WORKING PAPERS ON
CAPITAL AS POWER

No. 2016/04

Energy and Institution Size

Blair Fix

June 2016

http://www.capitalaspower.com/?p=2065

Energy and Institution Size

Blair Fix *

June 17, 2016

Abstract

Why do institutions grow? Despite nearly a century of scientific effort, there remains little consensus on this topic. This paper offers a new approach that focuses on energy consumption. A systematic relation exists between institution size and energy consumption per capita: as energy consumption increases, institutions become larger. I hypothesize that this relation results from the interplay between technological complexity and human biological limitations. I also show how a simple stochastic model can be used to link energy consumption with firm dynamics.

1 Introduction

Throughout the last century, there has been a recurrent desire to connect human social evolution to changes in energy consumption [20, 38, 66, 73]. The motivation is simple: the laws of thermodynamics dictate that any system that exists far from equilibrium must be supported by a flow of energy [48]. Since human societies are non-equilibrium systems, it follows that energy flows ought play an important part in social evolution. However, it has proved difficult to move from grand pronouncements based on the laws of thermodynamics to a *quantitative* understanding of the relation between energy use and social evolution [1]. This paper offers a contribution to such a quantitative understanding.

This paper is concerned with one particular aspect of social change: the growth in the size of *institutions*. I present evidence documenting a pervasive link between energy use per capita and institution size: as energy consumption increases, institutions tend to become *larger*. This relation is important for two reasons: (1) it provides strong evidence that social evolution does have an energetic basis; and (2) it represents a *quantitative* regularity that occurs amidst the myriad of complex, qualitative changes associated with social change.

*York University, blairfix@gmail.com

I pursue two avenues for understanding the relation between energy and institution size. The first approach draws on the rich history of stochastic modelling within firm size theory. Stochastic (random) models have been successfully used to link firm *dynamics* to the overall firm size distribution. Yet there is little understanding of what drives variations in firm dynamics. Using data on firm age and firm size to constrain a stochastic model, I demonstrate that firm dynamics are likely related to rates of energy consumption, and I offer a prediction of what this relation should look like.

The second approach aims to offer a more general explanation of why institution growth is related to rates of energy consumption. I argue that increases in energy consumption require an increase in the scale and complexity of energy conversion technology. This, in turn, requires an increase in the scale of social coordination. However, anthropological evidence suggests that humans have a limited ability to maintain social relations [23]. I hypothesize that institutions allow a way around this limitation. Because of their internal hierarchical nature (featuring a nested chain of command), institutions allow the coordination of large groups of people without a corresponding increase in the number of inter-human relations [72]. Thus, I argue that increases in institution size constitute an investment in social hierarchy, an investment that is necessary to mobilize increasing flows of energy.

1.1 Theories of Institutional Size

Theories of institution size can be divided into two classes: those that concern themselves with the *causes* of institutional growth ('why' theories) and those that do not ('how' theories). 'How' theories have met with great empirical success, while 'why' theories have struggled to offer explanations that are testable.

All 'how' theories of institutional size can be traced back to the work of the French economist Robert Gibrat, who discovered that the rate of growth of business firms seemed to be *independent* of their size [33]. While later investigation found this 'law of proportional effect' to be only approximately true — growth rate variance tends to decline with size [40, 45, 65] — it has led to a rich history of stochastic firm growth models [21, 70]. The basic principle is that firm growth is treated probabilistically. Each firm is submitted to a series of random shocks that make it grow (or shrink) over time. When applied to large numbers of firms, the result is a firm size *distribution*. The surprising finding is that these purely random models can very accurately predict the functional form of real-world firm size distributions (see Appendix F).

Despite their success, 'how' theories are not particularly satisfying because they do not explain *why* institutions grow. Unfortunately, theories that *do* attempt to explain the cause of institution growth often rely on unmeasurable variables, and as a result, are untestable.

The theory of the firm has been dominated by Ronald Coase's *transaction cost* approach. According to Coase, "... a firm will tend to expand until the costs of organizing an extra transaction within the firm become equal to the costs of carrying out the same transaction by means of an exchange on the open market or the costs of organizing in another firm" [19]. Unfortunately, transaction costs have been notoriously difficult to define (let alone measure), rendering Coasian theory untestable [31, 56]

Other theories propose that management talent is the driver of firm growth. For instance, Robert Lucas assumes that the firm size distribution results from "allocat(ing) productive factors over managers of different ability so as to maximize output" [51]. Yet Lucas concedes that the causal factor in this model — the talent of managers — is "probably unobservable". Despite this problem, Lucas's theory remains popular [35, 60].

Still other theories propose that firm growth is the result of a resource-driven competitive advantage [8, 59]. Unfortunately, this approach has struggled to stipulate exactly how a particular resource is transformed into a value-creating competitive advantage. Priem and Butler argue that the 'resource-based view' advances a theory of value that is tautological — resources create *value* because they are (among other things) *valuable* [61].

In terms of measurability, theories of government size have faired no better than theories of firm size. One approach is to apply the rational-choice model to the behavior of voters. Government size is treated as a reflection of the preferences of utility maximizing voters [54, 58]. However, without an objective measure of individuals' internal preferences, this theory is untestable.

Another approach is to assume that government bureaucracies (or government as a whole) are self-serving entities that attempt to maximize their budgets, but are restrained by voters and/or an institutional framework such as the constitution [15, 55]. While maximizing behavior is one of the fundamental postulates of neoclassical economics, the hypothesis that humans maximize external pay-offs has been falsified [43].

The lack of measurable variables has consistently plagued 'why' theories of institution size. If a new theory is to be successful, it must demonstrate a connection between institution size and some universally measurable quantity. Energy consumption is just such a quantity.

2 Energy and Institution Size: Empirical Evidence

To study the relation between energy and institution size, I compare variations in energy use per capita to variations in the size of public and private institutions over both space and time. For the private sector, I investigate how changes in the base, tail and mean of the firm size distribution are related to changes in energy use per capita. I use self-employment data to investigate the base of the firm size distribution — relying on the assumption that self-employer firms are very small. To investigate the tail of the firm size distribution, I look at the employment share of the largest firms.

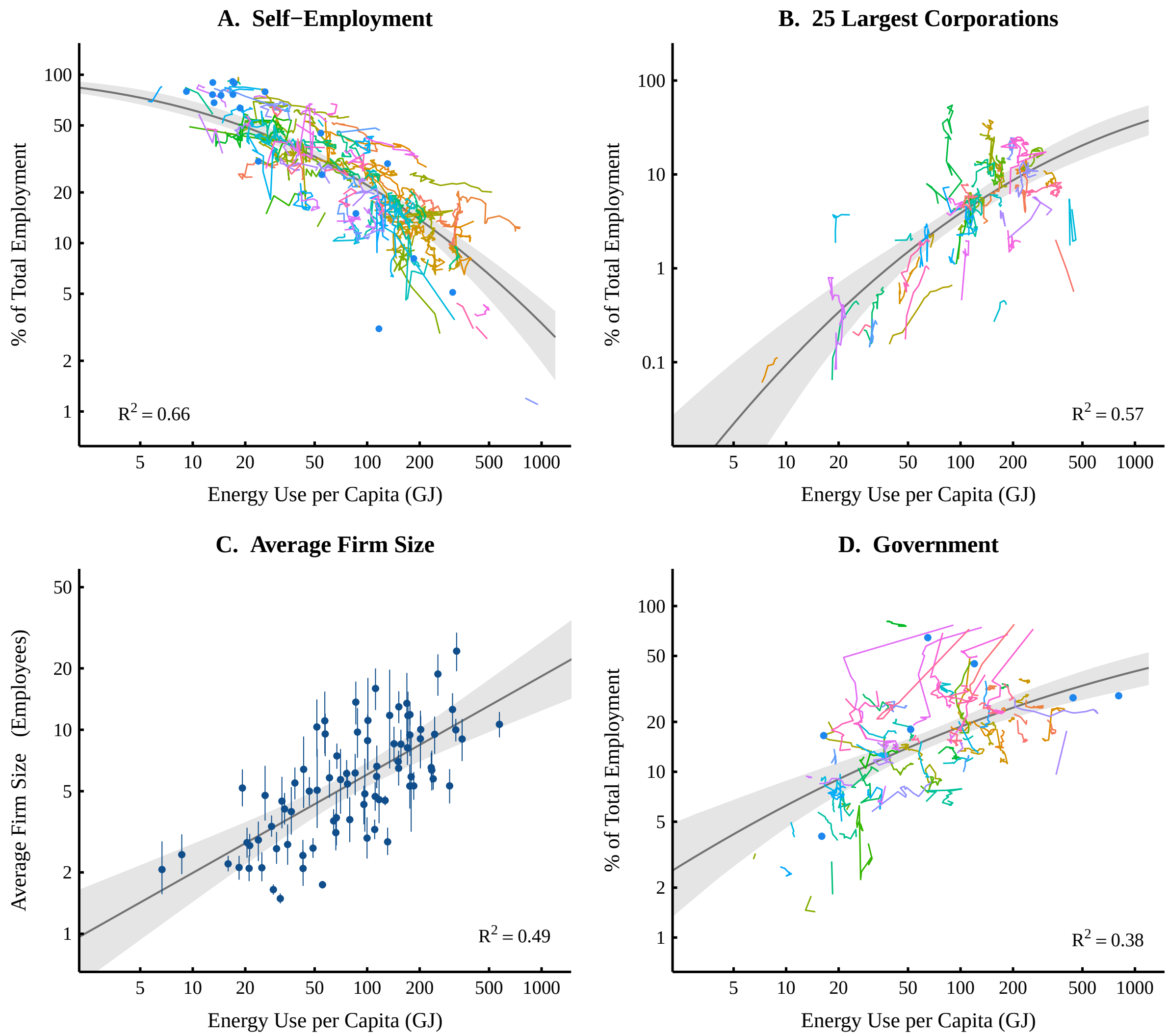

Figure 1: Institution Size vs. Energy Use per Capita at the International Level

This figure shows how different metrics of institution size vary with energy consumption per capita. Panels A-C analyze variations in firm size by looking at the base, tail, and mean of the firm size distribution. Panel C analyzes variations in government size. In order to show as much evidence as possible, panels A, B and D are a mix of time series and scatter plot. Lines represent the path through time of individual countries while points represent a country with a single observation. Error bars in panel C represent the 95% confidence interval of mean firm size. Variations in self-employment, large-firm, and government employment share vs. energy are modelled with log-normal cumulative distribution functions. Mean firm size vs. energy is modelled with a power law. Grey regions indicate the 99% confidence region of each model. For sources and methodology, see Appendix A.

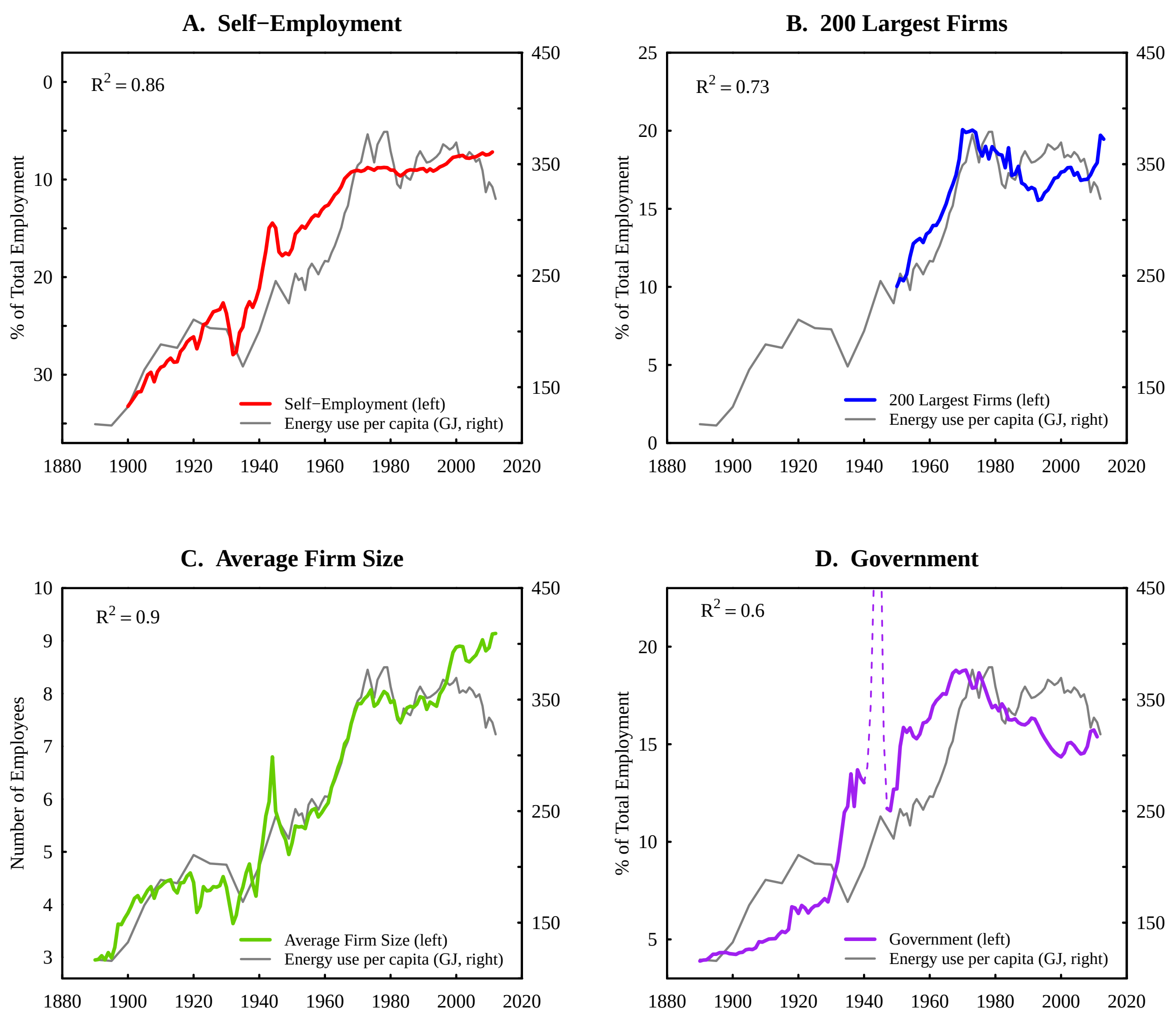

Figure 2: Institution Size vs. Energy Use per Capita in the United States

This figure shows the trends for various measures of institution size in the United States over the last century. Trends mirror those found at the global level. As energy consumption per capita increases, self-employment rates decline (panel A, note reverse scale), the large firm employment share increases (panel B), mean firm size increases (panel C), and the government employment share increases (panel D). Note that government regressions exclude World War II (dotted line). For sources and methodology, see Appendix A.

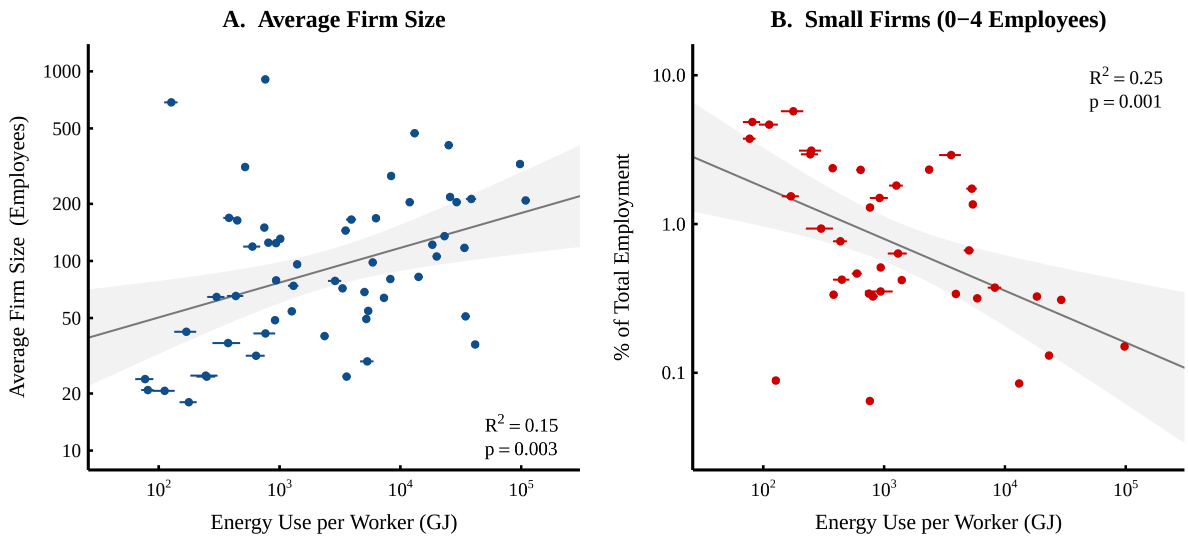

Figure 3: Firm Size vs. Energy Use per Worker in US Manufacturing Sub-Sectors

This figure shows data for US Manufacturing sub-sectors in 2010. As with the global and national level, mean firm size tends to increase with energy use per worker, while the small firm employment share declines. Trends are modelled with a power law. Grey regions indicate the 99% confidence region of each model. While the correlation is smaller than at the national level, it is still significant at the 0.5% level. Data for the large firm employment share (more than 10,000 employees) showed no significant correlation with energy consumption (and is therefore not shown here). However significant 'noise' is added to large firm data by the US census in order to protect firm anonymity. For sources and methodology, see Appendix A.

Comparison of these institution size metrics with energy use per capita are shown in Figures 1-3. Figure 1 shows international trends (each colored line represents the path through time of a specific country). Figure 2 shows time-series data for the entire United States, while 3 looks at sub-sectors of the US manufacturing sector. Together, the evidence in Figures 1-3 suggests the following 'stylized' facts. As energy use per capita increases:

1. The small firm employment share *declines*;
2. The large firm employment share *increases*;
3. The mean firm size *increases*;
4. The government employment share *increases*.

Findings 1-3 suggest that increases in energy consumption are associated with a shift in employment from small to large firms. This indicates that the firm size distribution becomes more *skewed* as energy consumption increases. In Appendix C, I demonstrate that this shift can be accurately modelled in terms of the changing exponent of a power law distribution.

Assuming a correlation between energy use and GDP, then this evidence is consistent with previous research that has focused on the relation between firm size and GDP per capita [9, 10, 34, 51, 60]. However, my focus on energy use is part of a larger effort to study biophysical (rather than monetary) phenomena in economics [6, 18, 27, 37].

Following the long-standing division in institution size theory, I adopt two separate approaches for understanding the relation between institution size and energy consumption. The first approach deals with the 'how' question: *how* exactly do these changes occur. To answer this question, I use a stochastic model to illuminate the relation between energy use and firm dynamics. The second approach deals with the more difficult 'why' question: *why* is institution size related to energy consumption. To answer this question, I investigate the relation between energy, technology, and social coordination.

3 The 'How' Question: Energy and Firm Dynamics

Beginning with the work of Gibrat [33] and later Simon and Bonini [64], stochastic models have been successfully used to explain the functional form of the firm size distribution in terms of firm *dynamics*. The implication of these models is that changes in average firm size occur through changes in firm dynamics. Given the connection between energy consumption and firm size, it follows that firm dynamics ought to vary with changes in energy consumption.

Ideally, we would look at this relation directly by investigating international variations in the firm growth rate distribution and comparing them to variations in energy consumption. Unfortunately, data constraints make such a comparison difficult. Calculating international firm growth rate distributions would require longitudinal data for a large, representative sample of firms in many countries. I am not aware of the existence of any such data at the present time. However, we can use what little data *is* available to make inferences about the relation between energy and firm dynamics.

Firm age data provides an indirect window into firm dynamics. If we assume that new firms start at a small size, then we can infer the historic rate of growth of any firm, given its current age and size (i.e. a new, large firm likely grew rapidly, while an old, small firm likely grew slowly). Figure 4A shows how firm age is related to rates of energy consumption per capita. As energy use per capita increases, the fraction of firms that are under 42 months of age *declines*. This data clearly hints at a systemic relation between energy consumption and firm dynamics. We can use a stochastic model to make specific predictions about the form of this relation.

3.1 A Stochastic Model

The essence of any stochastic firm model is that growth is treated probabilistically. Each firm begins with some arbitrary initial size L_0. After every discrete time interval, the firm is subjected to a series of random 'shocks' (x_i) that perturb it from its initial size. In our model, these shocks are drawn randomly from a Laplace distribution. At any point in time, each firm's size $L(t)$ is equal to the initial size times the product of all shocks (Eq. 1). If the time interval is years, then each shock can be interpreted as the annual growth rate (in fractional form).

$$L(t) = L_0 \cdot x_1 \cdot x_2 \cdot ... \cdot x_t \tag{1}$$

This basic Gibrat model is unstable unless additional stipulations are added (see Appendix E). I add a reflective lower bound that disallows firms from shrinking below the size $L = 1$ (this is sometimes called the Keston process [11, 29, 47]). As long as firm growth rates have a *downward* drift, the model will produce a stable firm size distribution. Using this model requires the following assumptions:

1. The firm size distribution is a power law.
2. Firm growth rates are independent of size.
3. New firms are all born at size $L = 1$.
4. The firm birth rate is equal to the firm death rate.
5. Firm growth rates come from a *Laplace* distribution.
6. The firm size distribution exists in an equilibrium.

Assumption 1 is necessary because the model produces a power law distribution (see Appendix F). Recent studies have found that firm size distributions in the United States [5] and other G7 countries [30] are approximately power laws. Less is known about other countries. In Appendix C, I demonstrate that the international data shown in Figure 1 is largely consistent with variations in a power law distribution, as are variations in the US firm size distribution over the last century.

Assumption 2 is a property of most stochastic firm growth models, and dates back to the work of Gibrat. Many empirical studies have concluded that firm growth rate variance is *not* independent of size; rather, variance tends to decline with firm size [40, 45, 65]). In Appendix D, I confirm this finding using the Compustat database. However, this decline is of importance for only a small minority of firms. I therefore ignore it here.

Assumptions 2 and 3 give meaning to the reflective lower bound. We can interpret this boundary as a firm birth/death zone. Any firm that passes below $L = 1$ is assumed to have 'died'. The reflection then represents the 'birth' of a new firm of size $L = 1$. This interpretation allows firm age to be defined as the period since the last reflection. In the real world, new firms are obviously not all born at size one; however, evidence suggests that they are much smaller than established firms [3, 16].

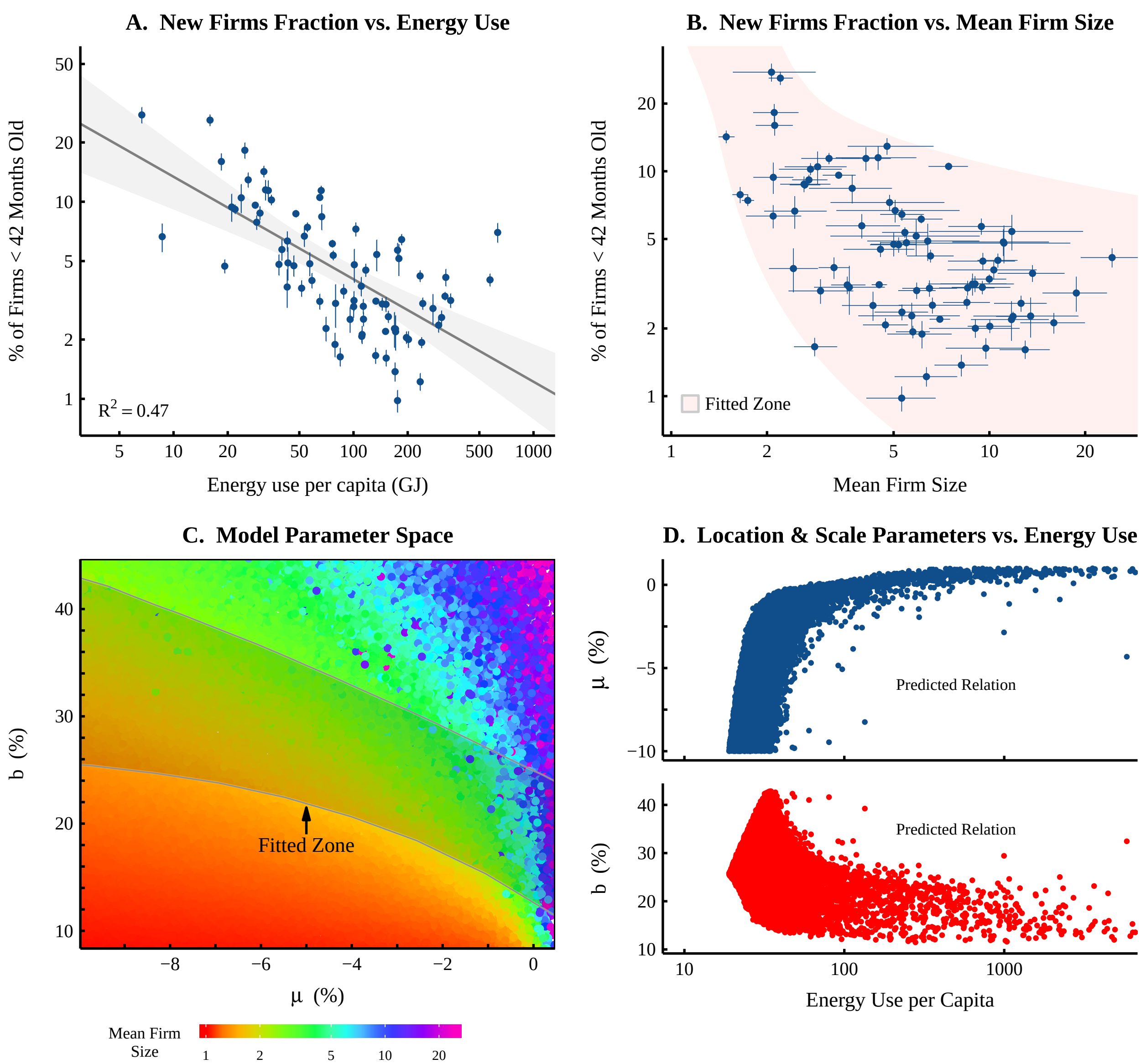

Figure 4: Using Firm Age Data to Estimate International Firm Dynamics

This figure demonstrates how firm age and mean size data can be used to restrict the parameter space of a stochastic model. This allows predictions to be made about the relation between energy use and firm dynamics. Panel A shows the country-level relation between the fraction of firms under 42 months old vs. energy use per capita (the grey region indicates the 99% confidence region of the regression). Panel B shows the country-level relation between the fraction of firms under 42 months old and mean firm size (error bars indicate 95% confidence intervals). The 'Fitted Zone' in Panel B shows the age-size relation produced by a stochastic model with a parameter range specifically chosen to capture the empirical data. Panel C shows the model's parameter space with the resulting mean firm size indicated by color. Using the regressed relation between mean firm size and energy use per capita (Fig. 1C), modelled mean firm size is then transformed into an estimate for energy use per capita. The resulting relation between μ and b vs. energy use per capita (for data in the fitted zone only) is plotted in panel D.

Regarding assumption 4, it is well established that the firm growth distribution has a tent-shape that can be modelled with the Laplace distribution [14, 67]. A Laplace (or double exponential distribution) has a sharper peak and fatter tails than a normal distribution. Various theories have been proposed to explain this phenomenon [13, 28]; however the causes of this growth rate distribution are exogenous to the current model.

Assumption 5 justifies testing the model against empirical data. Given some arbitrary initial conditions, the model will always approach a stable firm size distribution that is a function of only the growth rate distribution (provided that the stability conditions are met). Prior to arriving at equilibrium, there is *no relation* between the growth rate distribution and the firm size distribution (since any initial condition is possible). The equilibrium assumption justifies the link between growth rates and the firm size distribution.

3.2 Estimating Variations in Firm Dynamics

The goal of this analysis is to estimate how firm dynamics (i.e. growth rate distributions) change with levels of energy consumption per capita. This estimation involves three steps. First, we must use appropriate empirical data to restrict the parameter space of the model. Second, we analyze how this parameter space relates to mean firm size. Finally, we extrapolate, from mean firm size, the relation between model parameters and energy use per capita.

Modelled growth rates are determined by the Laplace probability density function below, where μ and b are the location and scale parameters, respectively.

$$p(x) = \frac{1}{2b} e^{-|(x-\mu)/b|} \tag{2}$$

The parameter μ indicates the most probable growth rate, while b corresponds to growth rate volatility (larger b indicates greater volatility). Because μ and b are *free* parameters, we must use appropriate empirical data to restrict their range.

To do this, I use the empirical relation between the proportion of firms under 42 months of age and mean firm size (Fig. 4B). A range of model parameters is chosen so that the resulting stochastic model produces the 'fitted zone' in Figure 4B. This zone covers approximately 90% of empirical data. The corresponding parameter space of the model is shown in Figure 4C, with fitted zone parameters indicated by the shaded region. Equilibrium mean firm size for each μ and b coordinate is indicated by color.

The final step in the analysis is to use the regressed relation between mean firm size and energy use per capita (Fig. 1C) to estimate energy consumption levels from modelled mean firm sizes (for data within the fitted zone only). We can then plot the resulting predicted relation between model parameters and energy use per capita (Fig. 4D).

Our restricted stochastic model predicts the following: (1) μ should *increase* non-linearly with energy consumption; and (2) b should *decrease* non-linearly with energy consumption. In general terms, the model predicts that average firm growth rates should increase with energy consumption, while volatility should decline. This result represents a definitive prediction about how firm dynamics should vary with rates of energy consumption. Future empirical work can determine if this prediction is correct.

4 The 'Why' Question: Energy, Technology and Hierarchy

Any attempt to explain why institutions grow must first settle on the appropriate scale: do we attempt to explain why *individual* institutions grow, or do we concern ourselves only with changes in *average* size? The former is almost certainly a futile task, much like offering a general theory to explain why individual species go extinct. The answer is almost certainly, "It is complex". Species go extinct because of the complicated relation between their physiological characteristics and their environment. Likewise, individual institutions grow/shrink because of the complex relation between their characteristics and their environment (both biophysical and social).

The very success of stochastic firm growth models — in which *randomness* is the explanatory mechanism — suggests that the individual institution is not the appropriate domain for a 'why' explanation. Rather, we should be concerned with *groups* of institutions. This decision effectively bars the traditional toolbox of economic theory, which is to construct models based on simple postulates about the behavior of individual entities (consumers, firms, governments, etc.). Instead, we must rely on qualitative reasoning, tested against quantitative empirical evidence.

My explanation of the relation between institution size and energy consumption is based fundamentally on the human species' *limited* ability to maintain social relations. I argue that institutions are defined by their hierarchical nature, and this hierarchy serves to limit the number of human relations among members, allowing the scale of social coordination to far exceed what would be possible through egalitarian relations. It is this fact, coupled with the increasing scale of social coordination required to build and maintain more complex technology, that I use to explain the relation between institution size and energy consumption per capita. I formalize this causal chain in the following joint hypotheses:

A. Increases in per capita energy consumption are accomplished (in part) by increasing the *complexity* of energy conversion technology.

B. Increases in technological complexity requires increases in *social coordination*.

C. Humans have a *limited* capacity to maintain social relations. Hence, egalitarian social coordination has strict limits.

D. *Social hierarchies* allow the scale of social coordination to grow without a corresponding increase in the number social relations.

E. *Institutions* are dedicated hierarchies.

In the following sections, I review the empirical evidence in support of each of these hypotheses.

4.1 Energy, Technological Scale and Social Coordination

Even the most terse reading of industrial history suggests that increases in energy consumption are associated with immense technological change that tends towards increasing complexity. Cars are almost certainly more complex than horse drawn carts, thermonuclear power plants are more complex than early steam engines ... the list goes on. But it is not clear why this increase in complexity necessitates an increase in institution size.

To *quantitatively* show why this might be the case, we must first define *exactly* what we mean by technological complexity. This necessarily involves moving from the grand scale to a very specific (narrow) metric for technological complexity. I choose to measure technological complexity in terms of the scale (capacity) of electricity generation/conversion technology. This is by no means the only metric for technological complexity; it is however, the *easiest* to measure.

Over the course of the 20th century, the average size of new US power plants increased enormously, and it did so in tight correlation with the increased demand for electricity (Fig. 5A). At least in the electricity generation sector, there is a trend towards technological bigness that mirrors increases in energy consumption. But how does this trend relate to social coordination?

Hypothesis B predicts that increases in power plant size should require an increase in the scale of social coordination. Testing this hypothesis requires a very specific, quantifiable definition of social coordination. I define and measure social coordination in terms of the labor time necessary to construct a power plant. Again this definition is mostly one of convenience: construction labor time is fairly easy to estimate.

I estimate labor construction time using plant *costs* (since data for direct measures of labor time is quite scarce). By the rules of double-entry accounting, all costs eventually become someone's *income*. If we assume that all income accrues to labor (i.e. we neglect capitalist income) we can divide the total cost of a project by average per capita income to give a rough estimate of the total labor time involved:

$$\text{Labor Time} \approx \frac{\text{Total Cost}}{\text{GDP per capita}} \tag{3}$$

Using this methodology, I find a strong correlation between plant capacity and construction labor time (Fig. 5B). This scaling relation is not unique to power plants.

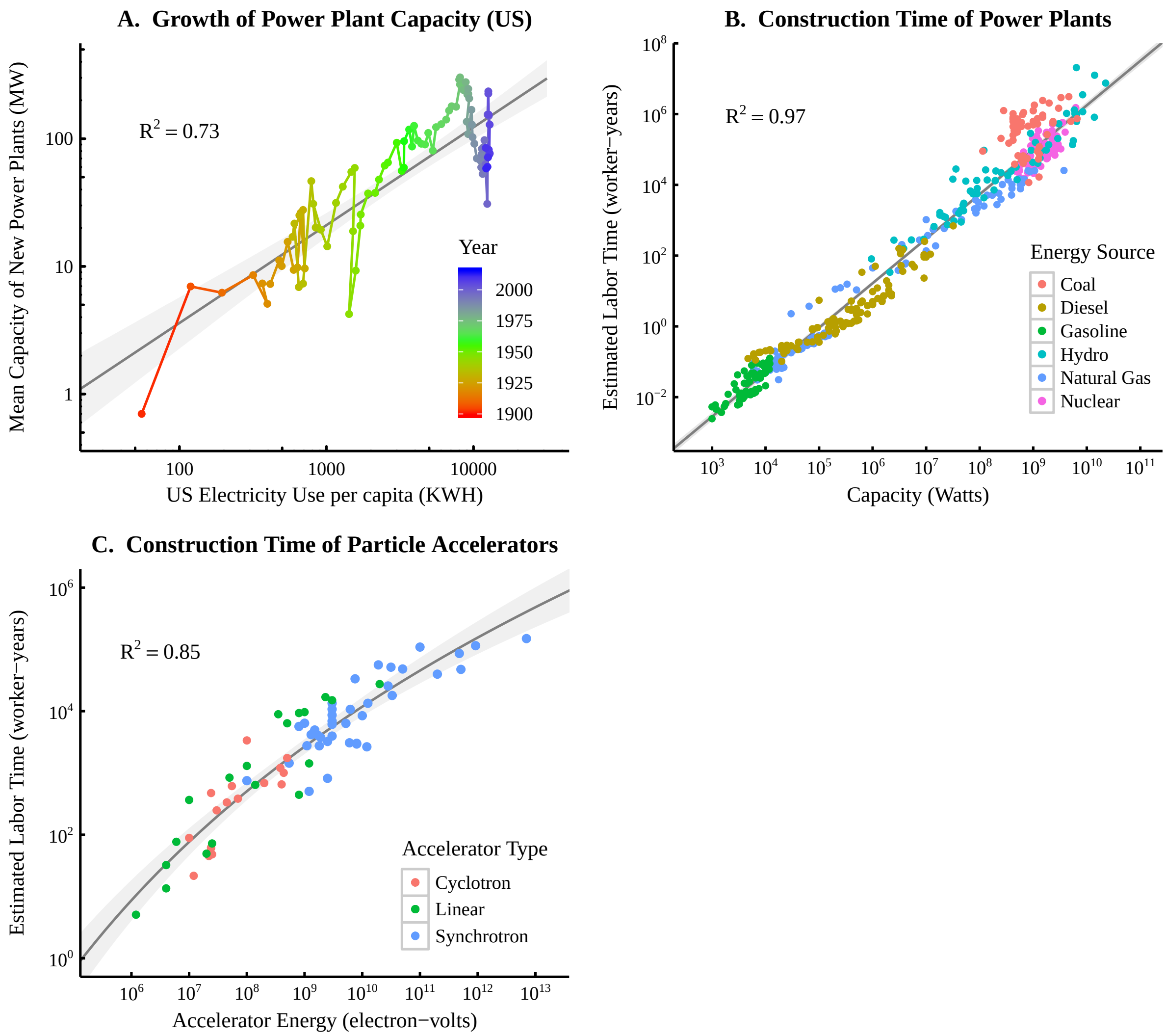

Figure 5: Social Coordination as a Function of Technological Scale

Panel A shows the time-series relation between the mean capacity of new US power plants and US electricity use per capita. Power plants tend to get larger as electricity consumption increases. Panel B shows the relation between power plant capacity and the estimated construction labor time. The entire range of electricity generation technology is included in this plot — from the smallest gasoline generators to the largest hydroelectric power plants. Different primary energy sources are indicated by color. Panel C shows a similar relation between particle accelerator energy and construction labor time. Accelerator energy refers to the maximum kinetic energy achieved by a particle in the accelerator. Different accelerator types are indicated by color. Data in panel A and B are modelled with a power law. Data in panel C is modelled with the function $y = a \log^b x$. Grey regions indicate the 99% confidence region of each model. For sources and methodology, see Appendix A.

It can also be found for particle accelerators — another type of electricity conversion technology (Fig. 5C). Both technologies exhibit scaling behavior that is consistent with hypothesis B: increases in technological scale are associated with an increase in the scale of social coordination.

An obvious critique of this investigation is that power plant (and particle accelerator) construction constitutes a tiny fraction of all economic activity. Increases in power plant scale can hardly explain the increase in institution size over the entire economy. This is certainly true. However, the goal of this investigation is only to test the *principles* advanced in hypothesis A and B (not to account for all institutional change). The empirical evidence reviewed here is consistent with these principles.

4.2 Social Coordination and Human Biology

The increased social coordination associated with increased technological complexity could conceivably be achieved in many different ways (customs, markets, institutions, etc.). Thus, an increase in social coordination does not necessarily imply an increase in institution size. Hypotheses C-E propose a chain of causation explaining why institutions are the most effective way of organizing large groups of people. The key to this reasoning is hypothesis C: humans have a *limited* a ability to maintain social relations.

The evidence for this hypothesis comes from the work of anthropologist Robin Dunbar [23] who demonstrates that primate brain size is highly correlated with mean group size (where brain size is defined as the relative size of the neocortex). Because maintaining social relations is a computationally intensive task, Dunbar hypothesizes that brain size *limits* group size. When Dunbar's evidence is extrapolated to humans, it predicts a mean group size of 150 (sometimes called Dunbar's number). While this number should be considered exploratory, it supports the hypothesis that humans have a limited capacity to maintain social relations.

One way of increasing group size beyond Dunbar's number is to organize groups in a way that *limits* human interaction. Turchin and Gavrilets [72] note that this is a key feature of social hierarchies. Within a hierarchy an individual must maintain social relations only with his direct superior and direct inferiors. Thus, hierarchy allows group size to grow without any corresponding increase in the number of human relations (hypothesis D).

4.3 Hierarchy and Institution Size

Hypothesis E proposes that institutions are simply dedicated hierarchies [62]. This explains why increases in social coordination lead to increases in institution size: as social hierarchies, institutions are the most effective way of organizing large groups of people. This hypothesis implies that an increase in institution size constitutes an investment in social hierarchy.

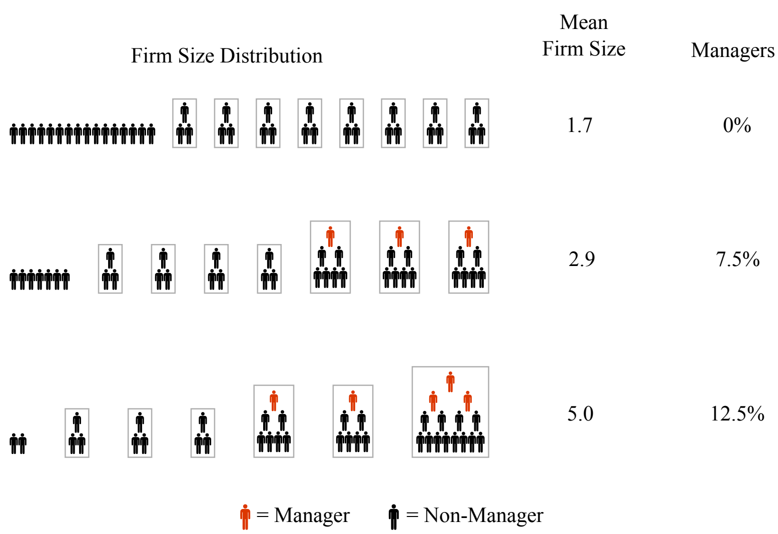

Figure 6: The Growth of Management as a Function of the Firm Size Distribution

This figure graphically demonstrates how the management fraction increases with firm size (assuming firms are 'ideal hierarchies'). Firms are indicated by boxes (with the exception of single-person firms) with a worker's hierarchical position shown vertically. The span of control — defined as the size ratio between adjacent hierarchical levels — is constant for all firms. In this picture, the span of control is 2. Managers (red) are assumed to be all individuals in and above the third hierarchical level. To maintain simplicity, this graphic does not use a power law firm size distribution.

Many studies have confirmed the basic hierarchical structure of individual firms [4, 7, 22, 25, 32, 50, 71]. But is there a way to test hypothesis E on a grander scale? I do so here by noting that hierarchies tend to become more top heavy as they become larger — the fraction of individuals in the upper echelons tends to grow as the size of the hierarchy increases. If we treat the upper echelons across all firms as a *single* cohort of people, we should expect an increase in mean firm size to be reflected in a growth in the relative size of this cohort.

Since the upper echelons of a hierarchy are almost exclusively involved in managing the activities of other people, it seems sensible to use the *management* profession as a metric for the size of this top cohort. Thus, we expect that increases in mean firm size should be associated with an increase in the relative employment share of managers.

To refine this prediction, I develop a hierarchical firm model (Fig. 6) based on the following assumptions:

1. All firms are 'ideal' hierarchies with a single span of control.

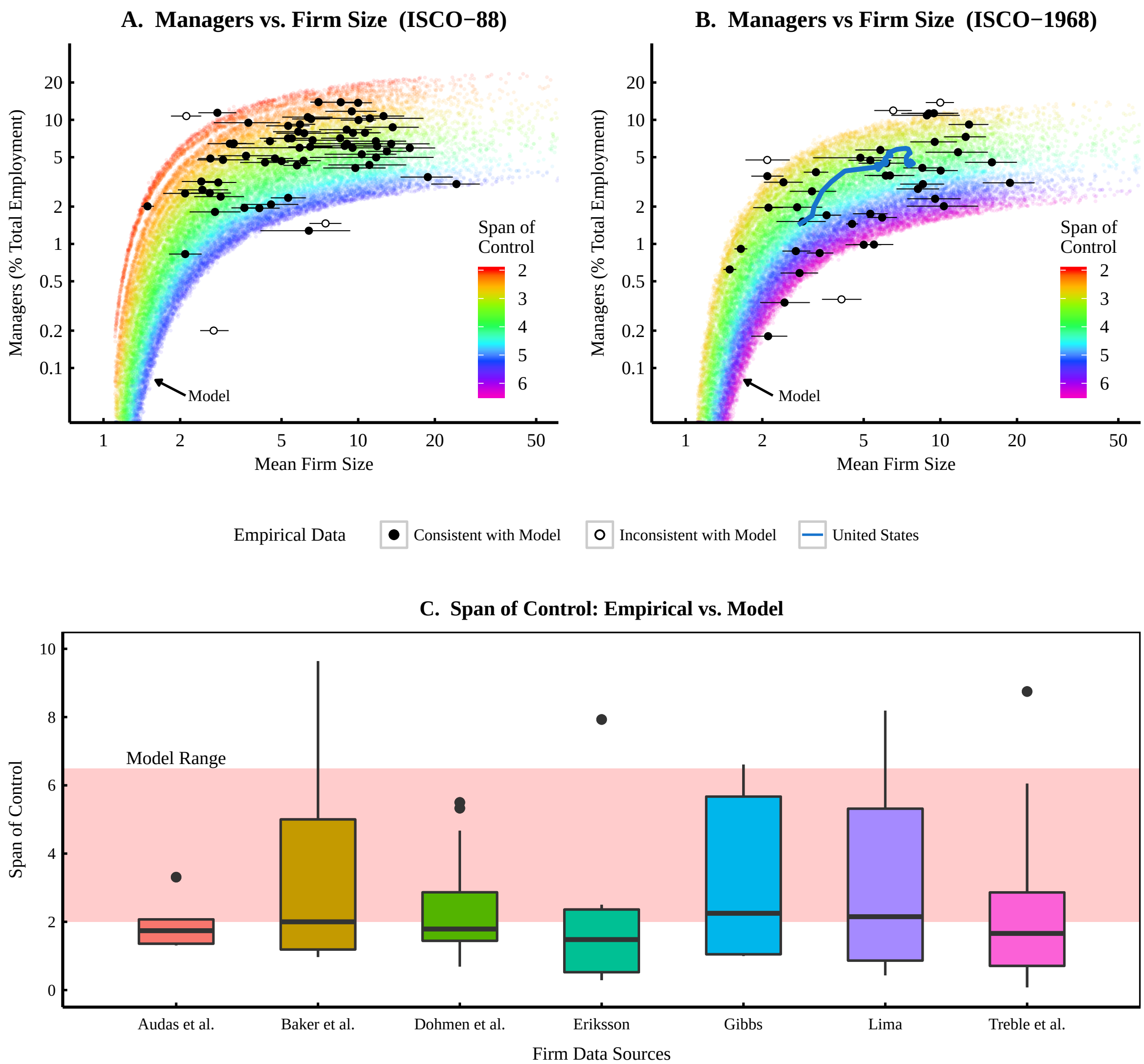

Figure 7: Testing the Hierarchical Model of the Firm Using Managment Rates

Panels A and B plot the country-level relation between the management fraction and mean firm size. Modelled data is also shown in the background, with the span of control indicated by color. Error bars indicate the 95% confidence intervals for mean firm size. Panel C compares the span of control range from the model to the span range for firm level empirical data. Each box plot shows the range of spans within a *single* firm (sometimes over multiple years). Panels A and B use different (incommensurable) classification methodologies for 'management'. Panel A uses ISCO-88 (which includes legislators, senior officials and managers) while panel B uses ISCO-1968 (which includes administrative and managerial workers). For sources and methodology, see Appendix A.

2. All individuals in and above the third hierarchical level are considered 'managers'.
3. The firm size distribution is a power law.

Why assume that management begins at the third hierarchical level? Obviously, individuals within the lowest hierarchical level have no management responsibilities. Those in the second hierarchical level can be thought of as 'working supervisors' — individuals who have some supervisory responsibilities but who spend a majority of their time engaged in 'production' [69]. I assume that individuals in and above the third hierarchical level are devoted mostly to managing the work of others.

This model predicts that the management fraction of employment should grow non-linearly with firm size, eventually approaching an asymptote defined only by the span of control. If the span of control is s, then the asymptote occurs at $1/s^2$ (see Appendix G for the details of this calculation).

In Figure 7 I test this model at the international level. Figure 7A and 7B plot the country-level relation between the management fraction of employment and mean firm size (the two plots show different occupation classification regimes). Empirical data is shown in black, while model predictions are shown in the background with the span of control indicated by color. Different mean firm sizes are produced by varying the exponent of the firm power law distribution (for a technical discussion of this model, see Appendix G).

While the model nicely fits the majority of empirical data, it is important to check that the modelled span of control range is consistent with the span range for *real* firms. Figure 7C shows data for seven studies that document the hierarchical structure of individual firms. Each box plot shows the distribution of spans within a single firm. The important result is that the model's span of control range is consistent with the range for real firms. Thus, the evidence does suggest that increases in mean firm size constitute an investment in social hierarchy.

5 Conclusions

A tacit assumption in most economic theory is that human social organization is to be understood in anthropocentric terms — that the motivations and desires of human beings are to be used to explain how societies evolve. This markedly contrasts with the biological sciences, where the motivations of animals (their desires) are paid little heed. Instead biologists look for deeper principles, often involving energetics.

All life on earth is united by a common struggle — a "struggle for free energy available for work" [12]. The ability to harness energy places key constraints on the structure of life, from the level of the cell [49], to the organism [2, 53], to the ecosystem [57]. Within this unifying context, it seems only fitting that the structure of human society ought to be related to its ability to harness energy — a hypothesis that has been repeatedly advanced by biophysically minded social scientists [20, 38, 66, 73].

The relation between institution size and energy use documented in this paper gives strong support to this hypothesis. Institution size reliably increases as rates of energy consumption grow. I have hypothesized that increases in institution size represent an investment in social hierarchy, an investment that is necessary to facilitate the increasing need for social coordination associated with more complex technology. Of course, this does not negate human agency; rather, it simply implies energetic and evolutionary constraints.

References

[1] ADAMS, R. N. Man, Energy, and Anthropology: I Can Feel the Heat, But Where's the Light? *American Anthropologist 80*, 2 (1978), 297–309.

[2] AHLBORN, B. K. Thermodynamic Limits of Body Dimension of Warm Blooded Animals. *Journal of Non-Equilibrium Thermodynamics 25*, 1 (2005), 87–102.

[3] ANGELINI, P., AND GENERALE, A. On the evolution of firm size distributions. *The American Economic Review 98*, 1 (2008), 426–438.

[4] AUDAS, R., BARMBY, T., AND TREBLE, J. Luck, effort, and reward in an organizational hierarchy. *Journal of Labor Economics 22*, 2 (2004), 379–395.

[5] AXTELL, R. L. Zipf distribution of US firm sizes. *Science 293* (2001), 1818–1820.

[6] AYRES, R., AND WARR, B. Exergy, power and work in the US economy, 1900-1998. *Energy 28*, 3 (2003), 219–273.

[7] BAKER, G., GIBBS, M., AND HOLMSTROM, B. The internal economics of the firm: evidence from personnel data. *The Quarterly Journal of Economics 109*, 4 (1994), 881–919.

[8] BARNEY, J. Firm resources and sustained competitive advantage. *Journal of management 17*, 1 (1991), 99–120.

[9] BECK, T., DEMIRGUS-KUNT, A., AND MAKSIMOVIC, V. Financial and legal institutions and firm size. *The World Bank Development Research Group 2997* (2003).

[10] BIGGS, T., AND OPPENHEIM, J. What drives the size distribution of firms in developing countries. *Employment and Enterprise Policy Analysis, Discussion paper*, 6 (1986).

[11] BIHAM, O., MALCAI, O., LEVY, M., AND SOLOMON, S. Generic emergence of power law distributions and Levy-Stable intermittent fluctuations in discrete logistic systems. *Physical Review E 58*, 2 (1998), 1352.

[12] BOLTZMANN, L. *Populare Schriften*. SEVERUS Verlag, 2011.

[13] BOTTAZZI, G., AND SECCHI, A. Why are distributions of firm growth rates tent-shaped? *Economics Letters 80*, 3 (Sept. 2003), 415–420.

[14] BOTTAZZI, G., AND SECCHI, A. Explaining the distribution of firm growth rates. *The RAND Journal of Economics 37*, 2 (2006), 235–256.

[15] BRENNAN, G., AND BUCHANAN, J. M. *The power to tax: Analytic foundations of a fiscal constitution*. Cambridge University Press, New York, 1980.

[16] CABRAL, L. M., AND MATA, J. On the evolution of the firm size distribution: Facts and theory. *American economic review 93*, 4 (2003), 1075–1090.

[17] CLAUSET, A., SHALIZI, C. R., AND NEWMAN, M. E. Power-law distributions in empirical data. *SIAM review 51*, 4 (2009), 661–703.

[18] CLEVELAND, C., COSTANZA, R., HALL, C., AND KAUFMANN, R. Energy and the US economy: a biophysical perspective. *Science 225*, 4665 (1984), 890–897.

[19] COASE, R. H. The nature of the firm. *Economica 4*, 16 (1937), 386–405.

[20] COTTRELL, F. *Energy & Society (Revised): The Relation Between Energy, Social Change, and Economic Development*. AuthorHouse, Bloomington, Nov. 2009.

[21] DE WIT, G. Firm size distributions: An overview of steady-state distributions resulting from firm dynamics models. *International Journal of Industrial Organization 23*, 5 (2005), 423–450.

[22] DOHMEN, T. J., KRIECHEL, B., AND PFANN, G. A. Monkey bars and ladders: The importance of lateral and vertical job mobility in internal labor market careers. *Journal of Population Economics 17*, 2 (2004), 193–228.

[23] DUNBAR, R. I. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution 22*, 6 (1992), 469–493.

[24] EFRON, B., AND TIBSHIRANI, R. J. *An introduction to the bootstrap*. CRC press, London, 1994.

[25] ERIKSSON, T. Executive compensation and tournament theory: Empirical tests on Danish data. *Journal of Labor Economics 17*, 2 (1999), 262–280.

[26] FEYNMAN, R. P., LEIGHTON, R. B., AND SANDS, M. *The Feynman Lectures on Physics, Desktop Edition Volume I*, vol. 1. Basic books, 2013.

[27] FIX, B. *Rethinking Economic Growth Theory from a Biophysical Perspective*. Springer, New York, 2015.

[28] FU, D., PAMMOLLI, F., BULDYREV, S. V., RICCABONI, M., MATIA, K., YAMASAKI, K., AND STANLEY, H. E. The growth of business firms: Theoretical framework and empirical evidence. *Proceedings of the National Academy of Sciences of the United States of America 102*, 52 (Dec. 2005), 18801–18806.

[29] GABAIX, X. Zipf's Law for Cities: An Explanation. *The Quarterly Journal of Economics 114*, 3 (1999), 739–767.

[30] GAFFEO, E., GALLEGATI, M., AND PALESTRINI, A. On the size distribution of firms: additional evidence from the G7 countries. *Physica A: Statistical Mechanics and its Applications 324*, 12 (June 2003), 117–123.

[31] GEROSKI, P. A. The growth of firms in theory and in practice. *Center for Economic Policy Research 2092* (1999).

[32] GIBBS, M. Incentive compensation in a corporate hierarchy. *Journal of Accounting and Economics 19*, 2 (1995), 247–277.

[33] GIBRAT, R. *Les inegalites economiques*. Recueil Sirey, 1931.

[34] GOLLIN, D. Nobody's business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics 55*, 2 (2008), 219–233.

[35] GOMES, P. M., AND KUEHN, Z. Human capital and the size distribution of firms. *IZA Discussion Papers Series*, 8268 (2014).

[36] GUPTA, H. M., CAMPANHA, J. R., DE AGUIAR, D. R., QUEIROZ, G. A., AND RAHEJA, C. G. Gradually truncated log-normal in USA publicly traded firm size distribution. *Physica A: Statistical Mechanics and its Applications 375*, 2 (2007), 643–650.

[37] HALL, C., AND KLITGAARD, K. *Energy and the Wealth of Nations: Understanding the Biophysical Economy*. Springer, New York, 2012.

[38] HALL, C., THARAKAN, P., HALLOCK, J., CLEVELAND, C., AND JEFFERSON, M. Hydrocarbons and the evolution of human culture. *Nature 426*, 6964 (2003), 318–322.

[39] HARRISON, J. M. *Brownian motion and stochastic flow systems*. Wiley, New York, 1985.

[40] HART, P. E. The size and growth of firms. *Economica 29*, 113 (1962), 29–39.

[41] Hart, P. E., and Prais, S. J. The analysis of business concentration: a statistical approach. *Journal of the Royal Statistical Society. Series A (General) 119*, 2 (1956), 150–191.

[42] Hasan, R., and Jandoc, K. R. The Distribution of Firm Size in India: What Can Survey Data Tell Us? *Asian Development Bank Economics Working Paper Series*, 213 (2010).

[43] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review 91*, 2 (2001), 73–78.

[44] Hipple, S. Self-employment in the United States: an update. *Monthly Lab. Review 127* (2004), 13.

[45] Hymer, S., and Pashigian, P. Firm size and rate of growth. *The Journal of Political Economy 70*, 6 (1962), 556–569.

[46] Kaldor, N. A model of economic growth. *The Economic Journal 67*, 268 (1957), 591–624.

[47] Kesten, H. Random difference equations and Renewal theory for products of random matrices. *Acta Mathematica 131*, 1 (Dec. 1973), 207–248.

[48] Kondepudi, D., and Prigogine, I. *Modern Thermodynamics: from Heat Engines to Dissipative Structures*. John Wiley & Sons, Chichester, 1998.

[49] Lane, N., and Martin, W. The energetics of genome complexity. *Nature 467*, 7318 (2010), 929–934.

[50] Lima, F. Internal Labor Markets: A Case Study. *FEUNL Working Paper 378* (2000).

[51] Lucas Jr, R. E. On the size distribution of business firms. *The Bell Journal of Economics 9*, 2 (1978), 508–523.

[52] Maddison, A. Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD, 2008.

[53] McNab, B. K. Resources and energetics determined dinosaur maximal size. *Proceedings of the National Academy of Sciences 106*, 29 (July 2009), 12184–12188.

[54] Meltzer, A. H., and Richard, S. F. A rational theory of the size of government. *The Journal of Political Economy 89*, 5 (1981), 914–927.

[55] Niskanen, W. A. *Bureaucracy and Representative Government*. Transaction Publishers, Piscataway, NJ, 1974.

[56] NITZAN, J., AND BICHLER, S. *Capital as Power: A Study of Order and Creorder*. Routledge, New York, 2009.

[57] ODUM, H. T. *Systems Ecology; an introduction*. Wiley, New York, 1983.

[58] PELTZMAN, S. The growth of government. *Journal of Law and Economics 23*, 2 (1980), 209–287.

[59] PETERAF, M. A. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal 14*, 3 (1993), 179–191.

[60] POSCHKE, M. The firm size distribution across countries and skill-biased change in entrepreneurial technology. *IZA Discussion Papers Series 7991* (2014).

[61] PRIEM, R. L., AND BUTLER, J. E. Tautology in the resource-based view and the implications of externally determined resource value: further comments. *Academy of Management Review 26*, 1 (2001), 57–66.

[62] RAJAN RAGHURAM, G., LUIGI, Z., RAJAN RAGHURAM, G., AND LUIGI, Z. The firm as a dedicated hierarchy: A theory of the origins and growth of firms. *The Quarterly Journal of Economics 116*, 3 (2001), 805–851.

[63] SANDEFUR, J. On the evolution of the firm size distribution in an African economy. *Center for the Study of African Economies Working Paper Series 5* (2010).

[64] SIMON, H. A., AND BONINI, C. P. The size distribution of business firms. *The American Economic Review 48*, 4 (1958), 607–617.

[65] SINGH, A., AND WHITTINGTON, G. The size and growth of firms. *The Review of Economic Studies 42*, 1 (1975), 15–26.

[66] SODDY, F. *Virtual Wealth and Debt: the Solution of the Economic Paradox*. George Allen & Unwin, London, 1926.

[67] STANLEY, M. H., AMARAL, L. A., BULDYREV, S. V., HAVLIN, S., LESCHHORN, H., MAASS, P., SALINGER, M. A., AND STANLEY, H. E. Scaling behaviour in the growth of companies. *Nature 379*, 6568 (1996), 804–806.

[68] STANLEY, M. H., BULDYREV, S. V., HAVLIN, S., MANTEGNA, R. N., SALINGER, M. A., AND STANLEY, H. E. Zipf plots and the size distribution of firms. *Economics letters 49*, 4 (1995), 453–457.

[69] STRAUSS, G. The Changing Role of the Working Supervisor. *The Journal of Business 30*, 3 (1957), 202–211.

[70] SUTTON, J. Gibrat's legacy. *Journal of Economic Literature 35*, 1 (1997), 40–59.

[71] TREBLE, J., VAN GAMEREN, E., BRIDGES, S., AND BARMBY, T. The internal economics of the firm: further evidence from personnel data. *Labour Economics 8*, 5 (2001), 531–552.

[72] TURCHIN, P., AND GAVRILETS, S. Evolution of complex hierarchical societies. *Social Evolution and History 8*, 2 (2009), 167–198.

[73] WHITE, L. A. Energy and the evolution of culture. *American Anthropologist 45*, 3 (1943), 335–356.

[74] WORLD BANK GROUP. About Us - Enterprise Surveys - World Bank Group, 2016.

Appendices

A Sources and Methodology

BEA US Bureau of Economic Analysis

BLS US Bureau of Labor Statistics

EIA US Energy Information Agency

HSUS Historical Statistics of the United States

ILO International Labour Organization

GEM Global Entrepreneurship Monitor

WBES World Bank Enterprise Survey

This paper has a companion website containing spreadsheets with all final results, as well as simulation code:

http://www.runmycode.org/companion/view/1709

Energy per capita (or per worker)

International energy use per capita data is from the World Bank (series EG.USE. PCAP.KG.OE). US total energy consumption is from HSUS, Tables Db155-163 (1890-1948) and EIA Table 1.3 (1949-2012). US population is from Maddison [52] (1890-2009) and World Bank series SP.POP.TOTL (2010-2012).

US manufacturing sub-sector energy use is from EIA Manufacturing Energy Consumption Survey Table 1.1 First Use of Energy for All Purposes (Fuel and Nonfuel), 2010. Sectoral employment is from Statistics of U.S. Businesses (US 6 digit NAICS for 2010).

Firm Age Composition

The fraction of firms under 42 months old is calculated from the GEM dataset aggregated over the years 2001-2011 (data series babybuso). Uncertainty in this data is estimated using the bootstrap method [24].

Firm size – International

International mean firm size data is calculated using the Global Entrepreneurship (GEM) database, aggregated over the years 2000-2011. In order to account for the over-representation of large firms, I remove firms with more than 1000 employees from the database (see Appendix B). To compare the resulting firm size observations with other time-based series, I use the average year of each country's aggregated data.

Firms with zero employees are assigned a size of 1. This is an attempt to deal with the ambiguity associated with incorporation. The owner of an incorporated sole-proprietorship is usually treated as an employee (by most statistical agencies), but the owner of an unincorporated sole-proprietorship is not. Both types of firms have a single member.

Uncertainty in mean firm size is estimated using the bootstrap method [24]. This involves resampling (numerous times, with replacement) the data for each country and calculating the mean of each resample. Confidence intervals are then calculated using the resampled mean distribution.

Firm size – US average

Average firm size data for 1977-2013 is calculated by dividing the number of persons engaged in production (BEA Table 6.8B-D) by the number of firms. The latter is calculated as the sum of all employer firms in US Census Business Dynamics Statistics plus the number of unincorporated self-employed individuals (BLS series LNU02032192 + LNU02032185).

Average firm size data for 1890-1976 uses firm counts from HSUS Ch408 (which excludes agriculture) and total private, non-farm employment from HSUS Ba471-473. To construct a continuous time-series, the two data sets are spliced together at US Census levels for 1977.

Firm size – US Manufacturing Sub-sectors

Mean firm size and small firm employment share (0-4 employees) are calculated using data from Statistics of U.S. Businesses (US 6 digit NAICS for 2010).

Government Employment

International government employment data is from ILO LABORSTA database (total public sector employment: level of government = Total, sex code = A, sub-classification = 06). Total employment in each country uses World Bank series SL.TLF.TOTL.IN. Note: government employment data points vary between the years 1995-98 and are matched with energy and labor force data for the appropriate year.

US government employment data is from HSUS Ba473 (1890-1928), Ba1002 (1929-40), and BEA 6.8A-D persons engaged in production (1940-2011). Total US employment is from HSUS Ba471 (1900-1928), Ba988 (1929-1940), and BEA tables 6.8A-D (1941-2011).

Large Firm Employment

The measurement of the large firm employment share is inspired by the work of Nitzan and Bichler [56]. Global data is from COMPUSTAT Global Fundamentals (series EMP). Total employment in each country uses World Bank series SL.TLF.TOTL.IN. US 200 largest firms (ranked by employment) from COMPUSTAT North America, series DATA29. Total US employment is from BEA tables 6.8A-D.

Management Fraction of Employment

Management fraction = management employment / total employment. International management employment is from the ILO LABORSTA database using ISCO-88 (Legislators, senior officials and managers) and ISCO-1968 (Administrative and managerial workers). Total employment is from World Bank series SL.TLF.TOTL.IN

US management employment is from BLS Occupational Employment Statistics (various tables, 1999-2014), ILO LABORSTA ISCO-88 (1970-1998) and HSUS Ba1037 (1860-1970). All series are spliced to BLS data. Total US employment is from HSUS Ba1033 (1860-1890), HSUS Ba471 (1900-1928), Ba988 (1929-1940), and BEA tables 6.8A-D (1941-2011). All series are spliced to BEA data.

Particle Accelerators

Data is compiled by the author. Data and sources are available at
http://www.runmycode.org/companion/view/1709.

Power Plant – Construction Time vs. Capacity

Data is compiled by the author. Data and sources are available at
http://www.runmycode.org/companion/view/1709.

Power Plant – Newly Installed Capacity

Plant nameplate capacity data is from 2014 version of EIA form 860 Schedule 3, *Generator Data* (Operable Units Only). This form counts generator capacity, not plant capacity. I treat multiple new generators in a single plant as one entity. Only generators in operation as of 2014 are counted.

Self-Employment

International self-employment data is from the World Bank, series SL.EMP.SELF.ZS. US self-employment data is from HSUS Ba910 (1900-1928), Ba988 (1929-1940) and BEA tables 6.7A–D (1941-2011). Total US employment is from HSUS Ba471 (1900-1928), Ba988 (1929-1940), and BEA tables 6.8A-D (1941-2011).

Span of Control

Empirical span of control data comes from [4, 7, 22, 25, 32, 50, 71]. The span is calculated as the employment ratio between adjacent hierarchical levels.

B Assessing Size Bias within Firm Databases

Like all scientific inquiry, the study of firm size distribution requires reliable data. Unfortunately, accurate firm-size data (with reasonable international coverage) is difficult to find. There are two primary data avenues available: government statistics (the *macro* level) and firm-level databases (the *micro* level). Each avenue has drawbacks.

The problem with relying on macro-level data is that it intrinsically limits the number of countries that can be studied. Apart from wealthy (OECD) nations, reliable macro statistics on firm size distribution are hard to find. This dearth of data often leads researchers to use micro-level databases instead.

The problem with using these databases to study firm size distribution is that they are rarely (if ever) designed to be *accurate* samples of the wider firm 'population'. As the analysis in this section demonstrates, firm-level databases typically under-represent small firms and over-represent large-firms. Thus, when using a micro database to study the firm size distribution, one must ask: is the database an accurate sample of the firm population? The question that immediately follows is: how do we know if the database is (or is not) biased?

In order to assess database bias, one must inevitably make comparisons to macro-level data. The key is to find macro data that is both relevant and *available* (the second criteria being the more difficult to fulfill). In the following sections I present and apply two methods for assessing firm-size bias within micro datasets.

Methods for Determining Firm-Size Bias within a Database

Method 1: Compare *macro* and *micro*-level average firm-sizes.

Method 2: Compare *micro*-level small-firm employment share to *macro*-level self-employment rates.

Method 1 is straightforward: it involves calculating the average firm-size within a *micro* database and comparing it to the average firm-size calculated from *macro* data. This approach is limited by the availability of macro data. For OECD countries, it is possible to directly compare firm-size averages between micro and macro data. I conduct such an analysis in Table 2 (visualized in Figure 10). Unfortunately, for most non-OECD countries, this approach is not feasible because relevant macro-level data does not exist (hence our need for micro data in the first place).

Method 2 is more indirect (and is dependent on some assumptions); however, its advantage is that self-employment data is readily available for most countries. The basic logic of method 2 is as follows:

1. Self-employed individuals work in *small* firms.

2. We can think of the self-employment rate as an indicator of the share of employment held by the smallest firms.
3. By comparing the self-employment rate to the small-firm employment share within a particular database, we can infer the degree of database bias.

As a starting point, I believe method 2 is more useful, since relevant data is more widely available. In Section B I apply method 2 to three databases: Compustat, the World Bank Enterprise Survey (WBES), and the Global Entrepreneurship Monitor (GEM). Figure 8 shows the firm size distribution within these three databases. The distributions are log-transformed in order to show the log-normal character of two of the three databases (Compustat and WBES).

While all three databases are global in scope, their respective firm size distributions are quite different (note the disparities in mean firm-size). Which database gives the most accurate picture of the underlying population of firms? Analysis reveals that the GEM database is the most consistent with available macro data. Based on these results, in Section B I then conduct a more detailed analysis of the GEM database (see Fig. 10).

Small Firm Employment Share as a Database Bias Test

The basic methodology of this test is to use macro-level self-employment rates as an indicator of the share of employment held by small firms. By comparing this rate to the small-firm employment share within a micro database, we can assess the level of bias.

To proceed with analysis, we must make assumptions about the size of self-employer firms. Let us imagine a boundary of size x, below which *all* firms are considered self-employers (and above which *no* firms are considered self-employers). While no such clear boundary exists in the real world, making this assumption greatly simplifies analysis. For the proceeding analysis, I choose $x = 5$: all firms with 5 or fewer members are considered 'self-employer firms'

The choice of 5 as the self-employment boundary may seem high, but this is because we typically think of self-employed individuals as *sole-proprietors* operating a firm with one member. However, the statistical category of 'self-employment' – as defined by the World Bank – is quite broad. It consists of the following sub-categories:

1. Employers
2. Own-account workers
3. Members of producers' cooperatives
4. Contributing family workers

The inclusion of contributing family workers is important, especially for developing countries where household production is still very common. A family of five engaged in a family business will all be considered self-employed. Thus, I believe the choice of 5 as a boundary point to be justifiable.

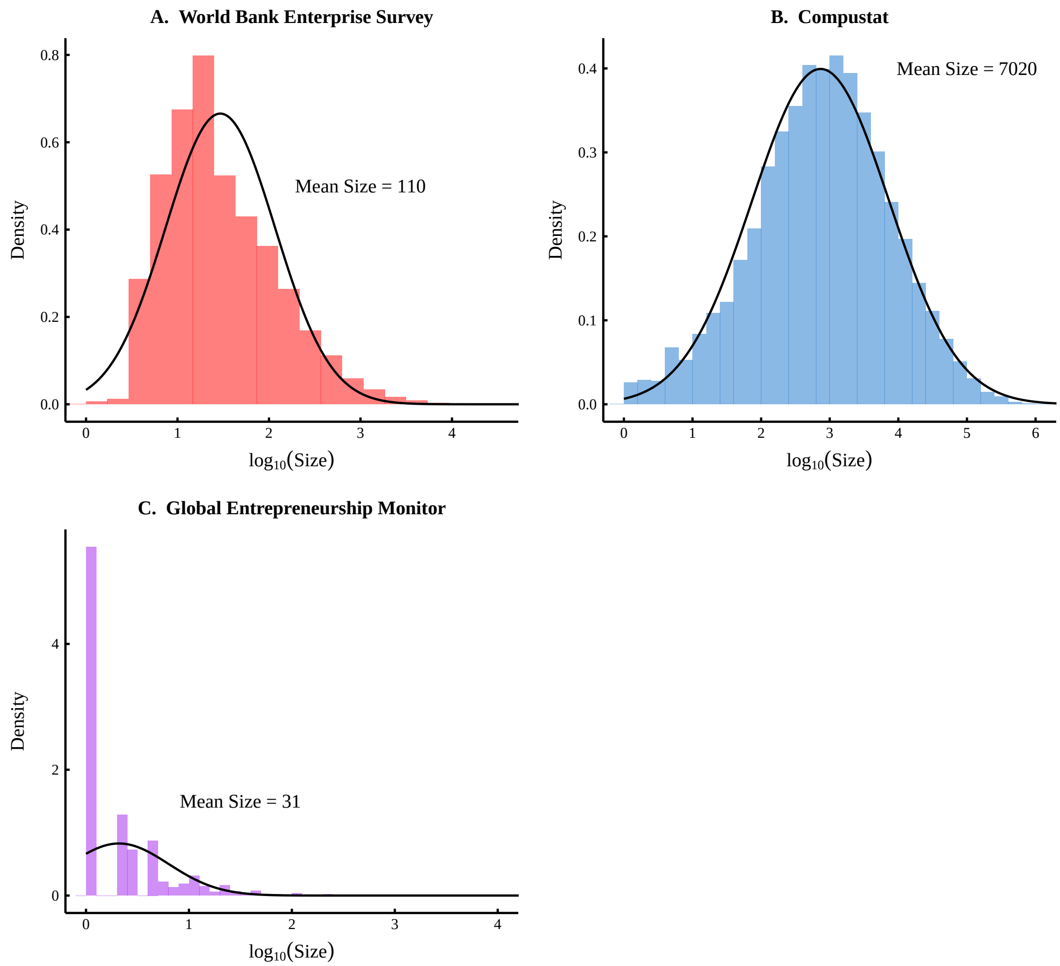

Figure 8: Firm Size Distributions in Selected Micro Databases

Histograms show the firm size distribution within each database (firm size = number of employees). Note that data is log-transformed. Black curves show the best log-normal fit. Panel A shows the firm size distribution of the entire World Bank Enterprise Survery database (for all years). Panel B shows the firm size distribution within the Compustat database (Compustat North America merged with Compustat Global – all available years). Panel C shows the firm size distribution of the Global Entrepreneurship Monitor (GEM) database (from 2000-2011). Note that the log-normal distribution fits both World Bank and Compustat data fairly well, but fits the GEM data very poorly.

Table 1: Small Firm Bias in 3 Selected Micro Databases

	(%)
Global Self-Employment Rate (WDI)	27
Compustat ≤ 5	0.001
WBES ≤ 5	0.3
GEM ≤ 5	5

Notes: This table assesses the relative bias within the World Bank Enterprise Survey (WBES), Compustat, and Global Entrepreneurship Monitor (GEM) databases. The share of employment held by firms with 5 or fewer employees is compared to self-employment levels within World Development Indicator data. Statistics shown are for *entire* databases, and are not weighted by country. The results indicate severe bias within the Compustat database (as expected) and significant bias within the WBES database. Bias within the GEM database appears moderate. Sources: Global self-employment data is for self-employed workers who are *non-employers*. This is calculated by subtracting *employer* rates (series SL.EMP.MPYR.ZS) from total self-employment rates (series SL.EMP.SELF.ZS).

The matter is complicated slightly by the inclusion of the *employers* sub-category within self-employment. Employers are those individuals whose income stems from profit and whose firms employ other people. Since all employer firms, large or small, are owned by *someone*, the employers category has no relation to any particular firm size. In order to focus on small firms only, we should exclude this category from analysis, leaving only self-employed *non-employers* (sub-categories 2–4). The rate of non-employers can be determined by subtracting the employers rate from the total self-employment rate.

The methodology is summarized below: at the micro level, the employment share of firms with 5 or fewer employees is compared to macro data on the self-employed (non-employer) rate.

$$\begin{aligned} \text{Micro} &\Leftrightarrow \text{Macro} \\ \text{Employment Share}_{\text{Firms}\leq 5} &\Leftrightarrow \text{Self-Employment Rate}_{\text{Non-Employers}} \end{aligned} \tag{4}$$

Table 1 applies this methodology to the Compustat, WBES and GEM databases. This analysis suggests that the Compustat database is extremely biased: the small-firm employment share is 4 orders of magnitude lower than the average self-employment rate within the World Development Indicator (WDI) database. Of course, this result is expected; the Compustat database maintains records only for public corporations (which are, by their nature, much larger than most private firms).

While the World Bank Enterprise Survery claims to be a "representative sample of an economy's private sector" [74], the evidence suggests otherwise. The WBES database undershoots WDI self-employment rates by a factor of 100. Only the GEM database gives small-firm employment share levels with the correct order of magnitude.

Assessing Firm-Size Bias Within the GEM Database

While sufficient to weed out extremely biased databases, the method used in Table 1 ignores the internal distribution of data within each database. In general, micro databases with global coverage do not contain equal sized samples for each country. Thus, a large, biased sample from one country could potentially skew the entire database, even if other samples are relatively unbiased. To further test database bias, it is important to group data at the national level. In this section I investigate national-level bias within the GEM database.

I begin with a continuation of the self-employment/small-firm method developed above. However, I now group all data at the national level. The GEM database contains firm samples from a total of 89 countries, 72 of which also have data available in the WDI database. For each country, the employment-share of firms with 5 or fewer employees is calculated (from GEM data) and compared to the WDI self-employed rate (non-employers only). This calculation is done for both the full GEM dataset, and a *truncated* version in which all firms with more than 1000 employees are excluded. This truncated version is tested on the hunch that the full GEM database still over-represents large firms (a hunch that is confirmed in Fig. 10).

The results of this analysis are shown in Figure 9. Both the full and truncated GEM databases have a small-firm employment-share distribution that is roughly equivalent to the WDI self-employment rate distribution. Of particular interest is the fact that the small-firm employment share within the truncated GEM database gives a nearly one-to-one prediction of WDI self-employment rates (see Fig 9A).

This analysis suggests that both the full and truncated GEM databases give a reasonably accurate sample of the international firm size distribution. In order to differentiate between the two, it is helpful to compare mean firm-size estimates with macro data. Due to macro data constraints, this must be done with a much smaller sample size than the 72 countries used above. Table 2 shows the 23 countries for which data is available.

Note that macro-level mean-size estimates are predicated on a few assumptions. Government published statistics usually include firm-counts for *employer* firms only (i.e. firms with employees). Non-employer firms are excluded. Thus, unincorporated self-employed individuals are typically not counted as 'firms' (incorporated self-employed workers are technically counted as employees of their business, and are thus employer firms). As a result, calculations done using official firm-counts only will give a mean firm-size that is disproportionately large. To account for this bias in macro data,

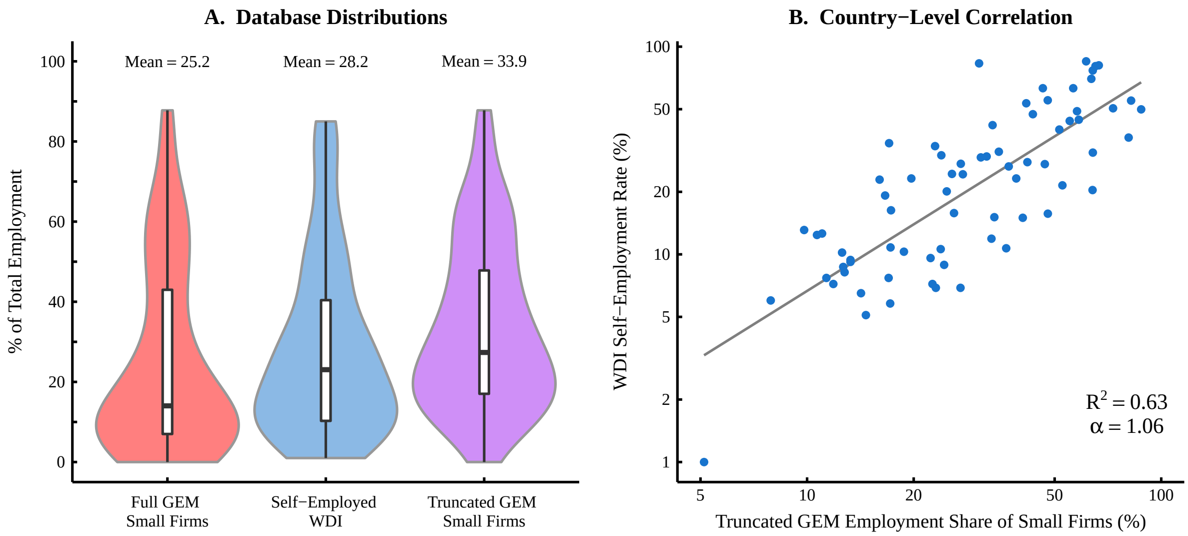

Figure 9: Assessing Small-Firm Bias in the GEM Database

Notes: This figure compares the employment share of small firms ($\leq$ 5 members) in the GEM database to the distribution of self-employment rates (non-employer firms only) within the WDI dataset. Only countries for which data is mutually available are shown (72 countries in total). Unlike Table 1, all data is aggregated at the national level (countries with small/large sample sizes are all weighted equally). Panel A shows how country-level data is distributed within each database. The 'violin' shows the distribution of data. The internal box plot shows the interquartile range (the 25th to 75th percentile), with the median marked as a horizontal line. Corresponding mean values are shown above. Panel B shows a scatter-plot of country-level data (each point is a country) for the self-employment rate vs. the small-firm employment share in the truncated GEM database. The line shows the best-fit power regression. Note that the regression exponent, α, is nearly 1. Thus, the relation between self-employment rates and small-firm employment share is roughly one-to-one. A similar regression for the non-truncated GEM database (not shown) gives $R^2 = 0.48$ and $\alpha = 0.54$, far from a one-to-one relation. This discrepancy between the full and truncated GEM dataset is the result of the over-representation of large firms within a handful of countries. This skews the small firm employment share downwards (note the low median for the full GEM database in Panel A). Thus, the truncated GEM database is more consistent with self-employment data, meaning we can infer that it has less of a firm-size bias.

Sources: Non-employer rates are calculated by subtracting employer rates (series SL.EMP.MPYR.ZS) from the total self-employment rate (series SL.EMP.SELF.ZS). WDI data is chosen for which the data year most closely matches the GEM year (which is calculated as the country-level mean year of all data entries from 2000-2011).

Table 2: Mean Firm-Size in the GEM Dataset vs. Macro Data

Country	Macro	GEM Trunc	GEM Full
Austria	7.6	11.7	12
Belgium	5.6	6.3	6
Czech Republic	3.5	13.5	30
Denmark	9.2	8.5	26
Finland	6.8	5.3	13
France	7.5	5.3	22
Germany	10.4	11.9	151
Hungary	5.7	6.1	8
Italy	3.5	2.8	17
Netherlands	6.1	10	27
Poland	4.7	2.9	16
Portugal	3.6	8.9	9
Russian Federation	18	9	16
Slovakia	4.2	11.8	17
Slovenia	4.9	13	19
Spain	5.5	4.5	10
Sweden	5.8	5.7	15
Switzerland	10.8	6.5	180
Turkey	3	9.5	18
United Kingdom	7.7	7	26
United States of America	9.1	10	164
India	2.6	5.2	6
Ghana	1.5	2.2	2
Mean	6.4	7.7	35.2

Notes: This table compares mean firm sizes within the GEM database to macro-level data. Data is shown for both the full GEM database, and its truncated version, which removes all observations of firms with more than 1000 employees. The rational for truncation is that large firms are over-represented within the dataset, skewing mean firm-size.

Sources and Methodology: Macro-level mean firm-size is calculated by dividing total employment by the number of firms. The number of firms N_{total} is calculated using Eq. (5), where N_{gov} is government data for the number of firms, S_T is the self-employment rate, S_E the self-employed *employer* rate, U is the fraction of self-employed firms that are unincorporated (hence not counted in official statistics), and L is the size of the labor-force.

$$N_{total} = N_{gov} + (S_T - S_E) \cdot U \cdot L \tag{5}$$

Data for S_T, S_E and L come from World Development Indicators (WDI) series SL.EMP.SELF.ZS, SL.EMP.MPYR.ZS, and SL.TLF.TOTL.IN, respectively. Data for the official number of firms comes from OECD Entrepreneurship at a Glance 2013. Due to lack of data, U is assumed to be 0.7, the level observed in the US [44]. For Ghana, all data comes from Sandefur [63], Table 1 and 2. For India, all data comes from Hasan and Jandoc [42], Table 1 and Table 3 (using the sum of the ASI and NSSO datasets). For US data sources, see Appendix A.

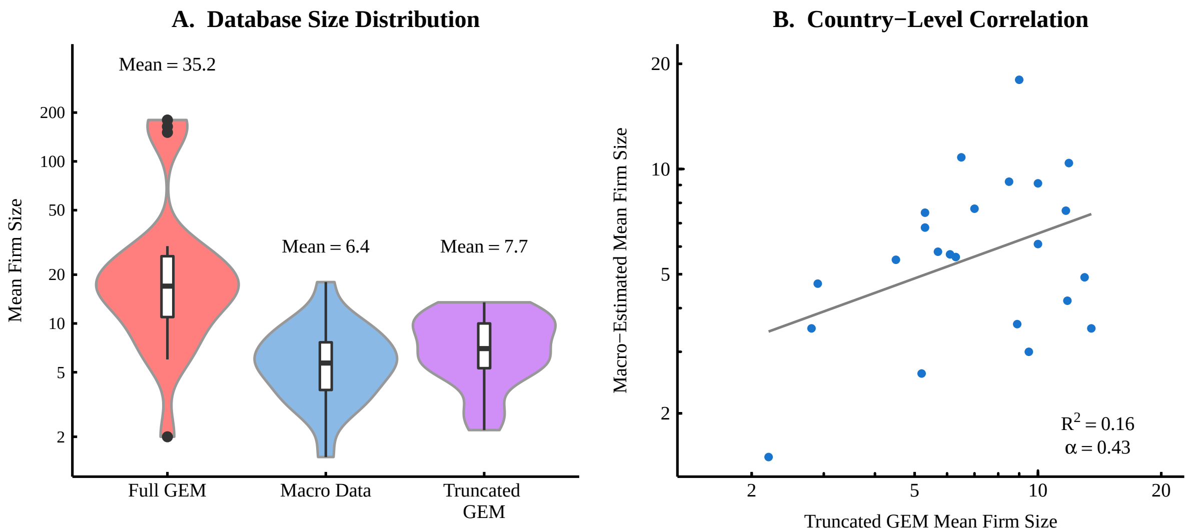

Figure 10: GEM Mean firm size distribution vs. Macro Data

Notes: This figure visualizes the mean firm-size data for the countries shown in Table 2. Panel A shows the mean firm-size distribution within each database. Relative to macro data, the full GEM database clearly over-represents large firms. The mean firm-size in the truncated GEM database is also slightly larger than the macro data, but given the small sample size, the difference is statistically insignificant ($p = 0.20$). Panel B shows the correlation between macro data and the truncated GEM data. A power regression gives an exponent $\alpha = 0.47$, below the desired one-to-one level that would indicate perfect agreement between the micro and macro data. Despite these shortcomings, the truncated GEM database appears to be a fairly accurate sample of the international firm-size distribution.

I adjust the official firm-count by adding an estimate for the number of self-employer firms (see the methodology in Table 2).

The results of this investigation are visualized in Figure 10. From this analysis, there is convincing evidence that the full GEM database over-represents large firms. For a few countries (Germany, Switzerland, and the US) this leads to a mean firm-size estimate that is a factor of 10 larger than macro estimates. Truncating the GEM database seems to effectively adjust for this bias.

Why is truncation effective (and is it justified)? The problem of firm-size bias is partially due to the extremely skewed nature of the firm size distribution. The presence of even a *single* extremely large firm can have a large effect on the mean of a sample. For instance, the GEM database contains roughly 170,000 observations. Suppose that the mean firm-size of these observations is 5. If we add a single observation of a Walmart-sized firm (2 million employees), the resulting average more than *triples* (to roughly 17). Of course, firms this large do exist, but the chance of observing one in a sample should be *extremely* small.

The fact that large firms are over-represented in the GEM database demonstrates a sampling bias. Discarding observations of very large firms is one method for dealing with this bias. Other methods are certainly possible, but I do not discuss them here.

Functional Form of the Firm Size Distribution

One of the first tasks for understanding an empirical distribution (of any kind) is to look for theoretical distributions that can be used to model it. Many observers have used the log-normal distribution to model firm size distributions [3, 16, 33, 36, 41, 68]. As shown in Figure 8, the log-normal distribution is a suitable model for the firm size distribution within the Compustat and WBES databases. However, the preceding analysis showed that these databases are rather poor representations of the actual global firm size distribution.

It may be that the use of the log-normal distribution is an artefact of researchers' reliance on biased micro databases [5]. For data that is more representative of the actual firm size distribution (i.e. the GEM dataset), a power law distribution is a much better fit. The characteristic feature of the log-normal distribution is that its logarithm is *normally distributed* (hence the reason for the log transformation in Fig. 8). A power law distribution, however, will not *not* appear normally distributed under a log transformation. Instead, it will decline monotonically as the GEM database does.

Unlike Compustat and WBES, the GEM database is much better fitted with a power law than with a log-normal distribution (see Fig. 11A). For firms under 10,000 employees, the GEM database is consistent with a power law with a scaling exponent $\alpha \approx 1.9$. Note that the tail of the GEM database is 'fatter' than expect for a power law (it is above the 99% confidence interval). This is consistent with our earlier conclusion that the GEM database over-represents large firms. Macro data from for the US firm size distribution is also consistent with a power law (Fig. 11B).

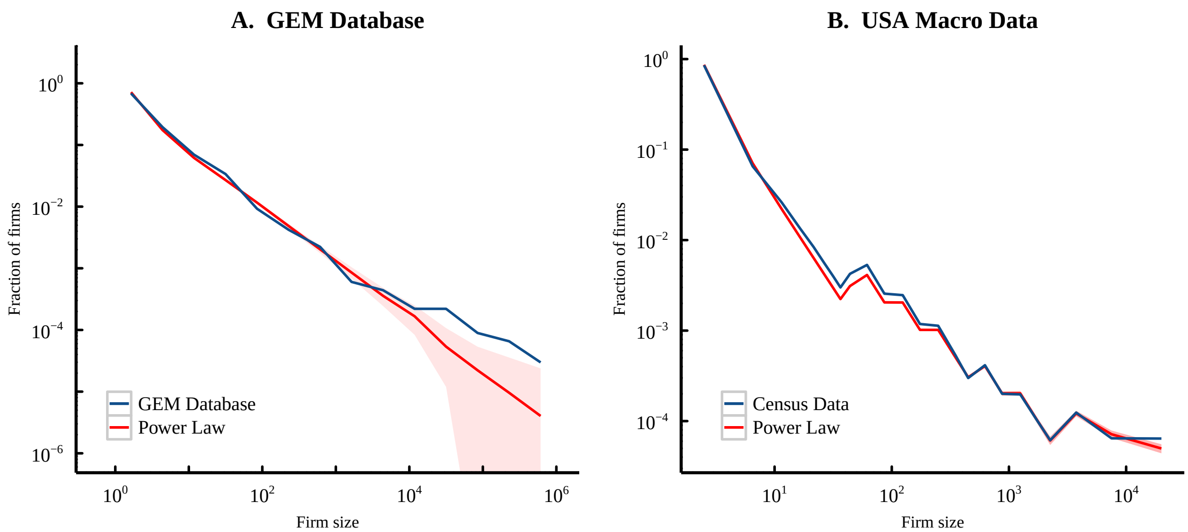

Figure 11: GEM and US Census Data are Consistent with Power Laws

Notes: Panel A shows the firm size distribution of the Global Entrepreneurship Monitor database (all years). For firms with less than 10,000 employees, the database is consistent with a discrete power-law distribution with exponent $\alpha \approx 1.9$. Panel B shows the US firm size distribution, which is consistent with a discrete power-law distribution with exponent $\alpha \approx 2$. Shaded regions show the 99% confidence interval for a simulated power law distribution with a sample size similar to each dataset.

Sources and Methodology: US 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. Both power-law distributions are simulated using the R poweRlaw package, and plotted with the same histogram bins used to plot empirical data. The GEM simulation uses 170,000 observations while the US simulation use 10 million observations.

Note: many readers will expect power law distributions to appear linear when plotted on a log–log scale. Departures from linearity shown in Panel B are artefacts of US census bin sizes (which do not always grow proportionately).

C The Firm Size Distribution as a Variable Power Law

Recent studies have found that firm size distributions in the United States [5] and other G7 countries [30] can be modelled accurately with a power law. Less is known about other countries. In this section, I test if country-level firm size distributions in the GEM database are consistent with a power law. I find that a power law distribution is favored over other heavy-tail distributions in the vast majority of countries. I also find that international variations in 3 summary statistics (mean, self-employment, and large firm employment share) are mostly consistent with a power law distribution.

Power Laws in the GEM Database

The firm size distribution in the *entire* GEM database is roughly consistent with a power law, although the end of the tail is slightly too heavy (Fig. 11A). In this section, I analyse the GEM firm size distribution at the *country* level to assess how well the data fit a power law distribution. I use the truncated GEM database, which contains only firms with fewer than 1000 employees. The rational is that the full GEM database slightly over-represents large firms (see Appendix B).

Historically, power law distributions have been fitted by using an ordinary least-squares (OLS) regression on the logarithm of the histogram. However, this approach is inaccurate, and it violates the assumptions that justify the use of OLS [17]. A more appropriate approach for fitting distributions is to use the *maximum likelihood method*. The likelihood function $\mathcal{L}$ assesses the probability that a set of data x came from a probability density function with the parameter(s) θ.

$$\mathcal{L}(\theta|x) = P(x|\theta) \tag{6}$$

The best fit parameter(s) θ_{mle} maximizes the likelihood function. Like any fitting method, the maximum likelihood indicates only the best fit parameters of the *specified* model, not the appropriateness of the model itself. To discriminate between two different models (1 and 2), we compare their respective maximum likelihoods in ratio form (Λ). The larger likelihood indicates the better fitting model.

$$\Lambda_{1,2} = \frac{\mathcal{L}_1(\theta_{mle}|x)}{\mathcal{L}_2(\theta_{mle}|x)} \tag{7}$$

It is often more convenient to use the log-likelihood ratio, $\log \Lambda$. The sign of $\log \Lambda$ indicates the preferred model (positive indicates that model 1 is better, negative indicates that model 2 is better). The magnitude of $\log \Lambda$ indicates the strength of this preference.

I use this method to assess if country-level firm size distributions in the GEM database are best modelled with a power law. I compare the likelihood of a power

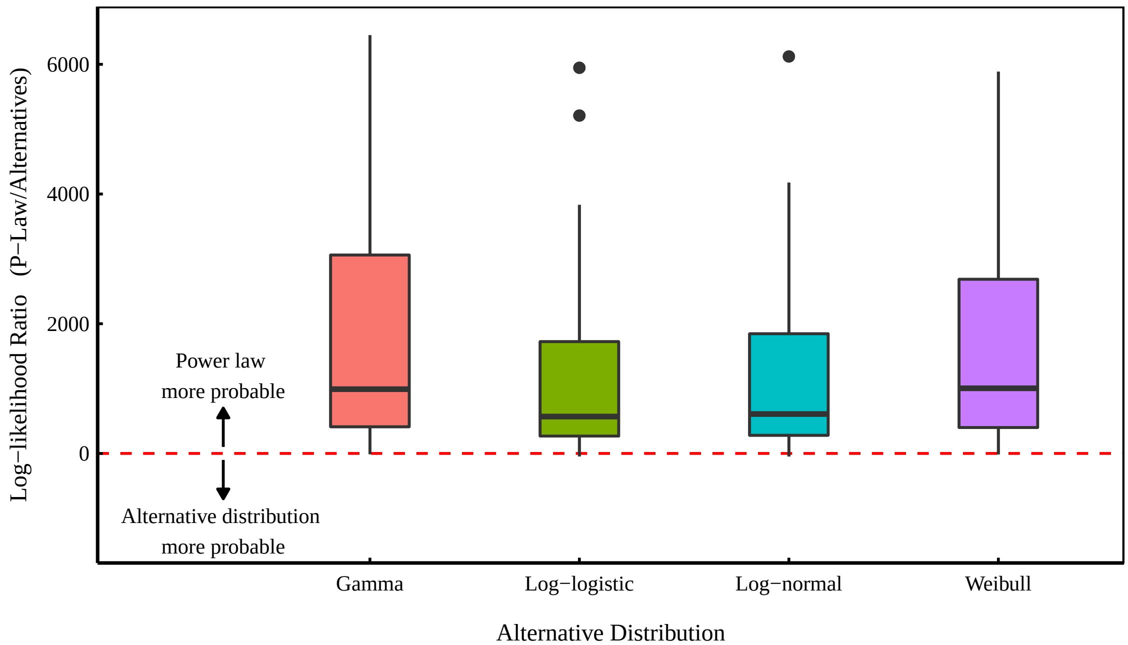

Figure 12: Comparing the Power Law to Alternatives in the GEM Database

Using country-level firm size distributions from the GEM database, this figure assesses the goodness of fit of a power law relative to four other heavy-tail distributions. The firm size distribution in each country in the GEM database is fitted with a power law, gamma, log-logistic, log-normal, and Weibull distribution. For each country, the log-likelihood ratio is computed between the power law and the four alternative distributions. The box plots display the resulting range of ratios. A positive ratio indicates that the power law is more probable, while a negative ratio indicates that the alternative distribution is more probable. In order to better display the majority of data, several large outliers favoring a power law are not shown. For all but 3 countries, a power law distribution is the best fit.

Notes: This figure shows the mean log-likelihood ratios for 100 re-samples (with replacement) of each country. Maximum likelihoods are calculated using the R packages 'poweRlaw' (for a power law) and 'fitdistrplus' (for alternative distributions). Although empirical data is discrete, all models used here are continuous.

law distribution to the likelihood of four other heavy-tail distributions: gamma, log-logistic, log-normal, and Weibull. The resulting range of log-likelihood ratios (one for each country in the GEM database) is shown in Figure 12. A power law distribution is favored over other distributions in the vast majority of countries (97%).

International Summary Statistics

Firm size summary statistics can be used as another way to test if the firm size distribution is consistent with a power law. This has the advantage of broadening the evidence to include more data sources (I combine GEM, World Bank, and Compustat data). My method is to pair two statistics and test if the resulting empirical relation can be reproduced by simulated samples from a power law distribution. I look at two pairings: (1) the self-employment rate vs. mean firm size; (2) the large firm employment share vs. mean firm size.

Self-Employment vs. Mean Firm Size

The rational for looking at the self-employment rate is that it indicates the relative share of employment held by small firms. Figure 13A shows the empirical relation between self-employment rates and mean firm size (black dots). The simulated relation is shown in the background, where the power law exponent α is indicated by color. Creating this simulation requires making assumptions about the size of self-employer firms. I assume that *all* firms below the size boundary L_s are considered self-employer firms. The simulated self-employment rate then consists of the fraction of employment held by firms with employment less than or equal to L_s.

To account for international variation in the size of self-employer firms, I let the boundary point vary randomly over the range $1 \leq L_s \leq 10$. In Figure 13A, $L_s = 1$ corresponds to the bottom of the coloured region, and $L_s = 10$ to the top. Why choose the upper bound to be so large? My reasoning is based on the definition of 'self-employment', which consists of 3 sub-categories: own-account workers, cooperatives, and family workers.[1] Especially in developing countries, where household production is still common, a self-employer 'firm' is synonymous with a *family*. A size of 10 seems a reasonable upper limit on the size of family. Given this assumption, a majority of countries (75%), as well as the entire time series for the United States, have a self-employment vs. mean firm size relation that is consistent with a power law.

[1] Most statistical databases add a fourth category of 'employers' (i.e. capitalists). Because this category is not related to small firms, I remove it from analysis.

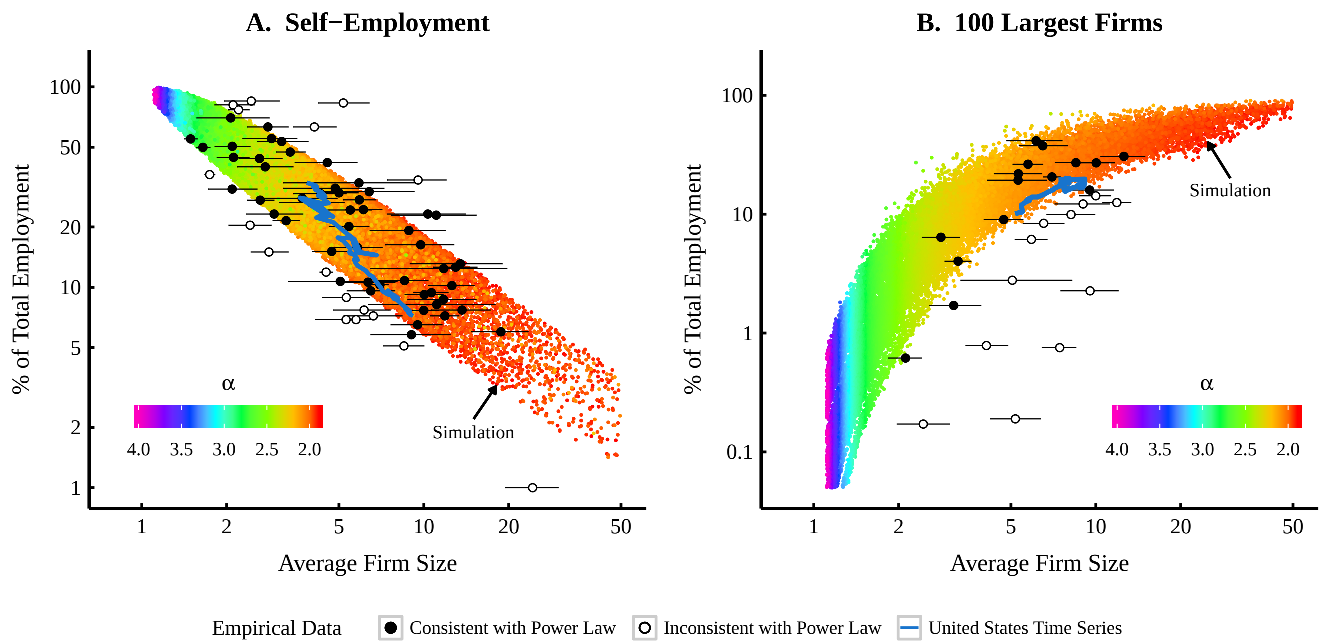

Figure 13: International Summary Statistics, Empirical vs. Power Law

This figure compares pairings of summary statistics for empirical and simulated data. Empirical data is at the country level. Simulated data is randomly generated from a power law distribution (the exponent α is indicated by color). Panel A shows self-employment rates vs mean firm size while panel B shows large firm employment share vs. mean firm size. Self-employment rates are modelled as the employment share of all firms less than the size L_s, which varies randomly over the range $1 \leq L_s \leq 10$. Uncertainty in mean firm size (95% level confidence intervals) is indicated by horizontal lines. Empirical data is judged to be consistent with a power law when the error bar is within the 99% range of simulated data. For data sources, see Appendix A.

Large Firm Employment Share vs. Mean Firm Size

To test if variations in the large firm employment share are consistent with a power law distribution, I use the same method as above: I plot the employment share of the 100 largest firms against mean firm size (Fig 13C). I then compare this relation to the one predicted by simulated power law data. To allow for the effects of differing country size, simulation sample sizes vary over the range of national firm populations (which are estimated by dividing the labor force by the mean firm size).

A slight majority of countries (56%), as well as the entire time-series for the United States, have a large firm employment share vs. mean firm size relation that is consistent with a power law distribution. Note that all data points that are *not* consistent with a power law lie *below* the simulation zone (rather than above). This could indicate that these countries have firm size distributions with a tail that is thinner than a power law, but it could also indicate a problem with the data. I have assumed that the 100 largest firms in the Compustat database are actually the largest firms in each nation. There is no guarantee that this assumption is true: the Compustat database may not give complete coverage of the largest firms, especially if a country has many large *private* companies. Further research is needed to determine if these findings indicate a departure from a power law distribution, or if they are artefacts of incomplete data.

D Testing Gibrat's Law Using the Compustat Database

Gibrat's 'law' states that firm growth rates are *independent* of firm size. To what extent is this supported by empirical evidence? I investigate here using the Compustat US database. My results are consistent with previous analysis of the Compustat database: growth rates are approximately Laplace distributed, and volatility declines with firm size [67]. However, I show that this decline is of importance to only a small subset of firms.

Analysis

Rather than directly calculate the mean and variance of Compustat firm growth rates, I fit the growth rate distribution with a truncated Laplace density function (growth rates less than -100% are rounded to -100%). I then investigate how the parameters of this function change with firm size (Fig. 14). The advantage of this approach is that it is not biased by large outliers, and it allows a direct comparison of empirical data to modelled data (where firm growth rates are drawn from a Laplace distribution).

To estimate the Laplace parameters, I fit the histogram of simulated data to the histogram of empirical data (using a Monte Carlo technique that minimizes the absolute value of the error). The results are displayed in Figure 14C-D. The location parameter (μ) shows no significant relation to firm size. However, growth rate volatility (the scale parameter, b) declines monotonically with firm size.

Interestingly, the location parameter is always less than zero, meaning the most probable rate of growth is *negative*. This finding is consistent with the conditions predicted by a stochastic model with a reflective lower bound. Such a model will be stable only when there is a net *negative* drift to firm size (Appendix E). In Appendix F I reproduce the US firm size distribution using a model with a location parameter of -1%, which is consistent with Compustat data.

Extrapolating to the Entire Economy

Because the Compustat database contains data only for publicly traded firms, it is not an accurate sample of the wider US firm population (see Appendix B). However, based on the assumption that the US firm size distribution is a power law, we can estimate how the volatility-percentile relation shown in Figure 14D might look for the economy as a whole. The method for this process is shown in Table 3.

The first step is to generate a US firm sample from a power law distribution that best fits empirical data (I use $\alpha = 2.01$ here), and then compute size percentiles. Next, we select a particular percentile (the green cell) and note the corresponding firm size in the Compustat database (left pink cell). We than find all firms within the power-law sample that have the same size (right pink cells). The scale parameter for the

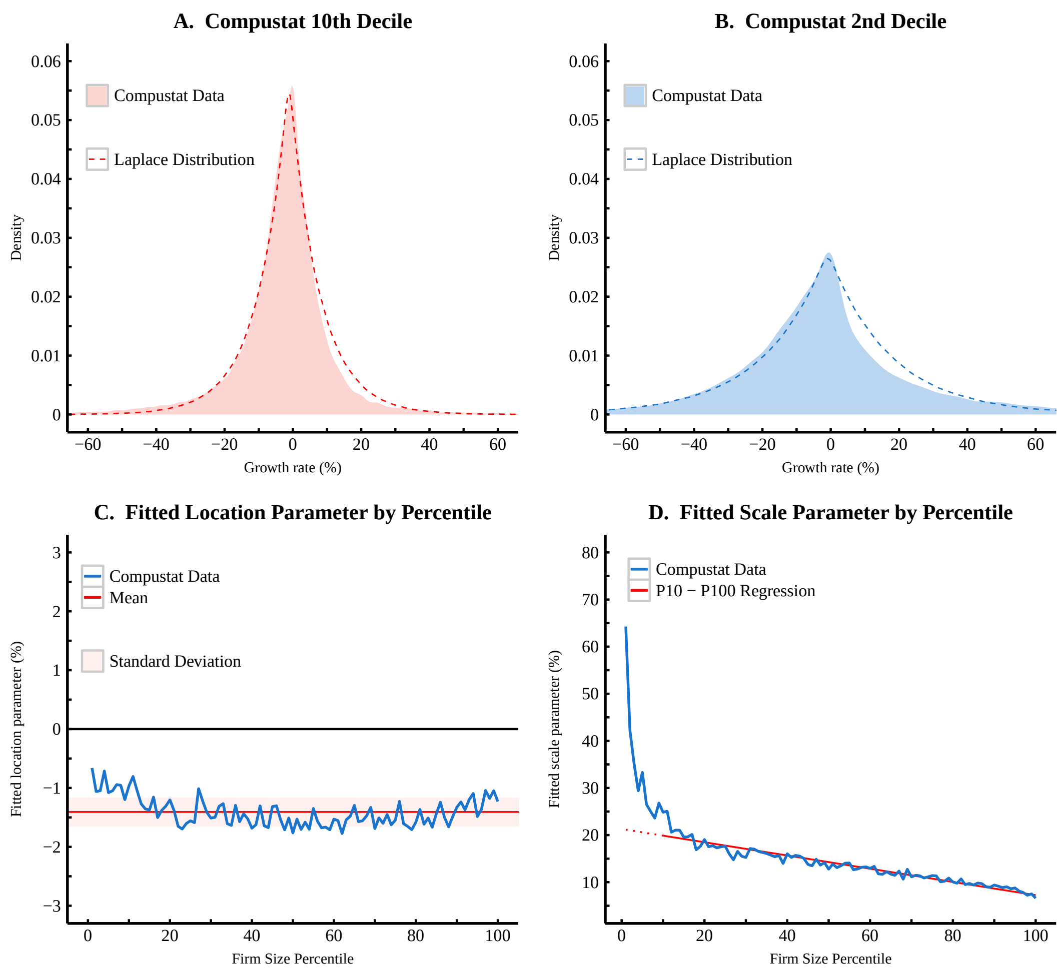

Figure 14: Firm Growth Rate Distribution in the Compustat US Database

This figure analyses firm growth rates (by employment) within the Compustat US database from 1970 to 2013. Panel A shows the growth rate distribution for firms in the 10th (top) decile, while Panel B shows the distribution for firms in the 2nd decile. Dotted lines indicate the best-fit Laplace distribution. Panel C and D show the results of Laplace regressions at the percentile level. Panel C shows the estimated location parameter (μ), while Panel D shows the estimated scale parameter (b). Laplace distributions are fitted using a Monte Carlo method. This analysis indicates that growth rate volatility is a function of firm size, while the growth rate mode is not. Given the firm-size bias of the Compustat database, results for lower percentiles (i.e. P1-P10) should be treated with scepticism.

Table 3: Method for Transforming Compustat Scale Parameter Regressions

Percentile	Compustat Firm Size	Scale	Power Law Firm Size	Transformed Scale
1	1	60	1	60
2	3	50	1	60
3	–	–	1	60
4	–	–	1	60
5	–	–	1	60
6	–	–	2	50

This table demonstrates the method for transforming the Compustat scale-percentile relation to an estimated relation for the whole economy. The first step is to select a percentile (the green cell P1 is selected here). We then match the Compustat firm size of this percentile to the equivalent power law firm size (pink cells). The Compustat scale parameter is then mapped onto all power law percentiles with matching firm sizes, resulting in a transformed scale function (purple cells).

---

selected Compustat percentile (left purple cell) is then mapped onto these firms, and their corresponding percentiles. The result (right purple cells) is a transformed relation between firm percentile and scale parameter that serves as our economy-wide estimate.

The results of this transformation are shown in Figure 15. Two different estimates are shown. The blue curve shows results using the raw data shown in Figure 14D, while the red dotted curve shows results using a linear regression for P10-100, extrapolated over all percentiles.

Why two different methods? The bias in the Compustat increases as firm size decreases: coverage for large firms is nearly complete, while coverage of small firms (under 10) is extremely limited. Thus, it is quite possible that the large increase in volatility for firm percentiles 1–10 may be an artefact of this bias. By using the linear regression of P10-P100, we remove this potential artefact. We can think of the two curves in Figure 15 as representing a plausible range for the US economy. The stochastic model used to reproduce the US firm size distribution (Fig. 16), has a location parameter of 34%, which is much nearer the lower bound of our Compustat estimates.

This analysis suggest that declines in growth rate volatility are important only to a small minority of firms.

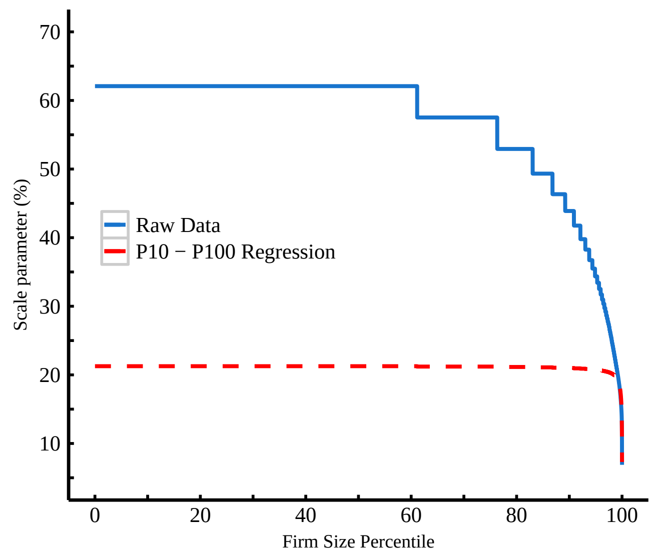

Figure 15: Scale Parameter vs. Percentile, Economy-Wide Estimates

This figure shows a transformation of the Compustat scale-percentile regressions (Fig. 14D) to a form that is consistent with the firm size distribution of the entire US economy. The US distribution is modelled with a power law ($\alpha = 2.01$). The blue curve shows the relation that would result from using the entire range of the Compustat regressions (P1-100). The step-wise pattern is a result of discrete data (steps correspond to a change in firm size by 1). The red dotted curve shows the relation resulting from using a linear regression of Compustat P10-100 (red line in Fig. 14D), extrapolated over P1-10.

E Instability of the Gibrat Model

The Gibrat model assumes that firm growth is a stochastic, multiplicative process. If L_0 is the initial firm size and x_i the annual growth rate, then firm size at time t is given by:

$$L(t) = L_0 \cdot x_1 \cdot x_2 \cdot ... \cdot x_t \;\; = \;\; L_0 \prod_{i=1}^{t} x_i \tag{8}$$

The instability of this model was first noted by Kalecki [46]. It stems from the model's diffusive nature: the resulting firm size distribution tends to *spread* with time. This tendency can be understood by relating the model to the classic example of diffusion: the one-dimensional random walk.

In a random walk model, a particle is subjected to a series of random additive shocks (y_i) that cause its position to change over time. At any given time, the particle's displacement from the initial position $d(t)$ is simply the sum of all of these shocks:

$$d(t) = y_1 + y_2 + ... + y_t \;\; = \;\; \sum_{i=1}^{t} y_i \tag{9}$$

In order to intuitively understand how this leads to diffusion, let us suppose that the shocks y_i are drawn from the uniform distribution $\{-1, 1\}$. At any given time, we can ask: what is the maximum possible displacement? In this case, it is exactly equal to t (the number of time intervals that have passed). When we introduce many randomly moving particles, some may attain this maximum displacement (however unlikely it is). Since the maximum grows with time, we can conclude that the displacement distribution must spread with time.[2]

The Gibrat model shares this property, except that the diffusion is *exponential*. To see this, we take the logarithm of Eq. 8, which allows us to express the growth rate product as a *sum*.

$$\begin{aligned} \log(L(t)) &= \log(L_0) + \log(x_1) + \log(x_2) + ... + \log(x_t) \\ &= \log(L_0) + \sum_{i=1}^{t} \log(x_i) \end{aligned} \tag{10}$$

We then exponentiate to get:

$$L(t) = L_0 e^{\sum_{i=1}^{t} \log(x_i)} \tag{11}$$

[2]For a step size drawn from the uniform distribution $\{-1, 1\}$, the standard deviation of the displacement is equal to $\sqrt{t}$. For a good derivation, see Feynman [26] Ch. 6.

By setting $\log(x_i) = y_i$, we can see that Eq. 11 is just Eq. 9 in exponential form: our firm growth model is a one-dimensional, *exponential* random walk. The resulting firm size distribution will therefore spread rapidly with time – a fact that is inconsistent with available evidence. For instance, we know that the US firm size distribution has changed little since 1970 (see Fig. 2).

The second problem with this model is that it gives rise to a *log-normal* distribution, contradicting our finding that most firm size distributions are best described by a power law. The proof that this model leads to a log-normal distribution is straightforward. For a sufficiently large number of iterations, the Central Limit Theorem dictates that the sum of independent, random numbers will be *normally* distributed. Thus, for a large number of random walkers, the displacement $d(t)$ will be normally distributed (so long as the distribution of y_i satisfies certain conditions). Because Eq. 11 is the exponential form of Eq. 9, the *logarithm* of $L(t)$ will be normally distributed – the defining feature of the log-normal distribution.

Adding a Reflective Lower Bound

One simple way to reform this model is to add a *reflective lower bound* that stops firms from shrinking below a certain size [47, 11, 29]). This slight change will cause the model to generate a power law, rather than a log-normal distribution. It also leads to model stability (under certain conditions).

Why does the introduction of a reflective boundary lead to a power law distribution? One way of understanding this is to relate back to the additive random walk. If a reflective barrier is added to a one-dimensional random walk, it will no longer tend towards normal distribution; rather, it will tend towards an *exponential* distribution (see [39], p 15 for a proof).

Recall that a multiplicative process can be transformed into an additive process by taking the logarithm. Therefore, for a multiplicative firm model with a lower bound, the logarithm of firm size (L) will be exponentially distributed. Thus, the firm size distribution $p(L)$ is given by Eq. 12, which reduces to a power law (where C is the normalizing constant, and α is the scale parameter).

$$\begin{aligned} p(L) &= Ce^{-\alpha \cdot \log(L)} \\ &= CL^{-\alpha} \end{aligned} \tag{12}$$

For a firm size distribution, the obvious choice for a minimum lower bound is $L = 1$ (a sole-proprietor with no employees). In the proceeding model, I implement this reflection through the following conditional statement, which is evaluated at every time interval:

$$\text{if} \;\; L(t) < 1, \;\; \text{then} \;\; L(t) = 1 \tag{13}$$

Introducing a reflective lower bound also solves the instability problem, but only when growth rates have a negative 'drift'. Why? Intuitively, we can state that a model will be stable if it is not possible for a firm to shrink or grow *forever*. Introducing a lower bound automatically stops firms from shrinking forever, but it does nothing to stop the possibility of unending growth.

However, if firm growth rates have a net *downward* drift, all firms will tend towards a size of 0, given enough time. This downward drift occurs when the *geometric* mean of the growth rate distribution is less than 1. We can draw an analogy with gas particles moving in a gravitational field on earth. The particles move randomly, but there must be a small net downward drift due to the force of gravity. The result is a stable distribution of particles. If we remove gravity, the particles are free to diffuse forever. Similarly, if we remove the downward bias to firm growth rates, the distribution becomes unstable.

F Properties of Stochastic Models

Despite their simplicity, stochastic models of firm growth are able to replicate many important properties of the real world. I review three such properties here. Stochastic models can be used to:

1. Generate a firm size distribution that is consistent with empirical data;
2. Reproduce the relation between firm size and firm age;
3. Simulate new firm survival rates over time.

Modelling the US Firm Size Distribution

The model used here assumes scale-free growth with a reflective lower bound at a firm size of one. Growth rates are drawn from a Laplace distribution that is truncated by rounding all (fractional form) growth rates less than 0 to 0. In order to maintain a discrete distribution, firms with non-integer size are rounded to the nearest integer (after the application of each growth rate).

This simple model can be used to replicate the US firm size distribution (Fig. 16). In this case, model parameters $\mu = 0.99$ and $b = 0.34$ are used. The model shows the distribution of 1 million firms after 100 time iterations. In order to capture fluctuations around the equilibrium, the model is run 100 times, with the shaded region showing the resulting range of outcomes.

Firm Age vs. Firm Size

Firm age is calculated as the time since a firm's last 'reflection'. The model described above can be used to replicate the size-age relation of firms in the World Bank Enterprise Survey (WBES) database (Fig. 17A). The fitted parameters are $\mu = 0.97, \;\; b = 0.55$. Note that the model diverges from WBES data for firms with fewer than 10 employees. Due to the size bias within the WBES database (see Appendix B), it is not clear if this divergence is significant, or an artefact of database bias.

Firm Survival Rates

The survival rate of new firms tends to decline exponentially over time (Fig. 17B). To replicate this behavior, we give our stochastic model an initial firm size distribution and then track firm survival over time. A firm 'dies' when it is reflected for the first time. At any given time, the firm survival rate is given by the fraction of firms that have never been reflected.

In order to model firm survival rates, we must choose an initial distribution of firms. We can make guesses about this distribution based on BLS *establishment* data.

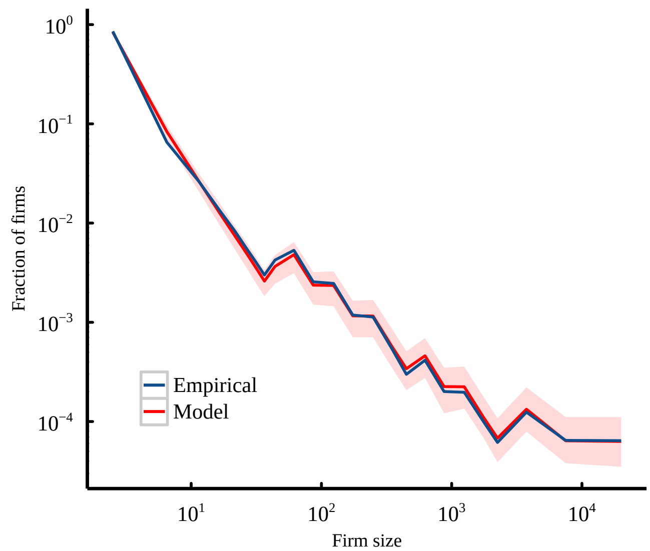

Figure 16: A Stochastic Model of US Firm Size Distribution

The US firm size distribution is shown for the year 2013 (blue line), along with a stochastic model (red) of 1 million firms with growth rates drawn from a truncated Laplace distribution with parameters $\mu = 0.99, \quad b = 0.34$. The shaded region indicates the 90% confidence region of the model. US 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. The model histogram uses Census bins to allow direct comparison.

In 1994 — the first year the BLS tracked survival rates — the average size of new establishments was 7.3. In the same year, the average size of *all* US establishments was 16.9 (using data from Census Business Dynamics Statistics. It seems reasonable to assume that the average size of new *firms* might also be about half the average for all firms. It also seems reasonable to assume that the distribution of new firms can be modelled with a power law. Using these assumptions, I model the initial firm size distribution with a power law of $\alpha = 2.1$. This gives a mean size of close to 5 (about half the US average).

The empirical data shown in Figure 17 comes from the US Bureau of Labor Statistics (BLS). A caveat is that this data is for *establishment* (not *firm*) survival rates. An establishment refers to a specific business location, while a firm is a legal entity that may contain multiple establishments. For modelling purposes, I ignore this distinction here and assume that establishments are equivalent to firms.

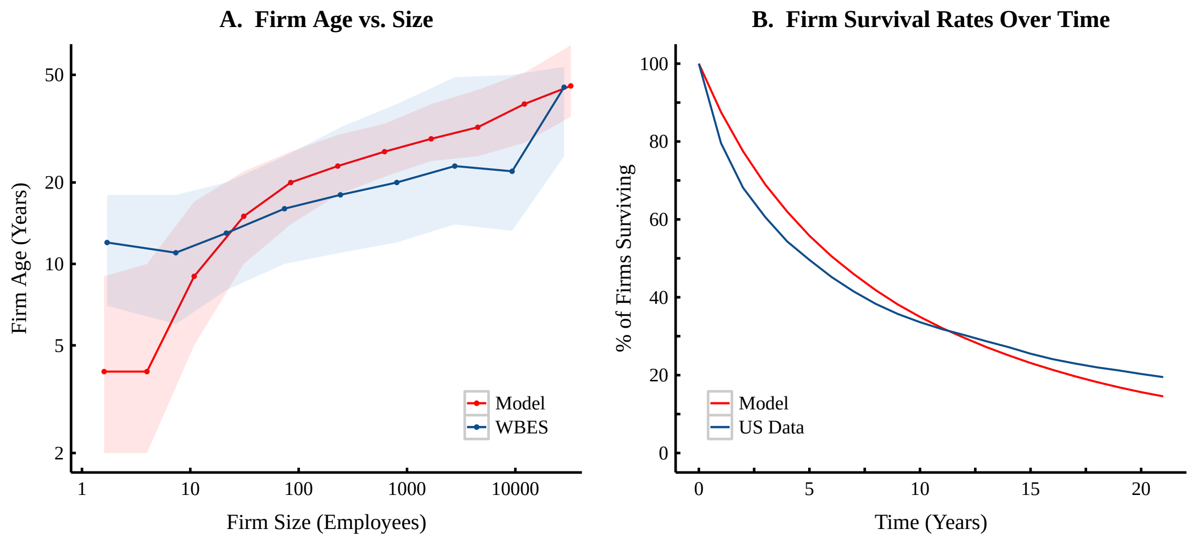

Figure 17: Stochastic Models Can Reproduce Firm Age/Survival Data

Panel A shows the relation between firm size and firm age within the World Bank Enterprise Survey (WBES) database (blue). A stochastic model (red) with growth rates drawn from a truncated Laplace distribution with parameters $\mu = 0.97, \;\; b = 0.55$ produces a similar firm size-age relation. Lines indicate medians and shaded regions indicate the interquartile range. Logarithmic bin locations are indicated with points. Panel B shows the survival rates of new firms over a period of 21 years. Empirical data (blue) is from the BLS Business Employment Dynamics database, Table 7, Survival of private sector establishments by opening year. The model (red) draws growth rates from a truncated Laplace distribution with parameters $\mu = 0.99, \;\; b = 0.35$.

Empirical and modelled survival rates are shown in Figure 17B). The survival rate model parameters ($\mu = 0.99, \;\; b = 0.35$) are nearly identical to the parameters ($\mu = 0.99, \;\; b = 0.34$) used to replicate the US firm size distribution (Fig. 16). These parameters are also consistent with the range estimated from Compustat data (Appendix D).

G Hierarchical Model of the Firm

An 'ideal' hierarchy has a constant span of control throughout — meaning the employment ratio between each consecutive hierarchical level is *constant* (Fig. 18). This property allows total employment to be expressed as a geometric series of the span of control s. If the number of individuals in the top hierarchical level is a, and h_t is the total number of hierarchical levels, then total employment L is given by the following series:

$$L = a\left(1 + s + s^2 + ... + s^{h_t - 1}\right) \tag{14}$$

Using the formula for the sum of a geometric series, Eq. 14 can be rewritten as:

$$L = a\frac{1 - s^{h_t}}{1 - s} \tag{15}$$

We make the assumption that individuals in and above the hierarchical level h_m are considered *managers*. The number of managers M in a firm with h_t levels of hierarchy is equivalent to the employment of a firm with $h_t - h_m + 1$ levels of hierarchy:

$$M = a\frac{1 - s^{h_t - h_m + 1}}{1 - s} \tag{16}$$

We can use Eq. 16 and Eq. 15 to express management as a fraction of total employment (M/L):

$$\frac{M}{L} = \frac{1 - s^{h_t - h_m + 1}}{1 - s^h} \tag{17}$$

Asymptotic Behavior of the Management Fraction

The management fraction tends to grow with the number of hierarchical levels, but only to a certain point (Fig. 19). For $h_t > 10$ the management fraction approaches an asymptotic limit that depends only on the span of control s. Finding the asymptotic behavior of M/L requires evaluating the following limit:

$$\lim_{h_t \to \infty} \frac{M}{L} = \lim_{h_t \to \infty} \frac{1 - s^{h_t - h_m + 1}}{1 - s^{h_t}} \tag{18}$$

To evaluate this limit, I use L'Hospital's Rule, which states that $\lim \frac{f(x)}{g(x)} = \lim \frac{f'(x)}{g'(x)}$. We first rewrite Eq. 18 in a differentiable form, with a base e exponent:

$$\lim_{h_t \to \infty} \frac{M}{L} = \lim_{h_t \to \infty} \frac{1 - e^{\log(s)\cdot(h_t - h_m + 1)}}{1 - e^{\log(s)\cdot h_t}} \tag{19}$$

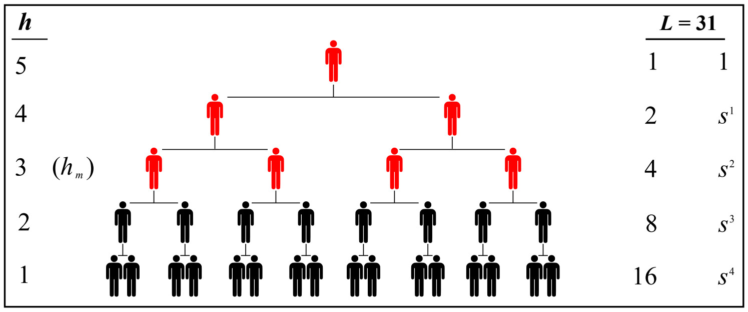

Figure 18: A Perfectly Hierarchical Firm

Within a perfectly hierarchical firm, the number of individuals in adjacent hierarchical levels differs by a factor of the span of control s (in this diagram, $s = 2$). This characteristic allows total employment L to be expressed as a geometric series of s. Managers (red) are defined as all individuals in and above level h_m (which equals 3 here).

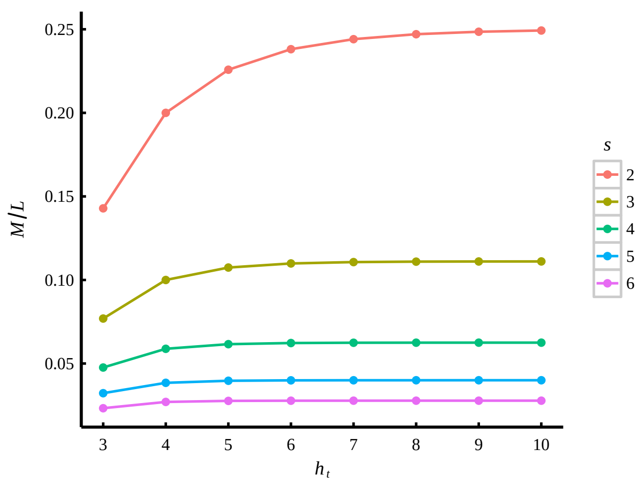

Figure 19: Asymptotic Behavior of the Management Fraction

This figure shows a plot of Eq. 17 for $h_m = 3$ and various s. As the total number of hierarchical levels (h_t) increases, the management fraction (M/L) within a firm grows rapidly, but soon reaches an asymptotic limit. This asymptote is a function of the span of control s, and the choice of h_m (the definition of where management begins).

Applying L'Hospital's Rule, we take the derivative (with respect to h_t) of both the numerator and the denominator in Eq. 19, giving:

$$\lim_{h_t \to \infty} \frac{M}{L} = \lim_{h_t \to \infty} \frac{-\log(s) \cdot e^{\log(s) \cdot (h_t - h_m + 1)}}{-\log(s) \cdot e^{\log(s) \cdot h_t}} \tag{20}$$

This simplifies to:

$$\lim_{h_t \to \infty} \frac{M}{L} \quad = e^{\log(s)(-h_m+1)} \quad = s^{-h_m+1} \tag{21}$$

Therefore, the asymptotic behavior of the management fraction depends only on the span of control, and our definition of management.

An Algorithm for Creating Hierarchies

The management model uses a power law simulated firm size distribution. In order to calculate the number of managers, each firm must be organized into hierarchical levels. I have developed the following algorithm to carry out this process.

Having selected a firm, we know its employment L and its span of control s; however, the total number of hierarchical levels h_t is unknown. To calculate h_t, we assume, for the moment, that the size of the top hierarchical level is one. Therefore, h_t must satisfy:

$$L = \frac{1 - s^{h_t}}{1 - s} \tag{22}$$

Solving for h_t gives:

$$h_t = \frac{\log\left[1 + L(s-1)\right]}{\log(s)} \tag{23}$$

Since h_t must be discrete, we round the solution to the nearest integer. My method is then to 'build' the hierarchy from the bottom up. If the bottom hierarchical level contains b workers, then L is defined by the series:

$$L = b\left(1 + \frac{1}{s} + \frac{1}{s^2} + ... + \frac{1}{s^{h_t - 1}}\right) \tag{24}$$

Using the formula for the sum of a geometric series, this becomes:

$$L = b\,\frac{1 - 1/s^{h_t}}{1 - 1/s} \tag{25}$$

At the moment, L is known but b is unknown. We therefore solve for b (and round the answer to the nearest integer):

$$b = L\,\frac{1 - 1/s}{1 - 1/s^{h_t}} \tag{26}$$

Once we have b, we can differentiate the firm into hierarchical levels by dividing b by powers of s (Eq. 24). Due to rounding errors, the sum of the employment of all hierarchical levels may differ from the original firm size L. Any discrepancies are added (or subtracted) to the base level to give the correct firm size. The number of managers M is then simply the sum from hierarchical level h_m to h_t.