

Blair Fix*

October 29, 2019

Abstract

This paper investigates a new approach to understanding personal and functional income distribution. I propose that hierarchical power — the command of subordinates in a hierarchy — is what distinguishes the rich from the poor and capitalists from workers. Specifically, I hypothesize that individual income increases with hierarchical power, as does the share of individual income earned from capitalist sources. I test this idea using evidence from US CEOs, as well as a numerical model that extrapolates the CEO data. The results indicate that income tends to increase with hierarchical power, as does the capitalist composition of income. This suggests that hierarchical power may be a determinant of both personal and functional income.

JEL Subject Codes: D31 (Personal Income, Wealth, and Their Distributions); D33 (Factor Income Distribution); B5 (Current Heterodox Approaches)

Keywords: hierarchy; power; functional income distribution; personal income distribution; inequality; capital as power; class

*blairfix@gmail.com

1 Introduction

Let me tell you about the very rich. They are different from you and me.

— F. Scott Fitzgerald (1926)

Yes, they have more money.

— Ernest Hemingway (1936)

What makes the rich different? Why do they earn more money than other people? And why does this money tend to come from property, not wages? These are the questions that underpin the study of personal (size-based) and functional (class-based) income distribution.

Responses to these questions have fallen roughly into two camps. Neoclassical economists argue that the rich are different because they are more productive. That this productivity comes largely from property then signals that property is itself productive. Marxists, in contrast, argue that the rich are different because they exploit workers. By owning the 'means of production', the rich are able to extract a surplus from workers. It is this exploitation, Marxists believe, that explains the greater income of the rich. And it is the ownership of the means of production that explains why the rich earn income largely from property.

After nearly a century of debate between these two schools of thought (with little conciliation), I believe it is time to look for an alternative approach. What makes the rich different, I propose, is not the productivity of their property or their exploitation of workers. Instead, I propose that the rich are different because of their greater control of subordinates — what I call 'hierarchical power'. I hypothesize that individuals with more hierarchical power tend to earn more income, and have a larger share of this income come from capitalist sources (Section 2).

To test this hypothesis, I look for a correlation between income size and hierarchical power, on the one hand, and a correlation between income composition and hierarchical power on the other. After discussing methods (Section 3), I analyze the income of US CEOs. I find that the relative income of US CEOs increases with their hierarchical power, as does the capitalist fraction of this income (Section 4). I then use a model to test the generality of this CEO evidence. This model suggests that in the United States, there is a three-way relation between income size, income composition and hierarchical power (Sections 5 and 6).

In short, I find evidence that hierarchical power is connected to both income size and income composition. At present, this is a correlation only. But it hints that hierarchical power may be an important determinant of both personal and functional income distribution.

1.1 The Motivating Problem

This paper is motivated by a long-standing problem. Despite more than a century of effort, we do not have a theory that adequately explains both income size (how much one earns) and income class (the source of one's income). We have three types of income distribution theory (below). Each has problems.

1. **Core Theories**: Marxist and neoclassical political economy
2. **Stochastic Models**: Models that generate skewed distributions using random shocks to individual income
3. **Power Theories**: Mostly qualitative descriptions of how power affects income

The hallmark of our core theories — Marxist and neoclassical political economy — is that they both assume value is produced. Neoclassical economists think both laborers and capitalists produce value. Each 'factor of production' then earns its (marginal) contribution to output (Clark 1899, Wicksteed 1894). Marxists agree that laborers produce value, but have different ideas about capitalists. According to Marx (1867), capitalists earn income by exploiting workers.

There are many problems with our core theories of distribution that I will not review here.[1] Instead, my concern is what these theories conclude about personal income (i.e. the size distribution of income). Neoclassical theory and Marxist theory both agree that labor produces value, and that this productivity is the source of labor income. Using this reasoning, neoclassical economists (Becker 1962, Mincer 1958, Schultz 1961) and Marxists (Rubin 1973) have concluded the same thing: if workers earn different incomes, they must have different productivity.[2] If we generalize this reasoning, it implies that workers' productivity should be as unequally distributed as their income. Yet this is not true. When workers' productivity is measured objectively, it fails to explain differences in income (Fix 2018d). This

[1] For problems with marginal productivity theory, see Cohen and Harcourt (2003), Felipe and Fisher (2003), Harcourt (2015), Hodgson (2005), Nitzan and Bichler (2009), Pullen (2009), Robinson (1953), Sraffa (1960). For problems with Marxist theory, see Nitzan and Bichler (2009), Samuelson (1971).

[2] There is a subtle distinction between neoclassical and Marxist theory. Neoclassical theory attributes labor income *directly* to productivity. But Marxist theory attributes income to the *value* of labor power. The latter is the labor time required to reproduce labor power. Since the labor power of more productive workers tends to take more to reproduce, more productive workers tend to earn higher wages.

leaves the 'productivist' aspect of neoclassical and Marxist theories at odds with the evidence.

In hindsight, this empirical shortcoming is understandable. The facts of personal income distribution were discovered after the core theories of income distribution were developed. It was late in the 19th century when Vilfredo Pareto (1897) showed that personal income was skewed and followed a power-law distribution. By the time these facts became well known (in the 20th century), neoclassical and Marxist theories of income distribution were firmly established.

While political economists were slow to react to Pareto's discovery, mathematicians soon looked for processes that could generate skewed distributions. They found that a simple random process could do the trick. If individuals' income grew randomly over time, it tended to create a skewed distribution of income. This process became know as a 'stochastic model' of income. For early stochastic models, see Champernowne (1953), Simon (1955), and Rutherford (1955). For more recent work, see Gabaix *et al.* (2016), Nirei and Aoki (2016), and Toda (2012).

These stochastic models are important because they show how the dynamics of individual income can lead to income inequality. The problem, though, is that these models do not explain *why* individuals earn what they do. And since stochastic models deal only with isolated individuals (rather than groups of individuals), they are not helpful for understanding how income relates to social class.

That brings us to theories of income distribution based on power. These theories propose that income differences result from asymmetries in social relations. The general idea is that people tend to use their power — their influence over others — for personal gain. The result is that concentrated power leads to income inequality. An incomplete list of people who have linked income to power would include Berle and Means (1932), Brown (1988), Commons (1924), Dugger (1989), Galbraith (1985), Huber *et al.* (2017), Lenski (1966), Mills (1956), Munkirs (1985), Nitzan and Bichler (2009), Peach (1987), Sidanius and Pratto (2001), Tool and Samuels (1989), Tool (2017), Veblen (1904, 1923), Weber (1978), and Wright (1979).

I find the 'power-income' hypothesis compelling. It avoids the trap of attributing income to productivity. And it avoids the atomism of stochastic models. And because it is concerned with social asymmetries, the power-income hypothesis naturally lends itself to the study of social class. For these reasons, I believe that the power-income hypothesis may provide a way to link the study of personal and functional income distribution.

But while promising (in my view), power theories have been plagued by a sim-

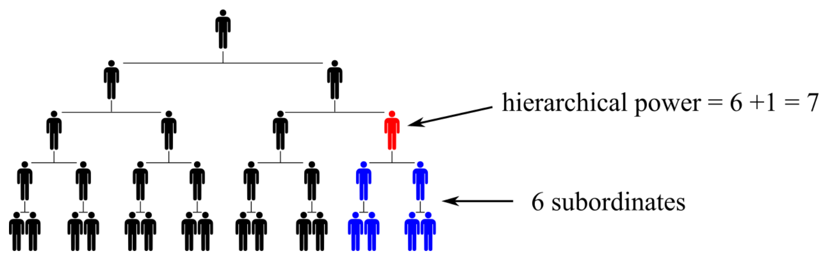

Figure 1: Measuring Hierarchical Power

This figure illustrates the calculation of hierarchical power. The red individual has 6 subordinates (blue). Using Eq. 1, the hierarchical power of this person equals 7.

---

ple problem. Because power has so many forms, it is difficult to measure. As a result, power theories of income distribution have tended to be qualitative. This, I believe is a major shortcoming.

The solution, I argue, is to reduce our scope and focus on power in a limited context. To measure power, I propose we focus only on 'hierarchical power' — the control of subordinates within a hierarchy.

2 A Theory of Income Distribution Based on Hierarchical Power

In this paper, I focus on a single dimension of power — the control over subordinates within a hierarchy. I call this 'hierarchical power'. I measure it as follows:

$$\text{hierarchical power} = \text{number of subordinates} + 1 \tag{1}$$

This definition of hierarchical power is structural, in that it deals only with one's abstract position within a hierarchy. I ignore other qualities (like charisma) that might affect power. I add one to the number of subordinates to signal that each individual has control over themselves. To count subordinates, we add both direct and indirect subordinates. Figure 1 shows an example calculation.

I propose that within hierarchies, individuals use their power to gain access to resources. The result is that income within the hierarchy should be proportional to hierarchical power. I call this the 'power-income hypothesis':

Power-Income Hypothesis: Within a hierarchy, individual income is proportional to hierarchical power

I have previously tested the power-income hypothesis using case studies of firm hierarchy. Fix (2018c) finds that in case-study firms, relative income tends to grow with hierarchical power (Fig. 2). Fix (2018d) further analyzes this data and finds

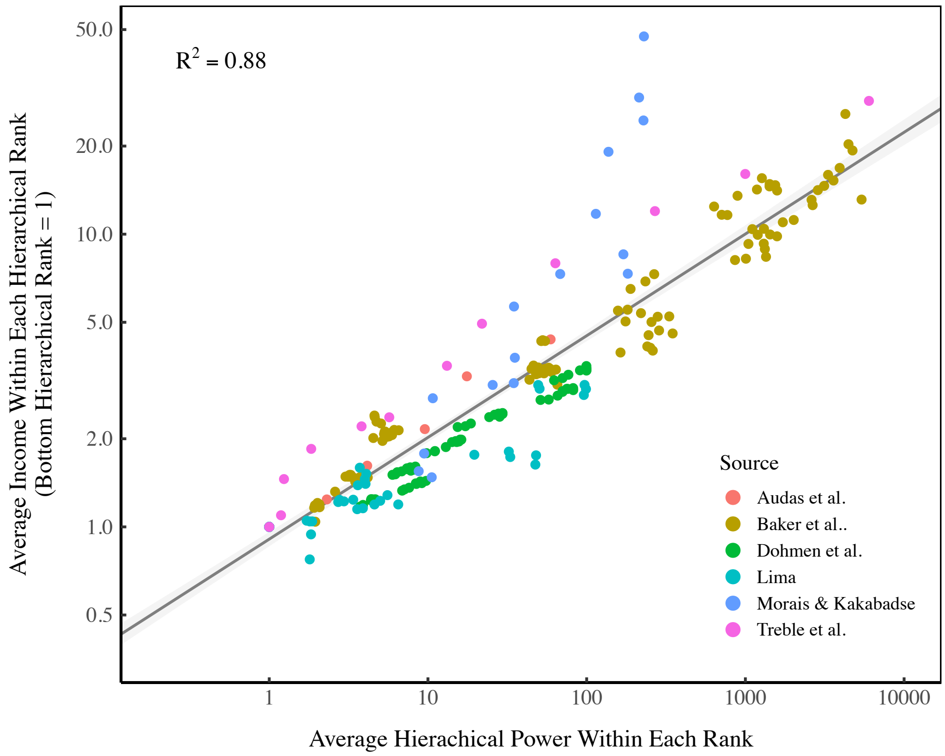

Figure 2: Average Income vs. Hierarchical Power Within Case-Study Firms

This figure shows data from six firm case studies (Audas *et al.* 2004, Baker *et al.* 1993, Dohmen *et al.* 2004, Lima 2000, Morais and Kakabadse 2014, Treble *et al.* 2001). The vertical axis shows average income within each hierarchical rank of each firm. Incomes are normalized so that income in the bottom rank equals one. The horizontal axis shows average hierarchical power of the individuals in each rank. Each point indicates a hierarchical rank in a given firm (in a given year). Grey regions around the regression indicate the 95% prediction interval. See Fix (2018c) for a detailed discussion of sources and methods.

---

that the income effect of hierarchical rank cannot be explained by well-known determinants of income like education, age, and firm experience.

2.1 Hierarchical Power and Income Class

This paper seeks to extend the power-income hypothesis to the study of class-based income. To do this, I draw on the work of Nitzan and Bichler (2009) and their concept of 'capital as power'.

Nitzan and Bichler observe that all societies have ideologies that legitimize power. These ideologies do three things. First, they legitimizes the power of rulers. Second, they justifies the rulers' income. Third, ideologies create a distinct income class for the rulers. The power of feudal kings, for instance, was legitimized by religion (the divine right of kings). This ideology then justified the king's income and gave it a distinct class — taxes.

Nitzan and Bichler argue that much remains the same in capitalist societies, except that the ideology is no longer religion. It is *ownership*. The ideology of ownership justifies the power and income of capitalist rulers, and it gives these rulers a separate income class. Owners earn profit. Non-owners earn wages.

I take this thinking and apply it to hierarchies. I argue that ownership is an ideology for justifying hierarchical power. To understand how this might work, imagine you buy all the shares of a corporation. As the sole owner, you now command the corporate hierarchy. In effect, you purchased hierarchical power. Once in command of the hierarchy, you have the power to distribute resources within it. You can divide the firm's income stream, keeping some for yourself and giving the rest to your subordinates. Like the feudal king, your income then gets its own class. As owner, you earn profit (capitalist income). Those you command earn wages (labor income).

This reasoning leads to a simple model of how capitalist income might relate to hierarchical power. Shown in Figure 3, a single owner commands the hierarchy, and uses his/her authority to divide the firm's income stream. The owner earns capitalist income. Everyone else earns labor income. Expanding on Nitzan and Bichler's concept of 'capital as power', I represent capital as the commodified ownership of the hierarchy.

2.2 A Capitalist Gradient Hypothesis

Our model in Figure 3 is intuitive, but likely too simple. The problem is that there is only a single owner. In modern firms, partial ownership is the norm. This means that ownership is divided among many people.

Does partial ownership mean that capitalists no longer control the corporate hierarchy? Berle and Means (1932) thought so. They argued that diffuse ownership caused capitalists to cede control to professional managers. The problem with Berle and Means' 'separation thesis', however, is that it assumes a dichotomy between owners and non-owners. But the truth is that over the 20th century, accounting practices have become more complex. Many owners now pay themselves a salary — a non-ownership income. And many employees earn income from stock options — a form of ownership income.

Instead of a dichotomy, what if there is a *gradient* of ownership within firm hierarchies? This would look like Figure 4. Here the firm has many owners. But

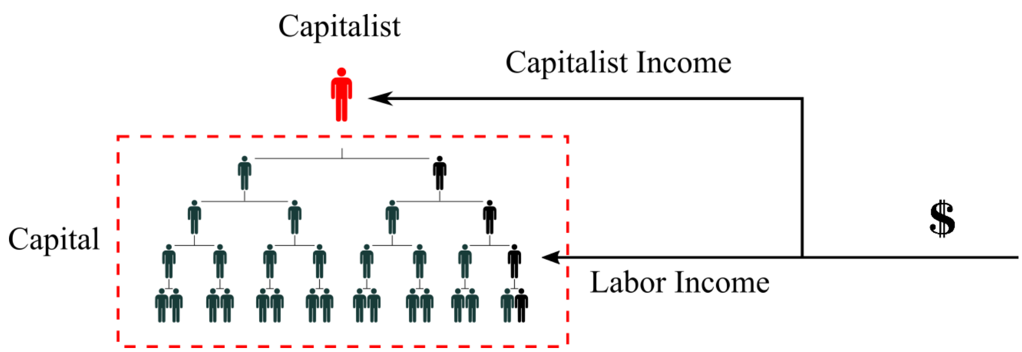

Figure 3: A Sole-Ownership Model of Capitalist Income in a Hierarchy

This figure shows how class-based income might relate to hierarchical rank. We suppose that a capitalist is the sole owner of a firm. This gives the capitalist the legal right to command the firm hierarchy. From this position of power, the capitalist divides the firm income stream and pays himself/herself capitalist income (profit). Everyone else earns labor income. Expanding on Nitzan and Bichler's (2009) concept of 'capital as power', I treat 'capital' as the commodification of the owner's hierarchical power.

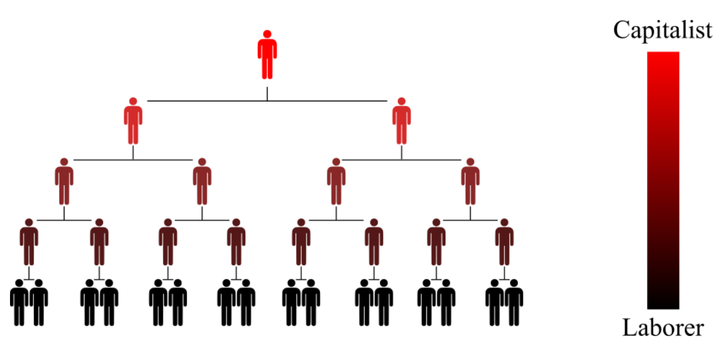

Figure 4: A Gradient Model of Capitalist Income in a Hierarchy

This figure shows a gradient model of class-based income. Ownership is distributed among many individuals but remains connected to hierarchical power. Top-ranked individuals have large ownership shares, while bottom-ranked individuals have small ownership shares. Thus capitalist income fraction increases as a function of hierarchical power. I call this the 'capitalist gradient hypothesis'.

ownership is still related to hierarchical power. Those at the top have a large ownership stake while those at the bottom have a small one. With this spread of ownership comes a spread of capitalist income. Those at the top still earn mostly capitalist income, and those at the bottom still earn mostly labor income. But in between, the lines are blurred. In short, income composition is (at least in part) a function of hierarchical power. I call this the 'capitalist gradient hypothesis':

Capitalist Gradient Hypothesis: The capitalist fraction of individual income increases with hierarchical power

We can interpret this hypothesis a few different ways. First, we could apply it to a single firm. But this is realistic only for firms that are fully employee owned. Such firms do exist, but are not the norm. Second, we could apply the capitalist gradient hypothesis to employee stock ownership plans. These give partial ownership to a firm's employees. The problem is that employee ownership makes up about 4% of total US market capitalization.[3] Thus it is not the main source of capitalist income.

I interpret the capitalist gradient hypothesis at the *societal* level. I admit that the ownership structure of any given firm is complex. I also admit that individuals earn capitalist income from a variety of firms. But at the societal level, I hypothesize that individuals with more hierarchical power (within a given firm) earn a larger fraction of their income from capitalist sources.

3 Methods: Classifying Income

To test the capitalist gradient hypothesis, we must put income into classes. How we do this depends on our ideas about property and ownership. I discuss here two ways of classifying income — one that makes sense from a neoclassical standpoint, and one that makes sense if we treat ownership as a tool for power.

In both approaches, labor income is defined the same way. It is income that does not come from ownership. The sticking point is capitalist income. Does capitalist income come from any form of ownership. Or just *some* forms? The answer depends on our preconceptions about property. If, like neoclassical economists, we think property is a thing that produces value, then all ownership is the same. We should use the income class system in Table 1 and treat all property income as capitalist.

[3] In 2017, employee ownership plans had total assets of roughly $1.3 trillion (NCEO 2017), while total US market capitalization was roughly $30 trillion, according to the Russel 3000 index.

The problem with this system is that it mixes two forms of ownership — *scalable* and *non-scalable*. Corporate ownership is scalable. Corporations range from tiny shell companies to giant firms like Walmart. They can be any size. But by a quirk of the law, this is not true for proprietor and rental ownership. These forms of ownership are non-scalable. By definition, rent can flow only to unincorporated individuals. And proprietors are mostly the self-employed. If either a landlord or a proprietor grows their business, they will incorporate and their income will be reclassified as profit. By legal quirk, then, landlord and proprietor forms of ownership are inherently small scale.

Why does the distinction between scalable and non-scalable ownership matter? In neoclassical theory it doesn't, because all property is treated as productive. But in my theory of income distribution based on hierarchical power, the distinction matters. I propose that capitalists are not simply those who own property. Instead, capitalists are those who own *hierarchy*. This means we want to distinguish between corporate and non-corporate ownership. Because corporate owners can own large firms, they can control large hierarchies. I therefore define corporate owners as capitalists. In contrast, non-corporate owners cannot control hierarchies because they own small firms. Thus I put non-corporate owners into separate category of small-scale ownership.

The three-class system shown in Table 2 is my preferred class system for testing the capitalist gradient hypothesis. Unfortunately the available empirical data does not (for the most part) fit cleanly into these categories.

3.1 Empirical Measures of Class-Based Income

Because of data constraints, I use the measures of capitalist income shown in Table 3.

When studying CEO income (Section 4), I measure capitalist income using the realized gains from stock options. This is not ideal because it excludes income from CEOs' personal investments. Still, stock options are a significant source of capitalist income. The realized gain is the difference between the option value and the market value at the exercise time. This gain is taxable income (Hopkins and Lazonick 2016).

To measure the distribution of US income by class (Section 5), I use data from the World Inequality Database (WID). The WID data has unparalleled depth, but comes with some caveats. WID income classes (Table 3) do not align with any of my own (Tables 1 and 2). The WID data divides proprietor income into capital and labor components. This means some proprietor income is classified as labor

Table 1: Income Classes if all Owners are Capitalists

Income Type	Symbol	Definition	Composition
Labor Income	L	monetary returns to *non-owners*	wages/salaries + pensions
Capitalist Income	K1	monetary returns to *owners*	distributed corporate profit + interest + rents + proprietor income + capital gains on all property

Table 2: Income Classes if Only Corporate Owners are Capitalists

Income Type	Symbol	Definition	Composition
Labor Income	L	monetary returns to *non-owners*	wages/salaries + pensions
Small-Scale Owner Income		monetary returns to *non-corporate* ownership	rents + proprietor income + capital gains on rental or proprietor property
Capitalist Income	K2	monetary returns to *corporate* ownership	distributed corporate profit + interest+ capital gains on corporate equity and bonds

'Distributed corporate profits' are paid to individuals. This includes dividends from M corporations and profit from S corporations.

income and some is classified as capitalist income. The rationale is that part of a proprietor's income comes from their property, and part comes from their labor. This method comes from Piketty *et al.* (2017b), who are the primary source of the WID data. Piketty et al. assign to proprietors the same capital-labor mix as the corporate sector.

The WID labor income series L* (Table 3) is like my definition L (Table 1), but with some proprietor income mixed in. Similarly, the WID capitalist income series K1* is like my definition K1, but with some proprietor income mixed in. When possible, I construct my own measure of capitalist income, K2 (Table 3). But due to WID data constraints, I can do this only for certain types of analysis.

As shown in Table 3, I use and compare different measures of capitalist income throughout the paper. This is by necessity. The available data relating hierarchical power to income size and income class is scarce. To test my hypotheses, I use

Table 3: Methods for Measuring Class-Based Income in the United States

Measure	Symbol	Source	Composition	Use
Labor Income	L*	World Inequality Database (WID)	wages/salaries + pensions + 'labor' portion of proprietor income	Fig. 11
Capitalist Income of CEOs	K_{CEO}	Execucomp	realized gains from stock options	Fig. 7, basis for hierarchy model
Capitalist Income (All Ownership)	K1*	World Inequality Database (WID)	distributed corporate profit + interest + rents + 'capital' portion of proprietor income + capital gains on all property	Figs. 12 and 14
Capitalist Income (Corporate Ownership Only)	K2	World Inequality Database (WID)	distributed corporate profit + interest + capital gains on corporate equity	Fig. 14

Notes: Pension income includes employee and employer contributions. It excludes asset income from pension investments. 'Distributed corporate profits' are paid to individuals. This includes dividends from M corporations and profit from S corporations. I focus on distributed corporate profits because I am interested in personal income. The other forms of profit (taxed profit and retained earnings) do not flow directly to individuals. For sources and methods, see the Appendix.

measures of class-based income that are non-ideal , and I compare data sets that are not perfectly compatible. Given the coarseness of these methods, we should consider the results preliminary.

4 A Case Study of US CEOs

To study how income size and income class relate to hierarchical power, I use data from US CEOs. I turn to CEOs because their role as corporate leaders provides a simple way to measure their hierarchical power. The hierarchical power of a CEO is equivalent to the size of the firm they command. For instance, if a CEO commands a firm with 100 employees, then 99 of them are subordinate to the CEO. So the CEO's hierarchical power is $99+1=100$. Figure 5 shows this equivalence between firm size and the hierarchical power of the CEO. I use this method to investigate how CEO income and the capitalist fraction of this income relate to hierarchical power.

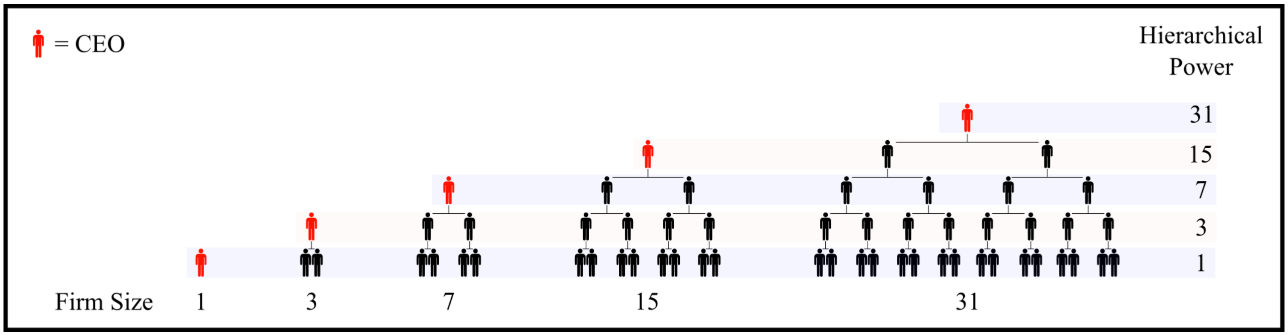

Figure 5: Using Firm Size to Measure the Hierarchical Power of CEOs

This figure illustrates how we can use firm size to measure the hierarchical power of CEOs. Each hierarchy represents a different firm, with the CEO at the top (red). If firm size is x, each CEO has $x-1$ subordinates. Since hierarchical power equals the number of subordinates plus one, CEO hierarchical power is equal to firm size x.

The CEO data comes from Compustat and Execucomp and covers the years 2006–2016. For sources and methods, see the Appendix.

4.1 CEO Pay vs. Hierarchical Power

To test how CEO income relates to hierarchical power, I regress the following equation onto CEO data

$$\ln(C_{f,t}) = \beta \ln(P_{f,t}) + \varepsilon \tag{2}$$

Here C is the CEO pay ratio — the ratio between CEO income and average pay within the firm. P is the hierarchical power of the CEO, measured using firm size. The parameter β, which I call the 'power-income exponent', measures how rapidly CEO income grows with hierarchical power. The parameter ε (which I do not report) indicates the value for $\ln(C)$ when $\ln(P) = 0$. The subscripts f and t indicate the firm and year, respectively.

I use Eq. 2 for three different regressions. I first regress CEO pay onto hierarchical power for all firm-year observations (all values of f and t). I report the resulting power-income exponent as β. Results are shown in Figure 6A. Each data point indicates a CEO observation in a given year.

Next, I regress Eq. 2 onto CEO data in each year t. The resulting power-income exponent β_t indicates how rapidly CEO income increases with hierarchical power across firms in a given year. Results are shown in Figure 6B. Between 2006 and 2016, β_t trended slightly upwards.

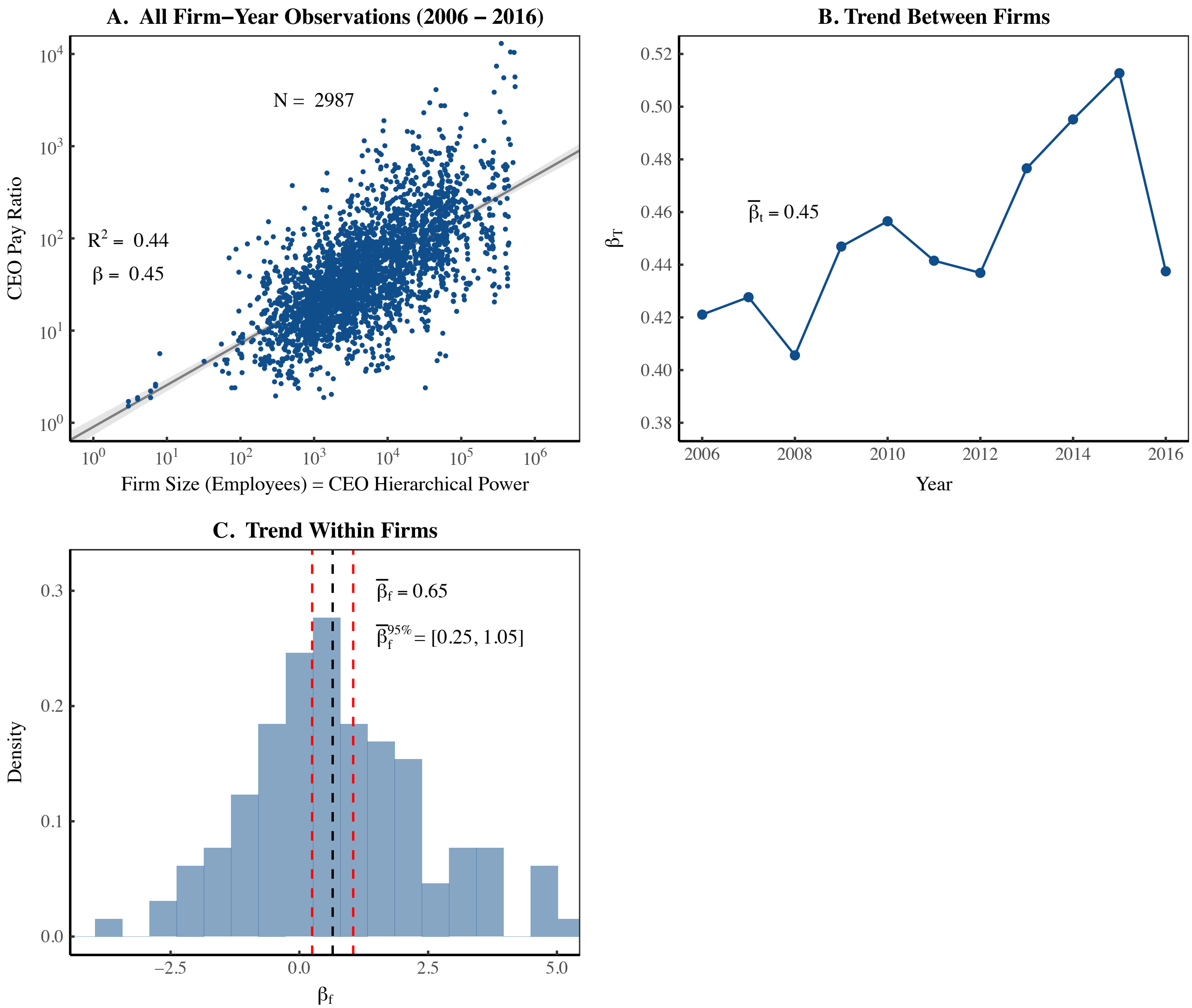

Figure 6: Relative Income vs. Hierarchical Power Among US CEOs

This figure analyzes how the relative income of US CEOs (as measured by the CEO pay ratio) relates to hierarchical power. The data includes roughly 3000 CEO observations over the years 2006–2016. Panel A shows data over all years. Panel B shows how the trend between firms has changed over time. β_t is the power-income exponent (between firms) in a given year. Panel C shows the trend within firms. β_f is the power-income exponent within a specific firm over time. The histogram shows the distribution of β_f for all firms. Vertical lines show the mean of β_f with the associated confidence interval. For data sources and methods, see the Appendix.

Table 4: Power-Income Exponents for US CEOs

	Statistic	Value	95% CI	R^2	p
All Compustat Firm-Year Observations	β	0.45	$0.43 < \beta < 0.47$	0.44	0
Annual Trend Between Compustat Firms	$\bar{\beta}_t$	0.45	$0.43 < \bar{\beta}_t < 0.47$	0.45 ($\bar{R}^2$)	1.1×10^{-23} ($\bar{p}$)
Trend Over Time Within Compustat Firms	$\bar{\beta}_f$	0.65	$0.25 < \bar{\beta}_f < 1.04$	0.23 ($\bar{R}^2$)	0.34 ($\bar{p}$)

Statistics are for the regression equation shown in Eq. 2. β is the power-income exponent across all data. β_t is the power-income exponent between firms in a given year. β_f is the power-income exponent within a firm (across time).

Lastly, I regress Eq. 2 onto CEO data while holding the firm f constant. This regression indicates the trend over time (within a single firm) between CEO pay and hierarchical power. I report the resulting power-income exponent as β_f. Results are shown in Figure 6C. The histogram indicates the distribution of β_f across all firms. The vertical lines show the mean of β_f with the 95% confidence interval.

Table 4 shows summary statistics for the analysis. The trend between firms (β_t) is statistically consistent with the trend within firms (β_t). Note, however that the trend within firms is, on average, not statistically significant. This is likely due to the short time period of the analysis.

To summarize, I find that among the US CEOs studied here, relative income increases consistently with hierarchical power. This appears to be true both between and within firms.

4.2 Capitalist Fraction of CEO Pay vs. Hierarchical Power

I define the capitalist fraction of CEO income as the portion of total compensation coming from stock options:

$$\text{Capitalist Fraction of CEO Income} = \frac{\text{Realized Gains from Stock Options}}{\text{Total Compensation}} \tag{3}$$

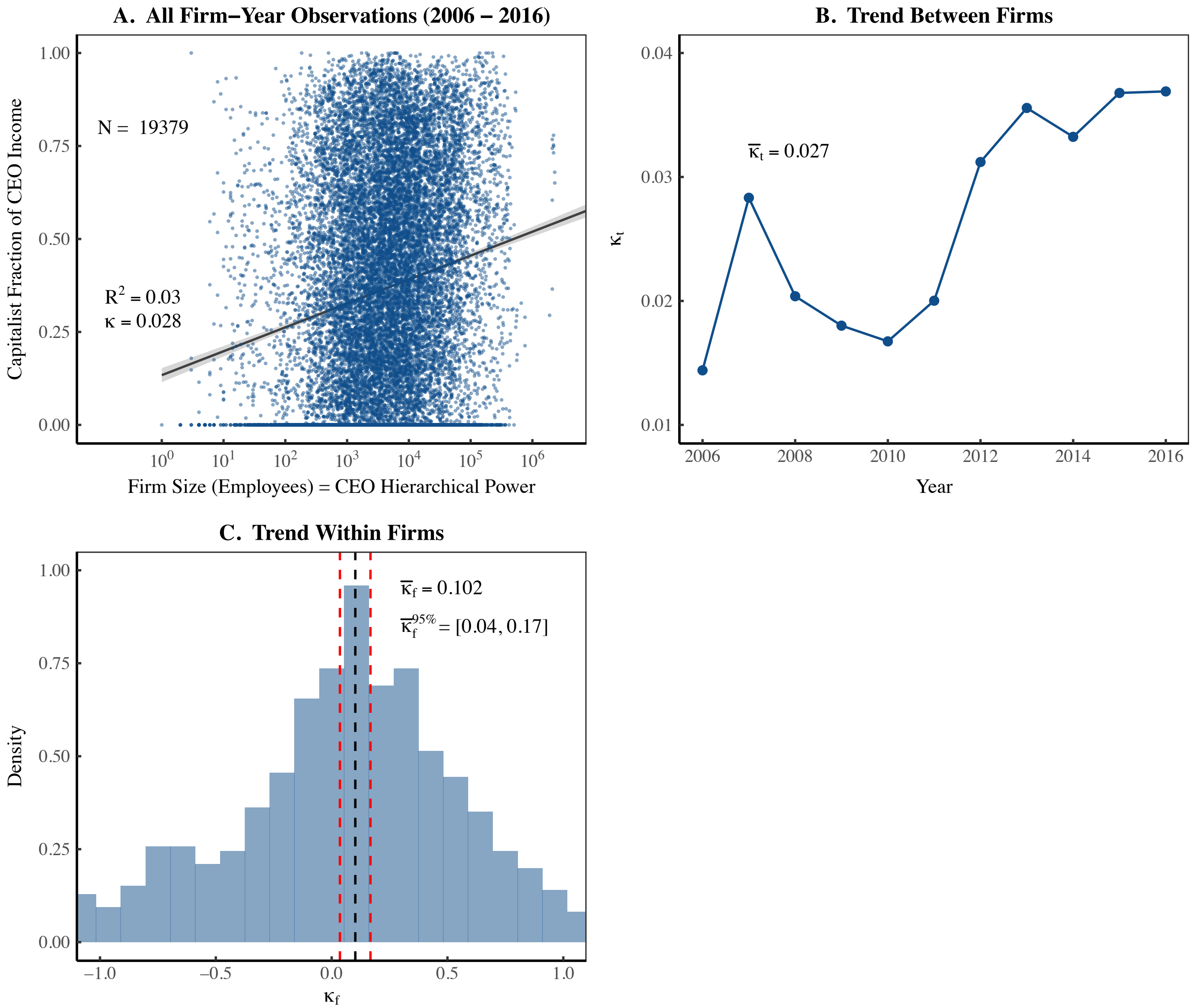

Figure 7: Capitalist Fraction of Income vs. Hierarchical Power Among US CEOs

This figure analyzes how the capitalist fraction of CEO income relates to hierarchical power. The data includes roughly 20,000 CEO observations over the years 2006–2016. Panel A shows data over all years. Panel B shows how the trend between firms has changed over time. κ_t is the capitalist-gradient slope (between firms) in a given year. Panel C shows the trend within firms. κ_f is the capitalist-gradient slope within a specific firm over time. The histogram shows the distribution of κ_f for all firms. Vertical lines show the mean of κ_f with the associated confidence interval. For data sources and methods, see the Appendix.

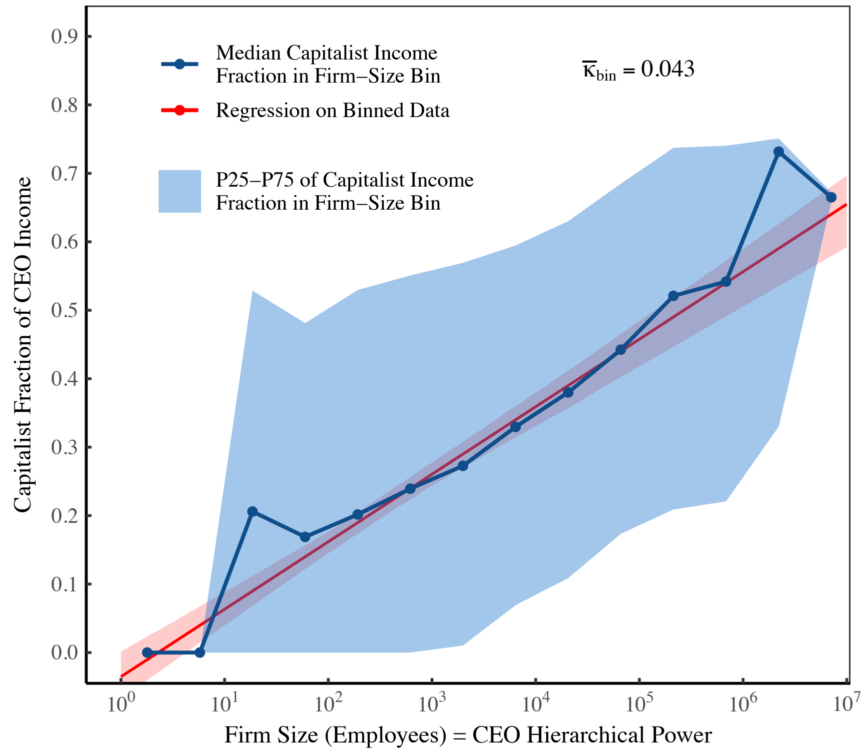

Figure 8: A Binned Analysis of the Capitalist Fraction of CEO Income

This figure shows a binned analysis of the relation between the capitalist fraction of CEO income and the hierarchical power of CEOs. I group the CEO data into firm-size bins that are log spaced. Within each bin, I plot the median of the capitalist fraction of income against the midpoint of the firm-size bin. The red line indicates the regressed trend on this binned data. To incorporate uncertainty caused by the choice of bin size, I regress Eq. 5 onto CEO data for a variety of different bin sizes. The red shaded region indicates the 95% confidence interval of the regression.

To test how the capitalist fraction of CEO income relates to hierarchical power, I regress the following equation onto CEO data:

$$K_{f,t} = \kappa \ln(P_{f,t}) + \varepsilon \tag{4}$$

Here K is the capitalist fraction of the CEO's income. P is the hierarchical power of the CEO, measured using firm size. The parameter κ, which I call the 'capitalist-gradient slope', measures how rapidly the capitalist fraction of CEO income grows with hierarchical power. The parameter ε (which I do not report) indicates the value for K when $\ln(P) = 0$. The subscripts f and t indicate the firm and year, respectively.

Table 5: CEO Capitalist Income Fraction vs. Hierarchical Power

	Statistic	Value	95% CI	R^2	p
All Compustat Firm-Year Observations	κ	0.028	$0.0258 < \kappa < 0.0301$	0.03	1.4×10^{-130}
Annual Trend Between Compustat Firms	$\bar{\kappa}_t$	0.027	$0.0216 < \bar{\kappa}_t < 0.0313$	0.03 ($\bar{R^2}$)	2×10^{-4} ($\bar{p}$)
Trend Over Time Within Compustat Firms	$\bar{\kappa}_f$	0.102	$0.0366 < \bar{\kappa}_f < 0.1678$	0.19 ($\bar{R^2}$)	0.366 ($\bar{p}$)
Binned Analysis of All Compustat Firm-Year Observations	$\bar{\kappa}_{bin}$	0.043	$0.0426 < \bar{\kappa}_{bin} < 0.043$	0.92 ($\bar{R^2}$)	1.9×10^{-5} ($\bar{p}$)

Statistics are for the regression equation shown in Eq. 4. κ is the capitalist-gradient slope across all data. κ_t is the capitalist-gradient slope between firms in a given year. κ_t is the capitalist-gradient slope within a firm (across time). κ_{bin} is the capitalist gradient slope for a binned analysis (by firm size) of CEO data.

I use Eq. 4 for three different regressions. I first regress the capitalist fraction of CEO pay onto hierarchical power for all firm-year observations (all values of f and t). I report the resulting capitalist-gradient slope as κ. Results are shown in Figure 7A. Each data point indicates a CEO observation in a given year.

Next, I regress Eq. 4 onto CEO data in each year t. The resulting capitalist-gradient slope κ_t indicates how rapidly the capitalist fraction of CEO income increases with hierarchical power across firms in each year. Results are shown in Figure 7B. Between 2006 and 2016, β_t trended slightly upwards.

Lastly, I regress Eq. 4 onto CEO data while holding the firm f constant. This regression indicates the trend over time (within a single firm) between the capitalist fraction of CEO pay and hierarchical power. I report the resulting capitalist-gradient slope as κ_f. Results are shown in Figure 7C. The histogram indicates the distribution of κ_f across all firms. The vertical lines show the mean of κ_f with the 95% confidence interval.

Table 5 shows summary statistics for the analysis. The trend between firms (κ_t) is smaller than the trend within firms (κ_f). Note, however that the trend within firms is, on average, not statistically significant. The greater value of κ_f is likely due, in part, to a secular increase in the capitalist fraction of CEO income over the time studied here.

As visualized in Figure 7A, the trend between the capitalist fraction of CEO income and hierarchical power is noisy. To get a better sense for this trend, I conduct a binned analysis of the CEO data. I group CEO data into firm-size bins that are log spaced. Within each bin, I calculate the median of the capitalist fraction of income, as well as the 25th and 75th percentiles. Results are shown in Figure 8.

Onto this binned data, I regress the following equation:

$$\tilde{K}_{bin} = \kappa \ln(P_{bin}) + \varepsilon \tag{5}$$

Here $\tilde{K}_{bin}$ is the median of the capitalist fraction of CEO income within each bin. P_{bin} is the hierarchical power (of the CEO) at the midpoint of the firm-size bin. To incorporate uncertainty caused by the choice of bin size, I regress Eq. 5 onto CEO data for a variety of different bin sizes. The regressed trend is shown in Figure 8 and summarized in Table 5.

This binned analysis indicates that the capitalist fraction of CEO income grows consistently with hierarchical power. Note that the capitalist-gradient slope for this binned analysis is steeper than the slope estimated from raw data. This is because the binned analysis gives equal weight to each firm-size bin. In contrast, the regression on raw data gives little weight to small firms, which are sparse in our data sample.

To summarize, the results suggest that among US CEOs, the capitalist fraction of income increases consistently with hierarchical power. This appears to be true both between and within firms. Furthermore, a simple log-linear trend emerges (between capitalist income fraction and hierarchical power) when we study binned data.

5 From CEOs to the General Population: Extending the Evidence

The CEO evidence suggests that income size, as well as the capitalist fraction of this income, increase with hierarchical power. My next step is to test if these trends extend to the general US population. To do this, I propose the following method:

1. Assume CEO trends extend to the general pubic
2. Simulate the implied distribution of income
3. Compare this simulation to US data

This test works by inference. We use a model to predict the distribution of US income that should occur if the CEO relation between income size, income class and hierarchical power extends to the general population. If the model's predicted distribution of income is consistent with the empirical data, we can then plausibly infer that the CEO trends are also found among the general US population. I use this model-based method because data for hierarchical power among the general US population is currently unavailable (to my knowledge).

5.1 Using A Model to Extrapolate CEO Trends

To extrapolate the CEO evidence to the general population, I use a numerical model developed in Fix (2018b). This model takes the CEO data (as well as firm case-study data) as inputs and then predicts the distribution of income that should occur if CEO trends extend to the US population. The model's parameters are determined from micro-level data (from firms) . I do not 'tune' the model to produce desired results. From the inputted micro-level data, the model then predicts the macro distribution of income.

The model has three steps, summarized below. (See the Appendix for technical details). In Step 1, the model simulates the hierarchical structure of Compustat firms (the sample of firms used to study CEOs). In Step 2, the model generalizes this simulation to a more representative size distribution of firms. The model then simulates individual income for people employed in the US private sector. In Step 3, the model simulates the class component of each person's income.

Step 1: Simulate the hierarchical structure of Compustat firms. To simulate the hierarchical structure of Compustat firms, the model first simulates the employment hierarchy in each firm. To do this, the model assumes that Compustat firms have the same hierarchical 'shape' as the average found in case-study firms. To simulate the pay hierarchy, the model assumes that income is proportional to hierarchical power. The model then uses the observed CEO pay ratio to infer the pay hierarchy within each Compustat firm. The result is a simulation of the hierarchical structure of each Compustat firm.

Step 2: Generalize the simulation to the US firm population. The next step is to generalize the Compustat simulation to a size distribution of firms that is more representative of the US. I simulate the size distribution of US firms using a power-law distribution. I then use the information from step 1 to simulate the hierarchical structure of these firms. The result is a simulation of the hierarchical structure of the US private sector. The model simulates personal income for roughly 20 million individuals.

Step 3. Simulate class-based income. To simulate the class composition of individual income, the model fits functions to the binned analysis of the capitalist

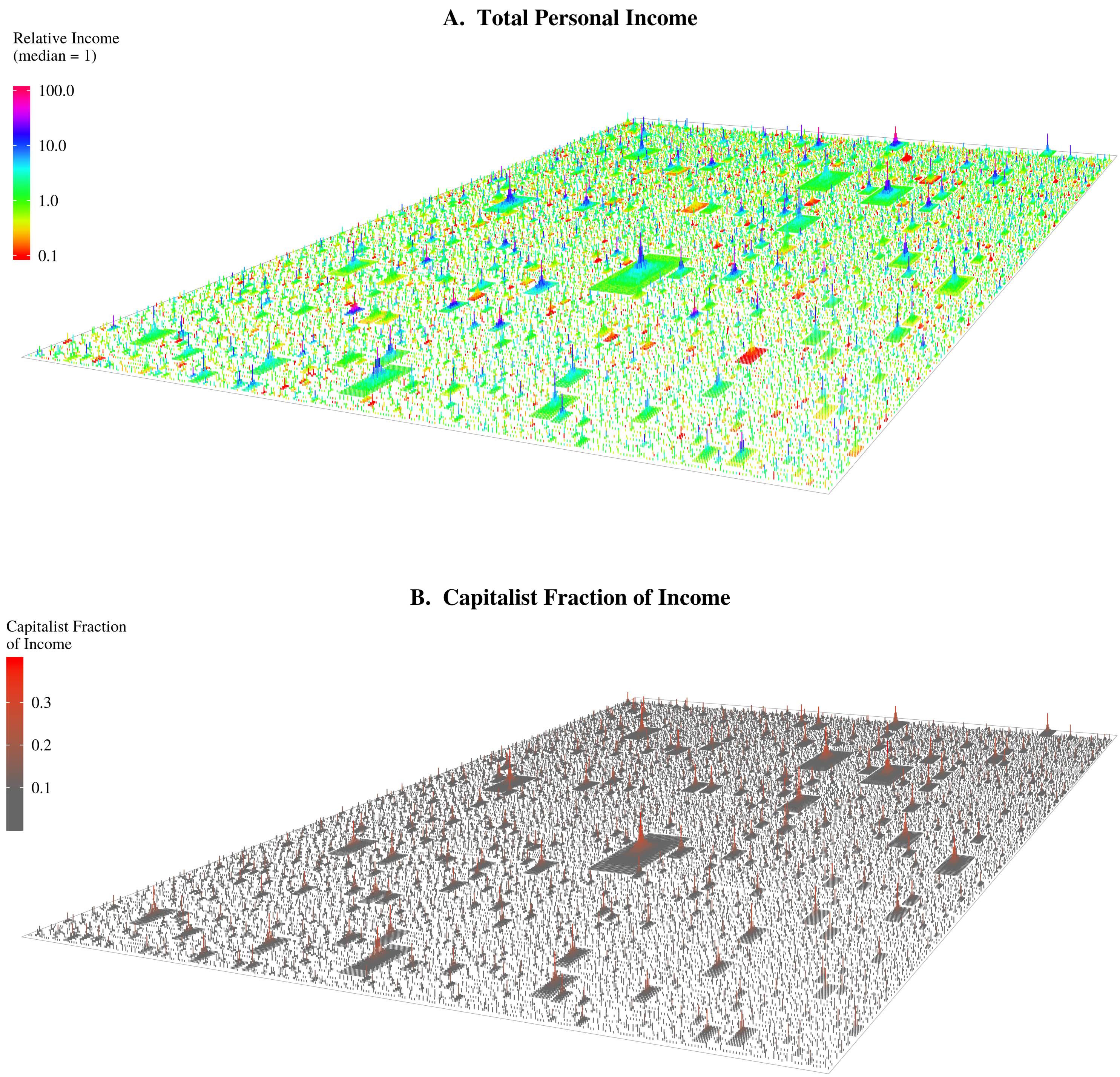

Figure 9: A Landscape View of the US Hierarchy Model

This figure visualizes the US hierarchy model as a landscape of three-dimensional firms. Each pyramid represents a single firm, with size indicating the number of employees and height corresponding to the number of hierarchical levels. Panel A uses color to indicate income relative to the median. Panel B uses shades of red to indicate the capitalist fraction of individuals' income. This visual shows 20,000 firms. The actual model uses 1 million firms to simulate the US firm population.

fraction of CEO income (Fig. 8). The model then uses these functions to simulate the class component of individual income as it relates to hierarchical power. This returns a capitalist and labor component for the income of each individual.

Visualizing the model. Figure 9 visualizes the model's output as a landscape, with firms displayed as pyramids. We can see here the main features of the model. Income increases with hierarchical rank (Fig. 9A), as does the capitalist fraction of income (Fig. 9B). We can also see the size distribution of firms. Most firms are small, but there are a few behemoths. Note that top earners are located mostly at the tops of large firms. These top-ranked individuals also have the largest capitalist component of income.

Testing the model. I test the model in two ways. I first compare to US data the model's predicted distribution of personal, labor and capitalist income (Section 5.2). I then compare to US data the model's predicted relation between income size and income composition (Section 5.3). The model is stochastic, so its results vary randomly over each iteration. To capture this variation, I run the model many times and measure both the average result and the range of variation.

5.2 Model Predictions for the Size Distribution of US Income By Class

Figures 10–13 show the model's predictions for the distribution of US personal, labor and capitalist income. I compare these predictions to US data over the years 2006–2014.

Summary Statistics

I use three different summary statistics (the Gini index, the top 1% share, and the power-law exponent of the top 1%) to compare the model to US data. Each statistic is sensitive to a different part of the income distribution.

The Gini index (Figs. 10A–12A) is sensitive to income dispersion in the 'body' of the distribution. The model accurately predicts the Gini index of US personal income (Fig. 10A) and capitalist income (Fig. 12A). However it overestimates the Gini index of US labor income (Fig. 11A).

The top 1 % income share (Figs. 10B–12B) is sensitive to inequality in the distribution tail. The model accurately predicts the top 1% share of personal income (Fig. 10A). However it overestimates the top 1% share of labor income (Fig. 11B) and capitalist income (Fig. 12B).

Table 6: Summary Statistics of Inequality — Model Predictions vs. US Data

Metric	Source	Labor Income	Personal Income	Capitalist Income
Gini Index (mean)	US	0.54	0.61	0.87
	Model	0.59	0.62	0.86
Top 1% Share (mean)	US	0.14	0.20	0.38
	Model	0.18	0.22	0.48
Power-Law Exponent (mean)	US	2.82	2.57	2.42
	Model	2.77	2.63	2.28

average model error =10.6%

This table compares model predictions to US data using the average of three different summary statistics of inequality. For the US, this is the mean between 2006–2014. For the model, it is the mean over many iterations. 'Average model error' indicates the mean percentage error of the model for all the statistics shown in this table.

The power-law exponent of the top 1% of incomes (Figs. 10C–12C) measures the 'fatness' of the distribution tail. A smaller exponent indicates a fatter tail. The model accurately predicts the power-law exponent of US personal income (Fig. 10) and labor income (Fig. 11C). However, the model underestimates the exponent of US capitalist income (Fig. 12C).

Table 6 shows the average values for these three summary statistics. Across all statistics and all income types, the model deviates from US data by about 11%, on average.

Income Probability Distributions

Panels D–G in Figures 10–12 show probability distributions for both the model and US data. To interpret these Figures, look at the overlap between the model and the US data. If the medians overlap, the model agrees with US data at the given point. If part of the shaded regions overlap, the model is consistent with *some* of the US data. If the shaded regions do not overlap, the model is inconsistent with the US data at the point in question.

Figures 10D –12D show the probability density of income, plotted on a linear scale. We can see that the model reproduces the general form of the US distribution of income. There are, however, deviations for small incomes. For instance, the model does not reproduce the spike of personal and labor incomes that are very close to zero. This spike likely represents unemployed and non-employed members of the US population. Since the model is based on corporate employment data (which is biased towards the full-time employed), it does not reproduce this spike in small incomes. For small capitalist incomes, the model's results are highly uncertain (indicated by the wide spread in the blue-shaded region in Fig. 12D).

While deviating from US data for small incomes, the model more accurately predicts the behavior of large incomes. This is best seen by looking at the tail of the complementary cumulative distribution ((Figs. 10F–12F)). For all three income types, the slope of the model's tail is close to the US data. The slope is closest for personal income and labor income. For capitalist income, the model slope deviates slightly from the US data. This shows visually what we found using the power-law exponent (Figs. 10C–12C), which quantifies the slope of the distribution tail.

Income Quantiles

Figure 13 compares the model to US data using a Q-Q (quantile-quantile) plot. Here, deviations from the 'perfect fit' line indicate a failure of the model. The results indicate that for all three income types, the model reproduces (with reasonable accuracy) the tail of the US distribution of income. However, the model fails for small incomes. The point of failure varies by income type. For personal income, the model fails below the 20th income percentile. For labor income, the model fails below the 10th percentile. For capitalist income, the model fails below the 35th percentile.

5.3 Model Predictions for Income Composition vs. Income Size

For both the model and US data, Figure 14 plots the capitalist composition of income against income size. I call this curve the capitalist income 'hockey stick' because of its similarity to the famous hockey-stick graph showing exploding temperatures in the 20th century (Mann *et al.* 1999). Here it is the capitalist fraction of income that explodes among top earners.

Figure 14 shows two measures of US capitalist income. I continue using capitalist income K1* (as in Figures 10–13), which includes income from all forms of property. But I also calculate capitalist income series K2, which includes only income from interest and corporate equity (see Table 3).

Consistent with the US data, the model predicts a rapid increase of the capitalist portion of income among the top 1%. The model results are closest to US capitalist income K2. This is significant because this series excludes proprietor income. The capitalist gradient hypothesis proposes that only income from corporate property

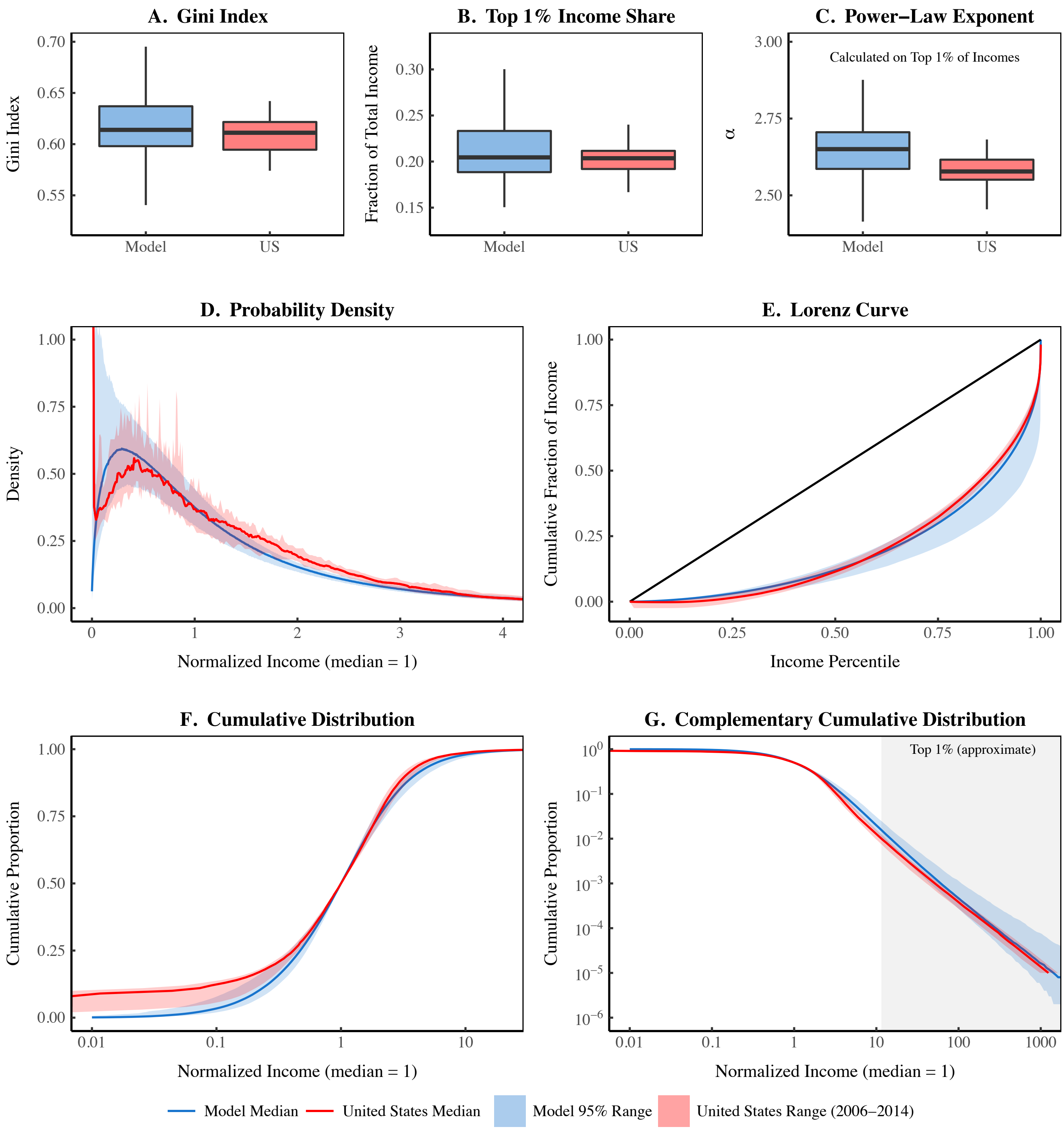

Figure 10: Personal Income Distribution —Model Predictions vs. US Data

This figure shows the distribution of total personal income. Each panel compares the hierarchy model's prediction to US data. Income in Panels D, F and G is normalized so the median is 1. In Panel G, the shaded region shows the approximate threshold for the top 1% of incomes. US data comes from the World Inequality Database. The hierarchy model is stochastic and varies between iterations. I show the model's 95% range. For sources and methods, see the Appendix.

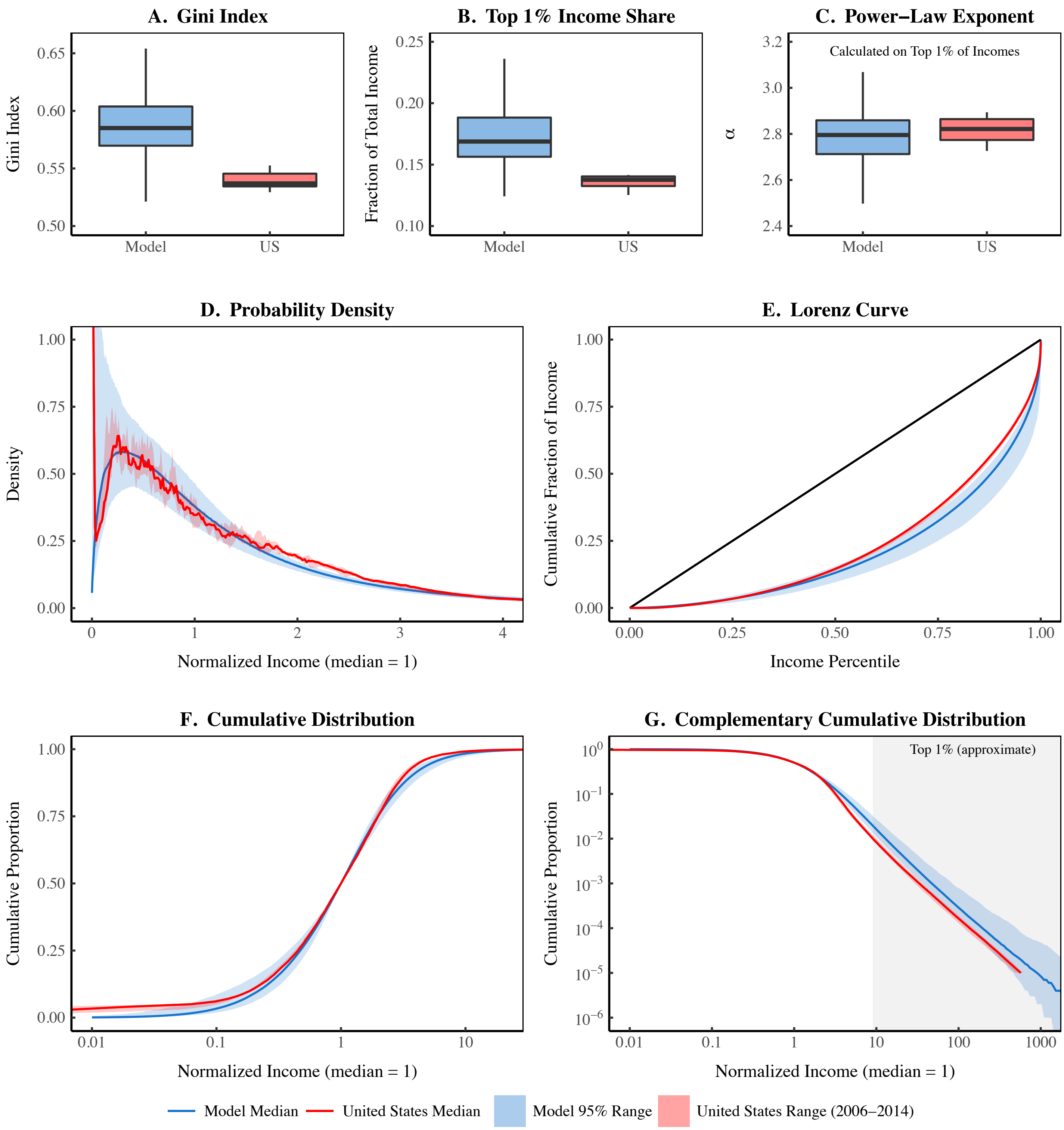

Figure 11: Labor Income Distribution — Model Predictions vs. US Data

This figure shows the distribution of labor income. Each panel compares the hierarchy model's prediction to US data. Income in Panels D, F and G is normalized so the median is 1. In Panel G, the shaded region shows the approximate threshold for the top 1% of incomes. US data comes from the World Inequality Database, using class definitions L1* (Table 3). The hierarchy model is stochastic and varies between iterations. I show the model's 95% range. For sources and methods, see the Appendix.

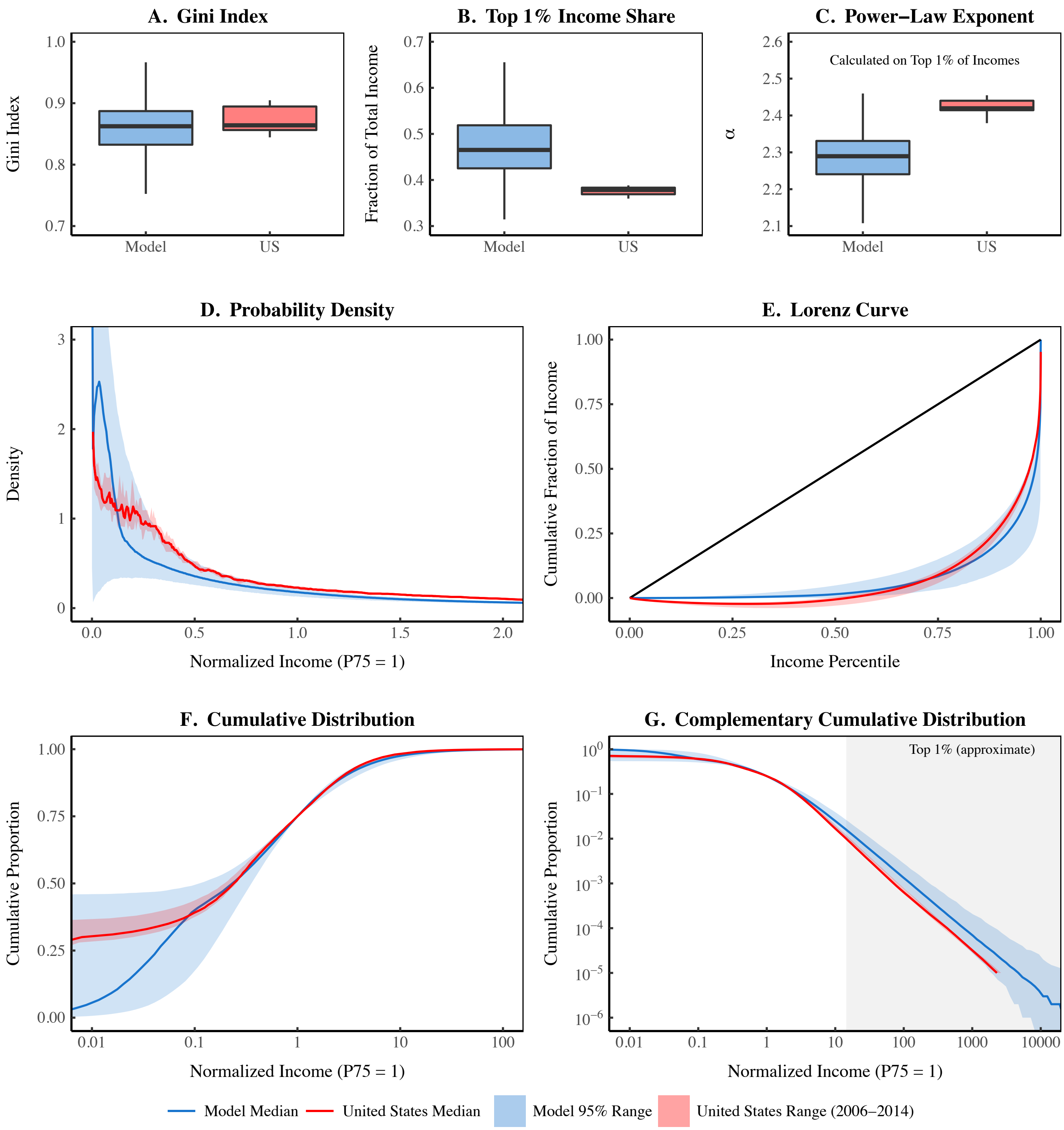

Figure 12: Capitalist Income Distribution — Model Predictions vs. US Data

This figure shows the distribution of capitalist income. Each panel compares the hierarchy model's prediction to US data. Income in Panels D, F and G is normalized so the P75 is 1. I normalize to P75 because the median capitalist income is sometimes zero. In Panel G, the shaded region shows the approximate threshold for the top 1% of incomes. US data comes from the World Inequality Database, using class definitions K1* (Table 3). The hierarchy model is stochastic and varies between iterations. I show the model's 95% range. For sources and methods, see the Appendix.

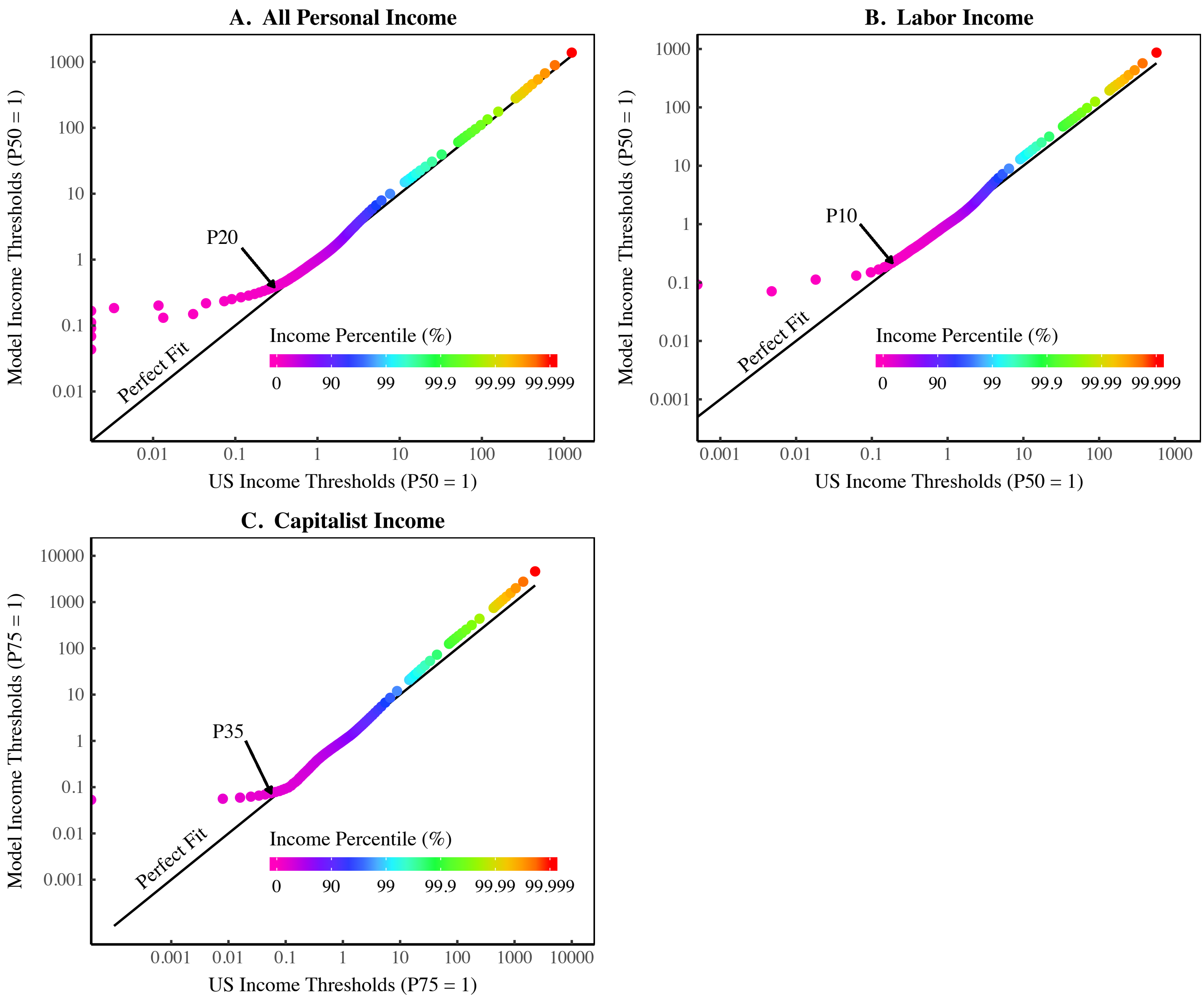

Figure 13: Q–Q Plot — Model Predictions vs. US Data

This figure uses a Q-Q (quantile-quantile) plot to compare the hierarchy model to US data. For a given income percentile, the Q-Q plot compares the model's income threshold (the income cutoff for this percentile) to the US income threshold. I compare here the medians of these thresholds. For the model, this is the median of all iterations. For the US, it is the median between 2006–2014. Income percentiles are shown in color. Deviations from the 'perfect fit' line indicate areas where the model fails. For each income type, I indicate the percentiles where the model starts to fail. For personal income (Panel A) and labor income (Panel B), income thresholds are normalized to the median. Because the median capitalist income is sometimes zero, I normalize capitalist income to P75. For sources and methods, see the Appendix.

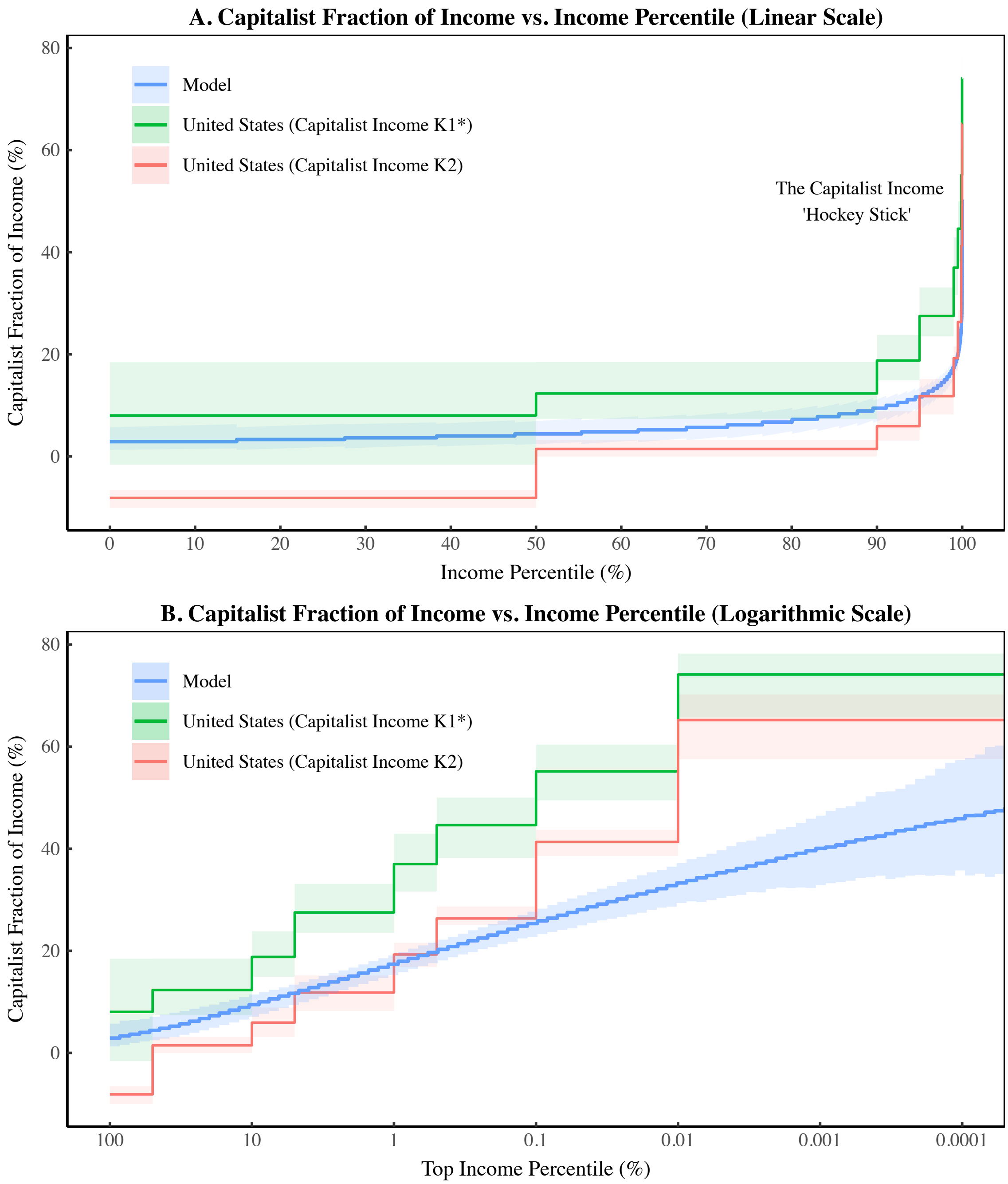

Figure 14: Income Composition vs. Income Percentile — Model Predictions vs. US Data

This figure shows, for both the model and US data, the relation between income composition and income percentile. I use two different measures of US capitalist income — K1* and K2 (see Table 3). Panel A uses a linear scale on the horizontal axis. Panel B shows the same data with a reverse logarithmic scale on the horizontal axis. Shaded regions indicate the 95% range of the data. Lines indicate the median. US data covers the years 2006–2014. For sources and methods, see the Appendix.

should relate to hierarchical power (Section 3). Including rent and proprietor income (in series K1*) does not significantly change the trend between the capitalist fraction of income and income size. This suggests that the explosion of capitalist income among top earners is driven mostly by income from equity and interest.

Compared to capitalist income K2, the model slightly underestimates the growth of the capitalist fraction of income among the top 1%. It also overestimates the capitalist fraction of income among the bottom 90%. Note that below the 50th income percentile, net capitalist income in series K2 is actually negative. This results from interest payments and capital losses on equity. The model does not allow negative capitalist income, and so does not reproduce this behavior.

5.4 Is the Model Valid for Inference?

There is no objective test that can determine if a model fits empirical data well enough to use for inference. How well the model must fit the empirical data depends, in part, on our goals. It also depends on systemic uncertainty in both the empirical data, and the data on which the model is based.

It is obvious, from Figures 10-14, that the model does not perfectly match the US data. A key point of failure is among small incomes. The question is, does this small-income failure matter for inferences about hierarchical power? I cautiously answer no. The reason is that a theory of income distribution based on hierarchy is primarily concerned with top incomes. Using a model similar to here, Fix (2018b) finds that the effects of hierarchy on income become important only among top earners. The implication is that what matters for inferences about hierarchical power is the accuracy of the model among top incomes. Fortunately, among top incomes the model is roughly consistent with the US data.

We should also remember that there is considerable systematic uncertainty in the data that underlies the model. The model draws conclusions about the shape of firm hierarchies from a handful of case studies. Also, the CEO data used to model the capitalist fraction of income is based on an incomplete accounting of capitalist income. Furthermore, the empirical measures of US capitalist income are based on accounting definitions that differ from those used to measure the CEO fraction of capitalist income. The goodness of the model's fit should be judged in the context of this systematic uncertainty.

My goal here is to make a rough first inference for how income size and income class relate to hierarchical power in the United States. For this purpose, I judge the model's results to be accurate enough. The model's predictions match the US data closely enough to infer that trends found among US CEOs plausibly extend to the general US population.

6 Inferences: Income Size, Income Class and Hierarchical Power in the United States

Using results from the model, Figure 15 shows the inferred three-way relation between income size, income composition and hierarchical power in the United States.

The model suggests that the vast majority of Americans have little hierarchical power and little capitalist income. But among the top 1%, the model suggests an explosion of hierarchical power and a corresponding explosion of capitalist income. This is consistent with Tim Di Muzio's (2015) distinction between the "1% and the rest of us". It also gives new meaning to Fitzgerald's assertion that the "rich are different". What makes the rich different, the model suggests, is hierarchical power.

6.1 Open Questions

The results in Figure 15, combined with the CEO evidence in Section 4, suggest that studying hierarchical power may shed light on the relation between personal and functional income distribution. But because this analysis is the first of its kind (to my knowledge), many questions remain.

First, there is the question of whether our model inference is accurate. Answering this question requires better data on firm hierarchy, which will hopefully become available in the future.

Second, there is the question of causation. Does hierarchical power cause the size and composition of income to change? If so, why? If not, what is the underlying connection? At present, data on firm hierarchy is too sparse to answer these causal questions. For this reason, I have focused on correlation only.

Third, there is the question of the 'boundaries' of hierarchy. I have focused here on hierarchy *within* firms. But hierarchies can also extend *between* firms — what Bichler and Nitzan (2017) call "meso hierarchies". This involves the use of partial ownership to wield control over firms. Recent work on corporate ownership shows that investment firms wield a surprising amount of power (Glattfelder 2010, Glattfelder and Battiston 2009, Vitali *et al.* 2011). Understanding how this network of ownership relates to hierarchical power is an important task for future research.

Fourth, do the results here extend beyond the United States? While direct evidence is not available, the modeling methods used here could be applied to other

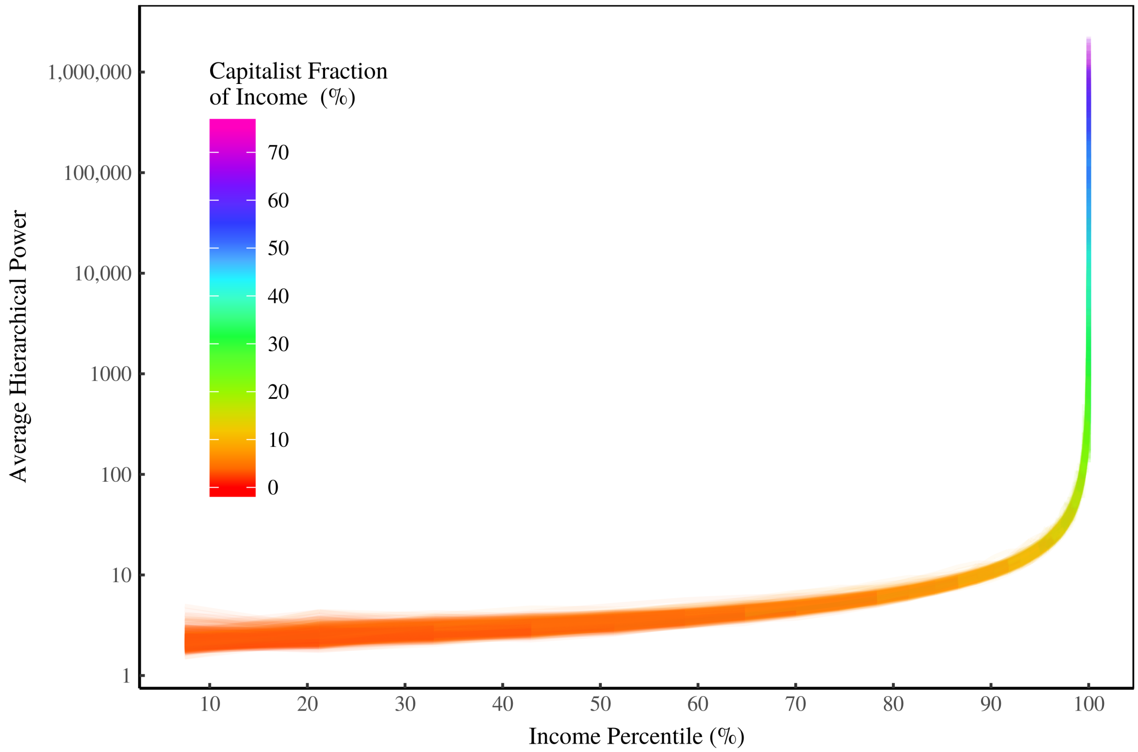

Figure 15: Model Inference for Income Size, Income Class and Hierarchical Power in the US

This figure shows the inferred three-way relation between income size, income class, and hierarchical power among the US public. Results are from the hierarchy model, which extrapolates CEO trends to the general public (Section 5). The plot shows the average hierarchical power of individuals, grouped by income percentile. The average capitalist fraction of income is indicated by color. Lines indicate different model iterations.

---

countries. Researchers interested in using this model can find the code in the Supplementary Material.

Fifth, has the three-way relation between income size, income class and hierarchical power changed over time? Given the drastic increase in US income inequality over the last three decades, this seems likely. Song *et al.* (2016) find that the recent growth of US top incomes is due mostly to increasing inequality *within* firms. Does this have to do with hierarchy? Possibly. Fix (2018a) finds that the growth of top US incomes may be due to a redistribution of income towards the tops of firm hierarchies.

Lastly, does the three-way relation between income size, income class and hierarchical power extend to non-capitalist societies? Unfortunately, we know little about hierarchy in pre-capitalist societies. However, using a similar model as here, Fix (2019) finds that the growth of hierarchy can possibly explain the origin and evolution of inequality. This suggests that income may increase with hierarchical power in pre-capitalist societies.

What about the relation between hierarchy and income class in pre-capitalist societies? Admittedly we have very little data on this topic. But consider how Reinhard Bendix describes the relation between authority, income and property rights in German feudal society:

> ... governmental functions were usable rights which could be sold or leased at will. For example, judicial *authority* was a type of *property*. The person who bought or leased that property was entitled to adjudicate disputes and *receive the fees and penalties* incident to such adjudication. (Bendix 1980, p. 149)[emphasis added]

If we paraphrase Bendix, we arrive at the same reasoning that I used to derive the capitalist gradient hypothesis. Building on the work of Nitzan and Bichler (2009), I argued that ‘capitalist authority’ is a ‘type of property’. The person who buys this property is ‘entitled’ to wield hierarchical power and ‘receive income’ in return. This reasoning led to the hypothesis that the capitalist portion of income should increase with hierarchical power. Bendix’s description of feudal authority hints that something like the capitalist gradient hypothesis (but for different income classes) may apply to feudal societies.

7 Conclusions

This paper has outlined a new way of studying personal and functional income distribution. Rather than appealing to productivity (like neoclassical economists) or to exploitation (like Marxists), I have proposed a theory of income distribution based on hierarchical power. It is the greater command over subordinates, I have argued, that distinguishes the rich from the poor and capitalists from workers.

The goal of this paper is primarily empirical. I have given theoretical reasons why hierarchical power might relate to income size and composition (Section 2). But for the most part, this proposed relation is more of a hunch than a rigorous prediction. This paper tests this hunch by looking for a correlation between income size and hierarchical power, on the one hand, and income composition and hierarchical power on the other.

Evidence from US CEOs suggests that this hunch is justified. I find that the relative income of CEOs increases with hierarchical power, as does the capitalist composition of this income (Section 4). A model that extrapolates this data suggests that the trend among CEOs plausibly extends to the general US population (Sections 5 and 6). In other words, it is plausible that there is a three-way relation between income size, income composition and hierarchical power in the United States.

While much remains unknown, the results here suggest that hierarchical power should be further investigated as a determinant of income size and income composition.

Acknowledgments

I thank Jonathan Nitzan for comments on an earlier draft of this paper.

Appendix

Supplementary materials for this paper are available at the Open Science Framework:

https://osf.io/wp8yu/

The supplementary materials include:

1. Source data;
2. Code for all analysis;
3. Hierarchy model code.

A US Class-Based Income

Data for US class-based income comes from the World Inequality Database (WID). My income measures are shown in Table 7. These are composed of the WID data series shown in Tables 8 and 9. I use two WID series to construct K1*, L1*, and T. This means I merge statistics from both WID series.

The WID data comes from Piketty *et al.* (2017b). For the methods of this study, see Piketty *et al.* (2017a). This is the most detailed study to date of US class-based income. However, it comes with some caveats. Piketty et al. subdivide proprietor income into capitalist and labor components. The capitalist component is series *fkbus*. The labor component is series *flmil* (Table 9). I cannot find, in Piketty's work, an explicit statement of the methods behind this split. But according to Rognlie (2016), Piketty assumes that proprietor income "has the same net capital share as the corporate sector".

This leads to a difference between my definitions of class-based income (defined in the main paper) and the empirical data (Table 7). My two-class definition of capitalist income (K1) includes *all* proprietor income. In contrast, the empirical measure K1* contains only a portion of proprietor income. My definition of labor income (L1) contains *no* proprietor income. In contrast, the empirical measure L1* contains a portion of proprietor income.

In addition to the capitalist income series provided by WID, I construct my own series K2 shown in Table 7. This includes equity and interest income (with capital gains).

Methods for Estimating Income Distribution Statistics

WID provides three types of data that I use to compute statistics:

1. Income percentile (bin)
2. Income share (by income percentile bin)
3. Income threshold (by income percentile bin)

As an example, the WID data may indicate that percentiles P99–P100 have an income share of 15%. This means the top 1% holds 15% of all income. The income threshold for this bin may be $200,000. This means that the lowest income of the top 1% is $200,000.

Gini Index: I estimate the Gini index by constructing a Lorenz curve from WID data. The Gini index equals the area between the Lorenz curve and the line of

Table 7: Measures of US Class-Based Income

Measure	Symbol	Composition
Capitalist Income (All Ownership)	K1*	both fkinc and pkinc
Capitalist Income (Scalable Ownership Only)	K2	fkequ + fkfix
Labor Income	L1*	both flinc and plinc
Total Income	T	both fainc and ptinc

Table 8: World Inequality Database Main Series

Series	Description	Composition
fainc	Personal factor income	flinc + fkinc
fkinc	Personal factor capital income	fkhou + fkequ + fkfix + fkbus + fkpen + fkdeb
flinc	Personal factor labor income	flemp + flmil + flprl
pkinc	Personal pre-tax capital income	fkinc + pkpen + pkbek
plinc	Personal pre-tax labor income	flinc + plcon + plbel
ptinc	Personal pre-tax income	plinc + pkinc

Table 9: World Inequality Database Component Series

Series	Description
fkbus	Business asset income
fkdeb	Interest payments
fkequ	Equity asset income
fkfix	Interest income
fkhou	Housing asset income
fkpen	Pension and insurance asset income
flemp	Compensation of employees
flmil	Labor share of net mixed income
flprl	Sales and excise taxes falling on labor
pkbek	Capital share of social insurance income
pkpen	(Minus) Investment income payable to pension funds
plbel	Labor share of social insurance income
plcon	(Minus) social contributions

perfect equality, divided by the total area under the line of perfect equality.

Top 1% Share: This is provided directly by the WID data.

Power-Law Exponent: I estimate the power-law exponent of the top 1% of incomes using income percentile and threshold data. I create binned data where we know the proportion of people in each bin, and the income boundaries of each bin. I then use the method discussed in Virkar and Clauset (2014) to estimate the power-law exponent from this binned data.

Probability Density: I estimate the probability density using income percentile and threshold data. I first normalize income threshold data so that the median equals 1. I then construct a cumulative distribution. This is the fraction of individuals below a given income. I estimate the probability density function from the *slope* of the cumulative distribution.

Lorenz Curve: The Lorenz curve is constructed from income percentile and income share data. It is the cumulative share of income vs. income percentile.

Cumulative Distribution: I construct the cumulative distribution from income percentile and threshold data. I normalize income threshold data so that the median equals 1.

Complementary Cumulative Distribution: I construct the complementary cumulative distribution (CCD) from the cumulative distribution (CD). The y-value for the CCD is 1 minus the corresponding y-value for the CD.

Income Quantiles: This data is provided directly by WID (reported as income thresholds by income percentile bin).

Capitalist Income Share vs. Percentile: Capitalist income share $\mathrm{K1}^*_{\mathrm{frac}}$ is calculated by merging two series:

$$\mathrm{K1}^*_{\mathrm{frac}} = \begin{cases} \text{fkinc / fainc} \\ \text{pkinc / ptinc} \end{cases} \tag{6}$$

Capitalist income share $K2^*_{\text{frac}}$ is calculated as:

$$K2^*_{\text{frac}} = (\text{fkequ} + \text{fkfix}) \;/\; \text{fainc} \tag{7}$$

B US CEO Data: The Compustat Firm Sample

Data for US CEOs (and their firms) comes from the Execucomp and Compustat databases. I will refer to this data as the 'Compustat firm sample'. I use this data for the case study of CEO pay and as the basis for the US hierarchy model. Methods are discussed below.

Finding the CEO

I identify CEOs using titles in the Execucomp series TITLEANN. I use a three-step algorithm:

1. Find all executives whose title contains one or more of the words in the 'CEO Titles' list in Table 10.
2. Of these executives, take the subset whose title does *not* contain any of the words in the 'Subordinate Titles' list in Table 10.
3. If this returns more than one executive per firm per year, chose the executive with the highest pay.

Table 10: Titles Used to Identify the 'CEO'

CEO Titles:	**Subordinate Titles**
president	vp
chairman	v-p
CEO	cfo
Chief Executive Officer	vice
chmn	chief finance officer
	president of
	coo
	division
	div
	president-
	group president
	chairmain-
	co-president
	deputy chairman
	pres.-
	Chief Financial Officer

Notes: Titles such as 'president-' and 'president of' are included in the subordinate list because they typically refer to a president of a division within the company: i.e. 'president of western division' or 'president-western hemisphere'.

Table 11: Execucomp Compensation Series

Series	Description
TOTAL_ALT2	SALARY + BONUS + OTHCOMP + NONEQ_INCENT + PENSION_CHG + OPT_EXER_VAL + SHRS_VEST_VAL
BONUS	The dollar value of a bonus earned by the named executive officer during the fiscal year.
SALARY	The dollar value of the base salary earned by the named executive officer during the fiscal year.
OTHCOMP	Other compensation received by the director including perquisites and other personal benefits, contributions to defined contribution plans (e.g. 401K plans), life insurance premiums, gross-ups and other tax reimbursements, discounted share purchases, consulting fees, awards under charitable award programs, etc.
NONEQ_INCENT	Value of amounts earned during the year pursuant to non-equity incentive plans.
PENSION_CHG	Composed of a) above-market or preferential earnings from deferred compensation plans, and b) aggregate increase in actual value of defined benefit and actual pension plans during the year.
OPT_EXER_VAL	Value realized from option exercises during the year. The value is calculated as of the date of exercise and is based on the difference between the exercise price and the market price of the stock on the exercise date.
SHRS_VEST_VAL	Value of restricted shares that vested during the year.

CEO Pay and Capitalist Income Fraction

Execucomp contains several different estimates of CEO pay. These differ primarily in the valuation of stock option compensation. Hopkins and Lazonick (2016) argue that we should use *actual realized gains*. This is the difference between the *market value* of the option and the *exercise value* at the time of exercise. Importantly, actual realized gains is the income recorded by the IRS for tax purposes. I measure CEO total pay and capitalist income fraction (K_{frac}) using the following series:

$$\text{Total Pay} = \text{TOTAL\_ALT2} \tag{8}$$

$$K_{\text{frac}} = \frac{\text{Actual Realized Gains from Stock Options}}{\text{Total Pay}} \tag{9}$$

$$K_{\text{frac}} = \frac{\text{SHRS\_VEST\_VAL} + \text{OPT\_EXER\_VAL}}{\text{TOTAL\_ALT2}} \tag{10}$$

Series descriptions are shown in Table 11.

CEO Pay Ratio and Firm Employment

I calculate the CEO pay ratio as:

$$\text{CEO Pay Ratio} = \frac{\text{CEO Pay}}{\text{Firm Mean Income}} \tag{11}$$

Firm mean income is calculated by dividing total staff expenses (Compustat Series XLR) by total employment (Compustat Series EMP):

$$\text{Firm Mean Income} = \frac{\text{Total Staff Expenses}}{\text{Total Employment}} \tag{12}$$

CEO pay ratio and firm mean income data are available for roughly 3000 firm-year observations from 2006-2016. Figure 16 shows summary statistics of this data.

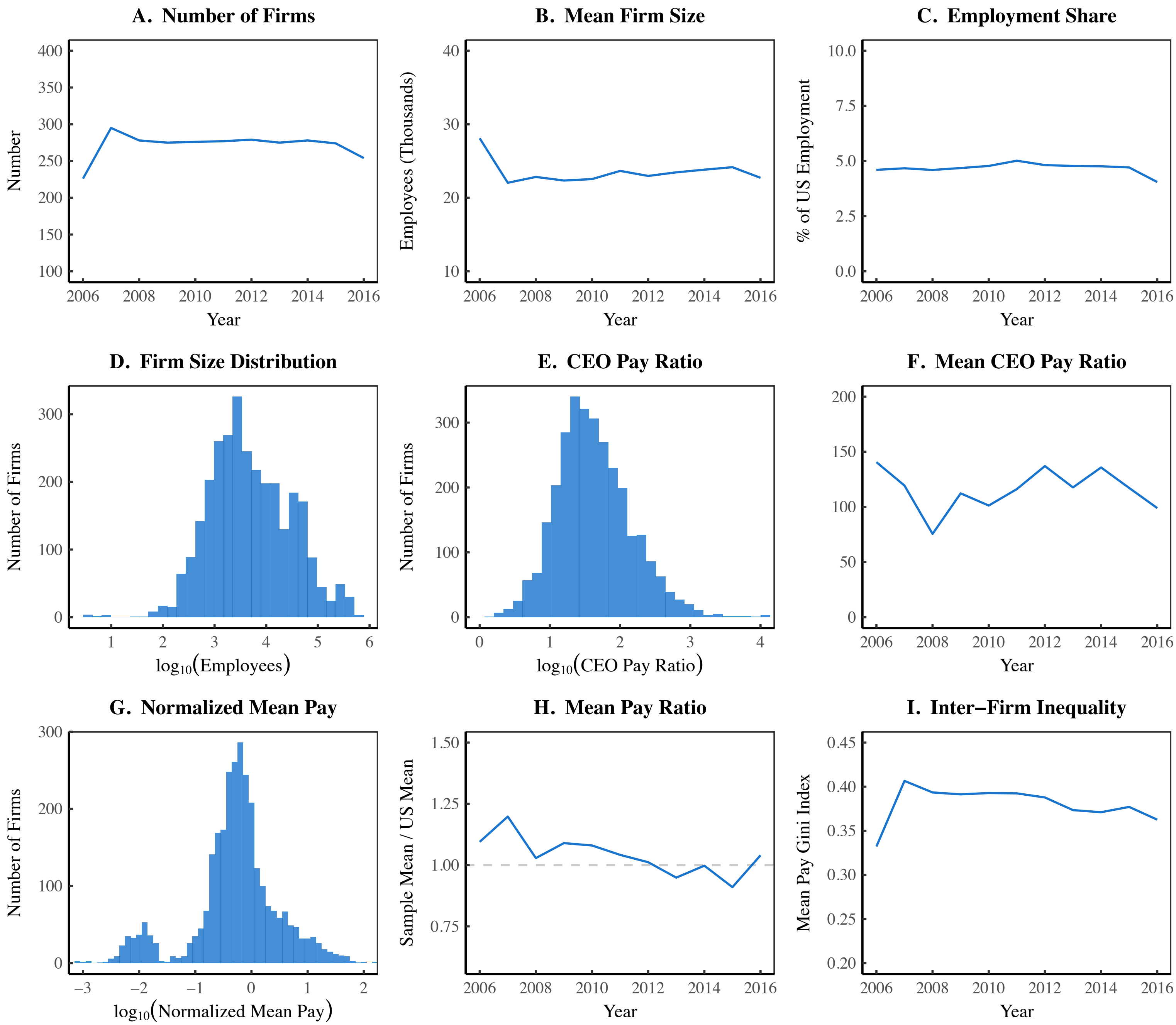

Figure 16: Statistics of the Compustat Firm Sample

This figure shows selected statistics of the Compustat firm sample. Panel A shows the number of firms in the sample over time, Panel B the average firm size, and Panel C the share of US employment held by these firms. Panel D shows the logarithmic distribution of firm size, and Panel E shows the logarithmic distribution of the CEO pay ratio. Panel F shows the mean CEO pay ratio of all firms over time. Panel G shows the logarithmic distribution of normalized mean pay (mean pay divided by the average pay of the firm sample in each year). Panel H shows the ratio of mean pay in the sample relative to the US average (calculated from BEA Table 1.12 by dividing the sum of employee and proprietor income by the number of workers in BEA Table 6.8C-D. Panel I shows the Gini index of firm mean pay over time.

Table 12: Model Notation

Symbol	Definition
a	span of control parameter 1
b	span of control parameter 2
C	CEO to average employee pay ratio
E	employment
ε	noise factor
h	hierarchical rank
I	income
n	number of hierarchical ranks in a firm
P	hierarchical power
s	span of control
S	number of subordinates
T	total for firm

C Hierarchy Model Equations

The hierarchy model assumes that firms are hierarchically structured, with a span of control that increases exponentially with hierarchical rank. The model simulates individual income as a function of hierarchical power. I discuss here the model's main equations. See Table 12 for notation.

C.1 The Employment Hierarchy

For each firm, the model generates an employment hierarchy using the span of control (s). This is the ratio of employment (E) between two consecutive hierarchical ranks (h). We let $h = 1$ be the *bottom* hierarchical rank. We define the span of control in rank 1 as $s = 1$. This leads to a piecewise function for the span of control:

$$s_h \equiv \begin{cases} 1 & \text{if } h = 1 \\ \dfrac{E_{h-1}}{E_h} & \text{if } h \geq 2 \end{cases} \tag{13}$$

Based on evidence from firm case studies (Fig. 18), the model assumes that the span of control increases exponentially with hierarchical rank, with a and b as free parameters:

$$s_h = \begin{cases} 1 & \text{if } h = 1 \\ a \cdot e^{bh} & \text{if } h \geq 2 \end{cases} \tag{14}$$

As one moves up the hierarchy, employment in each consecutive rank (E_h) decreases by a factor of $1/s_h$. This yields a recursive formula for calculating E_h:

$$E_h = \left\lfloor \frac{E_{h-1}}{s_h} \right\rfloor \qquad \text{for} \qquad h > 1 \tag{15}$$

The model assumes employment is a whole number and so rounds down to the nearest integer (notated by $\lfloor\rfloor$). By repeatedly substituting Eq. 15 into itself, we obtain a non-recursive formula for hierarchical employment:

$$E_h = \left\lfloor E_1 \cdot \frac{1}{s_2} \cdot \frac{1}{s_3} \cdot \ldots \cdot \frac{1}{s_h} \right\rfloor \tag{16}$$

In product notation, Eq. 16 becomes:

$$E_h = \left\lfloor E_1 \prod_{i=1}^{h} \frac{1}{s_i} \right\rfloor \tag{17}$$

Total employment E_T in the whole firm is the sum of employment in all hierarchical ranks. Defining n as the total number of hierarchical ranks, total firm employment is:

$$E_T = \sum_{h=1}^{n} E_h \tag{18}$$

Because the model builds the hierarchy from the bottom up, n is not known beforehand. The model defines n using Eq. 17. The model calculates employment in every hierarchical rank until it reaches a rank with zero employment. The top rank n is the highest rank with non-zero employment:

$$n = \{h \quad | \quad E_h \geq 1 \quad \text{and} \quad E_{h+1} = 0\} \tag{19}$$

To summarize, the employment hierarchy in each firm is determined by 3 free parameters: the span of control parameters a and b, and employment in the bottom rank, E_1. Code for this algorithm is located in `exponents.h` and `hierarchy.h` in the Supplementary Material.

C.2 The Pay Hierarchy

The model assumes that individual income is a function of hierarchical power:

$$I_{i,h,f} = \bar{I}_{1,f} \cdot (\bar{P}_{h,f})^{\beta_f} \cdot \varepsilon_i \tag{20}$$

Here $I_{i,h,f}$ is the income of the ith person in hierarchical level h of firm f. $\bar{I}_{1,f}$ is the average income in the bottom hierarchical level of the firm. $\bar{P}_{h,f}$ is average hierarchical power in level h of the firm. β_f is the power-income exponent of the given firm. Lastly, ε_i is a stochastic noise factor that adds dispersion to individual income.

In each firm, we define the average hierarchical power in level h as:

$$P_h = \bar{S}_h + 1 \tag{21}$$

Here $\bar{S}_h$ is the average number of subordinates per member of rank h:

$$\bar{S}_h = \sum_{i=1}^{h-1} \frac{E_i}{E_h} \tag{22}$$

C.3 Statistics

Mean Income in a Firm. Mean income in a firm ($\bar{I}_T$) is the average of mean income in each hierarchical rank ($\bar{I}_h$), weighted by the employment in each rank (E_h):

$$\bar{I}_T = \sum_{h=1}^{n} \bar{I}_h \cdot \frac{E_h}{E_T} \tag{23}$$

CEO Pay Ratio. The model defines the 'CEO' as the person in the top hierarchical rank, n. CEO pay is thus $\bar{I}_n$, average income in the top hierarchical rank. The CEO pay ratio (C) is defined as CEO income divided by average income in the firm:

$$C = \frac{\bar{I}_n}{\bar{I}_T} \tag{24}$$

Table 13: Parameters in the US Hierarchy Model

Parameter	Definition	Action	Scope
α	Power-law exponent for the size distribution of firms	Determines the skewness of the firm size distribution	—
a, b	Span of control parameters	Determines the shape of the firm hierarchy.	Identical for all firms.
E_1	Employment in base hierarchical level	Used to build the employment hierarchy from the bottom up. Determines total employment.	Specific to each firm.
β	Power-income exponent	Determines how rapidly income increases with hierarchical power.	Specific to each firm.
$\bar{I}_1$	Mean pay in base hierarchical level	Sets the base level income of the firm, which determines firm average pay.	Specific to each firm.
ε	Power-income noise factor	Adds stochastic noise to the relation between income and hierarchical power.	Identical for all firms.
μ_K, σ_K	Capitalist gradient parameters	Determine the capitalist fraction of individual income as a function of hierarchical power.	Identical for all individuals.

D The United States Hierarchy Model

The US hierarchy model uses the equations from Section C to simulate the hierarchical structure of the US private sector. The model's parameters are summarized in Table 13. I detail here how I restrict these parameters.

D.1 Simulating the Size Distribution of US Firms

Evidence suggests that the size distribution of firms in the US (and other G7 countries) roughly follows a power law (Axtell 2001, Gaffeo *et al.* 2003). In a power-law distribution, the probability of finding a firm of size x is:

$$p(x) \propto \frac{1}{x^{\alpha}} \tag{25}$$

Figure 17 compares the size distribution of US firms to a discrete power law. The inset plot shows the best-fit values for the power-law exponent α, fitted using the method described in Virkar and Clauset (2014).

To simulate the size distribution of US firms, I use a discrete power-law distribution of 1 million firms. In each iteration, the model sets the power-law exponent

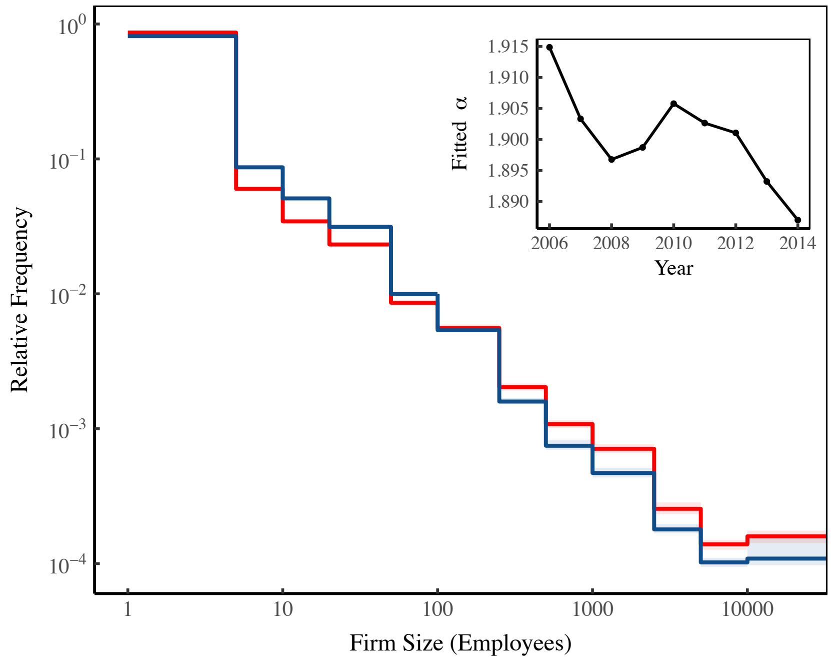

Figure 17: The Size Distribution of Firms in the United States

This figure compares the firm size distribution in the United States to a discrete power-law distribution. The 'steps' indicate the firm-size bins. The inset plot shows the best-fit power-law exponent (α) in each year. The US data combines 'employer' firms and unincorporated self-employed workers. Data for 'employer' firms is from the US Census Bureau, Business Dynamics Statistics. I augment this data with Bureau of Labor Statistics data for unincorporated self-employed workers (series LNU02032185 and LNU02032192). The histogram preserves firm-size bins used by the Census. I add self-employed individuals to the first bin. The last histogram bin contains all firms with more than 10,000 employees.

α by sampling from the set of fitted US values (Fig. 17, inset). To ensure that the simulation produces realistically sized firms, I truncate the power-law distribution at a maximum firm size of 2.3 million. This is the present size of Walmart, the largest US firm that has ever existed.

Code for the random number generator for the discrete power-law distribution can be found in `rpld.h`, located in the Supplementary Material. This code is an adaptation of Collin Gillespie's (2014) discrete power-law generator found in the R poweRlaw package. Gillespie's generator is, in turn, an adaptation of the algorithm outlined by Clauset *et al.* (2009).

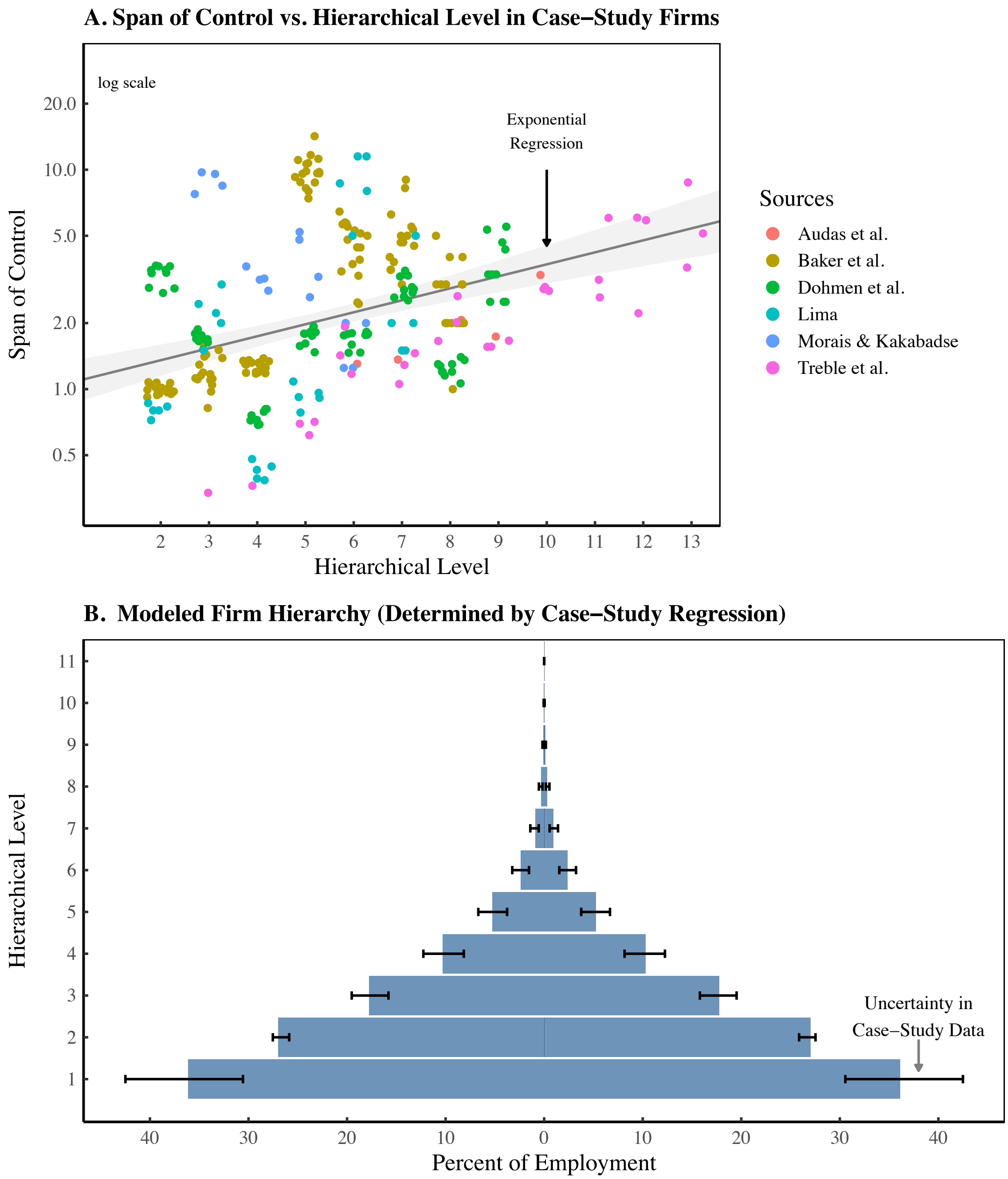

Figure 18: Idealized Hierarchy Implied by Firm Case Studies

Panel A shows how the span of control varies with hierarchical level in case-study firms (Audas *et al.* 2004, Baker *et al.* 1993, Dohmen *et al.* 2004, Lima 2000, Morais and Kakabadse 2014, Treble *et al.* 2001). The span of control is the subordinate-to-superior ratio between adjacent hierarchical levels. The x-axis corresponds to the *upper* hierarchical level in each corresponding ratio. Case-study firms are indicated by color. I have added horizontal 'jitter' to better visualize the data. The line indicates an exponential regression, with the grey region indicating the regression 95% confidence interval. Panel B shows the idealized firm hierarchy that is implied by the regression in Panel A. Error bars show the uncertainty in the hierarchical shape, calculated using a bootstrap resample of case-study data.

D.2 Span-of-Control Parameters a and b

The shape of the employment hierarchy in simulated firms is determined by the span-of-control parameters a and b. To set these parameters, I regress Eq. 14 onto span-of-control data from case-study firms (Fig. 18A). I then use Eqs. 14, 17, and 18 to create the employment hierarchy in each simulated firm. Note that all firms are assigned the same values for a and b.

The model incorporates uncertainty in a and b using the bootstrap method (Efron and Tibshirani 1994). I run the model many times, with each iteration regressing a and b on a bootstrapped sample of the firm case-study data. Figure 18B shows the shape of the modeled employment hierarchy for a generic large firm. Code implementing the bootstrap is located in `boot_span.h` in the Supplementary Material.

D.3 Employment in the Base Hierarchical Level (E_T)

Given span of control parameters a and b, each simulated firm hierarchy is constructed from the bottom hierarchical level up. To do this, we must estimate E_1, the employment in the base hierarchical level.

To estimate E_1 in each firm, I use the model to create a numerical function relating base level employment E_1 to total firm employment E_T. I input a range of different base employment values into equations 14, 17, and 18 and calculate total employment for each value. The result is a discrete mapping relating base-level employment to total employment. I then use the C++ Armadillo interpolation function to linearly interpolate between these discrete values. This creates a numerical function that returns E_1 when given total firm employment E_T and span-of-control parameters a and b.

Code implementing this method is located in `base_fit.h` in the Supplementary Material.

D.4 Power-Income Exponent β

The power-income exponent (β) determines the rate that income increases with hierarchical power in simulated firms (see Eq. 20). Unlike the span of control parameters, I allow β to vary between firms.

I restrict the variation of β using a two-step process. I first 'tune' the model to data from Compustat firms. This returns a distribution of β that is specific to

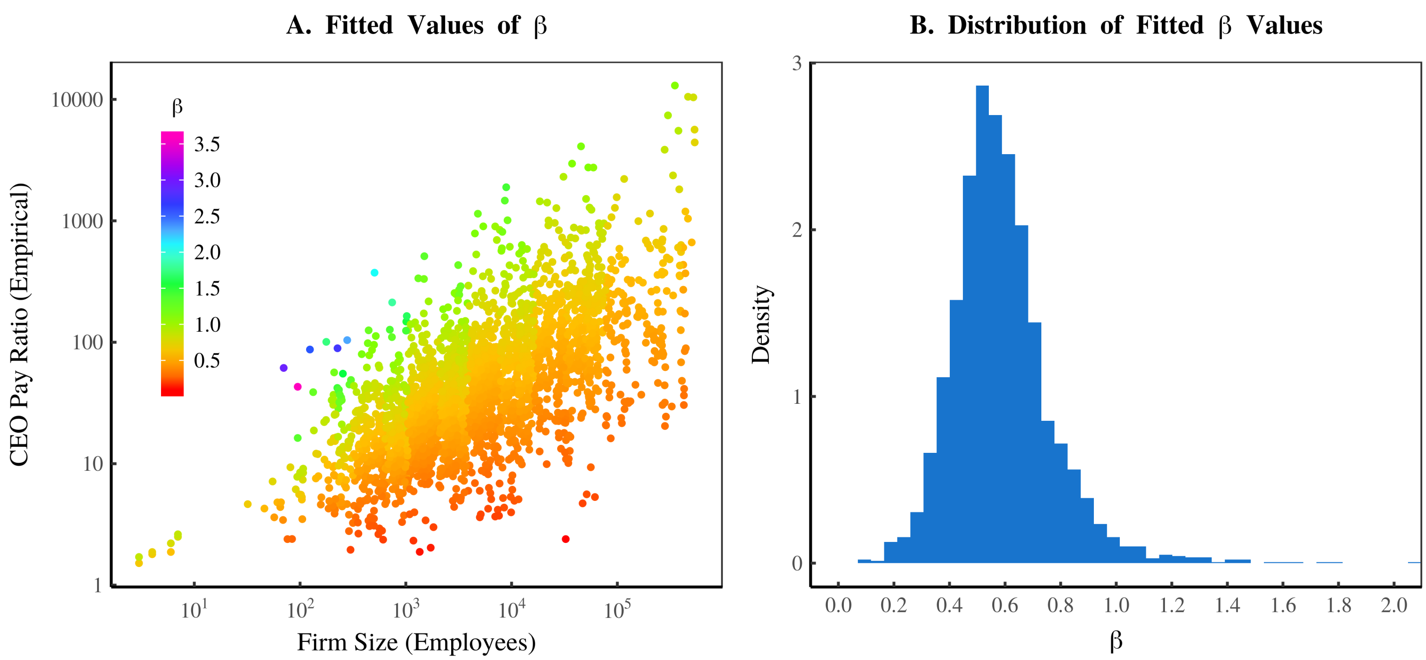

Figure 19: Estimating the Power-Income Exponent (β) Inside Compustat Firms

This figure shows the β values fitted to Compustat firms. Panel A shows how the fitted values of β relate to firm size and the CEO pay ratio. The discrete changes in color (evident as vertical lines) correspond to changes in the number of hierarchical levels within simulated firms. Panel B shows the distribution of the fitted values of β. Note that fitted values for β vary between model iterations.

Compustat firms. I then fit this data with a parameterized distribution, from which simulated values for β are randomly chosen.

D.4.1 Fitting Power-Income Exponent β to Compustat Firms

I fit β to Compustat firms using the CEO pay ratio (C). The first step of this process is to simulate the employment hierarchy for each Compustat firm using parameters a, b, and E_1 (the latter is determined from total employment). Given the simulated employment hierarchy, the CEO pay ratio in the modeled firm is uniquely determined by the parameter β. I choose β so that the model produces a CEO pay ratio that is equivalent to the Compustat data.

To find the best-fit value for β, I use numerical optimization (the bisection method) to minimize the following error function:

$$\varepsilon(\beta) = \left| C_{\text{model}} - C_{\text{Compustat}} \right| \tag{26}$$

Here C_{model} is the modeled CEO pay ratio, and $C_{\text{Compustat}}$ is the Compustat CEO pay ratio.

To ensure that there are no large errors, the model discards Compustat firms for which the best-fit β parameter produces an error larger than $\varepsilon = 0.01$. Figure 19 shows an example of the fitted β values for all Compustat firm-year observations. Code implementing this fitting method is located in `fit_beta.h` in the Supplementary Material.

D.4.2 Simulating the Distribution of β for All US Firms

Once we have estimated β for every Compustat firm, the next step is to fit a parameterized distribution to this data. For Compustat firms, the distribution of β is roughly lognormal, with dispersion that tends to decline with firm size

I model the distribution of β using a lognormal distribution with a constant location parameter μ and a scale parameter σ_E that varies with firm size:

$$\beta(E) = \ln \mathcal{N}(\beta; \mu, \sigma_E) \tag{27}$$

The location parameter μ is constant for all firms and is given by:

$$\mu = \overline{\ln(\beta_{\text{Compustat}})} \tag{28}$$

To estimate the scale parameter σ, I calculate the standard deviation of $\ln(\beta_{\text{Compustat}})$ within groups of firms binned by firm size E:

$$\sigma_E = \text{SD}\left[\ \ln(\beta_{\text{Compustat}})\ \right]_E \tag{29}$$

Figure 20A shows how σ_E varies with firm size. Each dot indicates σ_E calculated on a log-spaced firm-size bin. I model σ_E as a log-linear function of firm size:

$$\sigma_E = c_1 \ln(E) + c_2 \tag{30}$$

Once we have estimated the parameters μ and σ_E, we use Eq. 27 to generate β values for each simulated firm. Figure 20B shows how the modeled dispersion of β declines with firm size.

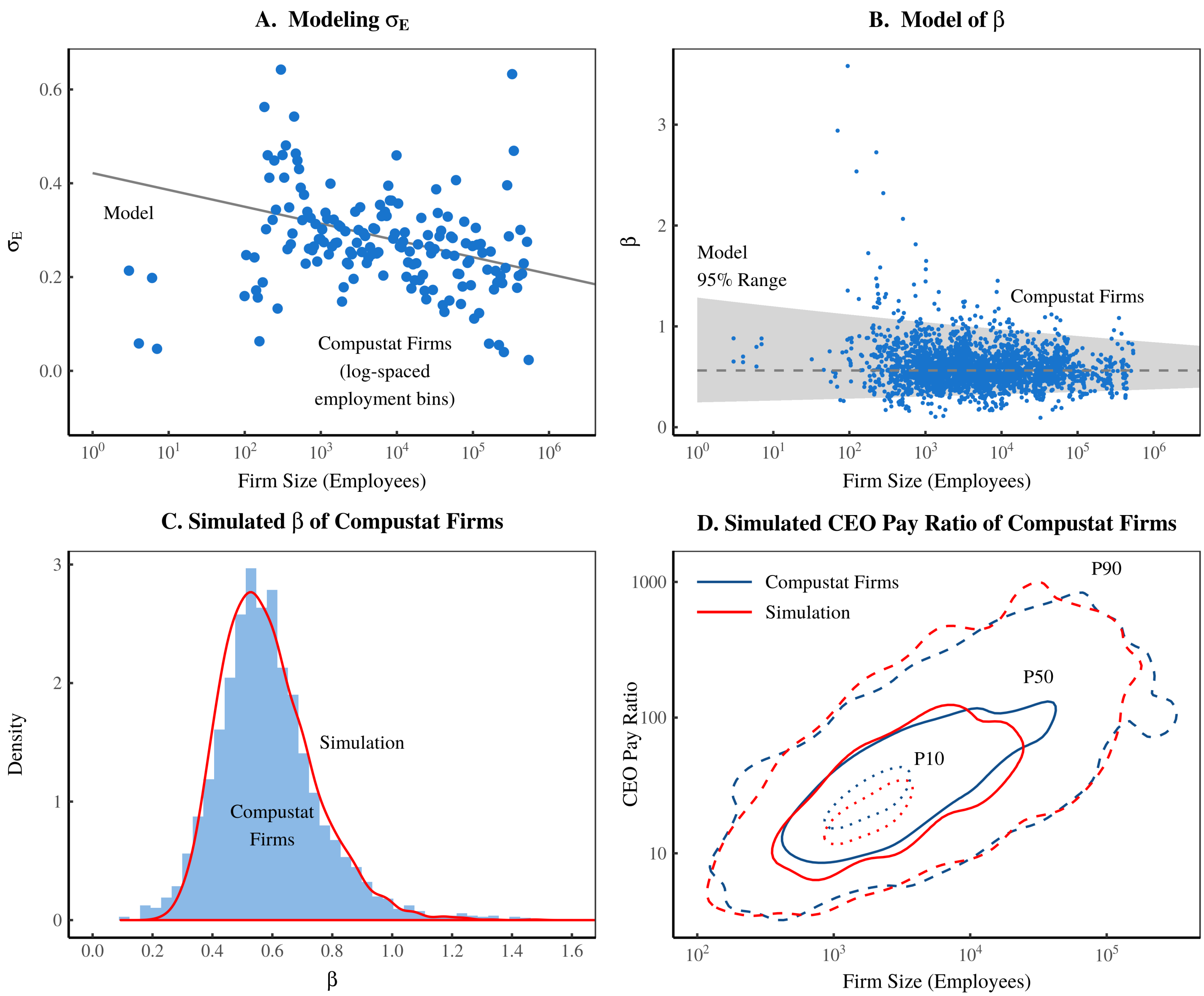

Figure 20: Modeling the Distribution of the Parameter β

This figure visualizes the algorithm used to simulate the distribution of the parameter β. This parameter determines how rapidly income increases with hierarchical power in a given firm. Panel A shows binned data for σ_E (the lognormal scale parameter) for Compustat firms. Each dot indicates σ_E for the given firm-size bin. The straight line indicates the modeled relation (Eq. 30). Panel B shows how the modeled dispersion of β decreases with firm size. Panel C compares the distribution of β for Compustat firms to the simulated distribution, created by injecting Compustat-sized firms into the model. Panel D uses the same method to compare the CEO pay ratio in Compustat firms to that produced by injecting Compustat-sized firms into the model. Contour P10 contains 10% of the data, contour P50 contains 50%, and contour P90 contains 90% of the data.

To test the above algorithm, I apply it back to Compustat firms. I 'inject' Compustat-sized firms into the model, and test if the properties of these simulated firms match the properties of the real-world Compustat firms. Figures Figure 20C and Figure 20D show the results. Figure 20C compares the simulated distribution of β to the values fitted to Compustat firms ($\beta_{Compustat}$). Figure 20D shows how the CEO pay ratio grows with hierarchical power. Instead of plotting raw data, this figure shows a contour of the data for various density thresholds. In both cases, the model reasonably approximates Compustat values.

D.5 Mean Pay in the Base Hierarchical Level ($\bar{I}_1$)

The base-level pay parameter ($\bar{I}_1$) determines average pay in simulated firms. As with β, I allow $\bar{I}_1$ to vary across firms. I restrict this variation using a two-step process. I first 'tune' the model to data from Compustat firms. This creates a distribution of base pay specific to Compustat firms. I then fit this data with a parameterized distribution, from which simulation parameters are randomly chosen.

D.6 Estimating Base Pay $\bar{I}_1$ in Compustat Firms

Having already fitted a hierarchical pay structure to each Compustat firm (in the process of estimating β), we can use this data to estimate base pay for each firm. To do this, we set up a ratio between base level pay ($\bar{I}_1$) and firm mean pay ($\bar{I}_T$) for both the model and Compustat data:

$$\frac{\bar{I}_1^{\text{Compustat}}}{\bar{I}_T^{\text{Compustat}}} = \frac{\bar{I}_1^{\text{model}}}{\bar{I}_T^{\text{model}}} \tag{31}$$

Because the Compustat data covers multiple years, I first adjust firm mean pay ($\bar{I}_T^{\text{Compustat}}$) for inflation. I normalize $\bar{I}_T^{\text{Compustat}}$ by dividing it by the average of firm mean pay for all firms in the given year.

The modeled ratio between base pay and firm mean pay ($\bar{I}_1^{\text{model}}/\bar{I}_T^{\text{model}}$) is independent of the choice of base pay. This is because the modeled firm mean pay is actually a function of base pay (see Eq. 23). If we run the model with $\bar{I}_1^{\text{model}} = 1$, then Eq. 31 reduces to:

$$\frac{\bar{I}_1^{\text{Compustat}}}{\bar{I}_T^{\text{Compustat}}} = \frac{1}{\bar{I}_T^{\text{model}}} \tag{32}$$

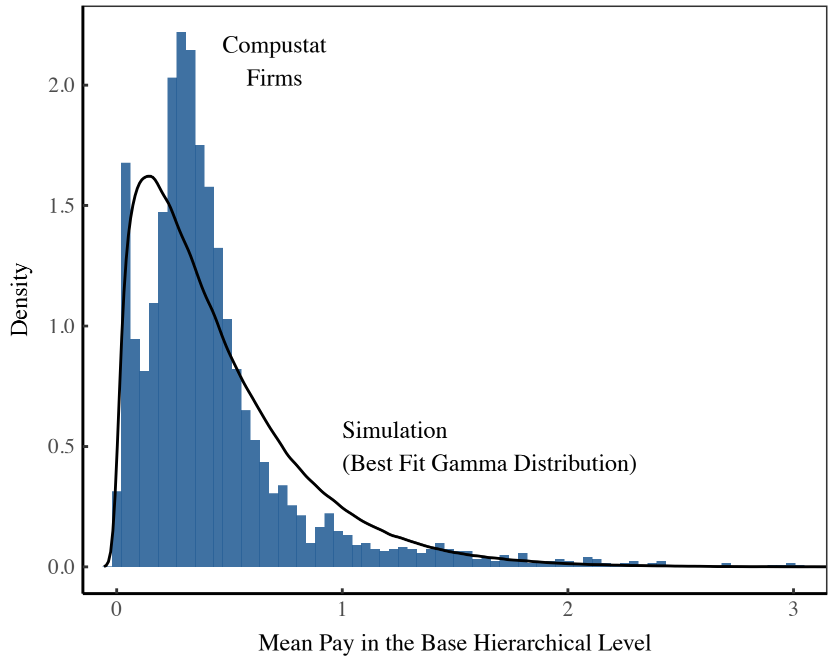

Figure 21: Modeling the Distribution of Base Pay in Compustat Firms

This figure shows the distribution of (fitted) mean pay in the base level of Compustat firms. Pay is normalized so that the average income in the Compustat sample (in each year) is 1. I model this data with a gamma distribution.

To solve for $\bar{I}_1^{\text{Compustat}}$, we rearrange Eq. 32 to get :

$$\bar{I}_1^{\text{Compustat}} = \frac{\bar{I}_T^{\text{Compustat}}}{\bar{I}_T^{\text{model}}} \tag{33}$$

The model uses Eq. 33 to estimate base pay $\bar{I}_1^{\text{Compustat}}$ for each Compustat firm. Code implementing this method is located in `fit_beta.h` in the Supplementary Material.

D.6.1 Simulating the Distribution of Base Pay $\bar{I}_1$ for All US Firms

Once each Compustat firm has a fitted value for base-level mean pay, we fit this data with a parametric distribution. I use a gamma distribution to model the distribution of base-level pay (Fig. 21).

Note that the distribution of base pay in Compustat firms has a bimodal structure. I do not try to replicate this structure because I feel it is not representative of the US firm population. The lower mode in the Compustat data is composed mostly of chain restaurants, which seem to be over-represented in the Compustat sample.

While the gamma distribution fits the Compustat data quite roughly, it fits better than other parameterized distributions.

Once we have fitted the Compustat data with a gamma distribution, we then sample from this distribution to simulate base-level pay in modeled firms. Code implementing this method is located in `base_pay_sim.h` in the Supplementary Material.

D.7 Power-Income Noise Factor

I model noise (ε) in the power-income relation using a lognormal random variate:

$$\varepsilon \sim \ln \mathcal{N}(\mu, \sigma) \tag{34}$$

The noise factor is designed to reproduce the average income dispersion within hierarchical ranks of case-study firms. I set the lognormal scale parameter (μ) so that the distribution of ε has a mean of 1:

$$\mu = \ln(1) - \frac{1}{2}\sigma^2 \tag{35}$$

To determine σ, we first calculate the mean Gini index ($\bar{G}$) of inequality within hierarchical ranks of case-study firms (Fig. 22). We then calculate σ using:

$$\sigma = 2 \cdot \text{erf}^{-1}(\bar{G}) \tag{36}$$

This equation is derived from the definition of the Gini index of a lognormal distribution: $G = \text{erf}(\sigma/2)$.

To incorporate uncertainty in the case-study data, each model iteration uses a different bootstrap resample to calculate $\bar{G}$. Code implementing this method is located in `boot_sigma.h` in the Supplementary Material.

D.8 Class Composition of Individual Income

I model the class composition of individual income as a function of hierarchical power. The capitalist fraction of income (K_{frac}) increases with the logarithm of hierarchical power (P), with some associated noise (ε):

$$K_{\text{frac}} \propto \ln(P) \cdot \varepsilon \tag{37}$$

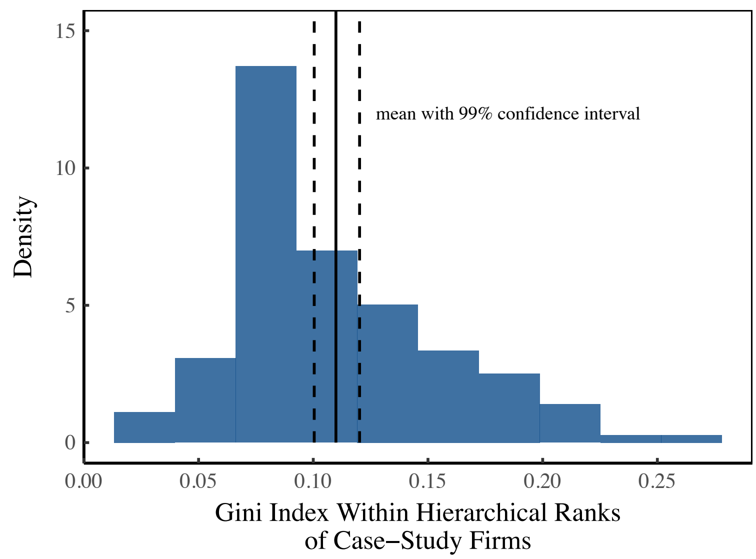

Figure 22: Determining the Power-Income 'Noise' Parameter

This figure shows the distribution of income dispersion (measured using the Gini index) within hierarchical ranks of case-study firms (Audas *et al.* 2004, Baker *et al.* 1993, Dohmen *et al.* 2004, Grund 2005, Lima 2000, Morais and Kakabadse 2014, Treble *et al.* 2001). I use the mean of this distribution (with associated uncertainty) to set the power-income noise parameter ε. For methods used to calculate within-rank income dispersion in the case-study data, see the Appendix in Fix (2018c).

The labor fraction of income (L_{frac}) is then the complement of the capitalist fraction:

$$L_{\text{frac}} = 1 - K_{\text{frac}} \tag{38}$$

Conceptually, then, class-based income is a simple function of hierarchical power. The complication, however, is that the dispersion ε in capitalist income fraction is not simple. To replicate dispersion in the capitalist fraction of CEO income, I model K_{frac} as a partially truncated normal distribution. I draw values from a truncated normal distribution with an upper bound of 1:

$$K_{\text{frac}} \sim \mathcal{N}(K_{\text{frac}}; \mu_K, \sigma_K); \quad K_{\text{frac}} \leq 1, \tag{39}$$

I then create a lower bound by setting to zero all randomly drawn values that are less than zero:

$$\text{if}(K_{\text{frac}}) < 0 \ \text{ then } \ K_{\text{frac}} = 0 \tag{40}$$

The parameters μ_K and σ_K are both functions of firm size. To model these parameters, I first fit a truncated normal distribution to CEO capitalist income fraction data, binned by firm size. Within each firm-size bin, I use numeric optimization to find the values for μ_K and σ_K that best reproduce the distribution of the capitalist fraction of CEO income. I then model μ_K as a log-linear function of firm size (Fig. 18A):

$$\mu_K = c_1 \ln(P) + c_2 \tag{41}$$

I model σ_K by first modeling the relative standard deviation $|\sigma_K/\mu_K|$ as a power function of hierarchical power (Fig. 18A):

$$RSD = \left| \frac{\sigma_K}{\mu_K} \right| = c_1 P^{c_2} \tag{42}$$

I then define σ_K as:

$$\sigma_K = RSD \cdot |\mu_K| \tag{43}$$

I test the above algorithm by applying it back to the CEO data. I 'inject' (into the model) individuals with the same hierarchical power as those in our CEO sample. I then simulate the capitalist component of their income. Figures 23C and 23D show how this simulation compares to the original data. Figure 23C compares the distribution of the capitalist fraction of income for all individuals. Figure 23D shows how the capitalist fraction of income grows with hierarchical power. In both cases, the model reproduces (with reasonable accuracy) the trends found in the empirical data.

To incorporate uncertainty, each model iteration uses different firm-size bins to estimate μ_K and σ_K. Code implementing this method is located in `k_func.h` in the Supplementary Material.

D.9 Summary of Model Structure

The model is implemented in C++ using a modular design. Each major task is carried out by a separate function that is defined in a corresponding header file. Table 14 summarizes the model's structure in the order that functions are called. In each step, I briefly summarize the action that is performed, and reference the section where this action is described in detail.

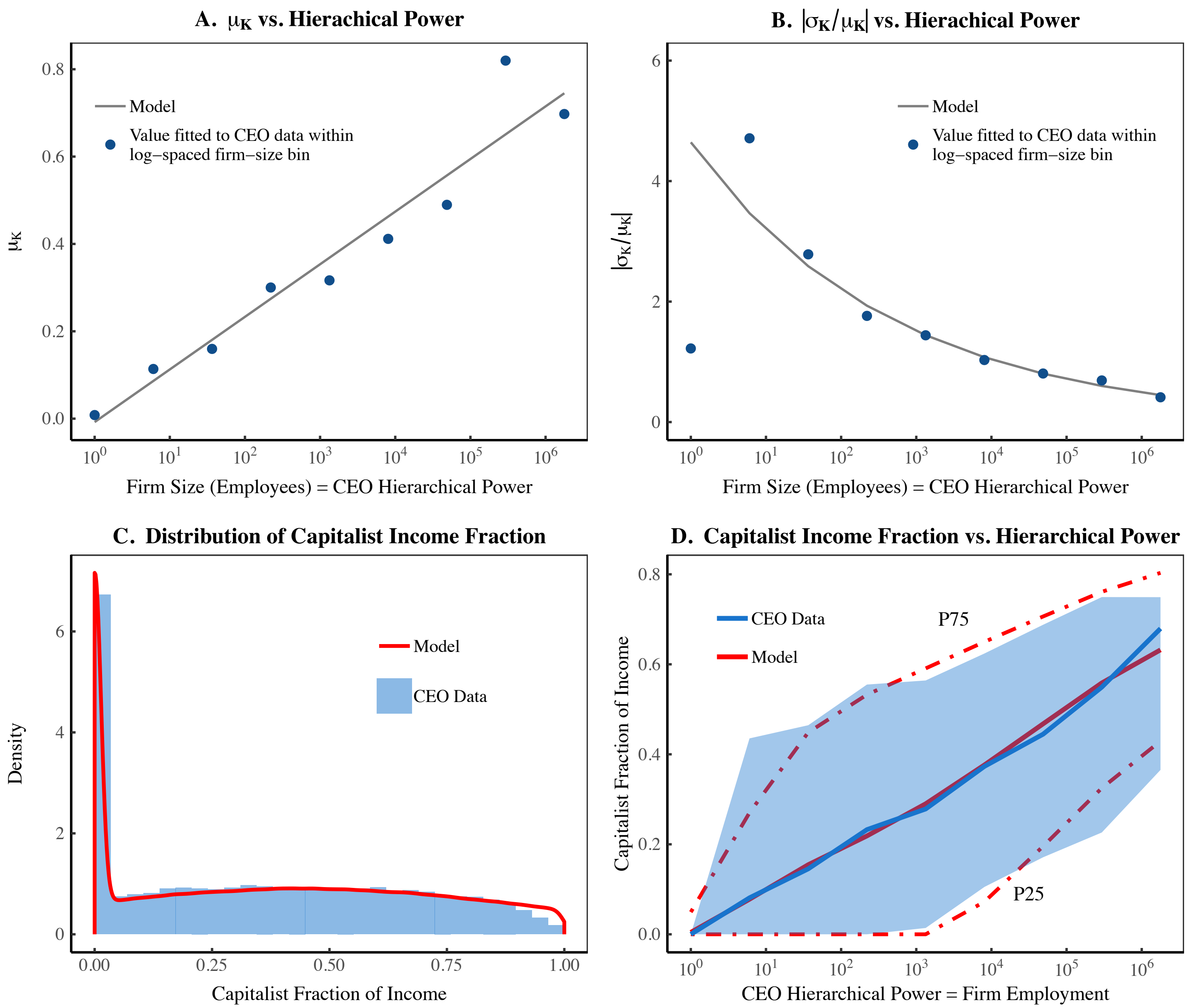

Figure 23: The Capitalist Gradient Model

This figure shows the steps used to simulate the relation between the capitalist fraction of income and hierarchical power. Using Eqs. 39 and 40, Panels A and B show how the parameters μ_K and σ_K vary with firm size. Each point represents the value fitted to binned CEO data. The line indicates the modeled relation. Panels C and D compare the CEO data to simulated data, created by injecting individuals with the same hierarchical power as our CEOs into the model. Panel C shows the distribution of the capitalist fraction of income. Using data binned by firm size, Panel D shows how the capitalist fraction of income changes with hierarchical power.

Table 14: Structure of the Hierarchy Model

Step	Action	Reference Section	Parameter(s)	File(s)
1	Bootstrap firm case-study data	D.2, D.7	a, b, σ	`boot_span.h` `boot_sigma.h`
2	Estimate employment in the base hierarchical level of each Compustat firm	D.3	E_1	`base_fit.h`
3	Fit power-income exponent to each Compustat firm	D.4.1	β	`fit_beta.h`
4	Estimate base-level pay in each Compustat firm	D.6	$\bar{I}_1$	`fit_beta.h`
5	Generate a firm-size distribution that follows a power law	D.1	α	`rpld.h`
6	Estimate base-level employment in each simulated firm	D.3	E_1	`base_fit.h`
7	Model the distribution of base-level pay. Assign a value to each simulated firm	D.6.1	$\bar{I}_1$	`base_pay_sim.h`
8	Model the distribution of the hierarchical pay-scaling parameter. Assign a value to each simulated firm	D.4.2	β	`beta_sim.h`
9	Run hierarchy model	C	all but capitalist gradient parameters	`model.h`
10	Assign class composition to individual income	D.8	μ_K, σ_K	`k_func.h`

Notes: Model code uses Armadillo, an open-source linear algebra library for C++ (Sanderson and Curtin 2016).

References

Audas, R., Barmby, T. and Treble, J., 2004. *Luck, effort, and reward in an organizational hierarchy*. *Journal of Labor Economics*, 22 (2), 379–395.

Axtell, R.L., 2001. *Zipf distribution of US firm sizes*. *Science*, 293, 1818–1820.

Baker, G., Gibbs, M. and Holmstrom, B., 1993. *Hierarchies and compensation: A case study*. *European Economic Review*, 37 (2-3), 366–378.

Becker, G.S., 1962. *Investment in human capital: A theoretical analysis*. *Journal of Political Economy*, 70 (5, Part 2), 9–49.

Bendix, R., 1980. *Kings or people: Power and the mandate to rule*. Berkeley: University of California Press.

Berle, A.A. and Means, G.C., 1932. *The modern corporation and private property*. London: Transaction Books.

Bichler, S. and Nitzan, J., 2017. *Growing Through Sabotage: Energizing Hierarchical Power*. *Capital as Power Working Papers*, 2017 (02), 1–59.

Brown, C., 1988. *Income Distribution in an Institutional World*. In G. Mangum and P. Philips (Eds.), *Three Worlds of Labor Economics*, volume 51, Armonk, N.Y.: M.E. Sharpe, pages 515–530.

Champernowne, D.G., 1953. *A model of income distribution*. *The Economic Journal*, 63 (250), 318–351.

Clark, J.B., 1899. *The Distribution of Wealth*. New York: Macmillan.

Clauset, A., Shalizi, C.R. and Newman, M.E., 2009. *Power-law distributions in empirical data*. *SIAM review*, 51 (4), 661–703.

Cohen, A.J. and Harcourt, G.C., 2003. *Retrospectives: Whatever happened to the Cambridge capital theory controversies?* *Journal of Economic Perspectives*, 17 (1), 199–214.

Commons, J.R., 1924. *Legal foundations of capitalism*. London: Transaction Publishers.

Di Muzio, T., 2015. *The 1% and the rest of us: A political economy of dominant ownership*. Zed Books Ltd.

Dohmen, T.J., Kriechel, B. and Pfann, G.A., 2004. *Monkey bars and ladders: The importance of lateral and vertical job mobility in internal labor market careers*. *Journal of Population Economics*, 17 (2), 193–228.

Dugger, W.M., 1989. *Corporate hegemony*, volume 97. Praeger Pub Text.

Efron, B. and Tibshirani, R.J., 1994. *An introduction to the bootstrap*. London: CRC press.

Felipe, J. and Fisher, F.M., 2003. *Aggregation in production functions: what applied economists should know*. *Metroeconomica*, 54 (2–3), 208–262.

Fitzgerald, F.S., 1926. *The rich boy*. Feedbooks.

Fix, B., 2018a. *The Growth of US Top Income Inequality: A Hierarchical Redistribution Hypothesis*. *SocArXiv*, doi:10.31235/osf.io/suqnk.

Fix, B., 2018b. *Hierarchy and the Power-Law Income Distribution Tail*. *Journal of Computational Social Science*, 1 (2), 471–491, doi:10.1007/s42001-018-0019-8.

Fix, B., 2018c. *Personal Income and Hierarchical Power*. *SocArXiv*, doi:10.31235/osf.io/pb475.

Fix, B., 2018d. *The Trouble with Human Capital Theory*. *Real-World Economics Review*, 86, 15–32.

Fix, B., 2019. *Energy, hierarchy and the origin of inequality*. *PLOS ONE*, 14 (4), e0215692, doi:10.1371/journal.pone.0215692.

Gabaix, X., Lasry, J.M., Lions, P.L. and Moll, B., 2016. *The dynamics of inequality*. *Econometrica*, 84 (6), 2071–2111.

Gaffeo, E., Gallegati, M. and Palestrini, A., 2003. *On the size distribution of firms: additional evidence from the G7 countries*. *Physica A: Statistical Mechanics and its Applications*, 324 (1–2), 117–123, doi:10.1016/S0378-4371(02)01890-3.

Galbraith, J.K., 1985. *The anatomy of power*. Boston, Mass.: Houghton Mifflin Co.

Gillespie, C.S., 2014. *Fitting heavy tailed distributions: the poweRlaw package*. *arXiv preprint arXiv:1407.3492*.

Glattfelder, J.B., 2010. *Ownership Networks and Corporate Control: Mapping Economic Power in a Globalized World - Google Scholar*. Ph.D. thesis, ETH Zurich.

Glattfelder, J.B. and Battiston, S., 2009. *Backbone of complex networks of corporations: The flow of control*. *Physical Review E*, 80 (3), 036104.

Grund, C., 2005. *The wage policy of firms: comparative evidence for the US and Germany from personnel data*. *The International Journal of Human Resource Management*, 16 (1), 104–119.

Harcourt, G.C., 2015. *On the Cambridge, England, Critique of the Marginal Productivity Theory of Distribution*. *Review of Radical Political Economics*, 47 (2), 243–255, doi:10.1177/0486613414557915.

Hemingway, E., 1936. *The Snows of Kilimanjaro*. *Esquire*, August.

Hodgson, G.M., 2005. *The fate of the Cambridge capital controversy*. In *Capital controversy, post Keynesian economics and the history of economic thought*, Routledge, pages 112–125.

Hopkins, M. and Lazonick, W., 2016. *The mismeasure of mammon: Uses and abuses of executive pay data*. *Institute for New Economic Thinking*, Working Paper No. 49, 1–60.

Huber, E., Huo, J. and Stephens, J.D., 2017. *Power, policy, and top income shares*. *Socio-Economic Review*, 0 (0), 1–23, doi:10.1093/ser/mwx027.

Lenski, G.E., 1966. *Power and privilege: A theory of social stratification*. Chapel Hill: UNC Press Books.

Lima, F., 2000. *Internal Labor Markets: A Case Study*. *FEUNL Working Paper*, 378.

Mann, M.E., Bradley, R.S. and Hughes, M.K., 1999. *Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations*. *Geophysical Research Letters*, 26 (6), 759–762, doi:10.1029/1999GL900070.

Marx, K., 1867. *Capital, Volume I*. Harmondsworth: Penguin/New Left Review.

Mills, C.W., 1956. *The power elite*. Oxford: Oxford University Press.

Mincer, J., 1958. *Investment in human capital and personal income distribution*. *The Journal of Political Economy*, 66 (4), 281–302.

Morais, F. and Kakabadse, N.K., 2014. *The Corporate Gini Index (CGI) determinants and advantages: Lessons from a multinational retail company case study*. *International Journal of Disclosure and Governance*, 11 (4), 380–397.

Munkirs, J.R., 1985. *The transformation of American capitalism: From competitive market structures to centralized private sector planning*. Armonk, N.Y.: ME Sharpe Inc.

NCEO, 2017. *ESOPs by the Numbers*.

Nirei, M. and Aoki, S., 2016. *Pareto distribution of income in neoclassical growth models*. *Review of Economic Dynamics*, 20, 25–42.

Nitzan, J. and Bichler, S., 2009. *Capital as Power: A Study of Order and Creorder*. New York: Routledge.

Pareto, V., 1897. *Cours d'economie politique*, volume 1. Librairie Droz.

Peach, J.T., 1987. *Distribution and economic progress*. *Journal of Economic Issues*, 21 (4), 1495–1529.

Piketty, T., Saez, E. and Zucman, G., 2017a. *Distributional national accounts: methods and estimates for the United States*. *The Quarterly Journal of Economics*, 133 (2), 553–609.

Piketty, T., Saez, E. and Zucman, G., 2017b. *Distributional National Accounts: Methods and Estimates for the United States Data Appendix. Online at: http://gabriel-zucman. eu/files/PSZ2018DataAppendix. pdf.*

Pullen, J., 2009. *The marginal productivity theory of distribution: a critical history*. London: Routledge.

Robinson, J., 1953. *The production function and the theory of capital*. *The Review of Economic Studies*, 21 (2), 81–106.

Rognlie, M., 2016. *Deciphering the fall and rise in the net capital share: accumulation or scarcity?* *Brookings papers on economic activity*, 2015 (1), 1–69.

Rubin, I.I., 1973. *Essays on Marx's theory of value*. Black Rose Books Ltd.

Rutherford, R., 1955. *Income distributions: a new model*. *Econometrica: Journal of the Econometric Society*, 23 (3), 277–294.

Samuelson, P.A., 1971. *Understanding the Marxian notion of exploitation: a summary of the so-called transformation problem between Marxian values and competitive prices*. *Journal of Economic Literature*, 9 (2), 399–431.

Sanderson, C. and Curtin, R., 2016. *Armadillo: a template-based C++ library for linear algebra*. *Journal of Open Source Software*.

Schultz, T.W., 1961. *Investment in human capital*. *The American Economic Review*, 51 (1), 1–17.

Sidanius, J. and Pratto, F., 2001. *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Simon, H.A., 1955. *On a class of skew distribution functions*. *Biometrika*, 42 (3/4), 425–440.

Song, J., Price, D.J., Guvenen, F., Bloom, N. and Von Wachter, T., 2016. *Firming up inequality*. Technical report, National Bureau of Economic Research.

Sraffa, P., 1960. *Production of commodities by means of commodities: Prelude to a critique of economic theory*. London: Cambridge University Press.

Toda, A.A., 2012. *The double power law in income distribution: Explanations and evidence*. *Journal of Economic Behavior & Organization*, 84 (1), 364–381.

Tool, M., 2017. *The discretionary economy: A normative theory of political economy*. New York: Routledge.

Tool, M.R. and Samuels, W.J., 1989. *The economy as a system of power*. New Brunswick, NJ: Transaction Books.

Treble, J., Van Gameren, E., Bridges, S. and Barmby, T., 2001. *The internal economics of the firm: further evidence from personnel data*. *Labour Economics*, 8 (5), 531–552.

Veblen, T., 1904. *The Theory of Business Enterprise*. New York: Martino Fine Books.

Veblen, T., 1923. *Absentee ownership: business enterprise in recent times: the case of America*. Transaction Pub.

Virkar, Y. and Clauset, A., 2014. *Power-law distributions in binned empirical data*. *The Annals of Applied Statistics*, 8 (1), 89–119.

Vitali, S., Glattfelder, J.B. and Battiston, S., 2011. *The network of global corporate control*. *PloS one*, 6 (10), e25995.

Weber, M., 1978. *Economy and society: An outline of interpretive sociology*. Univ of California Press.

Wicksteed, P.H., 1894. *An Essay on the Co-ordination of the Laws of Distribution (1932 Edition)*. London: London School of Economics.

Wright, E.O., 1979. *Class structure and income determination*, volume 2. New York: Academic Press.