

WORKING PAPERS ON CAPITAL AS POWER

No. 2018/05

The Growth of US Top Income Inequality A Hierarchical Redistribution Hypothesis

Blair Fix

July 2018

<http://www.capitalaspower.com/?p=2512>

The Growth of US Top Income Inequality: A Hierarchical Redistribution Hypothesis

Blair Fix*

July 23, 2018

Abstract

What accounts for the growth of US top income inequality? This paper proposes a hierarchical redistribution hypothesis. The idea is that US firms have systematically redistributed income to the top of the corporate hierarchy. I test this hypothesis using a large-scale hierarchy model of the US private sector. My method is to vary the rate that income scales with hierarchical rank within modeled firms. I find that this model is able to reproduce four intercorrelated US trends: (1) the growth of the top 1% income share; (2) the growth of the CEO pay ratio; (3) the growth of the dividends share of national income; and (4) the ‘fattening’ of the entire income distribution tail. This result supports the hierarchical redistribution hypothesis. It is also consistent with the available empirical evidence on within-firm income redistribution.

JEL Subject Codes: B5, C5; D31; D33;

Keywords: top income share; inequality; hierarchy; power; functional income distribution; personal income distribution;

*Author contact: blairfix@gmail.com

1 Introduction

There is now overwhelming evidence that US income inequality has increased in recent decades. Interestingly, much of this inequality growth has occurred in the *tail* of the income distribution. Top earners, it seems, have pulled away from the rest of the pack (Alvaredo et al., 2013a; Atkinson and Piketty, 2010; Piketty and Saez, 2001, 2006). But what explains this trend? There is no consensus on this issue.

This paper proposes a hierarchical redistribution hypothesis. The idea is that US firms have systematically redistributed income to the top of the corporate hierarchy. What suggests that this is true? The growth of the CEO pay ratio is one indicative trend. But there are others. In previous work (Fix, 2018b), I have found that the power-law scaling of top incomes might be due to hierarchical pay within firms. If so, changes in the distribution tail might be due to changes in the hierarchical pay structure of firms. There is also a curious correlation between the US top 1% income share and the dividends share of national income. I think this may also be related to firm hierarchy. In Fix (2018a), I find evidence that individuals' capitalist income fraction relates to hierarchical rank within firms. This suggests that hierarchical redistribution might lead to functional income redistribution.

My goal, in this paper, is to move beyond inductive reasoning based on suggestive facts. Instead, I test if an empirically constrained hierarchy model can reproduce the following four intercorrelated US trends:

1. The growth of the top 1% income share;
2. The growth of the CEO pay ratio;
3. The growth of the dividends share of national income;
4. The fattening of the entire income distribution tail
(as measured by a power-law exponent).

My method is to vary the rate at which income scales with hierarchical rank within the model. I then see if the model can reproduce the above trends. The results are conclusive. Although not tuned to do so, the US hierarchy model is able to replicate these trends with reasonable accuracy. This result supports the hierarchical redistribution hypothesis.

This finding has important implications for the study of income redistribution. First, firm hierarchical pay structure needs to be better studied. The results here suggest that hierarchical redistribution has driven recent trends in top income inequality. To understand this process better, we need more data. Second,

the results suggest that limiting executive compensation could mitigate top income inequality. Lastly, these results beg the question: *why* have US firms shifted income towards the top? This question likely has no simple answers.

This paper is laid out as follows. Section 2 briefly reviews some of the existing models of income redistribution. Section 3 outlines the US historical data that I aim to reproduce. Section 4 discusses the US hierarchy model and the methods used to test the hierarchical redistribution hypothesis. Section 5 reviews the model's results. Lastly, sections 6 and 7 discuss the significance of these results, and outline areas for future research.

2 Theories of Increasing Income Inequality

The well-documented rise in US income inequality has led to a deluge of theories seeking to explain this trend. I review only a small part of the literature here. It is useful to separate these theories into 'why' and 'how' explanations. 'Why' theories look for socio-economic causes that are driving the rise in income inequality. In contrast, 'how' theories focus on models that generate power-law distributions. These generation mechanisms are then used to understand distributional changes.

I begin with 'why' theories of increasing top income inequality. Perhaps the most well-known is Piketty's ' $r > g$ ' hypothesis. Piketty proposes that inequality will grow whenever the average return to capital (r) is greater than the rate of economic growth (g) (Piketty, 2014; Piketty and Zucman, 2015). Another hypothesis is that inequality has risen because top tax rates have declined (Hubmer et al., 2016; Piketty et al., 2014). Other authors propose that globalization may be driving increases in inequality (Feenstra and Hanson, 1996; Haskel et al., 2012). Yet another possibility is that there is a mismatch between the supply and demand of skilled workers (Autor, 2014). 'Why' theories are important because they seek socio-economic causes behind inequality trends. However, these theories do not usually explain the *specific* changes in the income distribution tail. This is what 'how' theories try to do.

Most 'how' theories of growing inequality are based on the stochastic (random) growth of individual income. The starting point is the power-law distribution of top incomes. Over a century ago, Pareto (1897) discovered that income distribution tails usually have a power-law form. This means that the probability of finding an individual with income x is roughly proportional to $x^{-\alpha}$, where α is the power-law exponent. Stochastic growth models are the most widely recognized mechanism for generating a power-law distribution. Therefore, they have

become the defacto standard for modeling changes in the income distribution tail.

The basic mechanism of stochastic growth models was first identified by [Champernowne \(1953\)](#). In this model, individual income is subjected to random, multiplicative shocks. This means that income growth rates are random. With a few other assumptions, the model produces a power-law distribution of income. For a good technical review of this stochastic mechanism, see [Kumamoto and Kamihigashi \(2018\)](#), [Mitzenmacher \(2004\)](#), and [Newman \(2005\)](#).

This type of 'how' theory is based on the truism that individuals' incomes *change over time*. Thus, it must be possible to explain the distribution of income in terms of the dynamics of individual income. The same must be true for *changes* in the distribution of income. A growing body of literature pursues this reasoning ([Benhabib et al., 2011, 2016](#); [Gabaix et al., 2016](#); [Jones and Kim, 2018](#); [Nirei, 2009](#); [Nirei and Aoki, 2016](#)).

A problem with this approach is that it is devoid of any institutional context. It deals with individuals in *isolation*. In the real world, individuals earn their income by working for (or owning) *institutions*. I am interested in tracing inequality trends to the changing institutional structure of society. My starting point is H.F. Lydall's (1959) work on *firm hierarchy*. Lydall showed that firm hierarchy could create a power-law distribution of income. The mechanism relies on the contrapuntal tendencies of hierarchical organization. First, employment within firms tends to *decrease* exponentially with hierarchical rank. Second, income tends to *increase* exponentially with hierarchical rank. When combined, these two contrapuntal tendencies produce a power-law distribution of income.

Lydall's 1959 paper was largely speculative. At the time, little was known about the hierarchical structure of firms. However, the empirical study of firm hierarchy has blossomed in the last two decades.¹ There is now enough evidence to explore the distributional consequences of hierarchy. Because the firm hierarchy data is still sparse, I use a model to fill in the empirical gaps. In [Fix \(2018b\)](#), I develop a hierarchy model of the United States. This model extrapolates the available firm-level data to simulate the hierarchical structure of the US private sector. Without tuning it to do so, the model reproduces the power-law scaling

of top US incomes.

To reiterate, the model suggests that firm hierarchy creates the power-law income distribution tail. This result begs a question. Can *changes* in hierarchical pay structure explain the *growth* of top income inequality? This is the hypothesis that I test.

3 The Historical Evidence

My analysis of US income redistribution focuses on the suite of intercorrelated trends shown in Figure 1. As the top 1% income share increased after 1980, the following occurred:

1. The CEO pay ratio increased;
2. The dividends share of national income increased;
3. The tail of the income distribution got ‘fatter’ (as measured by the decline in the power-law exponent).

I believe this evidence suggests that hierarchical redistribution within firms has occurred. I review my reasoning below.

Let’s begin with the growth of the CEO pay ratio. First, let’s assume that CEOs sit at the top of the corporate hierarchy. It follows that an increase in the CEO pay ratio means that top-ranked employees have increased their relative income. In other words, income has been *redistributed* towards the top brass (and away from bottom ranks). I think this is an obvious sign of hierarchical income redistribution.

The ‘fattening’ of the income distribution tail also hints at hierarchical redistribution (note that a smaller power-law exponent indicates a ‘fatter’ tail). Why might this be the case? A key finding in [Fix \(2018b\)](#) is that firm hierarchy can create the power-law distribution of top incomes. It follows that hierarchical redistribution could cause *changes* to the distribution tail.

The connection between CEO pay and income inequality has been widely discussed ([Conyon and Murphy, 2000](#); [Gabaix and Landier, 2008](#); [Krugman, 2005](#); [Mishel and Sabadish, 2012](#); [Piketty and Saez, 2006](#)). It is also well known that

¹For case studies of firm hierarchy, see [Audas et al. \(2004\)](#); [Baker et al. \(1993\)](#); [Dohmen et al. \(2004\)](#); [Grund \(2005\)](#); [Lima \(2000\)](#); [Morais and Kakabadse \(2014\)](#); [Treble et al. \(2001\)](#). For aggregate studies of firm hierarchy, see [Ariga et al. \(1992\)](#); [Bell and Van Reenen \(2012\)](#); [Eriksson \(1999\)](#); [Heyman \(2005\)](#); [Leonard \(1990\)](#); [Main et al. \(1993\)](#); [Mueller et al. \(2016\)](#); [Rajan and Wulf \(2006\)](#); [Tao and Chen \(2009\)](#). For a meta analysis of these studies, see [Fix \(2018c\)](#).

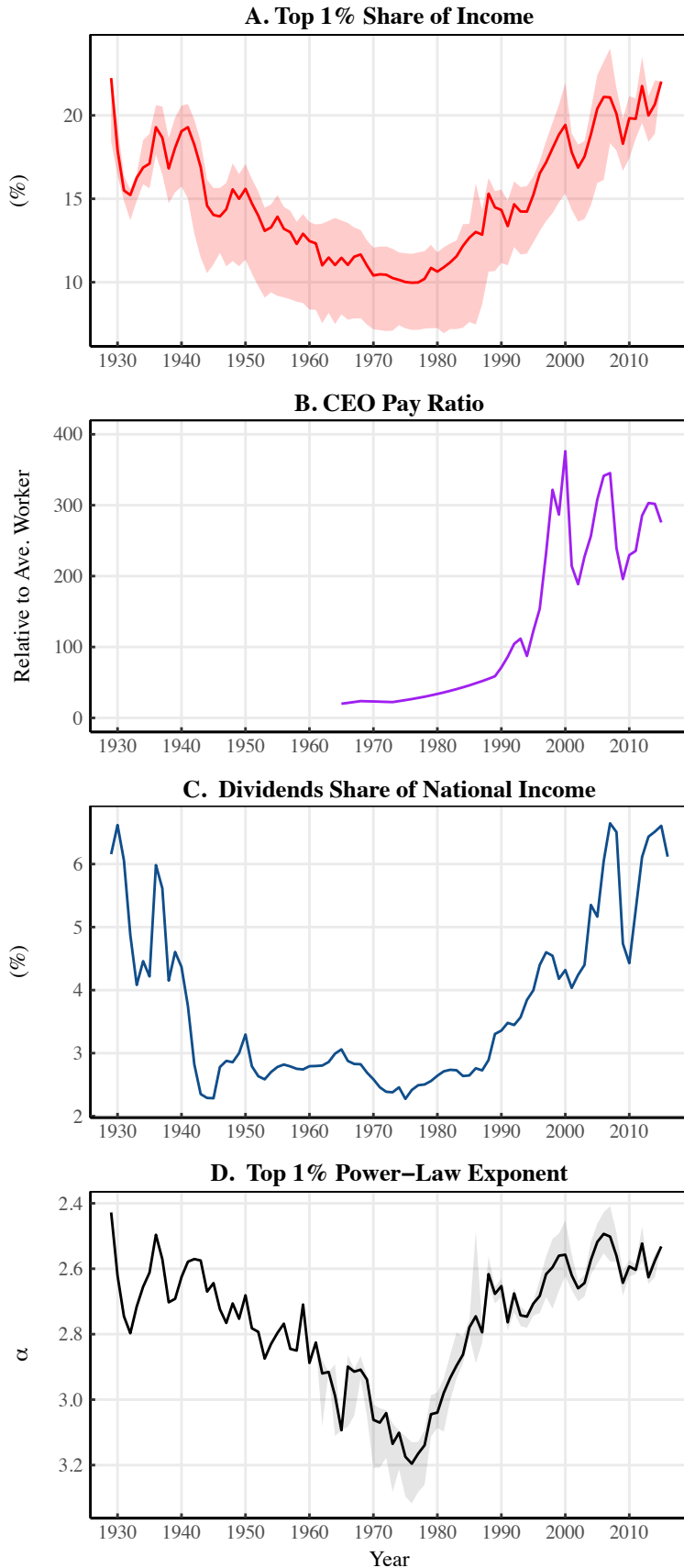


Figure 1: Historical Income Distribution Trends in the United States

This figure shows four intercorrelated trends in US income distribution. Panel A shows the trend in US inequality, as measured by the top 1% income share. Panel B shows the trend in the CEO pay ratio. This ratio is calculated using CEO income in the 350 largest US firms (ranked by sales), compared to the average income of workers in the firms' respective industry. The shaded region indicates the range of 17 different estimates for the top 1% income share. The line represents the median of these estimates. Panel C shows the trend in the dividend share of national income. Panel D shows the fitted power-law exponent for the top 1% of incomes. A smaller exponent indicates a 'fatter' tail. The grey region indicates the range of estimates (when different series are available). The line indicates the median estimate. For sources and methods see Appendix A.

growing top income shares correspond to a fattening of the income distribution tail (Atkinson, 2017; Gabaix et al., 2016; Silva and Yakovenko, 2004). But there is also a correlation between top income share and functional income. This relation has not been widely discussed. As shown in Figure 1, changes in the US dividend share of national income are strongly correlated with changes in the top 1% income share. The correlation coefficient ranges between 0.82 and 0.90, depending on the choice of data. Is this trend also related to hierarchical redistribution?

Evidence in Fix (2018a) suggests that it might be. The key finding here is that capitalist income can be modeled as a gradient function of *hierarchical power*. What is hierarchical power? I define it as proportional to the number of subordinates under an individual's control. Using this definition, I find evidence that the capitalist fraction of individual income increases with hierarchical power. This implies that hierarchical redistribution should change the capitalist share of total income. Why? Suppose that top-ranked individuals have a greater proportion of capitalist income, relative to bottom-ranked individuals. Let's assume that this proportion is fixed over time. If top-ranked individuals increase their share of the pie, the capitalist share of total income should increase as well.

To summarize, the trends in Figure 1 hint at a hierarchical redistribution of income within US firms. To test this hypothesis, I see if a hierarchy model of the US private sector can reproduce these trends.

4 Methods

In principle, we do not need a model to test the hierarchical redistribution hypothesis. If empirical data on firm hierarchy was plentiful, we could test this hypothesis directly. Unfortunately, the available evidence on firm hierarchical pay structure is sparse. This is why a model is useful. A good model can extrapolate the limited evidence and fill in the gaps. This allows us to investigate the distributional consequences of hierarchy using the limited data that exists.

In Fix (2018b), I develop a hierarchy model aimed at this task. The model extrapolates the available firm-level data to simulate the hierarchical structure of the US private sector. Without tuning it to do so, this model reproduces the power-law distribution of top US incomes. This model can be adapted to simulated *changes* in the income distribution. The model contains a single parameter that controls the rate that income increases with hierarchical rank. If we vary this parameter, we simulate a hierarchical redistribution of income. This leads to a very simple test of the hierarchical redistribution hypothesis. I vary the

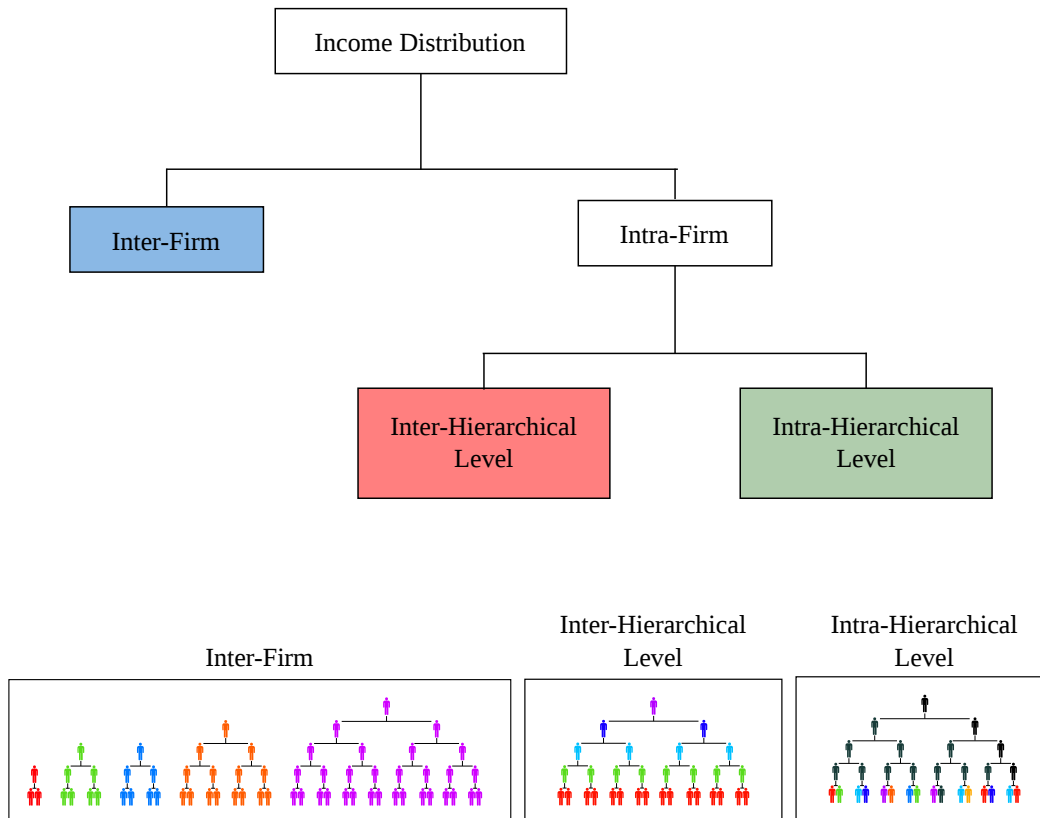


Figure 2: A Tripartite Division of Income Distribution

This figure illustrates the income distribution grouping scheme used by the hierarchy model. The model allows for three sources of income dispersion. *Inter-firm* dispersion consists of differences in (average) pay between firms. Within each firm, there are two further sources of dispersion. *Inter-hierarchical* level dispersion consists of differences in (average) pay between hierarchical levels, while *intra-hierarchical* level dispersion consists of differences in pay within each hierarchical level.

model's hierarchical pay-scaling parameter, and test if this reproduces the suite of intercorrelated trends shown in Figure 1.

4.1 A Hierarchy Model of the United States

The US hierarchy model simulates the hierarchical structure of the US private sector as it was (on average) between 1992 and 2015. The model has three

sources of income dispersion (see Fig. 2):

Source 1: Income dispersion *between* hierarchical levels of each firm;

Source 2: Income dispersion *within* hierarchical levels of each firm;

Source 3: Income dispersion *between* different firms.

The model's parameters are restricted entirely by firm-level empirical data. I do not tune the model to macro-level data. This leads to a simulation of the hierarchical structure of the US private sector as it was (on average) over the last two decades. For a technical discussion of the model, see the Appendix.

To get a sense for what the US hierarchy model 'looks' like, Figure 3A shows a landscape view of the model's structure. Each pyramid represents a different hierarchically organized firm. The size of each pyramid corresponds to the number of employees, height represents hierarchical level, and color represents relative income. To simulate hierarchical redistribution, I perturb this model by varying hierarchical pay. All other parameters remain unchanged.

So far we have simulated individual income *size* only. In [Fix \(2018a\)](#), I add a simulation of the *functional* income of individuals. I hypothesize that capitalist income is a gradient function of one's *hierarchical power*. I call this the 'capitalist gradient' hypothesis. The impetus for this hypothesis is twofold. First, it expands on the work of [Nitzan and Bichler \(2009\)](#), who propose that capitalist income stems from institutionalized power. Second, this hypothesis seeks to address the rise of partial ownership in firms. Who are capitalists? Where do they sit in the corporate hierarchy? In the 19th century, capitalists mostly sat at the top of firm hierarchies. But today, many individuals earn small amounts of capitalist income, if only from retirement investments. Given modern partial ownership practices, I propose that there is a gradient of ownership throughout the firm. This means that the capitalist income fraction should increase with hierarchical rank and power.

In [Fix \(2018a\)](#), I find evidence for this hypothesis. I define hierarchical power within a firm as proportional to the number of subordinates under an individual's control. I put this in formula form as:

$$\text{hierarchical power} = \text{number of subordinates} + 1 \quad (1)$$

The logic of this equation is that all individuals start at a baseline power of 1, indicating that they have control over themselves. Hierarchical power then increases linearly with the number of subordinates. Using this equation I find that the capitalist income fraction of US CEOs scales with hierarchical power. I then generalize this relation using the hierarchy model. Using regressions on

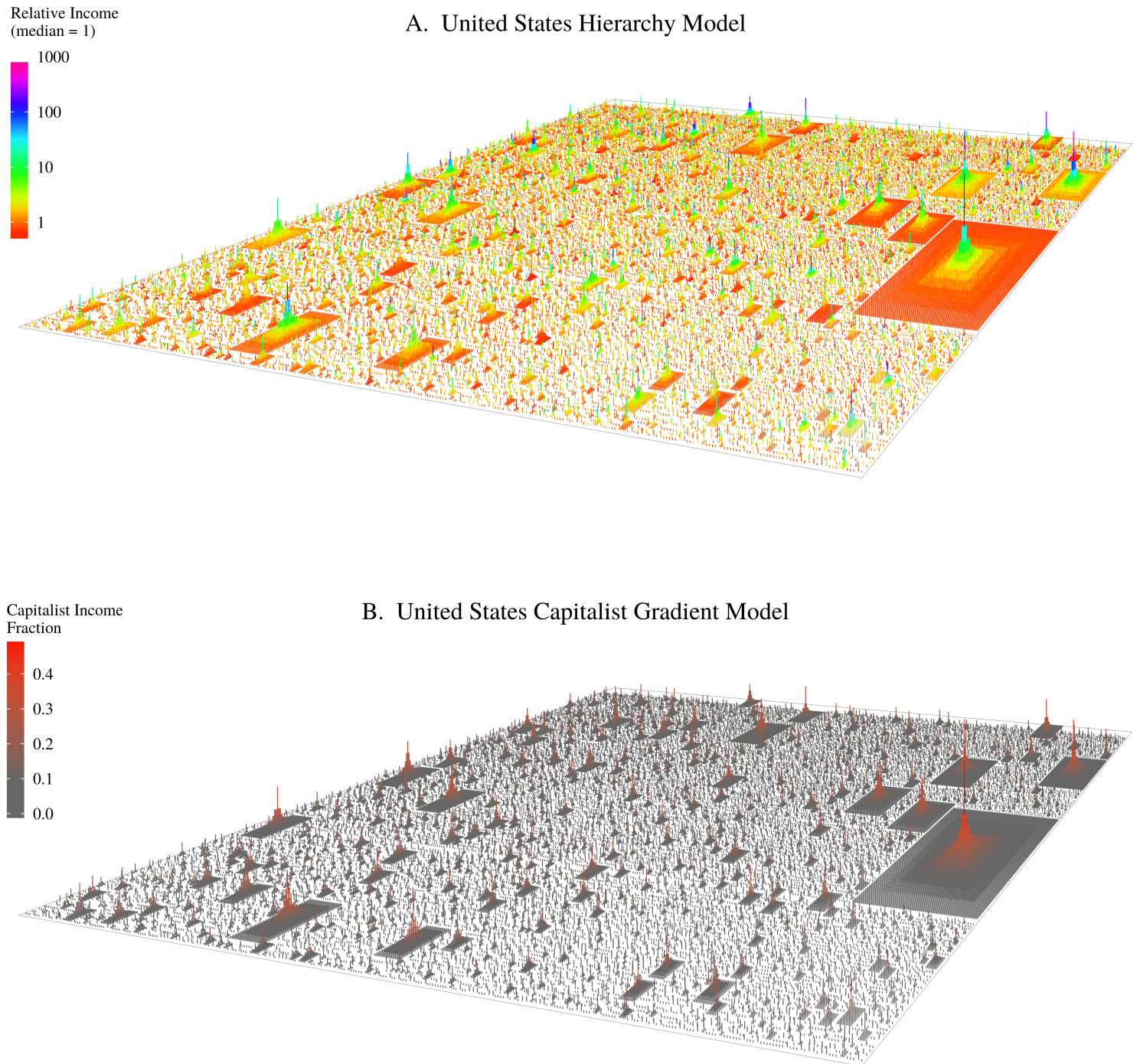


Figure 3: A Landscape View of the US Hierarchy Model

Panel A 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. If you look closely, you will see vertical lines corresponding to individuals. Income (relative to the median) is indicated by color. Panel B visualizes the US ‘capitalist gradient’ model. Capitalist income fraction is indicated by color. This visualization has 20,000 firms — a small sample of the actual model, which uses 1 million firms.

CEO pay data, I model the capitalist income fraction (K_{frac}) of individuals as a logarithmic function of hierarchical power (P):

$$K_{\text{frac}} = 0.05 \ln(P) \quad (2)$$

The ‘capitalist gradient’ model uses this function to assign a capitalist income fraction to each individual in the hierarchy model. In [Fix \(2018a\)](#) I show that this model reproduces key features of the US distribution of capitalist income. [Figure 3B](#) visualizes the model. High capitalist income fractions are concentrated at the tops of large firms. However, almost everyone earns some capitalist income.

4.2 Modeling Hierarchical Redistribution

To test the hierarchical redistribution hypothesis, I perturb the hierarchy model that is shown in [Figure 3](#). I take this model, and vary the hierarchical pay-scaling parameter. [Figure 4](#) shows the effect. The horizontal axis shows hierarchical rank. The vertical axis shows the average income of all the individuals of that rank. For reference, average pay in the bottom rank is indexed to one. Each data point represents a different model iteration. Colors show the average pay-scaling parameter.

When the pay-scaling parameter is small, relative pay increases slowly with hierarchical rank. But when the pay-scaling parameter is large, income increases rapidly with hierarchical rank. The particular values of the pay-scaling parameter are not important. (For a technical explanation of what this parameter does, see the Appendix). What matters is that changing this parameter creates a hierarchical redistribution of income. This leads to a very simple test of the hierarchical redistribution hypothesis. First, we randomly perturb the hierarchical pay-scaling parameter. Second, we see if the results of this perturbation match US historical trends. If the model results match, this supports the hierarchical redistribution hypothesis.

Time Scope

I restrict the scope of analysis to the years 1965 onward. I do this for two reasons. First, the CEO pay ratio data begins in 1965. Second, the model assumes an unchanging firm-size distribution. From the late 1960s onward this assumption is valid — the US firm-size distribution changed very little. But before the 1960s the US firm-size distribution changed rapidly ([Fix, 2017](#)). This change violates the model’s assumptions.

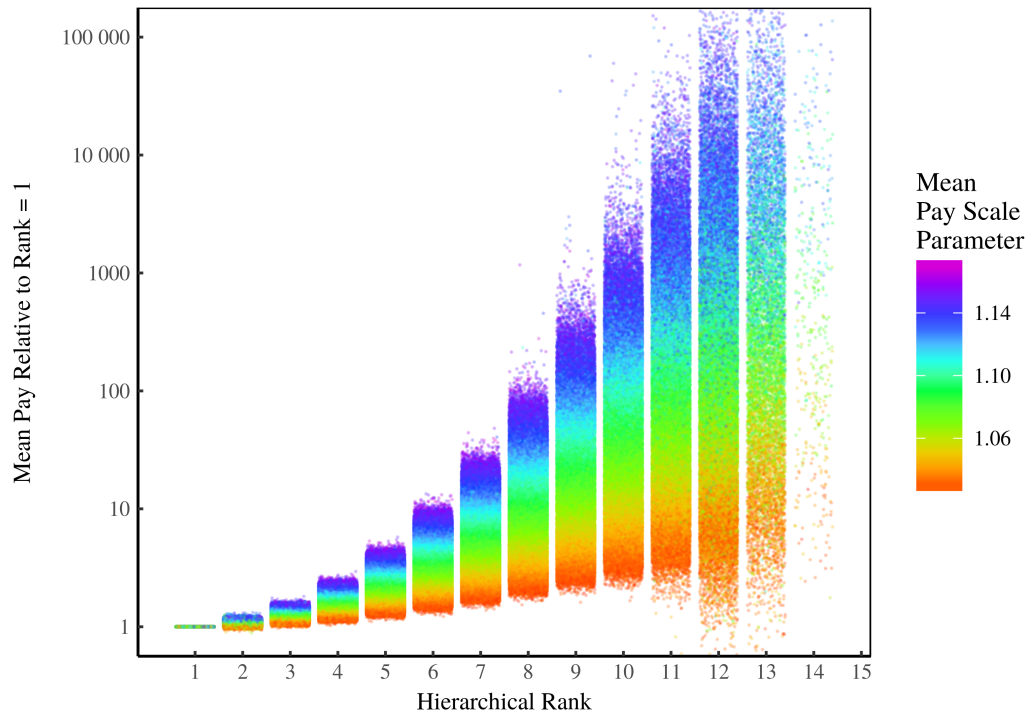


Figure 4: Varying How Income Scales with Hierarchical Rank

This figure shows the results of perturbing the hierarchy model’s hierarchical pay-scaling parameter. The horizontal axis shows hierarchical rank (rank = 1 is the bottom of the firm). The vertical axis shows the average income of all the individuals of that rank. For reference, average pay in the bottom rank is indexed to one. Each data point represents a different model iteration. Colors show the average pay-scaling parameter. Horizontal ‘jitter’ is added to increase the visibility of all data points. Note that the trend becomes increasingly noisy for the very top hierarchical ranks. This is because individuals with very high rank are extremely rare, so the mean encompasses relatively few individuals.

Metrics

I calculate four metrics on the model’s output. These are:

1. The top 1% income share;
2. The CEO pay ratio of the top 350 firms;
3. The dividends share of total income.
4. The power-law exponent of the top 1% of incomes;

A few assumptions are necessary to calculate the CEO pay ratio and the dividends share of income. The US CEO-pay-ratio data ranks firms by *sales* (Mishel and Schieder, 2016). Since the hierarchy model does not simulate sales, I rank

firms by *total payroll*. I do this because sales and total payroll are tightly correlated (see the Appendix).

Calculating the dividend share of national income also requires assumptions. The ‘capitalist gradient’ model predicts *total* capitalist income only. It does not differentiate between interest and dividends. To model the dividend share of national income, I assume that dividends constitute exactly half of capitalist income. (Note that I define capitalist income as the sum of dividends and interest. I exclude taxes on profit). This assumption is based on US historical evidence. The dividends share of capitalist income oscillated around the 50% mark for most of the 20th century. True, there have been important historical variations in the composition of capitalist income (Nitzan and Bichler, 2009). However, these are not included in the model.

5 Model Results

Results of my test of the hierarchical redistribution hypothesis are shown in Figure 5. Each panel shows both US empirical and model relations. As in Figure 4, variation in the hierarchical pay-scaling parameter is indicated by color. Because the model has no time element, I compare only the relation *between* intercorrelated trends. The top 1% share is the common x -axis in all panels.

The main finding is that the hierarchy model reproduces the empirical trends identified in Figure 1. These include:

1. the non-linear scaling of the CEO pay ratio with top 1% income share;
2. the linear scaling of the dividends share of national income with top 1% income share;
3. the non-linear scaling of the power-law exponent of the distribution tail with the top 1% income share.

To be sure, the model’s results are not perfect. The model tends to underestimate the US top 1% income share. This causes a leftward shift in the modeled relations (relative to the empirical relations) in Figure 5. This underestimate may be due to the data used to restrict the model’s parameters. The hierarchy model is heavily dependent on the Compustat database. Because Compustat data includes only publicly traded companies, it is biased towards large firms. This may cause the model to underestimate US inter-firm income dispersion (see Fix, 2018b).

Another problem is that the model’s dividends versus top 1% slope is not quite correct. This slope is heavily dependent on the function relating hierarchi-

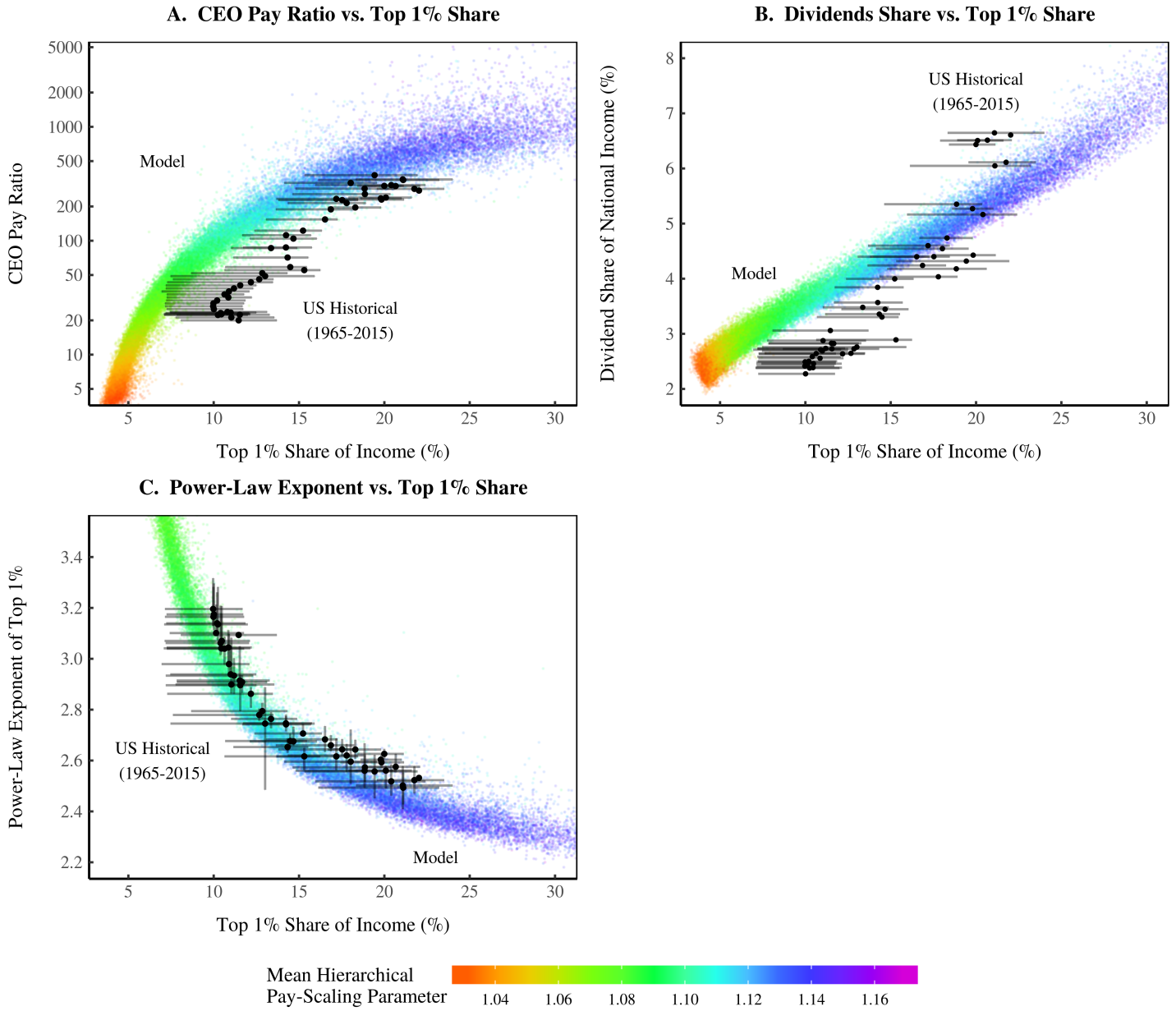


Figure 5: The Hierarchical Redistribution Model vs. US Data

This figure compares model results to historical trends in US income distribution. Model results are produced by varying the hierarchical pay-scaling parameter, indicated by color. Each colored point represents a single model iteration. US empirical data is shown in black, with horizontal error lines indicating the range of 17 different estimates for the top 1% income share. The point indicates the median of these estimates. Panel A plots the CEO pay ratio against the top 1% share. Panel B plots the dividend share of national income against the top 1% share. Panel C plots the fitted power law exponent of the top 1% of incomes against the top 1% income share. For sources and methods, see Appendix A.

cal power to capitalist income fraction. While the function that I use is based on empirical data, there is tremendous uncertainty in this relation. More empirical research is needed to understand the source of this model discrepancy.

Discrepancies aside, it is surprising that the model reproduces most of the historical trends. And this is accomplished by varying a single parameter that controls hierarchical pay. This result supports the hierarchical redistribution hypothesis.

6 Discussion

While we should always be cautious about drawing conclusions from a model, I want to offer my thoughts on the significance of these results. The model suggests that a single process can account for most of the intercorrelated US trends shown in Figure 1. According to the model, US firms have systematically redistributed income from the bottom to the top of the corporate hierarchy.

Figure 6 visualizes the model's estimate of this transformation. On top is a simulation of the United States in 1965. On the bottom is a simulation of the United States in 2015. The difference between the two versions of the United States is visible mostly at the the tops of large firms. Here we see an order of magnitude increase in relative pay. Note that within small firms, there is little discernible change.

This is consistent with the landmark findings of [Song et al. \(2016\)](#). Using a massive dataset from the Social Security Administration, Song et al. study how the pay structure within and between US firms has changed since 1981. They find that there has been a substantial increase in inequality within very large firms (those with more than 10,000 employees). In contrast, inequality within smaller firms has increased very little.

Song et al. also measure how changes in between- and within-firm income dispersion have affected US inequality. The results are surprising. They find a clear distinction between the bottom 99% of individuals versus the top 1%. For the bottom 99%, the growth of income inequality is mostly due to rising income dispersion *between* firms. However, for the top 1%, rising inequality is mostly due to rising income dispersion *within* firms.

This finding is consistent with the results in this paper. I find that increases in the top 1% income share are largely due to a hierarchical redistribution within firms. The work of Song et al. is also consistent with the findings in [Fix \(2018b\)](#). Here, I find a clear division between the *body* and *tail* of the US income distri-

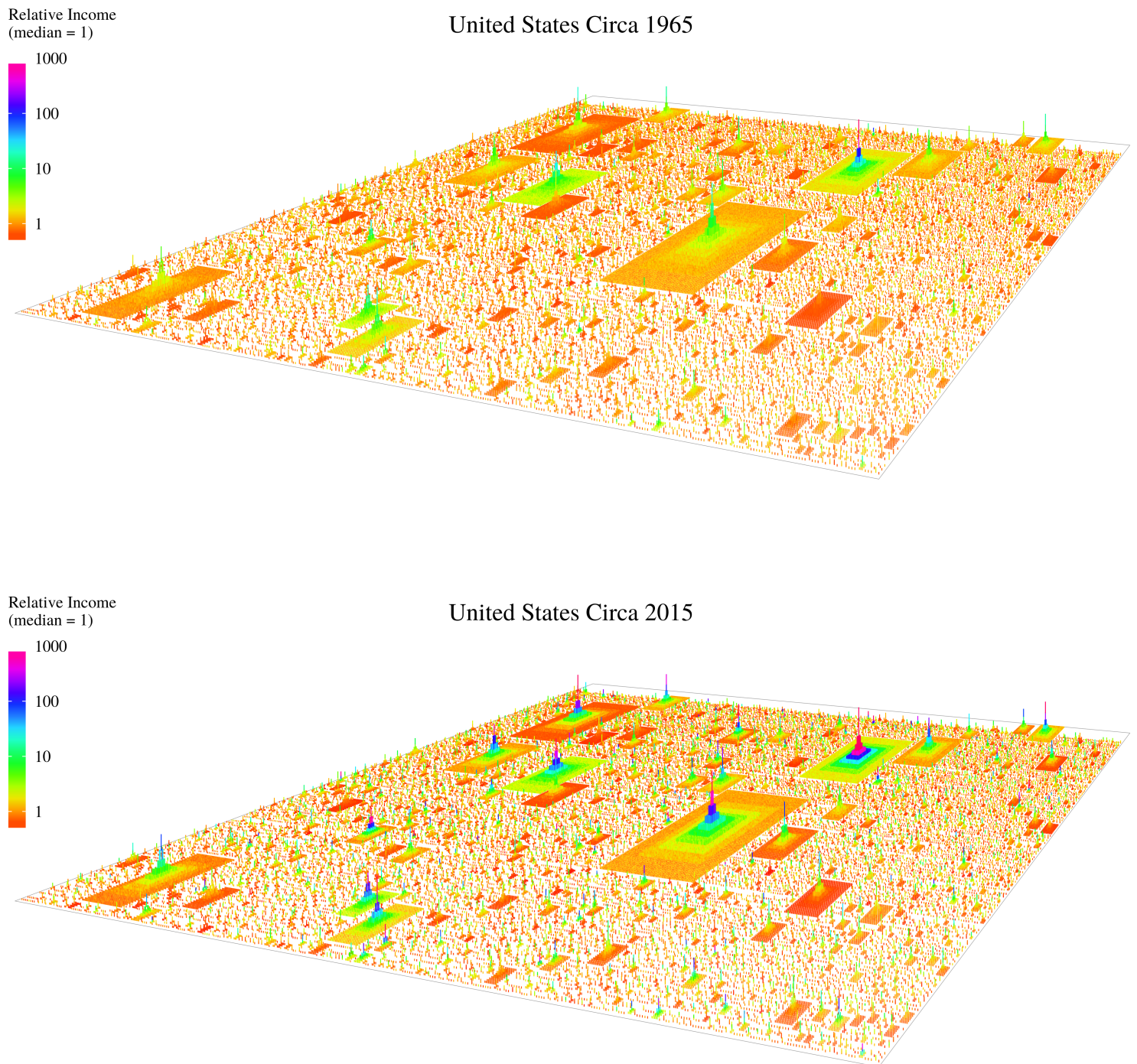


Figure 6: A Visualization of US Hierarchical Income Redistribution

This figure shows the model's representation of hierarchical income redistribution in the United States. The top model represents the US in 1965 while the bottom represents the US in 2015. I create these models by choosing the hierarchical pay-scaling parameter that best matches the US CEO pay ratio, top 1% and dividend share data in the year in question. The difference between the two models is mostly visible at the tops of large firms as an order of magnitude increase in the pay of top-ranked individuals.

bution. Within the body, I find that income dispersion is mostly due to income differences *between firms*. However, in the distribution tail (the top 1%), I find that income dispersion is mostly due to income differences *between hierarchical ranks*.

Song et al. also find a link between stock market returns and individuals' income rank within firms. For top-earning individuals, they find a correlation between the growth of income and S&P 500 returns. More importantly, Song et al. find that this correlation *increases* with income-rank within firms. This is consistent with the results of Mishel and Davis (2014), who note a strong correlation between CEO compensation and stock market returns.

All of this is consistent with the 'capitalist gradient' model. If individuals' capitalist income fraction increases with hierarchical rank, the earnings of top-ranked individuals should be correlated with stock market returns. Conversely, the income of low-ranked individuals should show no correlation with market returns. This raises the possibility of connecting income redistribution to capital accumulation, something that has been theorized by Bichler and Nitzan (2016).

To summarize, the hierarchy model suggests that US firms have systematically redistributed income to the tops of their hierarchies (since 1980). Because of the non-linear properties of hierarchy, this process has been most acute in very large firms. Of course, the model's success does not imply that hierarchical redistribution is all that has occurred. Nor does it imply that the model is correct. To better test the hierarchical redistribution hypothesis, more data is needed.

7 Conclusions

This paper has tested a 'hierarchical redistribution' hypothesis. The idea is that recent changes in US top income share are due to a hierarchical redistribution of income within firms. A model implementing this hypothesis reproduces the intercorrelated trends between the top 1% income share, the CEO pay ratio, the dividend share of national income, and the power-law exponent of the income distribution tail. This result supports the 'hierarchal redistribution' hypothesis.

If this finding is correct, it begs a question. *Why* has there been a hierarchical redistribution of income over the last three decades? A neoclassical interpretation would be that top-ranking individuals have increased their productivity (Gabaix and Landier, 2008). But I find this explanation implausible. A more reasonable explanation, offered by Piketty, is that top managers "have the power to set their own remuneration, in some cases without limit" (2014, p. 24). If so, it would seem that top managers have used this power to reward themselves

lavishly. But how have they gained this power? And how does this power relate to the wider political economy of the United States? These are questions worth asking. There are likely no simple answers.

Appendix

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

<https://osf.io/mkjtz/>

The supplementary materials include:

1. Raw source data;
2. R code for all analysis;
3. Hierarchy model code.

Notes: Appendix A documents the sources and methods used for US distributional data. The remainder of the Appendix discusses the technical details of the hierarchy model. Appendix B outlines the case study data on which the model is based. Appendix C outlines the US firm level data used to restrict model parameters. Appendix D outlines the model's equations, while Appendix E outlines the parameter restriction procedure.

A Sources and Methods

CEO pay ratio. US CEO pay ratio data comes from Mishel and Schieder (2016). This ratio is calculated using CEO income in the 350 largest US firms (ranked by sales), compared to the average income of workers in the firm's respective industry.

Power-law exponents. Power-law exponents for the top 1% of US incomes are estimated on binned data using the method outlined by Virkar and Clauzet (2014). I use income threshold data from the World Wealth and Income Database (see Table 1).

Top 1% income share. US top 1% income share data sources are shown in Table 2.

Dividend share of national income. The dividend share of national income is calculated using data from the Bureau of Economic Analysis, Table 1.12. (National Income by Type of Income).

Table 1: US Top 1% Power-Law Exponent Data Sources

Series	Description	Source
tfainc992j	Pre-tax factor income equal-split adults Threshold Adults constant 2015 local currency	Alvaredo et al. (2016)
tfiinc992j	Fiscal income equal-split adults Threshold Adults constant 2015 local currency	Alvaredo et al. (2016)
tfiinc992t	Fiscal income tax unit Threshold Adults constant 2015 local currency	Alvaredo et al. (2016)
tptinc992j	Pre-tax national income equal-split adults Threshold Adults constant 2015 local currency	Alvaredo et al. (2016)

Table 2: US Top 1% Income Share Sources

Series	Description	Source
sfainc992j	Pre-tax factor income equal-split adults Share Adults share of total (ratio)	Alvaredo et al. (2016)
sfainc996i	Pre-tax factor income individuals Share 20 to 64 share of total (ratio)	Alvaredo et al. (2016)
sfainc999i	Pre-tax factor income individuals Share All Ages share of total (ratio)	Alvaredo et al. (2016)
sfainc999t	Pre-tax factor income tax unit Share All Ages share of total (ratio)	Alvaredo et al. (2016)
sfiinc992j	Fiscal income equal-split adults Share Adults share of total (ratio)	Alvaredo et al. (2016)
sfiinc992t	Fiscal income tax unit Share Adults share of total (ratio)	Alvaredo et al. (2016)
sfiinc996i	Fiscal income individuals Share 20 to 64 share of total (ratio)	Alvaredo et al. (2016)
sfiinc999i	Fiscal income individuals Share All Ages share of total (ratio)	Alvaredo et al. (2016)
sfiinc999t	Fiscal income tax unit Share All Ages share of total (ratio)	Alvaredo et al. (2016)
sptinc992j	Pre-tax national income equal-split adults Share Adults share of total (ratio)	Alvaredo et al. (2016)
sptinc996i	Pre-tax national income individuals Share 20 to 64 share of total (ratio)	Alvaredo et al. (2016)
sptinc999i	Pre-tax national income individuals Share All Ages share of total (ratio)	Alvaredo et al. (2016)
sptinc999t	Pre-tax national income tax unit Share All Ages share of total (ratio)	Alvaredo et al. (2016)
sfiinc_z_US	World Top Incomes Legacy Series	Alvaredo et al. (2013b)
lakner	Calculated from micro data	Atkinson and Lakner (2017)
piketty_book_no_kgains	Legacy data from Capital in the 21st Century	Piketty (2014)
piketty_book_with_kgains	Legacy data from Capital in the 21st Century	Piketty (2014)

Firm Sales vs. Firm Payroll

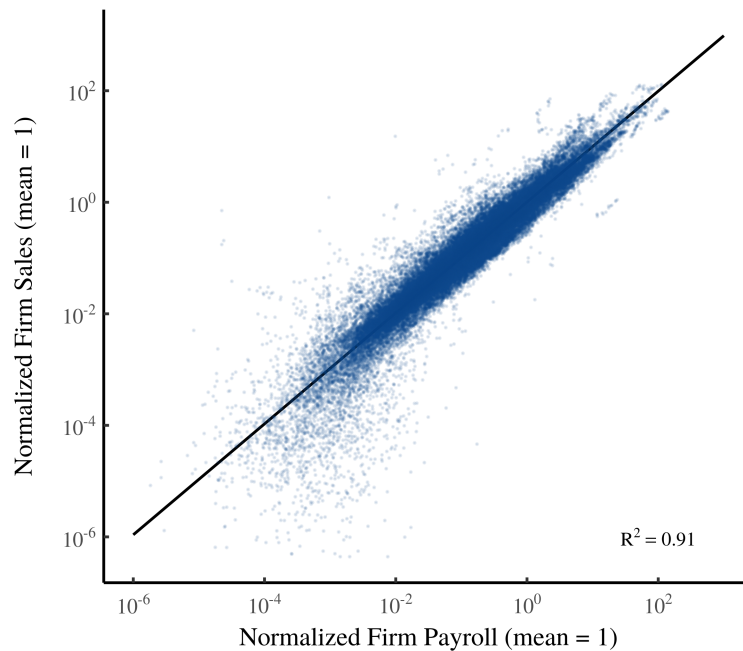


Figure 7: Firm Sales vs. Payroll in the Compustat US Database

This figure plots normalized firm sales against normalized firm payroll for every firm-year observation in the Compustat US database from 1950 to 2015. Each dot is a specific firm in a specific year. To adjust for inflation, I divide sales and payroll by the database averages in the respective year.

I use the hierarchy model to reproduce historical trends in the CEO pay ratio. The empirical data from Mishel and Schieder (2016) uses the CEOs in the top 350 US firms ranked by sales. Since the hierarchy model does not have sales, I calculate the CEO pay ratio by ranking firms by total payroll. Since payroll is highly correlated with firm sales (Fig. 7), the former is a good proxy for the latter.

B Case-Study Firms

In this section I review the case-study evidence that informs the hierarchy model. Table 3 summarizes the source data, while Figure 8 shows the hierarchical employment and pay structure of these firms. The firms remain anonymous, and are named after the authors of the case-study papers. Although the exact shapes vary, all the firms in this sample have a roughly pyramidal employment structure and inverse pyramid pay structure.

Figure 9 dissects these trends to allow further analysis. Figure 9A shows how the span of control (the employment ratio between adjacent ranks) changes as a function of hierarchical level. In these firms, the span of control is not constant, but instead tends to *increase* with hierarchical level. Similarly, Figure 9B shows the ratio of mean pay between adjacent levels. Like the span of control, the pay ratio tends to increase with hierarchical level. Lastly, Figure 9C shows income dispersion within hierarchical ranks of each firm (measured with the Gini index). Note that income dispersion within levels is quite low and there is no evidence of a trend.

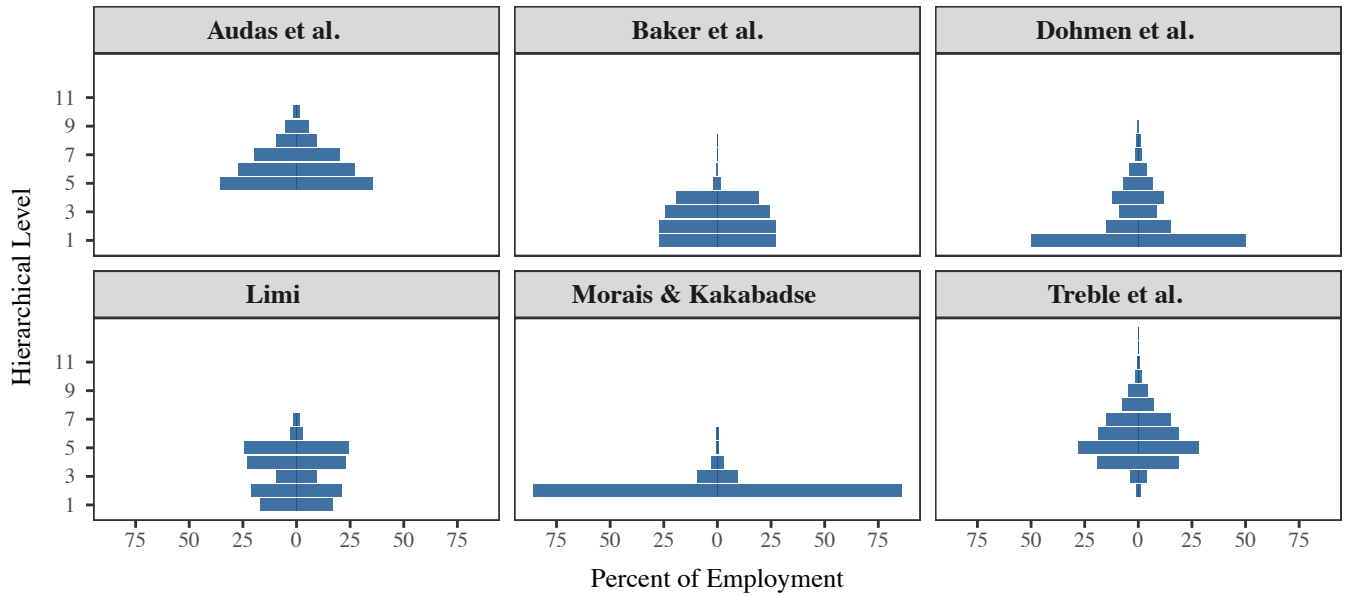
In addition to case-study data of single firms, several studies have reported the aggregate hierarchical structure of a sample of firms (see Table 4 and Figure 11). The data from these firms reveals the same general trends as the case studies. However, the aggregate data is less useful because these studies capture only the top few hierarchical ranks within firms.

The case-study data plays a central role in the hierarchical model developed in this paper. From the case-study evidence, I propose the following ‘stylized’ facts about firm employment and pay structure:

1. The span of control tends to *increase* with hierarchical level.
2. The inter-level pay ratio tends to *increase* with hierarchical level.
3. Intra-level income inequality is approximately *constant* across all hierarchical levels.

The case-study evidence informs the basic structure of the model, and also some of its key parameters. The ‘shape’ of modeled firm hierarchies is determined from the fitted span-of-control trend shown in Figure 9A. Figure 10 shows the idealized employment hierarchy that is implied by case-study data. Error bars indicate uncertainty, calculated using the bootstrap resampling method. Parameters for intra-level income dispersion are determined from the mean of data in Figure 9C. For a detailed discussion of the model algorithm and parameter-fitting procedure, see Sections D and E.

A. Firm Hierarchical Employment Structure



B. Firm Hierarchical Pay Structure

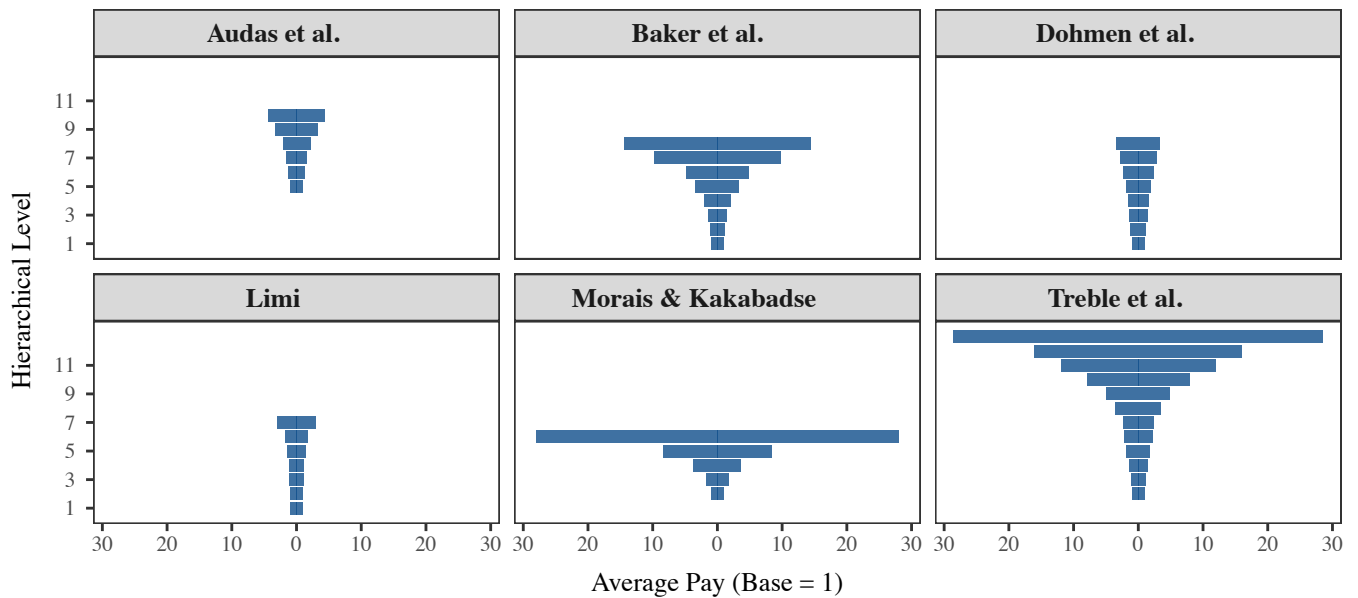


Figure 8: The Hierarchical ‘Shape’ of Six Different Case-Study Firms

This figure shows the hierarchical employment and pay structure of six different case-study firms. Panel A shows the hierarchical structure of employment, while Panel B shows the hierarchical pay structure.

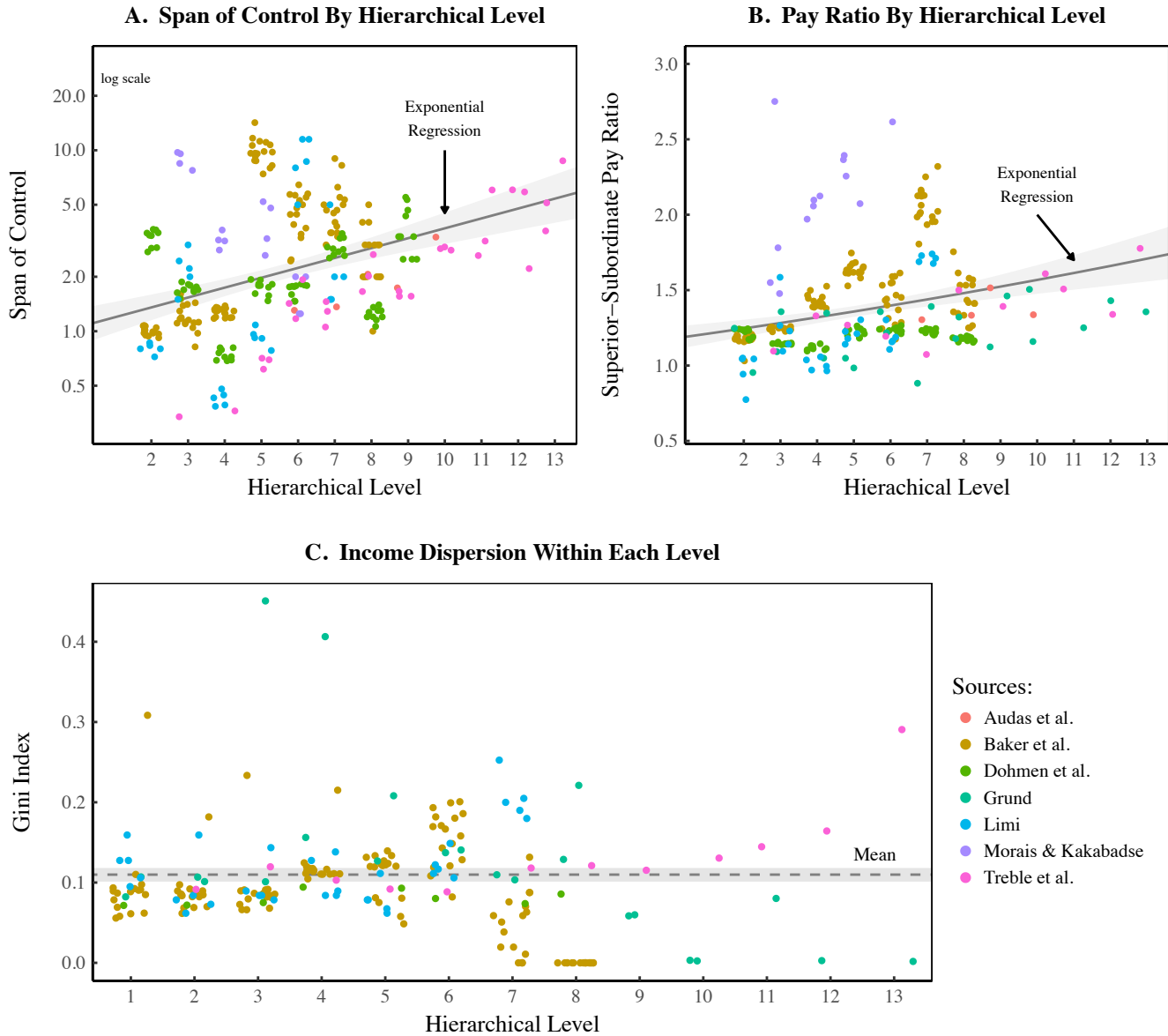


Figure 9: Analyzing the Hierarchical Structure of Case-Study Firms

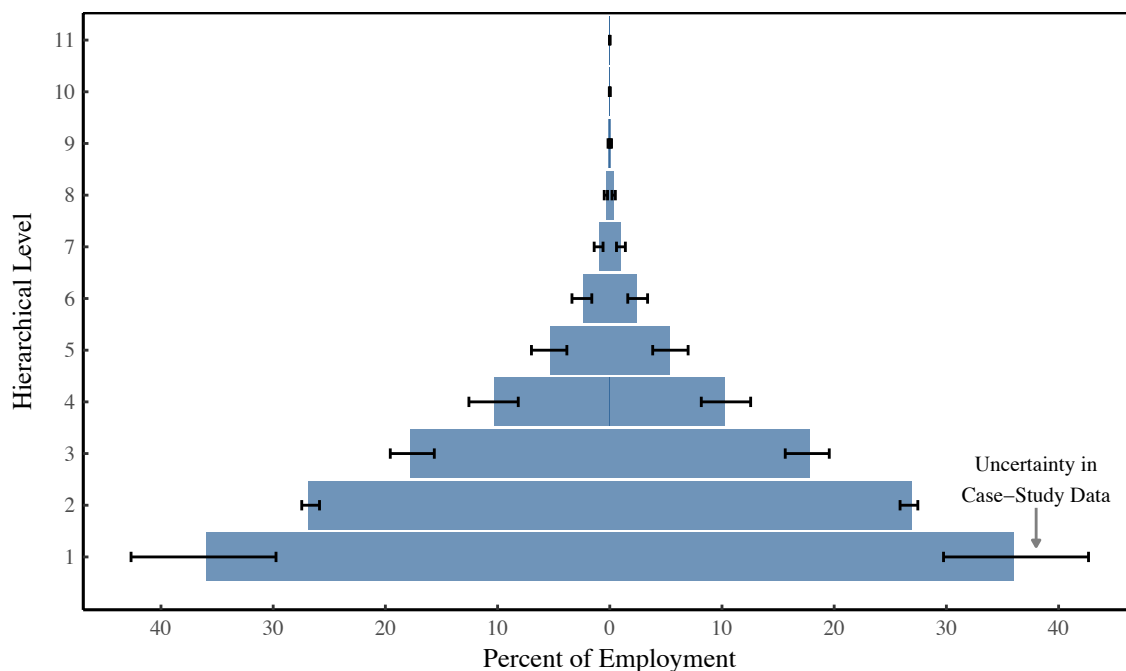
This figure shows data from 7 case-study firms. Panel A shows how the span of control (the subordinate-to-superior employment ratio between adjacent levels) varies with hierarchical level. Note the log scale on the y-axis. Panel B shows how the superior-to-subordinate pay ratio varies with hierarchical level. In Panels A and B, the x-axis corresponds to the *upper* hierarchical level in each corresponding ratio. Panel C shows the Gini index of income inequality within each hierarchical level. Different case-study firms are indicated by color, with names indicating the study author. Note that horizontal ‘jitter’ has been introduced in all three plots in order to better visualize the data (hierarchical level is a discrete variable). The lines in Panels A and B indicate exponential regressions, while the line in Panel C shows the average Gini index. Grey regions correspond to the 95% confidence intervals.

Table 3: Summary of Firm Case Studies

Source	Years	Country	Firm Levels	Span of Control	Level Income	Level Income Dispersion
Audas et al. (2004)	1992	Britain	All	✓	✓	
Baker et al. (1993)	1969-1985	United States	Management	✓	✓	✓
Dohmen et al. (2004)	1987-1996	Netherlands	All	✓	✓	✓
Grund (2005)	1995 & 1998	US and Germany	All		✓	✓
Lima (2000)	1991-1995	Portugal	All	✓	✓	✓
Morais et al. (2014)*	2007-2010	Undisclosed	All	✓	✓	
Treble et al. (2001)	1989-1994	Britain	All	✓	✓	✓

Notes: This table shows metadata for the firm case studies displayed in Fig. 9. The ‘Firm Levels’ column refers to the portion of the firm that is included in the study. ‘Management’ indicates that only management levels were studied.

*For the analysis conducted in this paper I discard (as an outlier) the bottom hierarchical level in Morais and Kakabadse’s data.

**Figure 10: Idealized Firm Employment Hierarchy Implied by Case Studies**

This figure shows the idealized firm hierarchy that is implied by fitting trends to case-study data (Fig. 9A). Error bars show the uncertainty in the hierarchical shape, calculated using a bootstrap resample of case-study data.

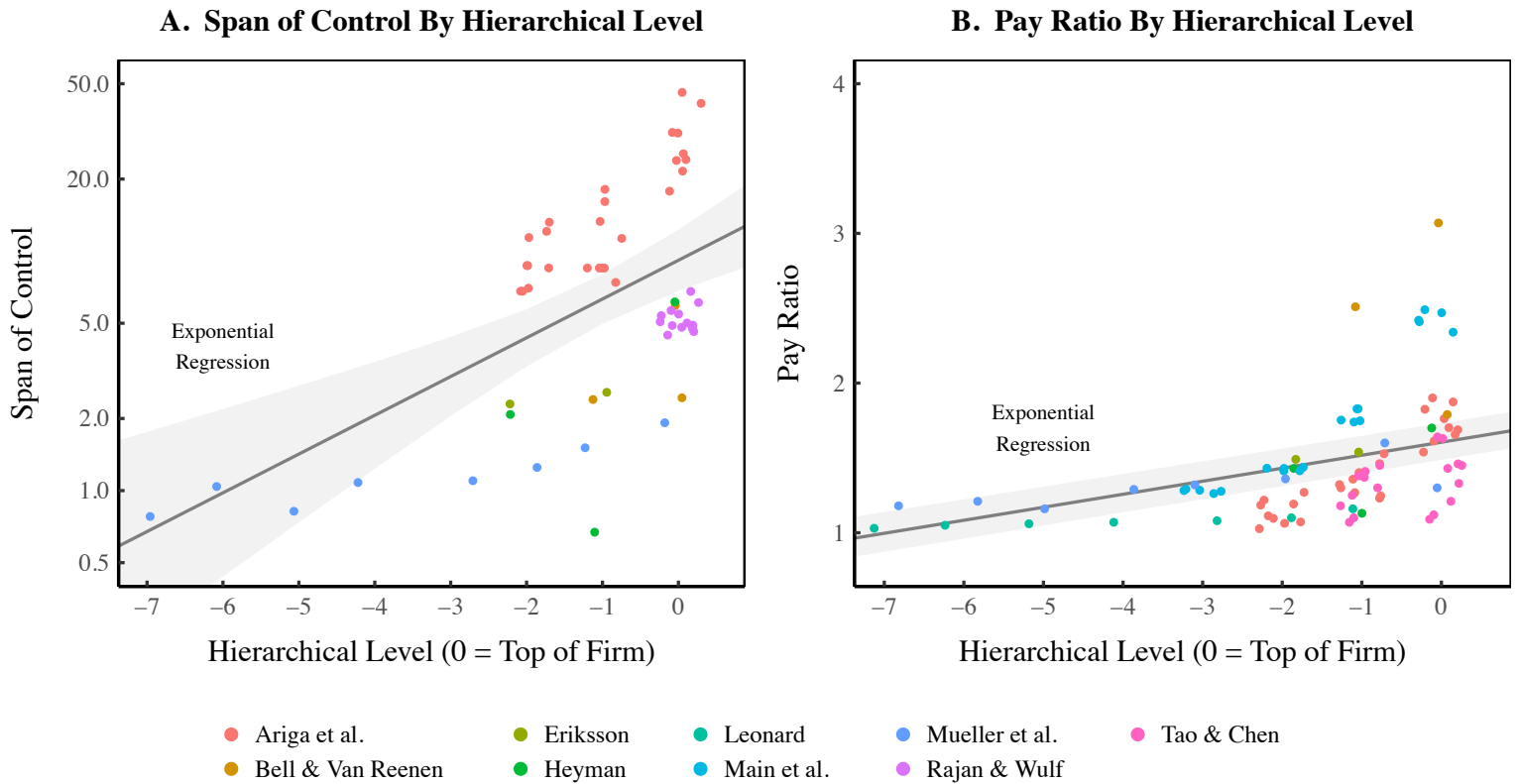


Figure 11: Aggregate Studies of Firm Hierarchical Structure

This figure shows data from 9 different aggregate firm studies. Most of these studies only survey the top several hierarchical levels in each firm. Because of this, I order hierarchical levels from the top down, where the CEO is level 0, the level below is -1, etc. Panel A shows how the span of control (the employment ratio between adjacent levels) relates to hierarchical level. Panel B shows how the pay ratio between adjacent levels varies with hierarchical level. In both plots, horizontal ‘jitter’ has been introduced in order to better visualize the data (hierarchical level is a discrete variable). Grey regions correspond to the 95% confidence interval for regressions.

Table 4: Summary of Firm Aggregate Studies

Source	Years	Number of Firms	Country	Firm Levels	Span of Control	Level Income
Ariga et al. (1992)	1981-1989	unknown	Japan	All	✓	✓
Bell et al. (2012)	2001-2010	552	United Kingdom	Top 3	✓	✓
Eriksson (1999)	1992-1995	210	Denmark	Management	✓	✓
Heyman (2005)	1991,1995	560	Sweden	Management	✓	✓
Leonard (1990)	1981-1985	439	United States	Top 9		✓
Main et al. (1993)	1980–1984	200	United States	Top 4		✓
Mueller et al. (2016)	2004-2013	880	United Kingdom	All	✓	✓
Rajan and Wulf (2006)	1986-1998	261	United States	Top 2	✓	
Tao and Chen (2009)	1986-1998	8101	Taiwan	Top 2		✓

Notes: This table shows metadata for the aggregate studies displayed in Fig. 11. The ‘Firm Levels’ column refers to the portion of the firm that is included in the study. ‘Top 2’, ‘Top 3’, etc. indicates that only the top n levels were included in the study (where the top level is the CEO).

C Compustat Data

This paper makes extensive use of the *Compustat* and *Execucomp* databases. Compustat contains data for most publicly traded US companies, while Execucomp contains data for executive compensation. Three key statistics used throughout this paper are calculated from this data: *firm mean income*, the *CEO-to-average-employee pay ratio*, and the *capitalist income fraction of executives*. I discuss the data and methods used for these calculations in the following sections.

C.1 Firm Mean Income

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

$$\text{Firm Mean Income} = \frac{\text{Total Staff Expenses}}{\text{Total Employment}} \quad (3)$$

C.2 CEO Pay Ratio

Throughout this paper, I use the term ‘CEO’ to refer to the executive at the top of the corporate hierarchy. I identify CEOs using the titles contained in the Execucomp series TITLEANN. Because titles vary greatly by company, identifying the top executive is not always a simple task. While a manual search would be most accurate, this is unrealistic given that the Execucomp database contains over 275 000 entries. Instead, I use the following three-step algorithm to identify the ‘CEO’:

1. Find all executives whose title contains one or more of the words in the ‘CEO Titles’ list (Table 5).
2. Of these executives, take the subset whose title does *not* contain any of the words in the ‘Subordinate Titles’ list (Table 5).
3. If this search returns more than one executive per firm per year, chose the executive with the highest pay.

After identifying the CEO (and matching CEO pay data with firm data contained in the Compustat database), I calculate the CEO pay ratio using the following equation:

$$\text{CEO Pay Ratio} = \frac{\text{CEO Pay}}{\text{Firm Mean Income}} \quad (4)$$

Table 5: 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: This table shows the Execucomp titles used to identify the CEO of each company. CEOs are deemed to be those whose title contains words in the left column, but not those in the right column. Titles such as ‘president-’ and ‘president of’ are included in the subordinate list because they typically refer to a president of a division with the company: i.e. ‘president of western division’ or ‘president-western hemisphere’.

CEO pay ratio and firm mean income data are collectively available for roughly 6000 firm-year observations over the period 1992-2016. I use this data to ‘tune’ my hierarchical model of the firm (see Section E) . Figure 12 shows selected summary statistics of this dataset.

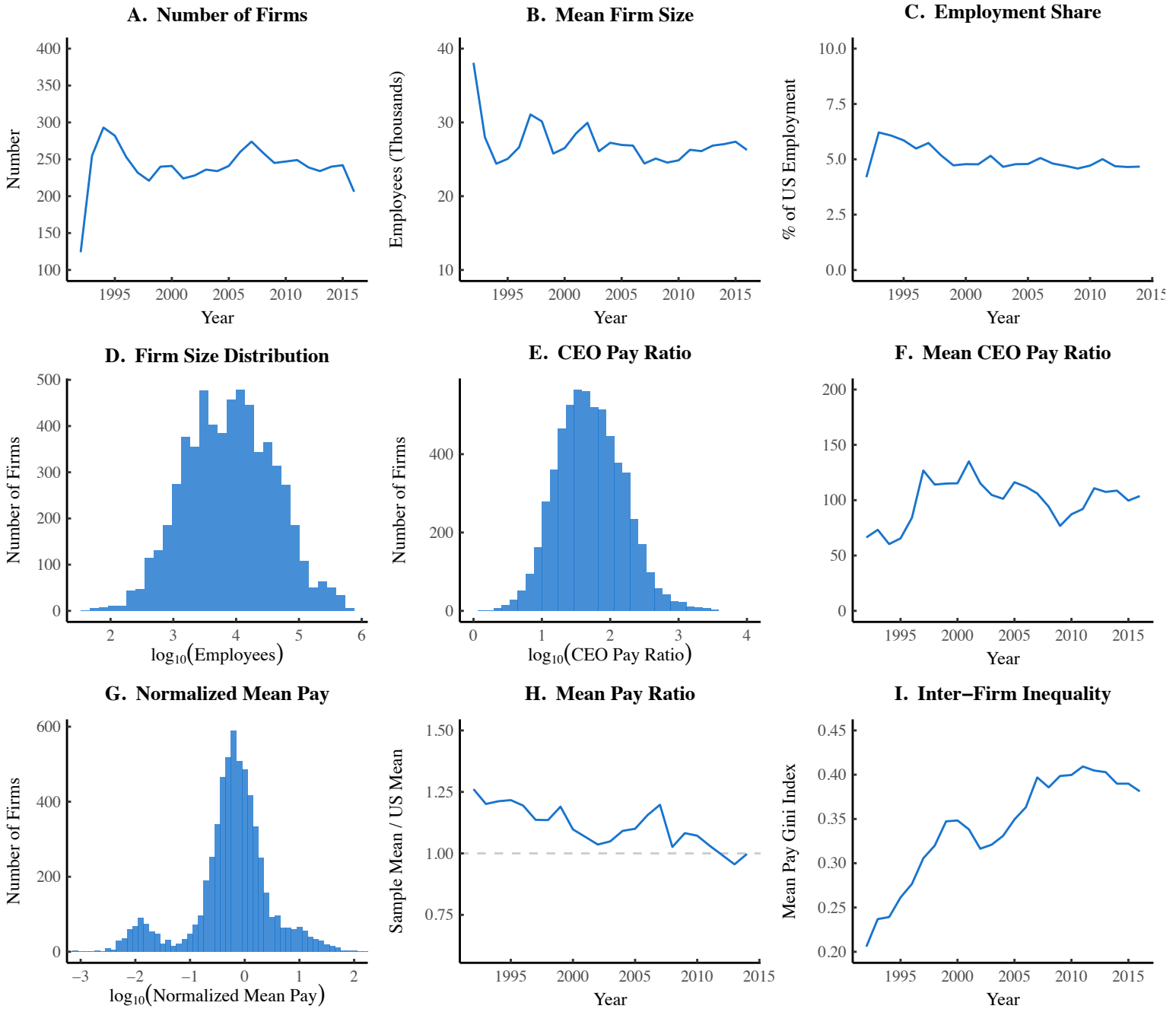


Figure 12: Selected Statistics from the Firm Sample Used for Model Tuning

This figure shows statistics for the Compustat firm sample used to tune my hierarchical model. 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 Compustat 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.

D Hierarchy Model Equations

In this section, I outline the mathematics underlying my hierarchical model of the firm. The model assumptions, outlined below, are based on the stylized facts gleaned from the real-world firm data in section B.

1. Firms are hierarchically structured, with a span of control that increases *exponentially* with hierarchical level.
2. The ratio of mean pay between adjacent hierarchical levels increases *exponentially* with hierarchical level.
3. Intra-hierarchical-level income is lognormally distributed and constant across all levels.

Using these assumptions, I first develop an algorithm that describes the hierarchical employment within a model firm, followed by an algorithm that describes the hierarchical pay structure.

Table 6: Notation

Symbol	Definition
a	span of control parameter 1
b	span of control parameter 2
C	CEO to average employee pay ratio
E	employment
F	cumulative distribution function
G	Gini index of inequality
h	hierarchical level
\bar{I}	average income
μ	lognormal location parameter
n	number of hierarchical levels in a firm
p	pay ratio between adjacent hierarchical levels
r	pay-scaling parameter
s	span of control
σ	lognormal scale parameter
T	total for firm
\downarrow	round down to nearest integer
\prod	product of a sequence of numbers
\sum	sum of a sequence of numbers

D.1 Generating the Employment Hierarchy

To generate the hierarchical structure of a firm, we begin by defining the span of control (s) as the ratio of employment (E) between two consecutive hierarchical levels (h), where $h = 1$ is the *bottom* hierarchical level. It simplifies later calculations if we define the span of control in level 1 as $s = 1$. This leads to the following piecewise function:

$$s_h \equiv \begin{cases} 1 & \text{if } h = 1 \\ \frac{E_{h-1}}{E_h} & \text{if } h \geq 2 \end{cases} \quad (5)$$

Based on our empirical findings in Section B, we assume that the span of control is *not* constant; rather it increases *exponentially* with hierarchical level. I model the span of control as a function of hierarchical level (s_h) with a simple exponential function, where a and b are free parameters:

$$s_h = \begin{cases} 1 & \text{if } h = 1 \\ a \cdot e^{bh} & \text{if } h \geq 2 \end{cases} \quad (6)$$

As one moves up the hierarchy, employment in each consecutive level (E_h) *decreases* by $1/s_h$. This yields Eq. 7, a recursive method for calculating E_h . Since we want employment to be *whole* numbers, we round down to the nearest integer (notated by \downarrow). By repeatedly substituting Eq. 7 into itself, we can obtain a non-recursive formula (Eq. 8). In product notation, Eq. 8 can be written as Eq. 9.

$$E_h = \downarrow \frac{E_{h-1}}{s_h} \quad \text{for } h > 1 \quad (7)$$

$$E_h = \downarrow E_1 \cdot \frac{1}{s_2} \cdot \frac{1}{s_3} \cdot \dots \cdot \frac{1}{s_h} \quad (8)$$

$$E_h = \downarrow E_1 \prod_{i=1}^h \frac{1}{s_i} \quad (9)$$

Total employment in the whole firm (E_T) is the sum of employment in all hierarchical levels. Defining n as the total number of hierarchical levels, we get Eq. 10, which in summation notation, becomes Eq. 11.

$$E_T = E_1 + E_2 + \dots + E_n \quad (10)$$

$$E_T = \sum_{h=1}^n E_h \quad (11)$$

In practice, n is not known beforehand, so we define it using Eq. 9. We progressively increase h until we reach a level of zero employment. The highest level n will be the hierarchical level directly *below* the first hierarchical level with zero employment:

$$n = \{h \mid E_h \geq 1 \text{ and } E_{h+1} = 0\} \quad (12)$$

To summarize, the hierarchical employment structure of our model firm is determined by 3 free parameters: the span of control parameters a and b , and base-level employment E_1 . Code for this hierarchy generation algorithm can be found in the C++ header files `hierarchy.h` and `exponents.h`, located in the Supplementary Material.

D.2 Generating Hierarchical Pay

To model the hierarchical pay structure of a firm, we begin by defining the inter-hierarchical pay-ratio (p_h) as the ratio of mean income (\bar{I}) between adjacent hierarchical levels. Again, it is helpful to use a piecewise function so that we can define a pay-ratio for hierarchical level 1:

$$p_h \equiv \begin{cases} 1 & \text{if } h = 1 \\ \frac{\bar{I}_h}{\bar{I}_{h-1}} & \text{if } h \geq 2 \end{cases} \quad (13)$$

Based on our empirical findings in Section B, we assume that the pay ratio increases *exponentially* with hierarchical level. I model this relation with the following function, where r is a free parameter:

$$p_h = \begin{cases} 1 & \text{if } h = 1 \\ r^h & \text{if } h \geq 2 \end{cases} \quad (14)$$

Using the same logic as with employment (shown above), the mean income I_h in any hierarchical level is defined recursively by Eq. 15 and non-recursively by Eq. 16.

$$\bar{I}_h = \frac{\bar{I}_{h-1}}{p_h} \quad (15)$$

$$\bar{I}_h = \bar{I}_1 \prod_{i=1}^h p_i \quad (16)$$

To summarize, the hierarchical pay structure of our model firm is determined by 2 free parameters: the pay-scaling parameter r , and mean pay in the base level (\bar{I}_1). Code for generating hierarchical pay can be found in the C++ header files `model.h`, located in the Supplementary Material.

D.2.1 Useful Statistics

Two statistics are used repeatedly within the model: mean firm pay, and the CEO-to-average-employee pay ratio.

Mean income for all employees (\bar{I}_T) is equal to the average of hierarchical level mean incomes (\bar{I}_h) weighted by the respective hierarchical level employment (E_h):

$$\bar{I}_T = \sum_{h=1}^n \bar{I}_h \cdot \frac{E_h}{E_T} \quad (17)$$

To calculate the CEO pay ratio, we define the CEO as the person(s) in the top hierarchical level. Therefore, CEO pay is simply \bar{I}_n , average income in the top hierarchical level. The CEO pay ratio (C) is then equal to CEO pay divided by average pay:

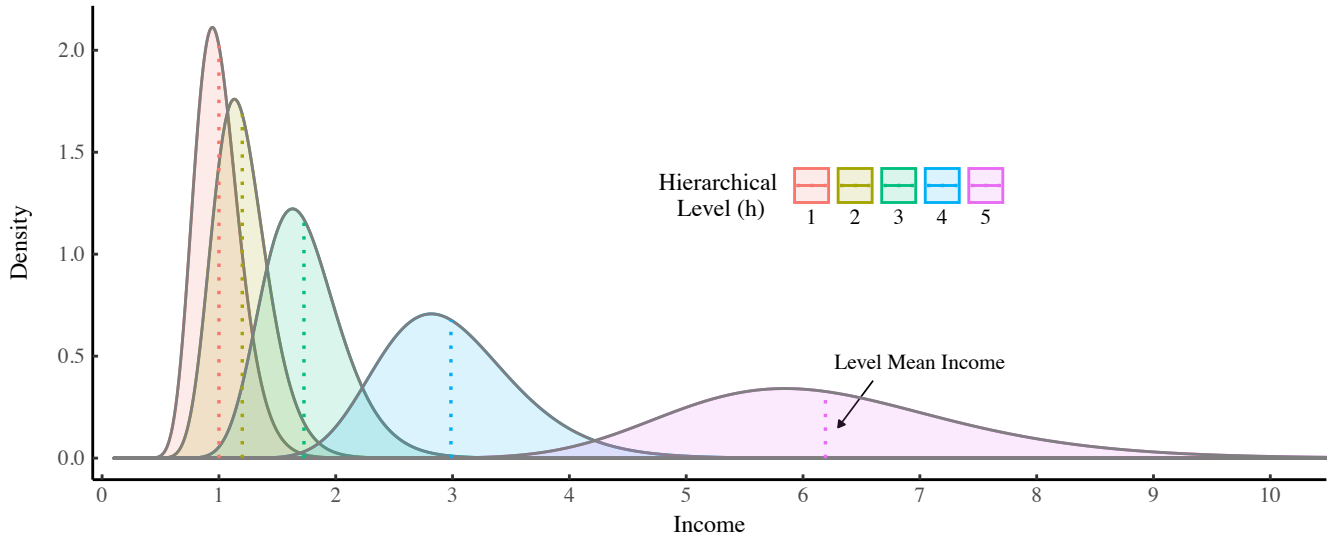
$$C = \frac{\bar{I}_n}{\bar{I}_T} \quad (18)$$

D.3 Adding Intra-Level Pay Dispersion

Up to this point, we have modeled only the *mean* income within each hierarchical level of a firm. The last step in the modeling process is to add pay *dispersion* within each hierarchical level.

I assume that pay dispersion within hierarchical levels is *lognormally* distributed. The lognormal distribution is defined by location parameter μ and scale parameter σ . Our empirical investigation of firm case studies indicated that pay dispersion with hierarchical levels is relatively constant (see Fig. 9C). Given this finding, I assume *identical inequality* within all hierarchical levels. This means that the lognormal scale parameter σ is the same for all hierarchical levels.

A. Adding Pay Dispersion Within Each Hierarchical Level



B. Relative Contribution to Intra-Firm Income Distribution

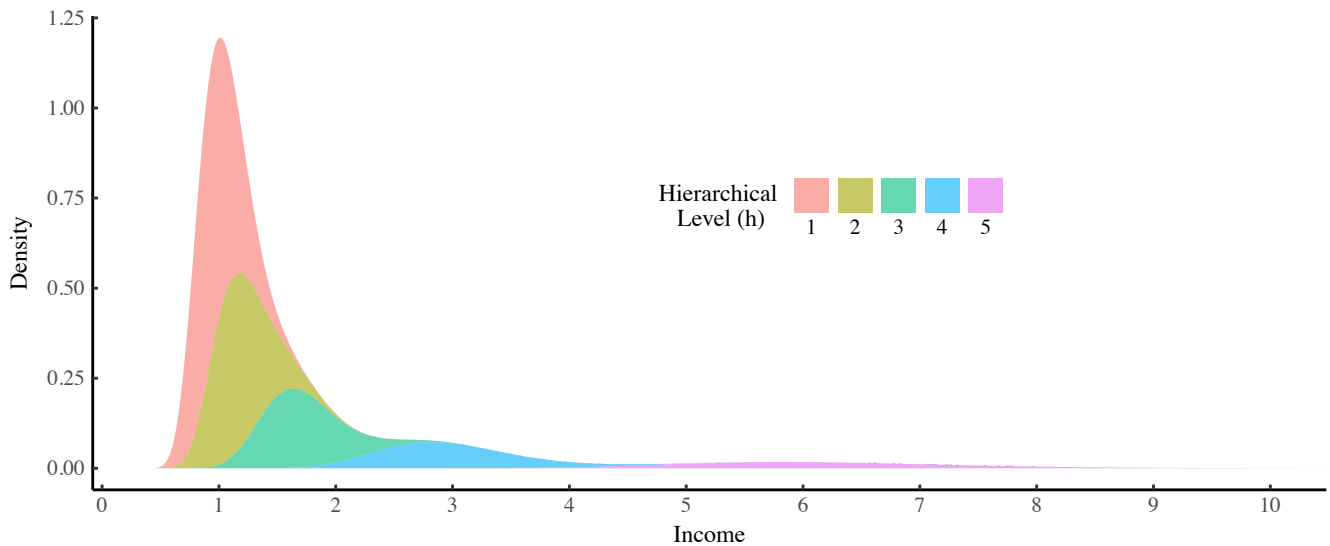


Figure 13: Adding Intra-Level Pay Dispersion to a Model Firm

This illustrates a model firm with lognormal pay dispersion in each hierarchical level. The model firm has a pay-scaling parameter of $r = 1.2$ and an intra-level Gini index of 0.13. Panel A shows the separate distributions for each level, with mean income indicated by a dashed vertical line. Panel B shows contribution of each hierarchical level to the resulting income distribution for the whole firm (income density functions are summed while weighting for their respective employment).

In order to add dispersion within each hierarchical level, I multiply mean pay \bar{I}_h by a lognormal random variate with an expected mean of one. Formally, this is represented by Eq. 19. Since the mean of a lognormal distribution is equal to $e^{\mu + \frac{1}{2}\sigma^2}$, I leave it to the reader to show that a mean of one requires that μ be defined by Eq. 20.

$$I_h = \bar{I}_h \cdot \ln \mathcal{N}(\mu, \sigma) \quad (19)$$

$$\mu = -\frac{1}{2}\sigma^2 \quad (20)$$

Given a value for σ (which is a free parameter), we can define the pay distribution within any hierarchical level of a firm. This process is shown graphically in Figure 13. Figure 13A shows the lognormal income distributions for each hierarchical level of a 5-level firm. Figure 13B shows the size-adjusted contribution of each hierarchical level to the overall intra-firm income distribution. Lower levels have more members, and thus dominate the overall distribution. The code implementing this method can be found in the C++ header file `model.h`, located in the Supplementary Material.

D.4 Calculating Hierarchical Power

I define an individual's hierarchical power as the number of subordinates (S) under their control, plus 1:

$$P = S + 1 \quad (21)$$

Because the hierarchy model simulates only the *aggregate* structure of firms (employment by hierarchical level), hierarchical power is calculated as an *average* per rank. For hierarchical rank h , the average hierarchical power (\bar{P}_h) is defined as the average number of subordinates (\bar{S}_h) plus 1:

$$\bar{P}_h = \bar{S}_h + 1 \quad (22)$$

Each individual with rank h is assigned the average power \bar{P}_h . The average number of subordinates \bar{S}_h is equal to the sum of employment (E) in all subordinate levels, divided by employment in the level in question:

$$\bar{S}_h = \sum_{i=1}^{h-1} \frac{E_i}{E_h} \quad (23)$$

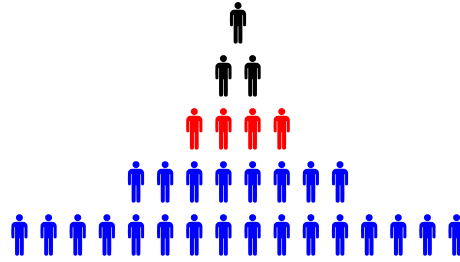


Figure 14: Calculating the Average Number of Subordinates

As an example, consider the hierarchy in Figure 14. The average number of subordinates below each individual in hierarchal level 3 (red) would be:

$$\bar{S}_3 = \frac{E_1 + E_2}{E_3} = \frac{16 + 8}{4} = 6 \quad (24)$$

Therefore, these individuals would all be assigned a hierarchical power of 7.

Table 7: Model Parameters

Parameter	Definition	Action	Scope
α	Firm-size distribution exponent	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.
r	Pay-scaling parameter	Determines the rate at which mean income (within a firm) increases by hierarchical level.	Specific to each firm.
\bar{I}_h	Mean pay in base hierarchical level	Sets the base level income of the firm, which determines firm average pay.	Specific to each firm.
σ	Intra-hierarchical level pay dispersion parameter	Determines the level of inequality within hierarchical levels of a firm.	Identical for all firms.

E Restricting Parameters

As discussed in section D, the hierarchy model has many ‘free’ parameters. Table 7 summarizes all of the parameters used in this model. While free to take on any value, I restrict these parameters exclusively using empirical data. In the following sections, I outline the methods used for this restriction.

E.1 Firm-Size Distribution

Recent studies have found that firm-size distributions in the United States (AxteLL, 2001) and other G7 countries (Gaffeo et al., 2003) can be modeled accurately with a power law. A power law has the simple form shown in Eq. 25, where the probability of observation x is inversely proportional to x raised to some exponent α :

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

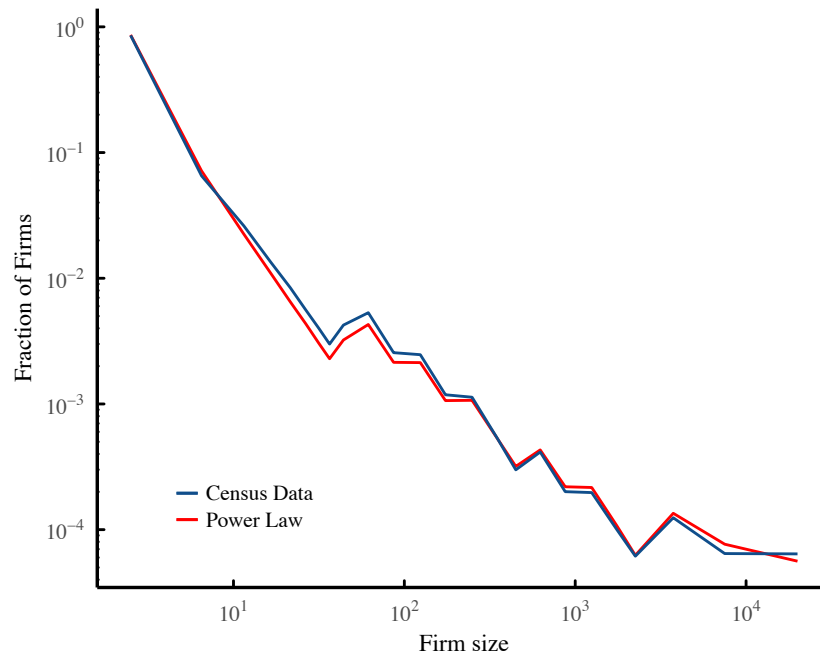


Figure 15: The United States Firm-Size Distribution

This figure shows the US firm-size distribution compared to a power-law distribution with exponent $\alpha = 2.01$ (a simulation with 15 million firms). The US histogram combines data for ‘employer’ firms with data for unincorporated self-employed workers. Data for ‘employer’ firms is from the US Census Bureau, Statistics of U.S. Businesses (using data for 2013). This data is augmented with Bureau of Labor Statistics data for unincorporated self-employed workers (series LNU02032185 and LNU02032192). The histogram preserves Census firm-size bins, with self-employed data added to the first bin. The last point on the histogram consists of all firms with more than 10,000 employees.

Figure 15 compares the US firm-size distribution with a power law of exponent $\alpha = 2.01$. Although not perfect, the fit is good enough for modeling purposes. I assume that the firm sizes can be modeled with a discrete power-law random variate. I model the US firm-size distribution with $\alpha = 2.01$.

A characteristic property of power-law distributions is that as α approaches 2, the mean becomes *undefined*. In the present context, this means that the model can produce firm sizes that are extremely large — far beyond anything that exists in the real world. To deal with this difficulty, I *truncate* the power-law distribution at a maximum firm size of 2.3 million. This happens to be the present size of Walmart, the largest US firm in existence.

Code for the discrete power-law random number generator can be found in

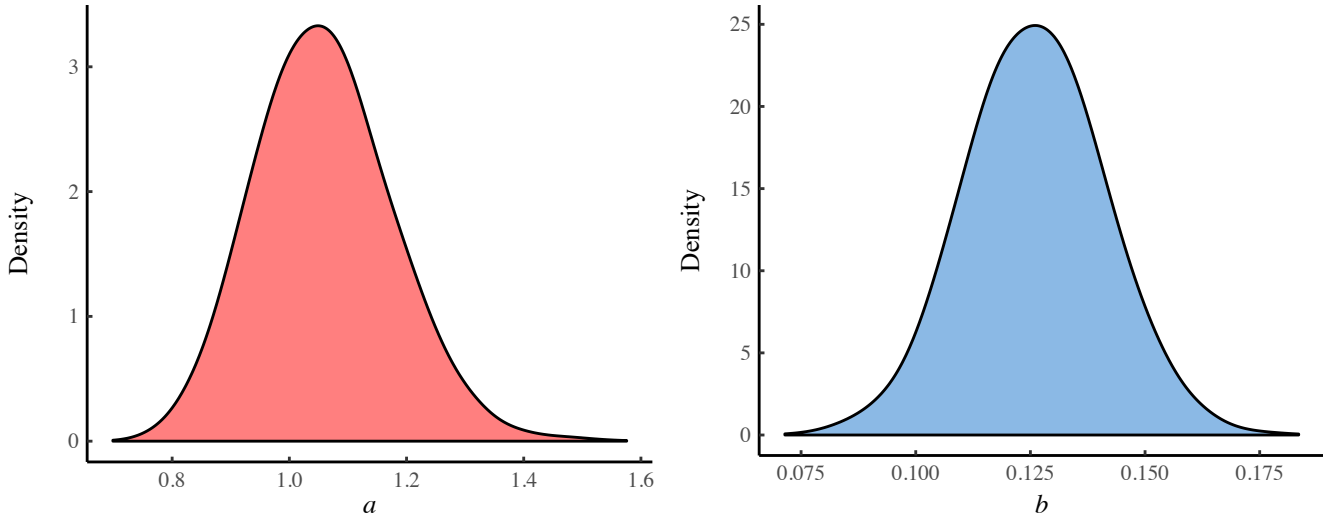


Figure 16: Density Estimates for Span of Control Parameters

This figure shows density estimates for the parameters a and b , which together determine the ‘shape’ of the firm hierarchy. These parameters are determined from regressions on firm case-study data (Fig. 9). The density functions are estimated using a bootstrap analysis, which involves resampling (with replacement) the case study data many times, and calculating the parameters a and b for each resample.

the C++ header file `rp1d.h`, located in the Supplementary Material. This code is an adaption of Collin Gillespie’s (2014) discrete power-law generator found in the R `powerLaw` package (which is, in turn, an adaption of the algorithm outline by Clauset et al. (2009)).

E.2 Span of Control Parameters

The parameters a and b together determine the shape of firm employment hierarchy. These parameters are estimated from an exponential regression on case study data (Fig. 9A). The model proceeds on the assumption that these parameters are *constant* across all firms.

Because the case-study sample size is small, there is considerable uncertainty in these values. I incorporate this uncertainty into the model using the *bootstrap* method (Efron and Tibshirani, 1994), which involves repeatedly resampling the case-study data (with replacement) and then estimating the parameters a and b from this resample. Figure 16 shows the probability density distribution resulting from this bootstrap analysis. I run the model many times, each time with a and b determined by a bootstrap resample of case-study data.

Code implementing this bootstrap can be found in the C++ header file `boot_span.h`.

E.3 Base Level Employment

Given span of control parameters a and b , each firm hierarchy is constructed from the bottom hierarchical level up. Thus, we must know base level employment. In practice, however, we don't know this value — instead we are given *total* employment for a particular firm. While it may be possible to use the equations in section D to define an analytic function relating total employment to base level employment, this is beyond my mathematical abilities.

Instead, I use the model to reverse engineer the problem. I input a range of different base employment values into equations 6, 9, and 11 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 allows us to predict base level E_1 , given total employment E_T . Code implementing this method can be found in the C++ header file `base_fit.h`, located in the Supplementary Material.

E.4 Pay-Scaling Parameter

The pay-scaling ratio r determines the rate at which mean pay increases by hierarchical level. Unlike the span of control parameters, the pay-scaling parameter is allowed to vary between firms. But how should it vary? I restrict the variation of this parameter in a two-step process. I first 'tune' the model to Compustat data. This results in a distribution of pay-scaling parameters specific to Compustat firms. I then fit this data with a parameterized distribution, from which simulation parameters are randomly chosen.

E.4.1 Fitting Compustat Pay-Scaling Parameters

I fit the pay-scaling parameter r to Compustat firms using the CEO-to-average-employee pay ratio (C). The first step of this process is to build the employment hierarchy for each Compustat firm using parameters a , b , and E_1 (the latter is determined from total employment). Given this hierarchical employment structure, the CEO pay ratio in the modeled firm is uniquely determined by the parameter r . Thus, we simply choose r such that the model produces a CEO pay ratio that is equivalent to the empirical ratio.

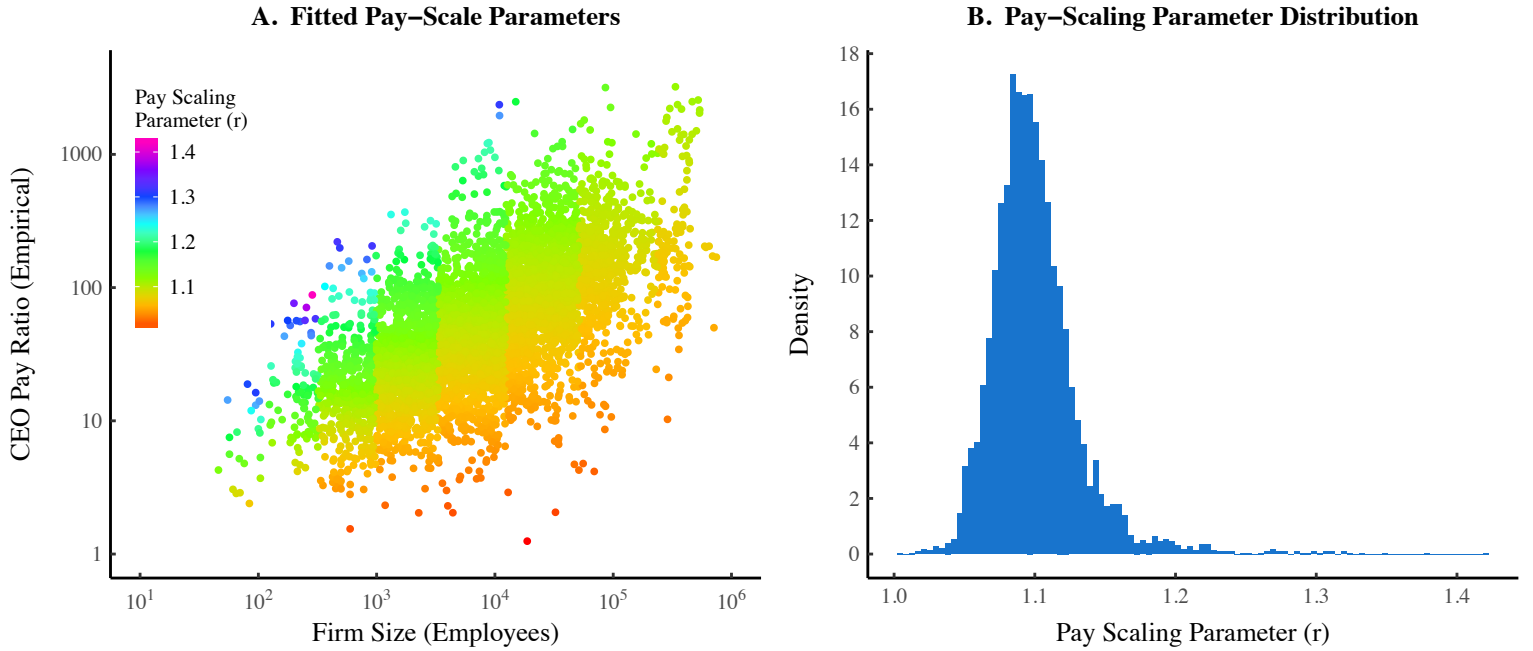


Figure 17: Fitting Compustat Firms with a Pay-Scaling Parameter

This figure shows the fitted pay-scaling parameters (r) for all Compustat firms. Panel A shows the relation between the CEO pay ratio and firm size, with the fitted pay-scaling parameter indicated by color. The discrete changes in color (evident as vertical lines) correspond to changes in the number of hierarchical levels within firms. The pay-scaling parameter distribution for all firms (and years) is shown in panel B.

To solve for this r value, I use numerical optimization (the bisection method) to minimize the error function shown in Eq. 26. Here $C_{\text{Compustat}}$ and C_{model} are Compustat and modeled CEO pay ratios, respectively.

$$\epsilon(r) = |C_{\text{model}} - C_{\text{Compustat}}| \quad (26)$$

For each firm, the fitted value of r minimizes this error function. To ensure that there are no large errors, I discard Compustat firms for which the best-fit r parameter produces an error that is larger than $\epsilon = 0.01$). Fitted results for r are shown in Figure 17. Code implementing this method can be found in the C++ header file `fit_model.h`, located in the Supplementary Material.

E.4.2 Generating a Pay Scaling Distribution

Once we have generated r parameters for every Compustat firm, the next step is to fit a parameterized distribution to this data. For Compustat firms, the dis-

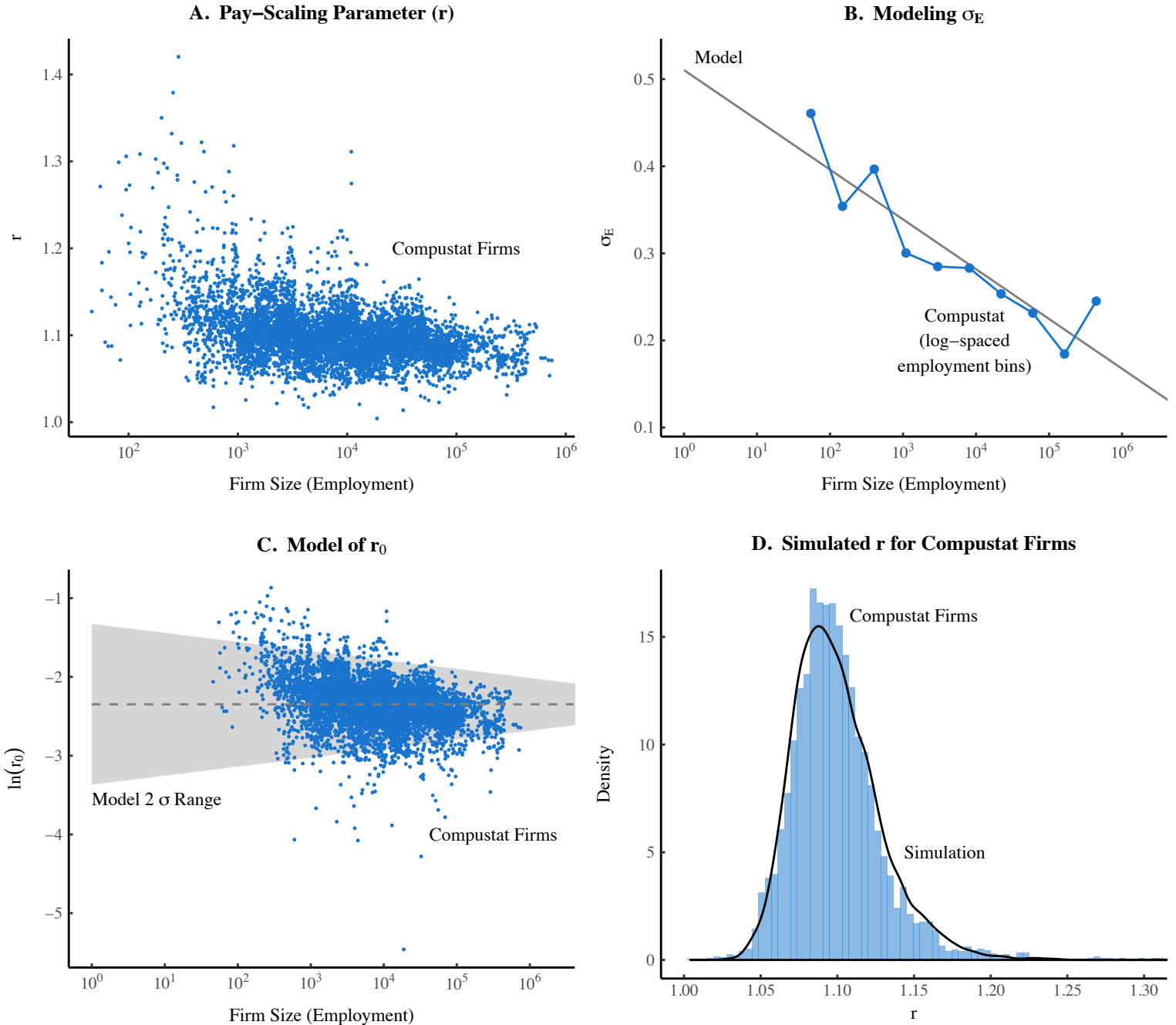


Figure 18: Modeling the Firm Pay-Scaling Distribution

This figure visualizes the model used to simulate firm pay-scaling parameters (r). Panel A shows the relation between r and firm employment for Compustat firms. For the simulation, the distribution of r is modeled with the lognormal variate r_0 . Panel B shows how the lognormal scale parameter σ_E (defined by Eq. 31) changes with firm size. The straight line indicates the modeled relation. Panel C shows how the modeled dispersion of $\ln(r_0)$ declines with firm size, and how this relates to Compustat r_0 data. The 2σ range indicates 2 standard deviations from the mean (on log-transformed data). Panel D shows how the distribution of r for Compustat firms compares to the *simulated* distribution achieved by applying the model to the same Compustat firms.

persion of r is approximately lognormal, and tends to decline with firm size (see Figure 18A). I model r as a shifted function of the lognormal variate r_0 :

$$r = 1 + \ln \mathcal{N}(r_0) \quad (27)$$

The lognormal variate r_0 is defined by location parameter μ and scale parameter σ . While μ is assumed to be constant for all firms, σ is a function of firm size E :

$$r_0(E) = \ln \mathcal{N}(r_0; \mu, \sigma_E) \quad (28)$$

I use the tuned Compustat data to solve for the parameters μ and σ . We first transform Compustat r values using Eq. 29 to get the Compustat distribution of r_0 :

$$r_0 = r - 1 \quad (29)$$

The best-fit value for μ is defined by taking the mean of $\ln(r_0)$:

$$\mu = \overline{\ln(r_0)} \quad (30)$$

Similarly, we can solve for the best-fit value for σ by taking the standard deviation of $\ln(r_0)$. However, unlike μ , the value σ will depend on the size range of firms (E):

$$\sigma_E = SD[\ln(r_0)]_E \quad (31)$$

Figure 18B plots σ_E vs. E for logarithmically spaced size groupings of Compustat firms. I model this relation using a log-linear regression. Figure 18C shows how the modeled dispersion in r_0 varies with firm size, and how this compares to Compustat data.

Once we have fitted the parameters μ and σ to the tuned Compustat data, we can generate r values for simulated firms using equations 27 and 28. Although the model is simple, it produces reasonably accurate results. To test this accuracy, we can apply the model to the same Compustat firms for which it is ‘tuned’. For each Compustat firm, we use the method outlined above to stochastically generate a pay-scaling value r . As Figure 18D shows, the resulting simulated distribution of r fairly accurately reproduces the original data.

When we move from simulating Compustat firms to a real-world distribution of firms, this model involves significant extrapolations for small firms. Why?

The Compustat firm sample has very few observations for firms smaller than 100. And those small firms that are included in the sample are likely *not* representative of the wider population, since they are small *public* firms. In the real world, virtually all small firms are *private*. As with all extrapolations, we simply do the best with the data that is available, while noting that better data might render the extrapolation moot. The code implementing this model can be found in the C++ header file `r_sim.h`, located in the Supplementary Material.

Varying Hierarchical Pay: When attempting to reproduce historical trends in US income inequality, I vary the mean of the pay-scaling distribution by multiplying the fitted lognormal component by a random constant c :

$$r = 1 + c \cdot \ln \mathcal{N}(r_0) \quad (32)$$

E.5 Base-Level Mean Pay

As with the pay-scaling parameter, base level mean pay varies across firms. How should it vary? Again, I restrict the variation of this parameter in a two-step process. I first ‘tune’ the model to Compustat data. This results in 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.

E.5.1 Fitting Compustat Base Level Pay

Having already fitted a hierarchical pay structure to each Compustat firm (in the process of finding r), 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:

$$\frac{\bar{I}_1^{\text{Compustat}}}{\bar{I}_T^{\text{Compustat}}} = \frac{\bar{I}_1^{\text{model}}}{\bar{I}_T^{\text{model}}} \quad (33)$$

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. 16 and 17). If we run the model with $\bar{I}_1^{\text{model}} = 1$, then Eq. 33 reduces to:

$$\frac{\bar{I}_1^{\text{Compustat}}}{\bar{I}_T^{\text{Compustat}}} = \frac{1}{\bar{I}_T^{\text{model}}} \quad (34)$$

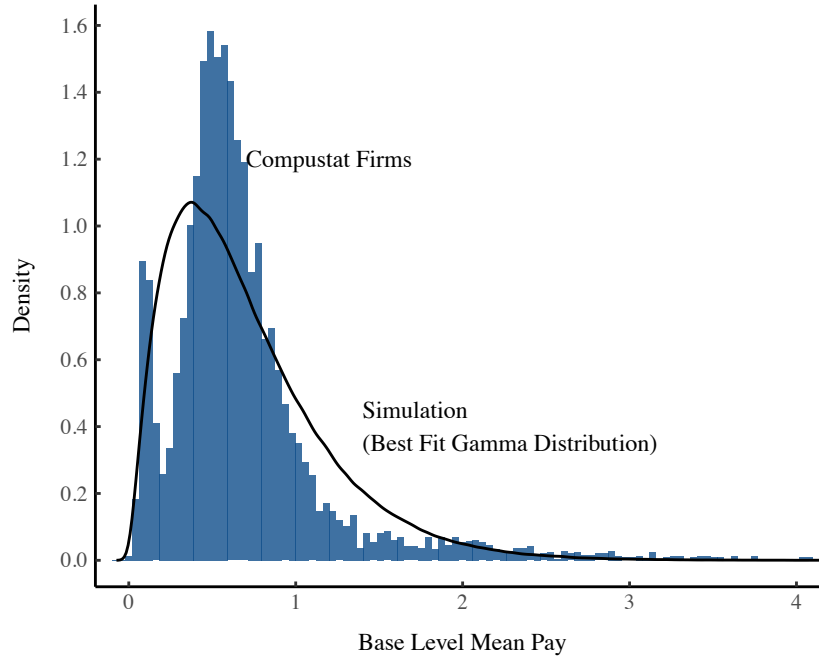


Figure 19: Modeling Firm Base Level Mean Pay

This figure shows the distribution of fitted base-level mean pay for Compustat firms. I model this data with a gamma distribution, from which simulated firm base-level mean pay is randomly drawn. Note that fitting the unimodal gamma distribution to the bimodal Compustat data means that the fit is not great. (The gamma distribution does fit the data better than other skewed distributions such as the Weibull or lognorma). The lower mode in the Compustat data is likely not representative of the general firm population. This lower mode is made up almost entirely of chain restaurants, which seem to be over-represented in this sample.

We can then rearrange Eq. 34 to solve for an estimated base pay for each Compustat firm ($\bar{I}_1^{\text{Compustat}}$):

$$\bar{I}_1^{\text{Compustat}} = \frac{\bar{I}_T^{\text{Compustat}}}{\bar{I}_T^{\text{model}}} \quad (35)$$

Code implementing this method is found in the C++ header file `fit_model.h`, located in the Supplementary Material.

E.5.2 Generating a Base Pay Distribution

Once each Compustat firm has a fitted value for base-level mean pay, we fit this data with a parametric distribution which is then used to stochastically generate

base-level mean pay for the simulation. Since Compustat data is comprised of observations over *multiple* years, in order to aggregate this data into a single distribution, we must account for inflation. Rather than use a price index like the GDP deflator, I divide all firm mean pay data by the average Compustat mean pay in the appropriate year. Since our simulation is concerned only with relative incomes (rather than absolute incomes) no pertinent information is lost in this process.

I model the Compustat firm base pay distribution with a gamma distribution (Fig. 19). Note that because the Compustat data has a bimodal structure (that I do not aim to replicate), the gamma distribution is not a particularly strong fit. Nonetheless the gamma model closely replicates the inequality of firm base pay (which has a Gini index of roughly 0.35). Code implementing this model can be found in the C++ header file `base_pay_sim.h` (in the Supplementary Material).

E.6 Intra-Hierarchical Level Income Dispersion

Intra-hierarchical level income dispersion is modeled with a lognormal distribution, with the amount of inequality determined by the scale parameter σ . I estimate σ from the case-study data shown in Figure 9C. This data uses the Gini index as the metric for dispersion.

To estimate σ , we first calculate the mean Gini index of all data (\bar{G}). We then use Eq. 36 to calculate the value σ , which corresponds to the lognormal scale parameter that would produce a lognormal distribution with an equivalent Gini index. This equation is derived from the definition of the Gini index of a lognormal distribution: $G = \text{erf}(\sigma/2)$.

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

The model proceeds on the assumption that σ is constant for all hierarchical levels within all firms. Because the case-study sample size is small, there is considerable uncertainty in these values. I quantify this uncertainty using the *bootstrap* method Efron and Tibshirani (1994), which involves repeatedly resampling the case-study data (with replacement) and then estimating the parameter σ from this resampled data.

Figure 20 shows the probability density distribution resulting from this bootstrap analysis. In order to incorporate this uncertainty, I run the model many times, with each run using a different bootstrapped value for σ . Code imple-

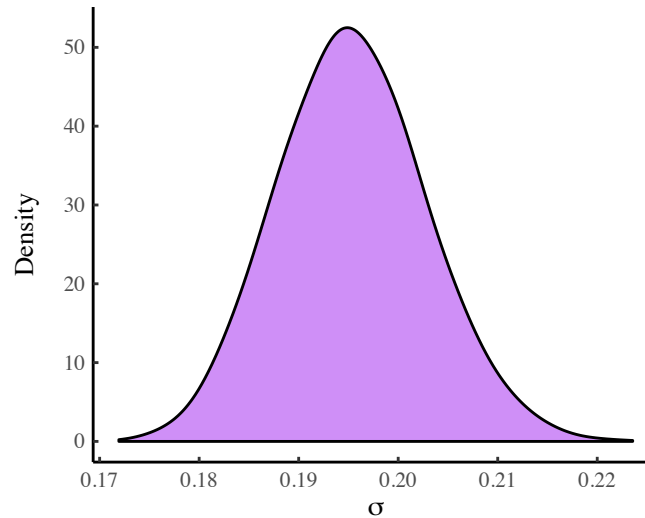


Figure 20: Density estimates for Intra-Hierarchical Level Pay Dispersion Parameter σ

This figure shows the distribution of the lognormal scale parameter σ , which determines pay dispersion within all hierarchical levels of all firms. The distribution is calculated using the bootstrap method.

menting this method can be found in the C++ header file `boot_sigma.h`, located in the Supplementary Material.

E.7 Counterfactual Models

To isolate the distributional effects of hierarchy, I create three counterfactual models, each with only one income-dispersion source. This is achieved as follows:

Inter-firm dispersion only: To create this model, I set the hierarchical pay-scaling parameter (r) to 1 for all firms (removing hierarchical pay-scaling) and set the intra-hierarchical dispersion parameter (σ) to zero (removing dispersion within hierarchical levels).

Inter-hierarchical dispersion only: To create this model, I set base-level pay (\bar{I}_1) in all firms to an identical constant (removing dispersion between firms), and set the intra-hierarchical dispersion parameter (σ) to zero (removing dispersion within hierarchical levels).

Intra-hierarchical dispersion only: To create this model, I set base-level pay (\bar{I}_1) in all firms to an identical constant (removing dispersion between firms), set the hierarchical pay-scaling parameter (r) to 1 for all firms (removing hierarchical pay-scaling).

E.8 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 8 summarizes this structure sequentially in the order that functions are called. In each step, I briefly summarize the action that is performed, giving reference to the section where this action is described in detail.

Table 8: Model High-Level Structure

Step	Action	Reference Section	Parameter(s)	Header File(s)
1	Bootstrap case-study data	E.2, E.6	a, b, σ	boot_span.h boot_sigma.h
2	Get Compustat base-level employment	E.3	E_1	base_fit.h
3	Fit Compustat pay-scaling parameters	E.4.1	r	fit_model.h
4	Get Compustat base-level mean pay	E.5.1	\bar{I}_1	fit_model.h
5	Generate power-law firm-size distribution	E.1	α	rp1d.h
6	Get simulation base-level employment	E.3	E_1	base_fit.h
7	Simulate pay-scaling parameter distribution by fitting Compustat data	E.4.2	r	r_sim.h
8	Simulate base mean pay distribution by fitting Compustat data	E.5.2	\bar{I}_1	base_pay_sim.h
9	Run hierarchy model	D	all	model.h

Notes: Model code makes extensive use of Armadillo, an open-source C++ linear algebra library ([Sanderson and Curtin, 2016](#)).

References

- Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez. 2013a. 'The top 1 percent in international and historical perspective.' *The Journal of Economic Perspectives* 27 (3): 3–20.
- Alvaredo, F., A. B. Atkinson, T. Piketty, E. Saez, and G. Zucman. 2016. 'The world wealth and income database.' Website: <http://www.wid.world>.
- Alvaredo, F., T. Atkinson, T. Piketty, and E. Saez. 2013b. 'The World Top Incomes Database.'
- Ariga, K., G. Brunello, Y. Ohkusa, and Y. Nishiyama. 1992. 'Corporate hierarchy, promotion, and firm growth: Japanese internal labor market in transition.' *Journal of the Japanese and International Economies* 6 (4): 440–471.
- Atkinson, A. and C. Lakner. 2017. 'Capital and labor: the factor income composition of top incomes in the United States, 1962-2006.' *World Bank Policy Research Working Paper* 8268.
- Atkinson, A. B. 2017. 'Pareto and the upper tail of the income distribution in the UK: 1799 to the present.' *Economica* 84 (334): 129–156.
- Atkinson, A. B. and T. Piketty. 2010. *Top incomes: A global perspective*. Oxford University Press.
- Audas, R., T. Barmby, and J. Treble. 2004. 'Luck, effort, and reward in an organizational hierarchy.' *Journal of Labor Economics* 22 (2): 379–395.
- Autor, D. H. 2014. 'Skills, education, and the rise of earnings inequality among the "other 99 percent".' *Science* 344 (6186): 843–851.
- Axtell, R. L. 2001. 'Zipf distribution of US firm sizes.' *Science* 293: 1818–1820.
- Baker, G., M. Gibbs, and B. Holmstrom. 1993. 'Hierarchies and compensation: A case study.' *European Economic Review* 37 (2-3): 366–378.
- Bell, B. and J. Van Reenen. 2012. 'Firm performance and wages: evidence from across the corporate hierarchy.' *CEP Discussion paper* 1088.
- Benhabib, J., A. Bisin, and S. Zhu. 2011. 'The distribution of wealth and fiscal policy in economies with finitely lived agents.' *Econometrica* 79 (1): 123–157.
- Benhabib, J., A. Bisin, and S. Zhu. 2016. 'The distribution of wealth in the Blanchard–Yaari model.' *Macroeconomic Dynamics* 20 (2): 466–481.

- Bichler, S. and J. Nitzan. 2016. 'A CasP model of the stock market.' *Real-World Economics Review* (77): 119–154.
- Champernowne, D. G. 1953. 'A model of income distribution.' *The Economic Journal* 63 (250): 318–351.
- Clauset, A., C. R. Shalizi, and M. E. Newman. 2009. 'Power-law distributions in empirical data.' *SIAM review* 51 (4): 661–703.
- Conyon, M. J. and K. J. Murphy. 2000. 'The prince and the pauper? CEO pay in the United States and United Kingdom.' *The Economic Journal* 110 (467): 640–671.
- Dohmen, T. J., B. Kriechel, and G. A. Pfann. 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.
- Efron, B. and R. J. Tibshirani. 1994. *An introduction to the bootstrap*. London: CRC press.
- Eriksson, T. 1999. 'Executive compensation and tournament theory: Empirical tests on Danish data.' *Journal of Labor Economics* 17 (2): 262–280.
- Feenstra, R. C. and G. H. Hanson. 1996. 'Globalization, outsourcing, and wage inequality.' Technical report, National Bureau of Economic Research.
- Fix, B. 2017. 'Energy and Institution Size.' *PLOS ONE* 12 (2): e0171823.
- Fix, B. 2018a. 'Capitalist Income and Hierarchical Power: A Gradient Hypothesis.' *SocArXiv* .
- Fix, B. 2018b. 'Hierarchy and the Power-Law Income Distribution Tail.' *Journal of Computational Social Science* (<http://dx.doi.org/10.1007/s42001-018-0019-8>).
- Fix, B. 2018c. 'Personal Income and Hierarchical Power.' *Journal of Economic Issues* (in press) (Preprint DOI: 10.31235/osf.io/pb475).
- Gabaix, X. and A. Landier. 2008. 'Why has CEO pay increased so much?' *The Quarterly Journal of Economics* 123 (1): 49–100.
- Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll. 2016. 'The dynamics of inequality.' *Econometrica* 84 (6): 2071–2111.

- Gaffeo, E., M. Gallegati, and A. Palestini. 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.
- Gillespie, C. S. 2014. 'Fitting heavy tailed distributions: the powerLaw package.' *arXiv preprint arXiv:1407.3492* .
- 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.
- Haskel, J., R. Z. Lawrence, E. E. Leamer, and M. J. Slaughter. 2012. 'Globalization and US wages: Modifying classic theory to explain recent facts.' *Journal of Economic Perspectives* 26 (2): 119–40.
- Heyman, F. 2005. 'Pay inequality and firm performance: evidence from matched employer–employee data.' *Applied Economics* 37 (11): 1313–1327.
- Hubmer, J., P. Krusell, and A. A. Smith Jr. 2016. 'The historical evolution of the wealth distribution: A quantitative-theoretic investigation.' Technical report, National Bureau of Economic Research.
- Jones, C. I. and J. Kim. 2018. 'A Schumpeterian model of top income inequality.' *Journal of Political Economy* .
- Krugman, P. 2005. 'For richer.' *Critical social issues in American education: Democracy and meaning in a globalizing world* 3.
- Kumamoto, S.-I. and T. Kamihigashi. 2018. 'Power Laws in Stochastic Processes for Social Phenomena: An Introductory Review.' *Frontiers in Physics* 6: 20.
- Leonard, J. S. 1990. 'Executive pay and firm performance.' *Industrial and Labor Relations Review* 43 (3): 13–28.
- Lima, F. 2000. 'Internal Labor Markets: A Case Study.' *FEUNL Working Paper* 378.
- Lydall, H. F. 1959. 'The distribution of employment incomes.' *Econometrica: Journal of the Econometric Society* 27 (1): 110–115.
- Main, B. G., C. A. O'Reilly III, and J. Wade. 1993. 'Top executive pay: Tournament or teamwork?' *Journal of Labor Economics* 11 (4): 606–628.
- Mishel, L. and A. Davis. 2014. 'CEO pay continues to rise as typical workers are paid less.' *Issue Brief* 380.

- Mishel, L. and N. Sabadish. 2012. 'CEO pay and the top 1%.' *Economic Policy Institute Issue Brief* 331: 1–7.
- Mishel, L. and J. Schieder. 2016. 'Stock market headwinds meant less generous year for some CEOs: CEO pay remains up 46.5% since 2009.' Technical Report 109799, Economic Policy Institute, Washington, D.C.
- Mitzenmacher, M. 2004. 'A brief history of generative models for power law and lognormal distributions.' *Internet mathematics* 1 (2): 226–251.
- Morais, F. and N. K. Kakabadse. 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.
- Mueller, H. M., P. P. Ouimet, and E. Simintzi. 2016. 'Within-Firm Pay Inequality.' *SSRN Working Paper* .
- Newman, M. E. 2005. 'Power laws, Pareto distributions and Zipf's law.' *Contemporary physics* 46 (5): 323–351.
- Nirei, M. 2009. 'Pareto distributions in economic growth models.' *IIR Working Paper WP#09-05* .
- Nirei, M. and S. Aoki. 2016. 'Pareto distribution of income in neoclassical growth models.' *Review of Economic Dynamics* 20: 25–42.
- Nitzan, J. and S. Bichler. 2009. *Capital as Power: A Study of Order and Creorder*. New York: Routledge.
- Pareto, V. 1897. *Cours d'economie politique*, volume 1. Librairie Droz.
- Piketty, T. 2014. *Capital in the Twenty-first Century*. Cambridge: Harvard University Press.
- Piketty, T. and E. Saez. 2001. 'Income Inequality in the United States, 1913-1998 (series updated to 2000 available).' Technical report, National Bureau of Economic Research.
- Piketty, T. and E. Saez. 2006. 'The evolution of top incomes: a historical and international perspective.' Technical report, National Bureau of Economic Research.
- Piketty, T., E. Saez, and S. Stantcheva. 2014. 'Optimal taxation of top labor incomes: A tale of three elasticities.' *American economic journal: economic policy* 6 (1): 230–71.

- Piketty, T. and G. Zucman. 2015. 'Wealth and inheritance in the long run.' In 'Handbook of income distribution,' volume 2, Elsevier, pp. 1303–1368.
- Rajan, R. G. and J. Wulf. 2006. 'The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies.' *The Review of Economics and Statistics* 88 (4): 759–773.
- Sanderson, C. and R. Curtin. 2016. 'Armadillo: a template-based C++ library for linear algebra.' *Journal of Open Source Software* .
- Silva, A. C. and V. M. Yakovenko. 2004. 'Temporal evolution of the "thermal" and "superthermal" income classes in the USA during 1983–2001.' *EPL (Europhysics Letters)* 69 (2): 304.
- Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter. 2016. 'Firming up inequality.' Technical report, National Bureau of Economic Research.
- Tao, H.-L. and I.-T. Chen. 2009. 'The level of technology employed and the internal hierarchical wage structure.' *Applied Economics Letters* 16 (7): 739–744.
- Treble, J., E. Van Gameren, S. Bridges, and T. Barmby. 2001. 'The internal economics of the firm: further evidence from personnel data.' *Labour Economics* 8 (5): 531–552.
- Virkar, Y. and A. Clauset. 2014. 'Power-law distributions in binned empirical data.' *The Annals of Applied Statistics* 8 (1): 89–119.