UNIVERSITÄT DUISBURG ESSEN

Open-Minded

ifso institute for socio-economics

Jakob Kapeller
Stefan Steinerberger

Why are there so many power laws in economics?

uni-due.de/soziooekonomie/wp

ifso working paper 2025 no.50

Why are there so many power laws in economics? *

Jakob Kapeller[1] and Stefan Steinerberger[2]

[1]Institute for Socio-Economics, University of Duisburg-Essen and
Institute for Comprehensive Analysis of the Economy (ICAE), Johannes Kepler University Linz,
jakob.kapeller@uni-due.de

[2]Mathematics Department, University of Washington, Seattle, steinerb@uw.edu

March, 2025

Abstract

Power law distributions are ubiquitous in socioeconomic contexts. While their general properties are well understood, it is often less clear why they regularly appear in empirical data. What are the generative mechanisms leading to power laws, how do they arise in the real world? This paper aims to partly fill this gap by discussing two candidate mechanisms that appear especially relevant for understanding the emergence of power laws in socioeconomic contexts. We identify core formal properties and potential real-world equivalents of these mechanisms. In addition, we explore the relation of power laws to indirectly related concepts relevant in heterodox economics, like path dependence, cumulative effects, power asymmetries or non-ergodicity.

Keywords: power law distributions, inequality, generative mechanisms, wealth inequality, firm size

JEL-Codes: B52, D00, D30

*This paper has profited from repeated discussions with Jonas Dominy, Andrea Ottolini, Miriam Rehm, Till van Treeck, Rafael Wildauer and Jan David Weber. Special thanks goes to the organizers of the *Young Economists' Conference 2023* in Linz (Austria), where a first, preliminary draft of this paper was presented to a larger audience.

1 Introduction

Power laws appear in economic data with surprising regularity. Among the variables that follow a power law distribution are key economic indicators such as income (Pareto, 2014), wealth (Gabaix, 2009), firm size (Axtell, 2001) or returns to innovation (Scherer, 1998). Moreover, one observes power law distributions in a series of other relevant socioeconomic contexts. These include notions as diverse as the size of cities (Gabaix, 1999), the returns and price fluctuations on financial markets (Lux and Alfarano, 2016; Gabaix, Gopikrishnan, et al., 2003), the distribution of carbon emissions across nations and households (Akhundjanov, Devadoss, and Luckstead, 2017; Oswald, Owen, and Steinberger, 2020), the number of interactions in social networks (Quelle and Bovet, 2025), the distribution of citations in the academic literature (Price, 1965) or the relative popularity of videos on YouTube (Kamiyama and Murata, 2019). This ubiquity of patterns that follow a power law distribution inherently raises the question why these recurring phenomena emerge in the first place. In this paper, we aim to revisit two well-known candidate mechanisms for the analysis and explanation of such patterns, namely cumulative advantage and random multiplication. Both of these mechanisms can be embedded into the broader interdisciplinary literature on generative mechanisms of power law distributions and have strong roots in (different traditions of) heterodox economics and political economy.

We first introduce some basic observations on power law distributions; afterwards, the goal is to make the similarity between a series of classical arguments and models in heterodox economics and their connection to two specific generative mechanisms as accessible as possible. We then proceed by discussing the general idea of what a generative mechanism is in section 2. In remainder of the paper we turn to the two main generative mechanisms mentioned above and detail both, key formal aspects and real-world examples and applications related to these. In doing so, we try to build intuition be recurring on simple simulation models that try to capture the core idea of the underlying mechanism.

A power law distribution is a heavy-tailed distribution: there is a small (but not extremely small) likelihood to observe exceptionally large values, values that are orders of magnitude greater than the mean or median values of the relevant variables. As a consequence, power law distributions describe situations coined by huge inequality, where most observations are 'dwarfs' as they lie far below the mean of the series, whereas some, rare observation are 'giants' situated far above the mean. The term 'heavy tail' typically refers to the latter part of the distribution that covers these rare and outstandingly large observations. As power law distributions also appear outside social systems, there exists a large interdisciplinary discourse on power laws.[1] However, as will be shown, *within* social systems power law distributions often reflect existing, real-world power structures.

An important caveat is that in some cases the distribution of such variables does not follow a power law over the entire range of parameters but accurately mimics one for some part of the distribution (often the upper segment). This limitation, when it applies, also comes with a distinct advantage from a socioeconomic perspective: focusing on the upper end of the distribution represents a first step towards understanding why this upper end emerges in the first place and why it has such peculiar properties. In addition, in economic contexts focusing on the upper tail often implies a focus on a the part of the problem of interest, that is both, sizable and important as well as complex and potentially idiosyncratic, i.e. structurally different from the typical observation represented in one's data.[2]

Any discussion of power laws will profit from the introduction of a more concise formal definition. In general, we define a *power law* as a function of the form

$$f(x) = a \cdot x^b \tag{1}$$

where $a, b \in \mathbb{R}$ are real numbers. Power laws arise in various disciplinary contexts as they provide a natural way to express the *scaling* of different variables when the relevant relationships are non-linear. Well-known examples for such a use of power laws are found in biology (doubling size of mammals will roughly increase their weight eight-fold)[3], geometry (doubling the side of a square increases the area it occupies four-fold) or in economic theories of production (e.g. increasing vs. decreasing returns to scale)[4].

When speaking about the fact that some random variable follows a *power law distribution*, we often have a specific formal definition in mind that builds on the general formulation in equation 1. The definition is specific insofar as the relevant scaling concerns the relation between the absolute

[1]Especially helpful works in this context are Newman (2005) for a general overview on the subject, Nair, Wierman, and Zwart (2022) for a more technical introduction to power laws and Wildauer and Kapeller (2021) for a quick guide to their empirical estimation.

[2]In a more abstract vein, we could invoke the central limit theorem to remind us that in the absence of interesting dynamics, one always expects the ever-prevalent Gaussian distribution. In other words, the presence of a heavy tail implies the presence of an interesting underlying dynamical behavior.

[3]Biology has a large number of power laws not all of which have b an integer. Kleiber's law says that an animals metabolic rate is proportional to its mass raised to the $3/4-$th power. This means that increasing mass by a factor of $2^4 = 16$ increases the metabolic rate by a factor of $(2^4)^{3/4} = 2^3 = 8$.

[4]Our examples employ power scalings with an exponent of 3 (biology), 2 (geometry) and around 1 (typical values for production functions). These exponents can also be said to represent the degree of homogeneity of the underlying functional relationship.

values taken by some variable and their relative position. More precisely, some variable x follows a power-law distribution, if the scaling relationship between the absolute values of x and the ranks of individual observations (as given by the descending ordering of x) can be expressed by some power law akin to the examples given above.[5] Here, the condition $b < 0$ always holds to account for the fact that ranks follow an inverse ordering as higher ranks are associated with lower numerical values.

Variables that follow such a power law distribution (for some upper segment) can be characterized by some general qualitative properties (see Figure 1). As already indicated, such a distribution describes a population with many 'dwarfs' – many observations in x will have a similar, comparatively low, value – and a few huge giants. As a consequence – and in contrast to the Gaussian standard case – mean and median can diverge significantly from each other as only the mean captures the impact of giants. For similar reasons, the distance between ranked observations will strongly increase in the *heavy tail* at the right side of the distribution. There we find the observations with the highest ranks while on the distribution's left side the density between observations increases strongly.

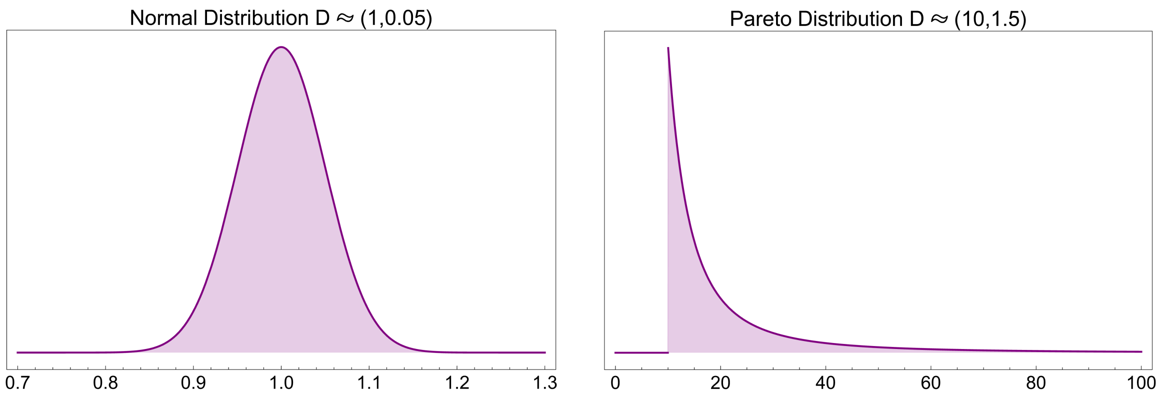


Figure 1: Left: A normal distribution. Right: A power law distribution.

Figure 1 shows a standard normal distribution next to a *Pareto distribution*, which is the most popular and well-known specification of a power law distribution in economics. Following Pareto (2014) it defines the relationship between ranks and observed values as

$$\mathrm{Rank}_i = N_i = c \cdot \frac{1}{x_i^{\alpha}}, \tag{2}$$

[5]In applied work ranks or relative positions can be expressed in different ways: while classic formulations often use absolute ranks in their formal specifications, these ranks can easily be transformed into more nuanced measuring of rankings. The most common transformation in applied work is to divide ranks by the number of observations to arrive at the *cumulative distribution function* (CDF) $F(x)$. From this the *complementary cumulative distribution function* $(1 - F(x))$ – the $CCDF$, sometimes also called the *survival function* – can be derived, which summarizes the relative position of observation on scale from 0 to 1. This transformation is also useful as a first step towards estimating the exact functional form of a supposed power law by means of OLS (Wildauer and Kapeller, 2021)

where c is a scaling constant and α represents the *shape parameter* of underlying Pareto type I distribution. This expression provides an intuitive way to more precisely assess how exactly the scaling relationship between absolute values and relative positions is conceptualized: as an increase in rank is given by N_i/N_{i+1}, it follows that what actually scales with α is the *relative distance* between ranks as expressed by $(x_{i+1}/x_i)^\alpha$, which also explains the differences in density of observations just mentioned.

To uncover a version of this formulation that is similar to equation 1, we can make use of the fact that the Pareto distribution is defined only for values greater than some $x_{\min}$. Hence, we can divide equation 2 by a version of itself that holds for observation the smallest observation n to arrive at a more general specification of the underlying power law distribution[6] in terms of the probability that a random variable exceeds a certain value x_i

$$\mathbb{P}(X \geq x_i) = \frac{N_i}{N_n} = \left(\frac{x_{\min}}{x_i}\right)^\alpha. \tag{3}$$

This formulation, which is identical to equation 1 for $a = x_{\min}^\alpha$ and $b = -\alpha$, represents a more general formulation of the *Pareto distribution* and has received wide-spread recognition, also outside of economics. In this context, it should be remarked that the *Pareto distribution* has a close cousin, if not to say twin, in the form of *Zipf's law* (Zipf, 1932), which is perhaps even better known across different sciences. In substance, Zipf's law is basically identical to the Pareto distribution, but posits the same relationship in inverse functional form, i.e. the rank N_i serves as the argument determining x_i. In other words, reasoning based on Zipf's law typically start from intuition of the form

$$x_i = \left(\frac{c}{N_i}\right)^{\frac{1}{\alpha}} \tag{4}$$

which is eventually identical to equation 2.[7] As a final preliminary remark, it should be noted that a log-transformation of equation 1 gives

$$\log(y) = b\log(x) + \log(a) \tag{5}$$

which illustrates the well-known results that applying the logarithm on both variables will return a linear function with a slope equal to b. This gives the characteristic log-log plot often spotted

[6] This formulation looks similar to the definition of a *CCDF* as given by $\mathbb{P}(X > x_i)$, but is not quite identical. See Wildauer and Kapeller (2021) for more details on this aspect.

[7] In the original application Yule asserted that $\alpha = 1.$, which leads to less general interpretation of *Zipf's law*.

in applied research related to variables that follow a power law distribution. Due to its linear features this transformation is also a typical starting point for the statistical estimation of power law distributions (Wildauer and Kapeller, 2021; Wildauer and Kapeller, 2022). Figure 2 exemplifies this by using a random sample from the distribution shown in the right panel of Figure 1.

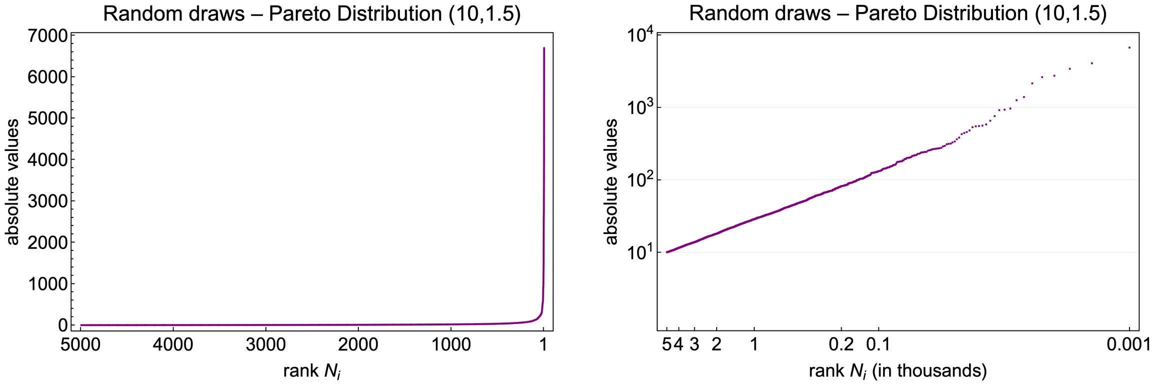


Figure 2: Left: Random draws from a Pareto distribution. Right: The same data in a log-log plot. The largest number is some orders of magnitude larger than the mean, median or any quantile.

2 Generative Mechanisms: a selective overview

Taking these preliminary considerations as our starting point we now turn to a more fine-grained discussion of how such power laws might arise in the first place as well as how specific assumptions about the processes underlying their genesis affect the general shape of the resulting distributions. In other words, we suggest studying generative mechanisms of power laws as a way of better understanding power law distributions in political economy. We aim to contribute to this discussion by synthesizing and integrating related arguments and models from different heterodox traditions, indicating ways to summarize key insights from that synthesis in simple formal models and illustrating how these mechanisms correspond to real-world cases and examples.

Given the prominence of power law distributions in real life data, it is not surprising that several interesting and informative reviews about generative mechanisms behind power law distributions already exist (Newman, 2005; Mitzenmacher, 2004). These contributions are often formulated in the language of the natural sciences, which makes it more difficult to apply the underlying intuitions to economic contexts or to interpret real-world patterns in the light of such generic mechanisms. Moreover, suggesting certain mechanisms for understanding real-world occurrences of power law distributions also brings further implications for theory, empirical methods and policy-advice that are not easy to anticipate. This latter argument gains additional weight given that (a) a variety of different generative mechanisms exist and (b) even within economics different mechanisms are applied in different fields of application.

Taking the classical contribution of Newman (2005) as a starting point, we find at least seven such mechanisms, including (1) the combination of exponentials, (2) taking inverses of quantities, (3) random walks, (4) the Yule process, (5) phase transitions, (6) self-organized criticality and (7) the multiplication of random numbers. The majority of these generative mechanisms is also used *somewhere* in economics but there is a little consensus and discussion across fields. While random walks and self-organized criticality pop up in finance, mainstream economic research relies mostly on the combination of exponentials for explaining the empirical occurrence of power laws (Jones, 2015). Indeed, this mechanism seems attractive from a mainstream perspective as it suggests that the exponent of the resulting power law distribution is stable, which implies a stable limit distribution of wealth and retains the possibility of a steady-state economy. In contrast, other generative mechanisms can be associated with structural instability, and ever-increasing inequality, which are unsatisfactory properties from a mainstream point of view. Hence, it comes as no surprise that the more specialized mainstream literature on the subject acknowledges the existence of these other approaches but typically amends them in a way that make the compatible with the standard mainstream vision of the economy (see Benhabib and Bisin, 2018, for various examples).

In contrast, we take a political economy perspective asserting that socio-economic provisioning systems can show different degrees of resilience or fragility – this is true for endogenous reasons as well as a consequence of how they are embedded in society and nature (Polanyi, 1957; Georgescu-Roegen, 1971). Taking such a view immediately suggests the importance of understanding *structural* sources of instability which leads to a focus different from the one applied in mainstream economics. One such potentially destabilizing force is found in distributional considerations as generative mechanisms quickly imply high and increasing levels of inequality, which may destabilize established social systems and institutions. At the same time such generative mechanisms often imply explosive growth for the highest strata of the distribution, which may, depending on the system in focus, cause instability if constraints are vulnerable, but essential, as in the case of planetary boundaries.

We focus mostly on distributional aspects as a core dimension for assessing the stability of socio-economic systems. In this context, openness for both – stable as well as unstable scenarios – is more in line with the empirical finding that distributional dynamics across history are not uniform: they strongly depend on technological and institutional conditions (Piketty, 2014; Pistor, 2020). Moreover, a political economy perspective is inherently interested in the dynamics of stratification and issues of path-dependent social polarization – these are sometimes diagnosed empirically (e.g. Balboni et al., 2021; Chetty et al., 2017), but only rarely related to generative mechanisms of power laws.

Following these intuitions, we focus on two modeling approaches that seem especially relevant from such a political economy perspective, namely cumulative advantage and the multiplication of random numbers. Our theoretical framing takes the core subject of inter-temporal wealth dynamics as a conceptual starting point. However, the underlying intuitions allow for a great variety of interpretations to that they can be legitimately projected on other, related areas of application that feature empirically observed power law distributions.

To allow for a concise, comparative illustration of the properties underlying these core intuitions and to ease the crafting of specific model variations or hybrid models we employ a shared notation and simulation framework across all applications. We start from a canonical representation of wealth dynamics, where the wealth of individual i at time t, we call it $w_{i,t}$, undergoes a change in time that can be written as

$$w_{i,t} = (1 + r_{i,t}) \cdot w_{i,t-1}. \tag{6}$$

Here, $r_{i,t}$ is the key variable governing wealth dynamics. It can be intuitively understood as a rate of accumulation and in what follows we use the terms 'growth rates' and 'accumulation rates' interchangeably. Interpreting r as an accumulation rate implies that r is also equal to rate of return, if the saving rate out of capital gains is equal to $s_\pi = 1$. A conceptual alternative is to interpret r as rate of return net of consumption out of capital income. By focusing on accumulation rates our core variable reflects a key property of capitalism as an economic system, which is then reflected in key economic variables following a power law distribution. However, similar accumulation dynamics can also be found in other contexts, especially in patterns of attention and interaction, which can be represented as network structures following some form of (non-linear) preferential attachment.

Our two key intuitions can be directly mapped unto this setup: first, in the case of cumulative advantage r depends on the relative status or 'fitness' associated with past wealth $w_{i,t}$. Secondly, random multiplication power laws can be seen to emerge from inherent randomness in r. As will be shown in section 4 even mild fluctuations in r can lead to astonishingly large differences in w_i if a reasonable amount of periods t has passed. To provide some intuition on the resulting dynamics we will repeatedly present simulation results dedicated to illustrate the impact of different (combinations of) generic mechanisms. Throughout, we will assume a population of $n = 5000$

individuals, a duration of $t = 1000$ discrete time steps and a starting wealth, where $w_{i,0}$ is equal or very close to 10. For all simulations, we test the plausibility of assuming a power law distribution for various cutoffs (Top1%, Top2%, Top5%, Top10 % and Top20 %).

3 Cumulative advantage as a first candidate mechanism

3.1 Formal basics and simulation models

Probably the simplest and most intuitive way to sketch a generic mechanism leading to a distribution for private wealth that follows a power law distribution is by assuming that those agents with greater net wealth will, on average, also achieve higher growth rates of wealth. This is typically achieved by assuming that richer households have a higher propensity to save (*differential saving rates* as in, e.g., Kaldor, 1955 or Pasinetti, 1962) and/or the ability to achieve higher rates of return (*differential rates of return* as in Kahn, 1959, Pasinetti, 1974, Pasinetti, 1983 or Piketty, 2014). It is easy to see that in such a constellation we will witness some form of emerging inequality as the *rich get richer* dynamics built into such a scenario will unfold without any dampening.

There exist a large number of modeling approaches that will produce an outcome in which the upper tail of the distribution follows a power law by assuming some form of 'Matthew effect', i.e. additional benefits for those, who are already better equipped. Obvious examples are given by replicator dynamics, (non-linear) preferential attachment models, models of technology adoption and diffusion that integrate positive feedback effects (Arthur, 1994) or by Piketty's 'first law of capitalism', where the latter suggests the existence of an ever-increasing capital income share as long as the return on capital surpasses economic growth rates (i.e., $r > g$; see Piketty (2014)). Notwithstanding, this great variety of different modeling-approaches, it should be emphasized that a basic similarity behind these models is that they assume some form of cumulative advantage and, hence, exhibit some form of power law distribution.

For choosing a setup that captures this core similarity in a way that is both, simple as well as general, we suggest using a stripped-down version of a multiplicative replicator equation (see Metcalfe, 1994; Stanley, 1998, for similar basic formulations). This setup gives rise to runaway growth for agents that pass a certain fitness threshold but leads to constant decay for others which encapsulates the key idea of cumulative advantage – that is, growth rates depending on existing levels – in a very simple framework. Building on equation 6 such a framework can be written as

$$w_{i,t} = (1 + r_{i,t}) \cdot w_{i,t-1} \qquad \text{with} \qquad r_{i,t} = \left(\frac{w_{i,t-1}}{k \cdot \bar{w}_{t-1}} - 1 \right) / m, \tag{7}$$

where $\overline{w_{t-1}}$ is the average wealth at time $t-1$, $k > 0$ is a parameter and $m > 0$ is a dampening factor. Note that apart from the initial distribution nothing here is random, the model is completely deterministic. This equation can be equivalently written as

$$w_{i,t} = \left(1 + \frac{1}{m} \left[\frac{w_{i,t-1}}{k \cdot \overline{w_{t-1}}} - 1 \right] \right) w_{i,t-1}. \tag{8}$$

The parameter $m > 0$ introduces a notion of scale into the process: a large value of m imply that the multiplicative factor is close to 1 and the system evolves changes slowly over time. A small value of m speeds up the system and leads to a faster change in the dynamics. The value k plays a more complicated role and it helps to start with $k = 1$. In that case, we are given a number of agents whose wealth is $w_{1,t-1}, w_{2,t-1}, \dots, w_{n,t-1}$. The ones whose wealth exceeds the average wealth $\overline{w_{t-1}}$ get a small boost in their wealth, where the exact size of the boost depends on how much they exceed the average. Conversely, the wealth of those below average will decrease over time because of negative accumulation rates. As shown in Figure 3, this setup will quickly lead to highly unequal distribution of wealth and produces a power law distribution (shaded areas) in the upper segment, if we allow for some variation in initial wealth.[8] For clarity, we provide two corresponding plots, one on absolute scale and another on relative (log-)scale.[9] More generally, the parameter $k > 0$ determines the exact threshold, where growth rates becomes either positive or negative and, hence, allow for regulating the share of the population that profits from over-proportional growth. For example, $k = 1.1$ would indicate that only agents whose wealth exceeds the current average by 10% would experience additional growth while $k = 0.9$ would mean that everyone whose wealth exceeds 90% of the current average will see a growth of wealth.

Such a model effectively captures the basic intuition of the 'rich-get-richer' dynamics associated with models of cumulative advantage. Variations of this general approach are easy to construct and can be used to further explore the main idea. A key notion of the above formulation is that some bifurcation arises between agents with rising and agents with decreasing wealth, which gives rise

[8] Specifically, we assume initial wealth is governed by $\mathcal{N}(10, 1)$.

[9] To prevent the creation of extreme outliers in long run simulation, it is necessary to cap extreme growth rates that over time will emerge at the upper and (potentially) the lower bound of the resulting distribution. For the outcome shown in 3 we capped growth rates at a maximum of 7%.

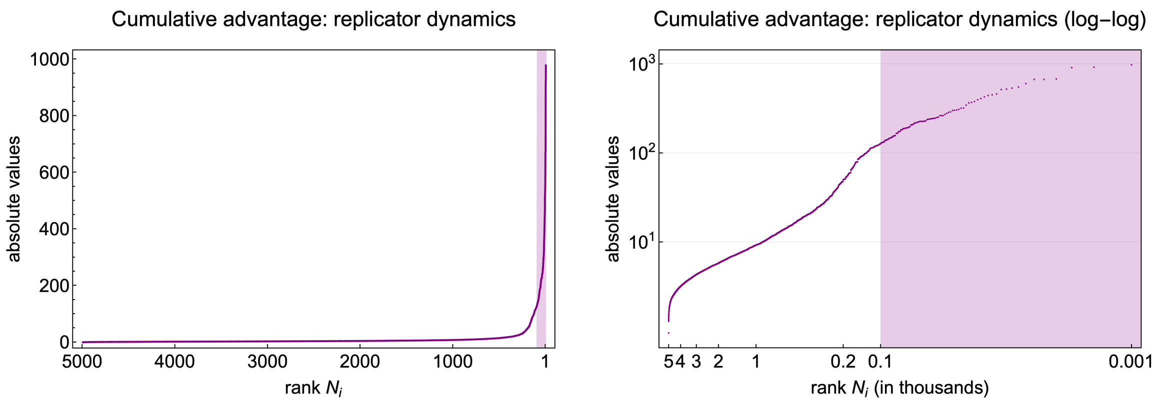


Figure 3: Left: Simulation outcome of a quasi-deterministic cumulative advantage model with $k = 1.2$ and $m = 400$. Right: The same outcome in a log-log plot. Shaded areas follow a power law distribution (top 2%).

to cumulative advantage in the first place. This not only creates large inequality overall, but also generates a Pareto distribution at the top, which mimics empirical observations. One can infer that the size of the segment that follows such a power law distribution is endogenously determined and can be moderated by choices on runtime as well as different settings of k.

Such a bifurcation of agents can also be reproduced with slightly different assumptions. An interesting variant is given by what we call the 'people have to eat' model. In this formulation we introduce a constant $c = w_0 \cdot \mu \cdot k$ that is subtracted from each agent's wealth in each period and that closely corresponds to the initial average growth rate of wealth, which is given by $w_0 \cdot \mu$ (with $\mu > 0$). Intuitively, c can be understood as subsistence costs expressed as a lump-sum that has to be supplied out of capital income.

To create a bifurcation scenario, again a minimal amount of heterogeneity is needed. Instead of a random assignment of starting values w_0 as used in our first application (Figure 3), we now apply minimal random variations on growth rates.[10] We can express this setup in a similar vain as in equation 7 by stating

$$w_{i,t} = (1 + r_{i,t}) \cdot w_{i,(t-1)} - c \qquad \text{with} \qquad r_{i,t} \sim \mathcal{N}(\mu, \sigma_{\text{small}}) \quad \text{and} \quad c = w_0 \cdot \mu \cdot k. \tag{9}$$

Due to our assumptions on c we would expect that, at least in the beginning, the cost of living is roughly matched by the typical growth of wealth. However, for people who were initially lucky,

[10]While this change takes us a little step towards models of multiplicative growth as discussed in section 4, it should be noted that the degree of randomness σ_{small} applied in this scenario is much smaller in this application as only a minimal amount of (random) heterogeneity is needed to generate qualitatively interesting results. Moreover, as it makes no qualitative difference for the outcomes, whether (random) heterogeneity is introduced in either starting values or growth rates it also helps to illustrate the robustness of the observed pattern.

the sustenance level starts having less of an impact, while for people who were initially unlucky, the sustenance becomes a bigger factor which introduces a bifurcation between winners and losers similar to equation 7. Again, the point of divergence can be moderated by using the parameter k, where $k = 1$ will cause the bifurcation to emerge exactly at the median, where values of $k > 1$ will drive this point up.

In our exemplary application we use the parameter specifications $w_0 = 10$, $\sigma_{\text{small}} = 0.001$, $\mu = 0.01$ and $k = 1.01$ (which gives $c = 0.101$ according to the definition given above). As in our first example we introduce a minimal amount of randomness to allow for endogenous development of cumulative advantage dynamics, which is structurally different from conceptualizing randomness as a source of power law patterns (as discussed in section 4. In addition, we introduce a minimum wealth of 1 to avoid negative values. Eventually, we arrive at the following outcome.

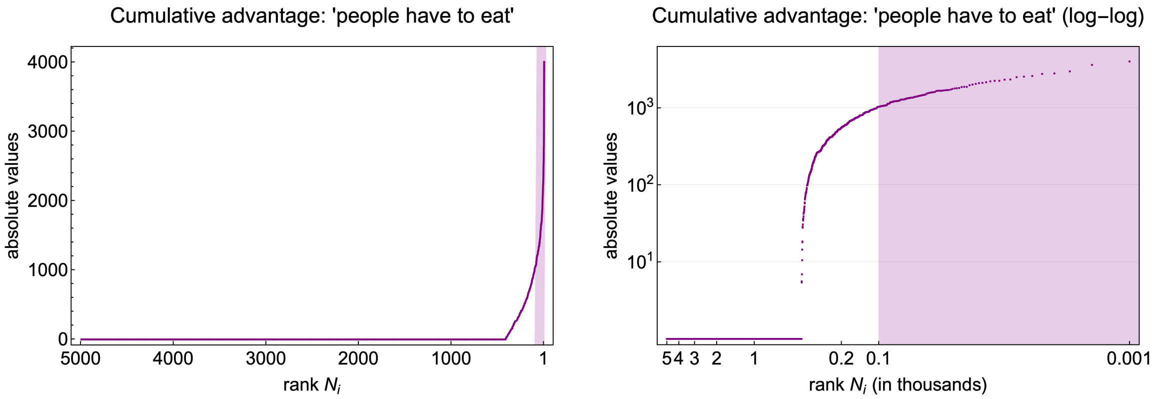


Figure 4: Left: Simulation outcome of a 'people have to eat' model. Right: The same outcome in a log-log plot. Shaded areas follow a power law distribution (top 2%).

The resulting divergence of levels of private wealth follows from the correspondence of c and $\mu \cdot w_0$, which renders the area in-between these two values a natural saddle-point. Agents that diverge from this point follow one of two different regimes: people for whom to return on investment is not enough to cover sustenance (which are essentially trapped to remain close to wealth 1) and people for whom sustenance costs are many orders of magnitude smaller than the return on investment. The second group gets to experience exponential growth of investment, the first group is stuck at the subsistence level. As Figure 5 indicates once wealth is below 10, there is typically exponential decay while, once wealth is above 10, we typically observe exponential growth. These path-dependent properties are also the reason, why it does not matter qualitatively, whether (random) heterogeneity is introduced via random starting values or random growth rates.

This interpretation obviously points to a bifurcation of individual trajectories of private wealth that are inherent in both models. Figure 5 below reproduces a small sample of these individual trajectories to allow for a broad comparison of the different bifurcation dynamic. It is clearly visible, that the constant, additive component c in the 'people have to eat model' creates sharper discontinuities than the simple replicator model is, where mean wealth as governing factor is an endogenous variable that to adjusts to reflect the overall growth of wealth in the model.

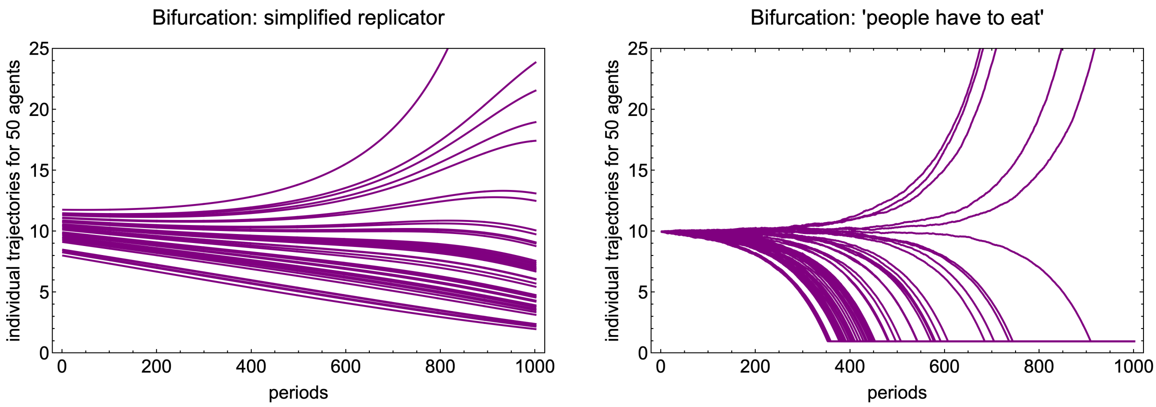


Figure 5: Left: Bifurcation process in the simple replicator model. Right: Bifurcation process in the 'people have to eat' model.

3.2 Applications to real-world examples and relations to established theoretical concepts

While the general principle of cumulative advantage can be represented by many different modeling approaches, most of these modeling approaches can be traced back to the core intuition of the Matthew principle ("For to every one who has will more be given", see Merton, 1968 or Rigney, 2010), sometimes also dubbed as 'rich-get-richer' dynamics (Newman, 2005). Such dynamics apply to variety of areas especially spheres concerning wealth and power (as in wealth rankings, network size or firm size) as well as fame and visibility (as in books sold, websites visited or citations received).

The models suggested here show explosive dynamics in the sense that inequality is ever-increasing over time. While such outcomes have often deemed unrealistic, long run time series on wealth inequality often exhibit bifurcation dynamics in the sense that wealth growth is negligible for large parts of the population while being rather pronounced for some upper segment of the distribution (Alfani, 2021), at least withing stable environments. As growth patterns in wealth follow cumulative dynamics wealth growth can even apply to a sizeable upper segment – as in the case of wealth growth in recent decades in Germany (Albers, Bartels, and Schularick, 2022), where a substantial share of the population already had some wealth to build on. Our modeling setup thereby captures the fact that such growth is most probably not proportional can accommodate variations in the size of the upper segment by different choices for k in the above setup. More importantly, these and related sources suggest that increasing inequality in terms of wealth is an endogenous feature of most recent economic systems and that different forms of cumulative advantage matter for this pattern (Alfani, 2021; Bowles and Fochesato, 2024). Hence, the question arises whether and which institution may modulate the impact of this endogenous to avoid economic inequality transgressing in to social instability (Scheidel, 2018; Bavel, 2016; Piketty, 2014). Conversely, it is suggested that increasing inequality repeatedly contributed to social tensions, revolutions, crises, war and other catastrophic events.

Key evidence for the importance of patterns of cumulative advantage in explaining power law distribution is supplied by anthropological research. It shows how the emergence of preservable and inheritable assets (that make multiplicative dynamics matter because endowments are not reset in every period but transmitted from the past) strongly contributes to increases in social stratification and creates more pronounced hierarchies (Mulder et al., 2009). Similarly, Smith and Codding (2021, p. 1) find that "the ability to control dense, predictable and highly clumped resource patches" is key for understanding the emergence of social hierarchies. Moreover, factors driving the emergence of inequality mentioned in anthropological research – like "storage, private property in land, and labor-saving innovation" (Bowles, 2023, p. 1195) – all relate more or less directly to cumulative advantage, which turns out to be a key tool for modeling dynamics of property, ownership, wealth and power.

In this context, one interesting feature of the models discussed in this section is that they not only exhibit the explosive aspects of wealth dynamics at the upper end of the distribution, but also indicate how agents can end up in poverty traps that makes any (further) accumulation of assets or reserves basically impossible. Such poverty traps are quite well documented in the empirical literature (Balboni et al., 2021) and basically refer to a lack of assets, that puts so much pressure in the income prospects of the respective agents that improving a desperate economic situation becomes virtually impossible (see also Mani et al., 2013). Our simulations capture these aspects very directly, also because they completely abstract from labor income, which renders social distinctions persistent once established.This general idea resonates with the concept of social classes, that take spatio-temporally specific forms and show corresponding variations in internal hierarchies. However, independent of their specific forms, the power asymmetries emerging from such distinctions then feed back unto the aggregate return structure. While this latter point is a commonplace in heterodox thought, it is also taken up in some mainstream models on financial returns and financial fluctuations (Gabaix, Gopikrishnan, et al., 2003).

The explosive properties of such simple models of cumulative advantage as given by equation 7 are often seen as an unwanted property.[11] This is motivated by both, theoretical consideration that like to posit a steady-state to ease long-run analysis, and normative commitments, that presuppose that the given system of provisioning is or must be stable in the long-run. While these lines are somewhat problematic in their own right as they aim to deduce an 'is' from an 'ought', many applications building on modeling approaches that leave room for cumulative advantage, include a built-in dampening factor that ensures some form of convergence for more pragmatic reasons. In models of technology diffusion, for instance, an exogenously given total demand typically serves as a dampening factor producing the famous s-curve of technology diffusion.

In classical or Harrodian models of economic dynamics these dampening factors are often given by to biological arguments as the rate of population growth as a key ingredient for identifying a 'natural' growth path. Similar, Wold and Whittle (1957) in a classic mainstream paper on the dynamics of wealth inequality, employ a birth-death process to act as a dampener for what is effectively a process of ever-increasing inequality. In other contexts, assumptions on relevant dampeners are located in the sphere of production where decreasing returns to scale (as in Ricardo, Malthus or Solow) or capacity constraints (as in post Keynesian economics; see Ederer and Rehm, 2018) are invoked to stabilize inequality and to derive steady state solutions for the corresponding models.

Finally, a key parameter determining model outcomes in our simple setups is k, which allows for calibrating the threshold where exactly individual trajectories diverge. In our application we set k to roughly emulate the 20%-own-80% pattern associated with the *Pareto principle* (Pareto, 2014). However, as the reader might have noted a consequence of this is that the results are highly sensitive to the choice of k – especially one quickly reaches constellations, where all individuals are experiencing growing or decreasing wealth. In our view, such extreme scenarios represent extreme case corresponding to existing ideal-type scenarios (e.g., ever growing wealth in a scenario, where there are no limits to private appropriation as in Locke and Laslett (1965) vs. ever-decreasing wealth as a scenario of ecological collapse).

In this spirit another advantage of k as a design parameter may lie in its capacity to rudimentary reflect shifts in historical conditions that impact on wealth formation and social mobility. A key

[11] This general verdict also applies to models of multiplicative growth discussed in Section 4).

example is given by the expansion of public infrastructures and the welfare state after WWII (Piketty, 2014). In our model setup, we can illustrate how the expansion of such basic public institutions affect the distributional dynamics of cumulative advantage, by assuming that such basic public institutions will ease the 'cost of living', which brings k closer to or even below 1 and, hence, reduces c (see Figure 6). In such an interpretation the main contribution of such public institutions is to shift the bifurcation point towards the middle of the distribution, which leads to the emergence of an hitherto unseen 'middle class' (see also Albers, Bartels, and Schularick, 2022, for comparison). However, the expansion of public infrastructures and services is by far not the example for such institutionally determined historical shifts in accumulation dynamics. Rather it could be argued that several forms of institutional 'checks and balances' have endogenously evolved over time as a response to dynamics of cumulative advantages, like jubilee days' and bankruptcy laws to cancel debts or interest bans and craft guild systems imposing limits to firm size to constrain endogenous concentration (see e.g. Graeber, 2011).

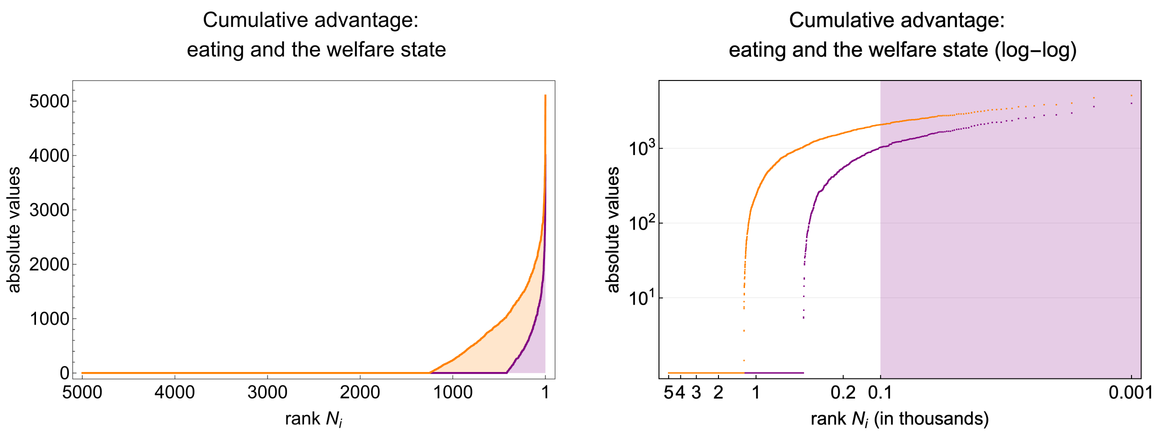


Figure 6: Left: Basic welfare state institutions (orange) in the simple 'people have to eat' model (purple) by changing $k = 1.01$ to $k = 1.005$. Right: Same as left in log-log scale. Both outcomes follow a power law distribution (top 2% for both scenarios; see shading in the right panel).

In such contexts, different measures of inequality may give conflicting evaluations of the change depicted in Figure 6, depending on how the respective algorithms weighs different parts of the distribution. The Gini coefficient, for instance, is smaller in the regime with basic public institutions, while many standard economic measures of inequality that focus more strongly on the dichotomies between upper and lower strata (like $\frac{\text{Mean}}{\text{Median}}$, $\frac{P90}{P50}$ or $\frac{P80}{P20}$), will show higher scores for the second regime. While the slope of the Pareto-tail is higher in the welfare state regime (which implies that wealth concentration is lower as the power law distribution is less skewed), the introduction of such a change will, on average, increase the values that are attained by the richest agents in the simulation.

As we have discussed the impact of cumulative advantage mainly against the backdrop of wealth dynamics it seems important to add that the general intuition presented here can be said to hold and apply in a variety of contexts, including, but not limited to, technology adoption and positive network externalities, cultural visibility and bandwagon effects, scientific authority and Matthew effects, social connectedness and preferential attachment or competitiveness in international trade and increasing returns to scale. It is a force that coins contemporary and past economic events, that also allows for explaining parts of what economic historians call the *Great Divergence*, especially why we have to expect a massive increase in international income disparities resulting from the long-term cumulative advantage associated with colonization and industrialization (Allen, 2009). The basic notion of increasing polarization explicated by both, this latter example as well as the simple models advanced in this section, is present in many socio-economic contexts, albeit to different degrees. While some of these contexts might require more careful considerations of relevant dampeners when compared to the case of wealth inequality, all those areas have repeatedly been associated with some forms of cumulative advantage that contribute to observed aggregate patterns in important ways.

4 Random multiplicative dynamics as a second candidate mechanism

4.1 Formal basics and a simulation model

We continue our discussion of generative mechanisms of power laws by focusing on randomness in multiplicative processes as a supposed source for power law distributions. We start our investigation by revisiting the classic formulation of random multiplicative dynamics incorporated in the *Gibrat model* of wealth reproduction (Gibrat, 1931). The Gibrat model draws on the generic formulation of wealth dynamics already presented in equation 6. In the simplest form of this model we complement this generic description with the assumption of random accumulation rates as in

$$w_{i,t} = (1 + r_{i,t}) \cdot w_{t-1} \qquad \text{with} \qquad r_{i,t} \sim \mathcal{N}\left(\mu, \sigma_{\text{large}}\right). \tag{10}$$

The behavior of such a process of accumulation clearly depends strongly on the random variable $r_{i,t}$. As will be shown in the standard scenario presented by the *Gibrat model* even mild fluctuations in r can lead to very large differences in w_i if a reasonable amount of periods t has passed. For simulation purposes we use our standard setup and assume that r is normally distributed with $\mathcal{N}(0, 0.05)$ (upper

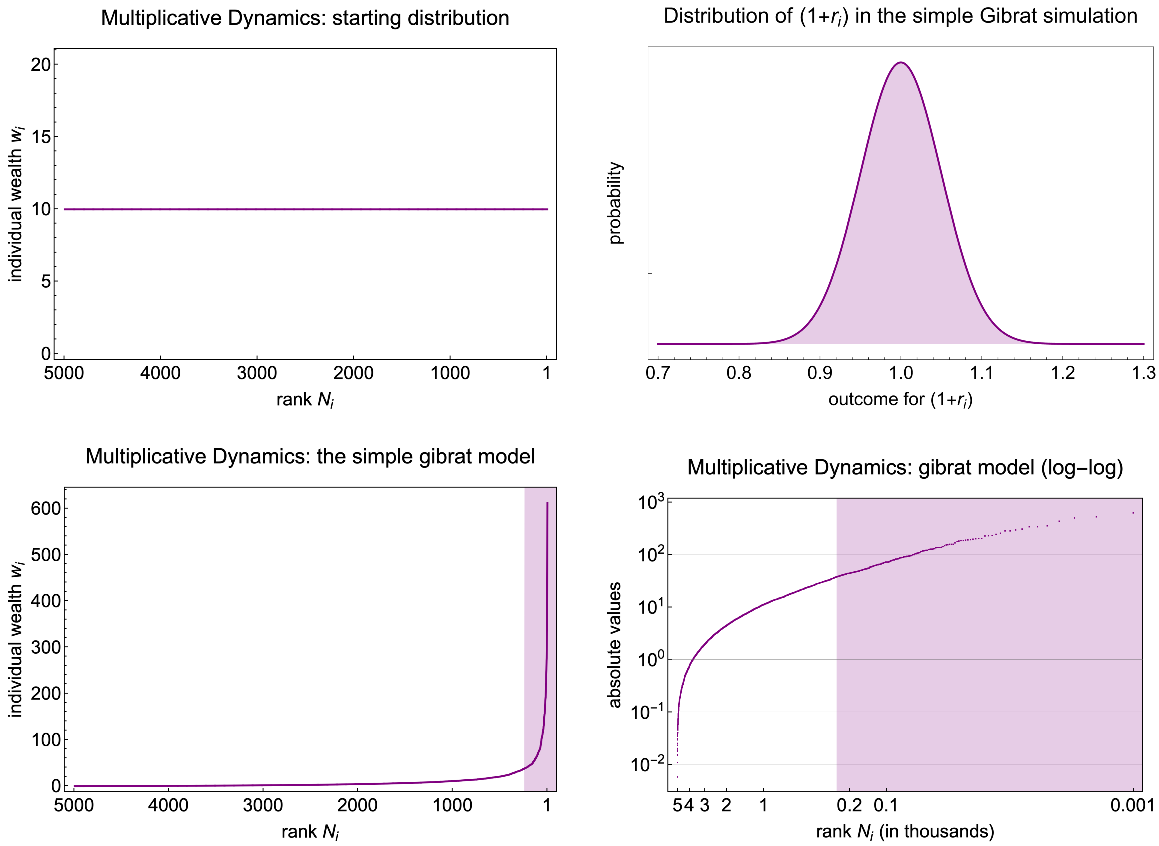


Figure 7: The lower panels will show the typical outcome of a Gibrat model after 1000 turns if random growth rates follow $\mathcal{N}(0, 0.05)$ (as illustrated in the upper right panel) and if all agents start with the same amount of wealth (here $w_{i,0} = 10\ \forall\ i$ as shown in the upper left panel).

right panel of Figure 7[12]). We find that the final outcome of this simulation is highly unequal: while a few agents have become rather wealthy, most agents see their wealth falling over time (see the figures in the lower panel of 7). Hence, a core intuition to be derived from the *Gibrat model* is that assuming multiplicative dynamics is, in many contexts, sufficient for extreme inequality to emerge from a situation of perfect equality – even if, the rules of the game are the same for every agent. And indeed, the resulting inequality follows a power law in the upper segments of the resulting wealth distribution (see the shaded areas in the lower panel of 7) for any large enough t. In a certain way, this general result is not too surprising as the non-ergodic dynamics associated with multiplicative growth are known to give rise to heavy-tailed distributions (Newman, 2005; Nair, Wierman, and Zwart, 2022). At the same this result merits explanation as it can be shown that the full, i.e. non-segmented, distribution that results from a *Gibrat model* follows a log-normal distribution. To see this clearly, go back to equation 10 and reiterate the rule to show that for every agent i

[12] As usual, we impose an artificial cutoff at 0 to avoid negative values – however, because of the tight concentration this is a very minor issue and not relevant in simulations.

$$w_t = (1 + r_t)w_{t-1} = (1 + r_t)(1 + r_{t-1})w_{t-2} = w_0 \prod_{s=1}^{t}(1 + r_s). \tag{11}$$

This looks complicated: a product of random numbers. Luckily, the logarithm turns products into sums and therefore

$$\log(w_t) = \log(w_0) + \sum_{s=1}^{t} \log\left(1 + r_s\right). \tag{12}$$

So while w_t might appear complicated, its logarithm is actually simply the sum of independent random variables. At this point, we can invoke the Central Limit Theorem[13] and deduce that, for large time t, we expect that

$$\log(w_t) \sim \mathcal{N}\left(t\mu, t\sigma^2\right) \tag{13}$$

where $\mu = \mathbb{E}\log(1 + r)$ is the expected logarithmic change of the process and $\sigma^2 = \mathbb{V}\log(1 + r)$ is the variance of the process. We emphasize that any single individual may gain or lose wealth in all sorts of different ways: at this point it makes sense to assume that a large number of individuals are independently subject to this dynamical system; then, the Central Limit Theorem applies. While we cannot say anything about any single individual, we can say something about what happens to the population at large, namely that *the logarithm of wealth is distributed like a normal distribution, i.e. wealth follows a log-normal distribution over the full population.* Unsurprisingly, this type of random variable is very well studied: if $X \sim \mathcal{N}(\mu, \sigma^2)$ is a Gaussian random variable, then $Z = e^X$ follows what is known as the lognormal distribution.[14]

There is an explicit expression for its density function: for two parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, the density is given by

$$p_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right). \tag{14}$$

It seems very clear that this distribution is not a power law $p(x) = a/x^b$. However, as indicated in Figure 8 it can behave like one over a suitable interval in its *upper segment.*

[13] This requires a minor degree of regularity on the random variable r_s. For example, we would like to ensure that $\varepsilon > -1$ to ensure that wealth stays positive. We will ignore these aspects for ease of exposition.

[14] This is the main reason why Steindl (1965) calls the lognormal distribution *Gibrat's law* (in contrast to *Pareto's law*).

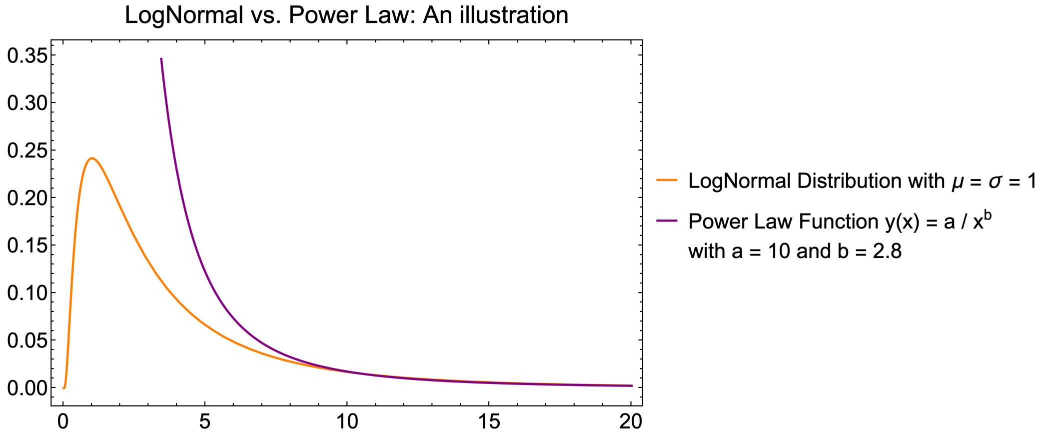


Figure 8: Comparing $p(x)$ for $\mu = 1 = \sigma$ (orange) following equation 14 and the power law function $y(x) = 11/x^{2.8}$ (purple). Both functions match reasonable well on the interval $[10, 20]$.

Put even more concisely, we can say that a function is a power law if and only if the logarithm is a linear function in $\log x$ since

$$y = ax^b \quad \Longleftrightarrow \quad \log y = b \cdot \log x + \log a. \tag{15}$$

We can check whether this is the case here and see that

$$\log(p_{\mu,\sigma}) = -\frac{1}{2\sigma^2}(\log x - \mu)^2 - \log(x) - \log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) \tag{16}$$

is clearly not a linear function in $\log x$. However it is a quadratic polynomial in $\log x$ and quadratic polynomials can be well approximated by linear functions on small intervals. In addition we observe that greater randomness (i.e. higher values for σ will translate into a flatter curvature making the assumption of log-linearity for some upper segment more plausible. These relationships explain why lognormal random variables can frequently emulate power laws in practice, especially if the upper segment of some distribution is concerned (see already Newman, 2005). The same type of argument also allows us to predict the long-time behavior. As is clear from the design of the process, one does not expect any type of limit to arise; for example, if the random variable r_i is always in the interval $[1.1, 1.2]$, then there will be exponential growth of wealth that just keeps increasing as time keeps increasing. However, the identity

$$\log(w_t) \sim \mathcal{N}\left(t\mu, t\sigma^2\right) \tag{17}$$

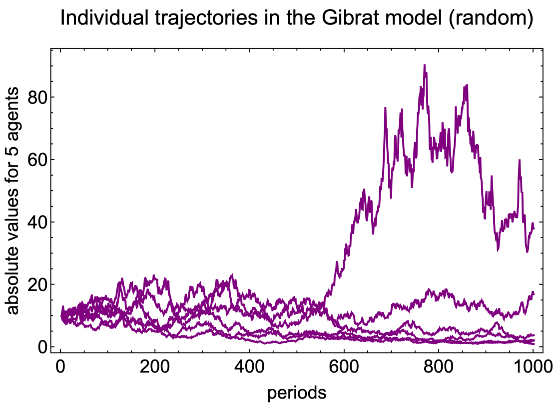
can be used to analyze the behavior at very large time scales. Using

$$\mathcal{N}\left(t\mu, t\sigma^2\right) = t\mu + \mathcal{N}\left(0, t\sigma^2\right) \qquad \text{and} \qquad \mathcal{N}\left(0, t\sigma^2\right) = \sqrt{t}\sigma \cdot \mathcal{N}\left(0, 1\right), \tag{18}$$

we deduce that wealth at time t is, approximately, described by

$$w_t \sim e^{t\mu} \cdot e^{\sqrt{t}\cdot\sigma\cdot W} \qquad \text{where} \qquad W \sim \mathcal{N}(0,1) \tag{19}$$

is a standardized normal random variable. We see that μ, the expected return, leads to general exponential growth while σ governs the disparity between different levels of wealth. If σ is small, then the amount of randomness is relatively small and it takes longer time (larger values of t) for randomness to impact on outcomes. Correspondingly, the shape of the resulting power law distribution is related to the key parameters in a clear-cut way, where both, rising σ as well as increasing the number of steps t induce greater levels of inequality, while the impact of μ is ambiguous for typical, that is, relatively defined, measures of inequality. The overall degree of inequality represented by such a system is determined mainly by the intensity of randomness (σ) as well as by its long-run stability (i.e. its age given by current t).



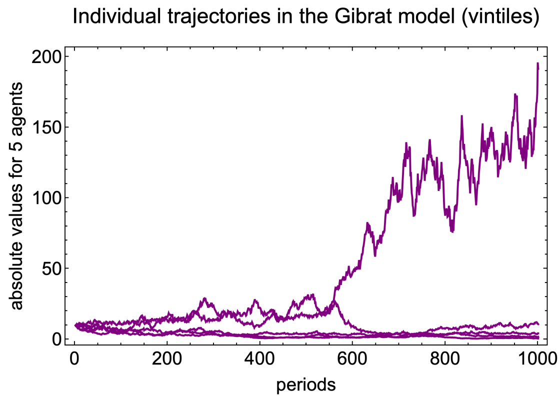


Figure 9: Left: The evolution of wealth for five random agents as sampled from the Gibrat model shown in Figure 7 Right: The evolution of wealth for those five agents that characterize the vintiles of the final distribution of the Gibrat model shown in Figure 7

The non-ergodic aspect of the underlying dynamics can be inferred from inspecting individual time series shown in Figure 9, which provide an exemplary illustration of the multiplicative random walks that are taken by individuals. These examples indicate that the typical individual time-series tends towards zero (which is indeed its formal limit). This creates a divergence between individual time-averages and the averages derived from analyzing the aggregate data at some period in time ('the ensemble'), where the latter is strongly influenced by a-typical outliers with very high wealth. This pattern signifies the case of broken ergodicity, i.e. a divergence between time- and ensemble averages, for the simple *Gibrat model* (Kirstein, 2019)). A key consequence is that typical measures – like averages measured at a certain point in time ('ensemble average') – are somewhat misleading (as, actually, no or only very few individual are close to the average) and that no stationary aggregate distribution exists in the limit, which has the key implication that the underlying economic system will never attain a steady-state.

In this context, it is noteworthy that the relationship between the *Gibrat model* and power law distributions in wealth is strengthened by assuming that individual wealth w_i will effectively stabilize at some minimum greater than zero for most individuals. Economically, this could be justified with reference to role of economic flows (e.g. wage payments) or, as in section 3.2, with reference to a basic social security system. Both would arguably help to stabilize wealth at some minimal level. Formally, assuming such a minimum level, however small, will act as a 'reflecting barrier' in the multiplicative process that effectively contributes to a heavier tail of the resulting distribution (Newman, 2005; Berman, Peters, and Adamou, 2021). Intuitively this stems from the fact that such a barrier provides a second chance for agents to achieve a winning strike associated with large increases in wealth. This would, in turn, base the *Gibrat model* on more realistic assumptions (by a rudimentary consideration of flows in addition to stocks) to also reach more accurate empirical predictions (the correspondence between tail wealth and a power law pattern increases).

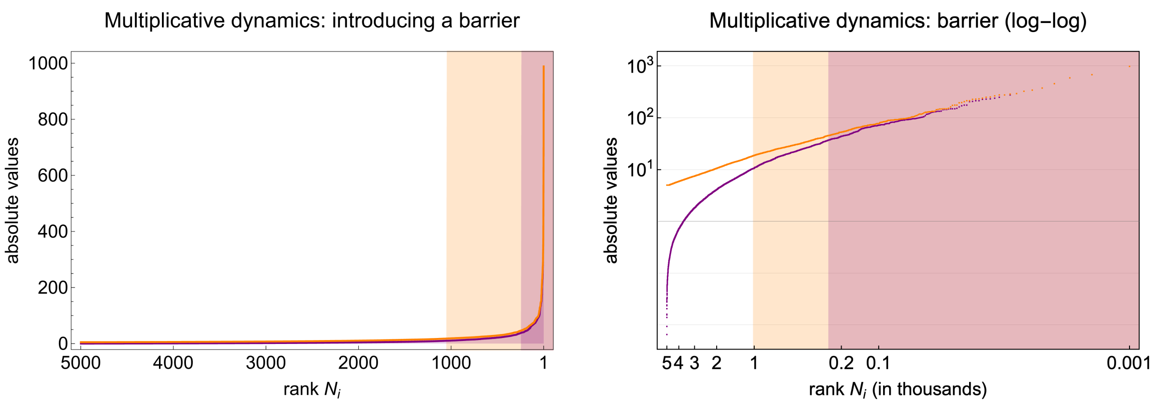


Figure 10: Comparison of the standard Gibrat model with (orange) and without (purple) a barrier at $w_{i,t} = 5$ (where all agents start with $w_{i,0} = 10$).

While this modification supposedly makes the model slightly more realistic in terms of both, assumptions and outcomes, the actual role of the barrier is to introduce a minimal form of differential returns, as those individuals very close to the barrier will experience a different return structure (with negative returns becoming less probable). However, skeptics will quickly note that this version of differential returns actually assumed here supposes – contrary to what is commonly observed and assumed – that the very poor, instead of the rich, have better prospects on their investments. In line with the observation that the so modified model leads to greater skewness the maximum of wealth observed will increase in the typical case[15], but to lower inequality overall as the barrier ensures a certain amount of wealth is also held by the bottom of the distribution.

4.2 Applications to real-world examples and relations to established theoretical concepts

In empirical terms, a key contribution of multiplicative dynamics is to provide a possible explanation for why large scales of inequality can quickly emerge from situation of relative equality. This can even happen in settings where the underlying dynamical evolution mechanism is treating every individual the same and is independent of the individual's wealth. It is noteworthy that this uneven distribution of resources arises despite the absence of adverse mechanism like exploitation or cumulative effects. In contrast, when asking for the emergence of such patterns in the first place, randomness is an often plausible but eventually incomplete explanation of long-run divergence. This is compatible with the anthropological research cited in section 3.2 as prior assignments of assets in accumulation processes can plausibly said to incorporate uncertainty and, hence, randomness in substantial degrees. In light of this we consider it a convenient finding that randomness is a sufficient assumption to endogenously explain the emergence of social classes which differ exactly with regard to their ability to control such resources.

A similar intuition can be advanced for the case of the case of endogenous concentration in markets, where the capturing of large markets shares by single or few firms can be delineated when modeling sales by means of random multiplication. The potential importance of randomness for understanding the growth and distribution of firms led Steindl (1965) to make the *Gibrat model* the starting point of his endeavors into understanding firm growth, which, e.g., undergoes random fluctuations in its population of customers. In such contexts, it could be argued that randomness alone will lead to oligopolistic market structures, which in turn creates the immanent opportunity for introducing strategic behaviors on the side of firms dedicated to capture certain markets, market shares or segments. In other words using randomness as a baseline model endogenously creates oligopolies, which,

[15] In the standard version of the *Gibrat model* many individuals quickly reach a w_i close to zero, which implies that they will not amass much wealth in the case of a 'lucky run' of successive good draws as their initial wealth has already shrunk so much. In the version with barrier, such a situation cannot occur, which means that there are more potential candidates that could gain from such a 'lucky run'.

in turn, endows some players with the ability and incentive to influence the further development of some market (Rothschild, 1947). This setup renders "real competition", that is "as different from so-called perfect competition as war is from ballet" (Shaikh, 2016, p. 263), the logical baseline case for further analysis.

Steindl (1965) asserts that firm size data conforms to what he calls *Gibrat's law*[16]. He also cautions that the simple model of multiplicative growth might not adequately reflect the nuances of the competitive process and, hence, requires some adaptions. Similarly, Champernowne (1953) uses the simple model of multiplicative growth as a vantage point for modeling the distribution of income thereby employing the modification that agents are grouped into different income classes. These cases suggest that the *Gibrat model* provides a solid shared basis to develop variations of these models that may account for the observed data patterns as well as the idiosyncrasies of socio-economic contexts. The latter argument is further reinforced by the observation that the classic preferential attachment model (Albert and Barabási, 2002) can be conceived as variant of the Gibrat model with strictly positive positive growth rates, that allows for the entries of new agents. Hence, preferential attachment can be interpreted as conceptualizing accumulation dynamics in (social) network contexts and it does so in a way that directly resonates with random multiplicative growth as a generic mechanism underlying the emergence of power laws.

A natural question is to what extent it makes sense to use the setting of multiplicative dynamics to model power laws without incorporating additional factors such as strategic behavior, network effects, increasing returns or other forms and sources of cumulative effects. Taking the intuition developed here as a starting point, the application of multiplicative growth to understand long-run dynamics is especially plausible when the underlying growth process is proceeding slowly (i.e. when $r_{i,t}$ is typically very close to zero). An example is given by city sizes for which multiplicative growth has repeatedly been suggested as a possible explanation (Gabaix, 1999). In contrast, in the setting of larger and more volatile rates of return multiplicative dynamics sheds light on the initial emergence of drastic inequality. Such dynamics can be assumed to play a role when novel assets or platforms emerge: one may think of the distribution of initial assets in local economies, the visibility on a newly created social platform or initial market shares in newly emerging sectors. These can be presumed to approximately follow random dynamics before additional factors influencing long-run dynamics kick in. Such an interpretation, where multiplicative dynamics are especially relevant for the formative phase of some phenomenon, is thereby akin to an analytical division repeatedly found

[16]Which he identifies with the lognormal distribution instead of the assumption of proportionate (growth) effects.

in theories of path dependency. In path-dependent phenomena one typically differentiates between an idiosyncratic and quasi-random path creation phase which is then complemented by a phase of path-dependency that shows much greater stability as growth rates become a function of the existing distribution.[17] Here, multiplicative dynamics are useful exactly because the absence of cumulative effects: they provide a simple minimal setting of plausible dynamics of inequality in capitalism. This has important implications for economic dynamics as it implicitly posits that economic systems will never reach a steady-state (a steady state would require the distributions for income and wealth to stay constant). In summary, the notion of multiplicative dynamics can be conceived as a possible paradigmatic example for understanding capitalist economies and accumulation processes as unstable, i.e. inherently crisis-prone, even in the absence of cumulative effects.

In terms of path-dependency, multiplicative dynamics also have implications for the study of social mobility, that is, the path-dependent character of social stratification as such. Notwithstanding the fact that the analytical limit is the same for all individuals (as already emphasized this limit is zero if $\mu \leq 0$), individual ranks or amounts of wealth can be predicted quite well by referring to past ranks or amounts of wealth: status differences are transmitted from past rounds and these differences vanish only very slowly, if at all, over any given finite time span. In practice this means that the stratification becomes more stable with increasing stability (i.e. with increasing t) and that economic hierarchies become subjectively more entrenched over time when considering that the lifetime of an individual has a fixed, absolute value. This more informal account of how multiplicative randomness relates to path-dependency and social mobility can be related to the formal question how past advantages (i.e., when $w_{i,t} > w_{j,t}$) impact future developments (i.e. the relation of wealth between i and j in some later period $t + z$). In this context it can be shown that player i will always enjoy an asymptotic advantage over player j as the probability that expected future values for $w_{i,t+z}$ are greater than corresponding values for $w_{j,t+z}$ is always greater than 50%.[18] Some

[17] The Polya urn model, the Yule model as well as Arthur's technology adoption model operate in a similar vain – see Section 4.3 for more details on this aspect.

[18] As we are interested in the ratio of wealth $w_{i,t}/w_{j,t}$, the value μ can be ignored. This simplifies the dynamics and we obtain $\log(w_t) = w_0 + \sqrt{t}\sigma \cdot \mathcal{N}(0,1)$ for both players i and j. The formal result that follows from this setup says that Player i is always more likely to have a larger wealth than j. However, provided sufficient amounts of time have passed, Player j can hope to approach or exceed the wealth of i with likelihood arbitrarily close to 1/2. Putting this down more succinctly leads to:

Theorem (Gibrat's advantage)**.** *Player i always retains an advantage*

$$\mathbb{P}(w_{i,t+z} \geq w_{j,t+z}) \geq \frac{1}{2}.$$

Moreover, we have in the asymptotic limit

$$\mathbb{P}(w_{i,t+z} \geq w_{j,t+z}) = \mathrm{CDF}\left(\frac{w_{i,t} - w_{j,t}}{\sqrt{2z}\sigma}\right),$$

where CDF *refers to the CDF of the standard Gaussian.*

This result can be related to the arguments of temporal persistence and observer-lifetime stated above: Assume applications, especially those situated in the econophysics tradition, also use multiplicative growth as a baseline model for the description of the long-run dynamics of the distribution of wealth. For instance, the standard perspective of econophysics on the distribution of wealth typically assumes some version of random multiplicative growth (Berman, Peters, and Adamou, 2021; Schulz and Milaković, 2021). An important example where this can be seen is in the respective approach to income distribution, which is assumed to consist of two blocks: an exponential bulk and a power law tail applying to top 1-5%. This mimics a labor vs. capital distinction where the latter can be derived from multiplicative dynamics (Silva and Yakovenko, 2004; Shaikh and Ragab, 2023). Such approaches can be justified in several ways: one view is argue that the dynamics of wealth follow a multiplicative growth process even if other factors – that are not represented in the simple *Gibrat model* — come into play. In that case multiplicative dynamics remains a valid baseline model than can be conveniently extended. Another, perhaps more pragmatic, perspective argues that multiplicative dynamics can easily be enriched by adding an additional additive term (leading to a reallocating geometric Brownian motion model).[19]

that $w_{i,j} = 1000$ and $w_{j,t} = 1$ and a standard fluctuation of 1%, say, $\sigma = 0.01$, then for player j to be reasonably likely to exceed i in wealth, we require, typically, something on the order of $k \geq 2.7 \cdot 10^{10}$ steps. This corresponds to a very long time period, much larger than anything typically observed in practice. To follow the derivation of the theorem use $X, Y \sim \mathcal{N}(0,1)$ to denote two independent standard random variables. It remains to analyze

$$\mathbb{P}\left(w_{i,t} + \sqrt{t}\sigma X \leq w_{j,t} + \sqrt{t}\sigma Y\right).$$

This likelihood can be reformulated as

$$\mathbb{P}\left(w_{i,t} + \sqrt{k}\sigma X \leq w_{j,t} + \sqrt{k}\sigma Y\right) = \mathbb{P}\left(\frac{w_{i,t} - w_{j,t}}{\sqrt{k}\sigma} \leq Y - X\right).$$

At this point we use a convenient fact: if X, Y are two independent standard random variables, then $X - Y = \sqrt{2} \cdot Z$ where $Z \sim \mathcal{N}(0,1)$ is yet another Gaussian random variable. Therefore

$$\mathbb{P}\left(\frac{w_{i,t} - w_{j,t}}{\sqrt{k}\sigma} \leq Y - X\right) = \mathbb{P}\left(\frac{w_{i,t} - w_{j,t}}{\sqrt{2k}\sigma} \leq Z\right)$$

[19]This notion has been taken up in a recent modification of the Gibrat model by (Berman, Peters, and Adamou, 2021), who propose a reacollation modification. Their model is continuous in time. We will, for simplicity of exposition, present a time-discrete model that shares all the essential features. We consider N agents each of which have wealth evolving according to a Gibrat model $w_{i,t} = (1 + r_{i,t}) \cdot w_{i,t-1}$ The Berman-Peters-Adamou model is based on the idea of a centralized pot such that each person pays a positive proportion $\tau > 0$ of their own wealth into a pot while receiving the same proportion $\tau > 0$ from the centralized pot. This updated model takes the form

$$w_{i,t} = (1 + r_{i,t}) \cdot w_{i,t-1} - \tau \cdot w_{i,t-1} + \frac{\tau}{N} \cdot \sum_{j=1}^{N} w_{j,t-1}.$$

The effect of this modification is as follows: if an individual agent has a wealth which by far exceeds the average wealth, they will distribute the surplus over the other agents. Conversely, agents with wealth below the group average will receive an additional income. This mechanism leads to a stabilization of wealth as long as $\tau > 0$ in the sense that wealth is more tightly concentrated around the group average (to an extent governed by the size of τ). The model shares some similarity with the 'people have to eat' model discussed in 3.2 and with the Ornstein-Uhlenbeck process.

4.3 Variations of multiplicative growth: persistent asymmetries

The models discussed in preceding chapters build on either cumulative advantage (in section 3) or random multiplication (in section 4) as the dominant generative mechanisms for power law distributions. A key difference is that the former mechanism delivers persistent asymmetries (i.e., when some agent i surpasses another agent j in terms of wealth, i will always stay richer), whereas the latter makes, as has been shown in footnote 18, persistence of asymmetries more probable, but not definitive. Moreover, for the Gibrat model relative asymmetries at any point in time are independent of the average growth rate (i.e., they only depend on variance and runtime), whereas in cumulative advantage model the average growth rate does impact on relative dispersion. Against this backdrop, it is natural to ask whether and to what extent this gap can be bridged. While in general different routes for doing so seem possible, in what follows we restrict ourselves to showing how this gap can be addressed by minor modifications of the Gibrat model. Proceeding this way leads also aligns well with the aim of theoretical integration as it spotlights two comparably simple candidates for such a conceptual bridge, that are already part of the broader heterodox canon. More specifically, these candidates are equivalent to two additional classic models in political economy and heterodox economics: the *Polya urn model* and a simplified version of the *Yule model*.

The first model can be found in (Eggenberger and Pólya, 1923) and has been brought to prominence in heterodox economics by various authors who proposed it as a useful null model for understanding path-dependency (see e.g. Arthur, Ermoliev, and Kaniovski, 1987; Dosi, Ermoliev, and Kaniovski, 1994; David, 2007). There are also several modifications to better align the model with the core idea of path dependence (Arthur, 1994; Page, 2006). The second model originates in work of the British statistician Udny Yule (Yule, 1925), which was carried out around the same time as the work of Eggenberger and Pólya (1923) and originally designed to model the growth of species over time. It has been brought to prominence in heterodox by, among others, Herbert Simon and Joseph Steindl (Simon, 1955; Simon, 1960; Steindl, 1965). Both models can be neatly illustrated with urn models.

Somewhat surprisingly, these two models are often conceived as unrelated to what has been discussed so far. However, both models can be understood as a variations of the Gibrat model. These variations use only positive growth rates, but retain the basic probabilistic notion of the Gibrat model and arrive at results of persistent asymmetry and dominance, without explicitly requiring notions like cumulative advantage. Formally, both models are equivalent simple modifications of the Gibrat model, where either the average growth rate (Polya) or its variance (Yule) decrease over time. This observation is highly relevant for the purpose of the present paper as it points to the deep conceptual connections and interdependencies of classic approaches to modeling generative mechanisms of power law distribution in heterodox economics and political economy. While the modeling approaches discussed in this paper are applied in many different branches and separate traditions of political economy research, their common denominators frequently remain unseen.

Before emphasizing how this similarity can be described on a formal level, we first explain and explore some core properties of the Polya and the simplified Yule model. We then assess the similarity of the three models under consideration employing a general formulation, that allows for explicating how different assumption on exact formation and process underlying growth dynamics modulate between the distinct patterns associated with these models.

In the standard *Polya urn model* we envision an urn containing balls of different colors. In the usual setting we have a relatively small number of colors, maybe only two. We close our eyes and draw a random ball from the urn. We open our eyes, find out what color the balls has and then return this ball together with a new ball of the same color back to the urn. In the simplest case, which is also covered in the upper panels of Figure 11, we have two colors and and will add one ball per turn dependent on which color has been drawn from the urn.[20] Hence, the probability of growth in a certain period directly depends on the existing stock relative to the total number of balls. There is also an intrinsic notion of stability that comes into effect: if we already have a large number of balls, say 60% of them being red and 40% of them being blue, and we draw 100 additional balls at random, chances are that around 60 of them are red and around 40 of them are blue: this means we add 60 new red balls and 40 new blue balls, thus preserving the overall ratio. Conversely, if the urn contains a small number of balls drawing an additional ball at random can quickly lead to strong shifts in probabilities for the next turn. The attentive reader may note that these dynamics correspond closely to those associated with preferential attachment (Albert and Barabási, 2002) – and, indeed, for a simple setup with two attracting nodes, an aggregate growth rate equal to 1 a preferential attachment model will give the same results as the simple Polya model, if we assume that newly entering nodes will not become attractors themselves.

The Polya urn model is arguably one of the most well studied stochastic processes and is very well understood (see, for example, Goldstein and Reinert, 2013; Mahmoud, 2008). For example, if we start with a red balls and b blue balls, then *the ratio of red to blue balls* is a random variable that follows a $\text{Beta}(a, b)$ distribution. A particularly nice case is $a = 1 = b$ in which case $\text{Beta}(1, 1)$ is

[20] As indicated, there exist many variations of this basic setup, which not only relate to the initial number of balls and colors and their distribution, but can also extend to varying the exact replacement mechanism (Arthur, 1994; Page, 2006).

simply the uniform distribution in $[0, 1]$: every asymptotic ratio is equally likely to appear. For general a, b, the distribution is more complicated to describe. As indicated, for larger initial values a, b the distribution becomes more stable, the initial ratio a/b is more likely to be approximately preserved.

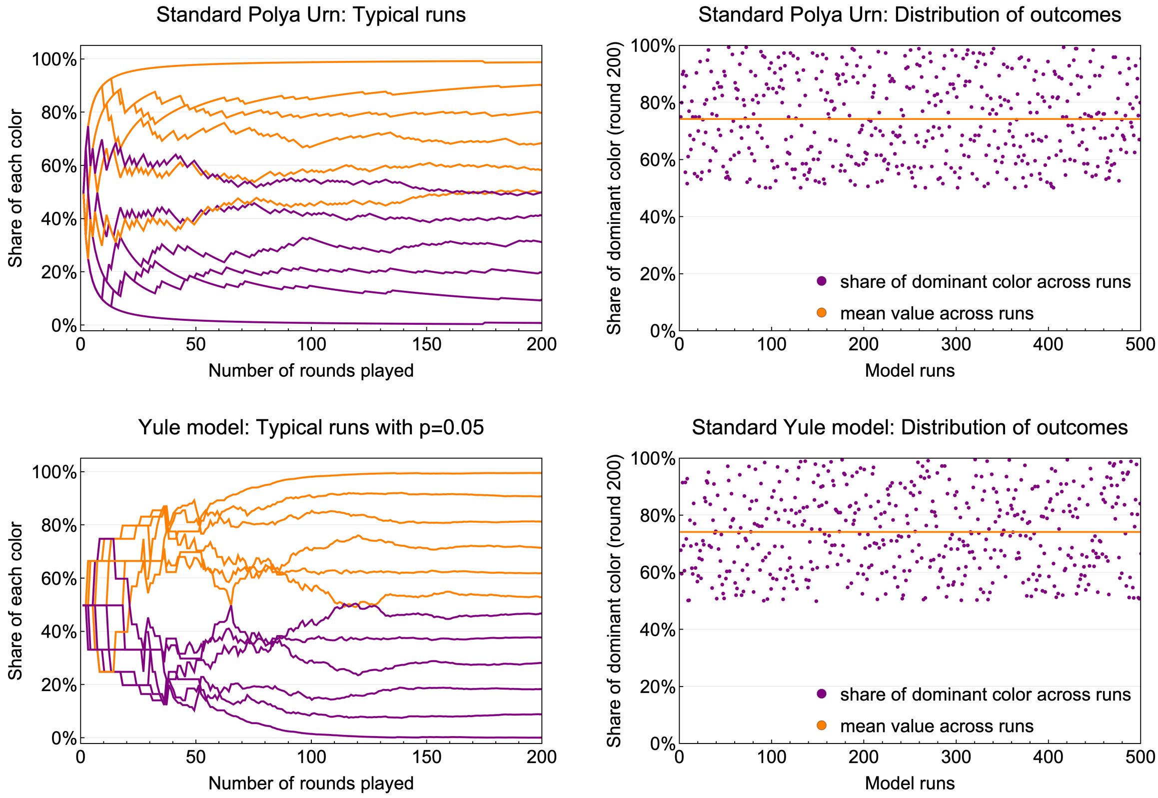


Figure 11: The upper panels show typical outcomes (left) as well as the distribution of outcomes across different runs (right) of the simplest Polya model (with two different colors), while the lower panels illustrate the same for the simplest Yule model (again with $n = 2$ and reproduction probability of $p = 0.05$ for each entity in each period). Typical values on left panels are chosen to represent the full spectrum of long-run outcomes for orange as a dominating color.

In this exposition we apply a simplified version of the original Yule model and again refer to an urn containing balls of different colors to illustrate its properties. In this setup, independently of how many balls of each color we have, at each point in time each individual ball creates a perfect duplicate of itself with likelihood $0 < p < 1$ (where p is the same for each ball). In Figure 11 we assume that $p = 0.05$ to illustrate some of its properties.[21]

By referring to the graphs we quickly note that this simplified Yule model exhibits some properties

[21] The key difference of this simplified Yule model to the original version is that the latter also allowed the emergence of new colors to also represent the birth of new species – an element that is deliberately suppressed in our simplified version.

that are similar to the Polya-urn model: it is again self-stabilizing in relative terms as the ratio of red and blue balls will tend to a (random) limit. If we have a large number of red and blue balls, then their populations in the next step will be approximately 5% larger but their ratio remains preserved. Outcomes for each colors in relative terms are approximately uniformly distributed across the range $[0, 1]$ for the examples given in Figure 11, which is another shared property between these two approaches. However, while for the Polya model this is a universal feature, in the Yule model this only holds more small p. For higher p boundary results become less probable and results will cluster in the center showing more balanced outcomes overall. Moreover, absolute growth is unbounded in the simplified Yule model , while in the simple Polya setting, which grows constantly by one ball per period, p emerges endogenously. Due to this latter feature the Yule model allows for higher overall rates of growth, while relative growth is diminishing over time in the Polya model.

The Yule model has achieved some visibility in economics due to the works of Herbert Simon (Simon, 1955; Simon, 1960) and Josef Steindl (Steindl, 1965), where the former proposes several variants of the model, including one version that is identical with the simplified Yule model employed in this chapter. Both authors use the underlying model as a general principle for explaining the emergence of heavy-tail distributions (Simon) as well as a way of thinking about firm growth and the evolution of market shares (Steindl). The resulting pattern can also guide an analysis that explains the distribution of income with recourse to emergent hierarchies in firms of different size (e.g. Fix, 2019). The defining characteristic of these two models is that the impact of randomness is reduced over time so that hierarchies can emerge and become more rigid and pronounced over time. Both models take different paths to achieve such persistence: in the *Polya urn model* total growth is constant so that the average growth rate decreases with rising stocks, while the *simplified Yule model* exhibits a decreasing variance over time due to the segmentation of the stochastic process, where the reproduction probability is evaluated sequentially for each element of the underlying population.

Notwithstanding this difference, there are remarkable similarities between the patterns produced by both models. The simplest case is that of two colors, two populations or, correspondingly, the wealth of two households or the size of two firms. Both variants show a high dispersion of results with rather egalitarian as well as highly concentrated outcomes (see left panels in Figure 11). However, this finding implies that the average share of the dominant color is about 75% (as indicated in the right panels of Figure 11, which means that such processes consistently produce a dominant fraction and, hence, stratification even among only two players. This basic property mimics most other results presented in this paper – namely, the endogenous emergence of asymmetry, dominance or inequality – although with much lower granularity as we only observe the evolution of two players.

Going back to the question how exactly these small variations in assumptions can be aligned in a general framework and, at the same time, produce such distinct patterns, we examine a general formulation that allows for capturing the similarity between the Gibrat model, the Polya model and the (simplified) Yule model. We rewrite the canonical representation from equation 6 as

$$w_{i,t} = (1 + r_{i,t}) \cdot w_{i,t-1} \qquad \text{with} \qquad \begin{cases} r_{i,t} \sim \mathcal{N}(\mu, \sigma) \\ \mathbb{E}(r_{i,t}) = \mathbb{E}(r_{j,t}) \ \forall \ i, j. \end{cases} \tag{20}$$

The above formulation implies that all three models approximately follow this general expression, which features an expectation value for the growth rate that is the same for all agents. For the Gibrat model this is straightforward as $\mathbb{E}(r_{i,t}) = \mu$ holds for all agents i over all periods t, regardless of their current state of wealth. For a Polya model, where growth has an absolute upper bound, expected growth rates are also equal for all agents in each period as the reproduction probability directly corresponds to the total counts of each color, which gives the same ratio for every color. As absolute growth is constant across periods the expected growth rate $\mathbb{E}(r_{i,t})$ diminishes over time, which reduces variation and induces persistence. As actual realizations in the Polya model are based on a Bernoulli-distribution (as returns always accrue to only one party), the formulation in equation 20 draws on the fact that we can approximately represent the latter by a normal distribution for large n. Finally, the Yule model is unbounded and, hence, the expected growth rate again stays constant and can be shown to be equal to the reproduction probability, i.e. $\mathbb{E}(r_{i,t}) = p$. However, the growth process unfolds differently as compared to the Gibrat-model because it is based on a partitioned, binary process – "partitioned", because there returns for each individual are evaluated sequentially before being added up to total population growth, and "binary", because possible outcomes for single evluations are 0 and 1, i.e. again follow a Bernoulli-process. In practice this implies that realized growth rates differ strongly from expected values for small n and growth is more erratic. However, as n grows, growth rates become more stable and more concentrated around their mean. In other words, the variance associated with $(r_{i,t})$ endogenously declines with rising n so that, again, existing asymmetries become persistent.

These two models build on slight variations of the standard description of random multiplication to arrive at patterns, where past outcomes create persistent contemporary asymmetries that impact all future outcomes. These models emulate a key feature of unconstrained cumulative advantage models without inheriting their explosive properties in terms of relative inequality. However, the connection of these models to power law distributions is still unclear. This is mainly because the canonical simplification to look at only two players (or two colors) constrains our perspective as it becomes difficult to speak about an overall distribution of outcomes across player (i.e. colors). To overcome this conceptual gap we increase the number of players (or colors) to $n = 1000$.

The results of this experiment are shown in Figure 12, which shows quite different dynamics for typical model runs (left panels). Both models lead to final distributions of balls over colors that eventually follow a power law at the top (see shaded areas in the right panels). Differences in typical trajectories can be traced back to the (un)boundedness of the aggregate growth rate – while in the Polya model total growth always equals one, for Yule this is given by p times the total number of balls (regardless of their specific color). It is obvious that the latter will at some point lead to explosive dynamics, even in cases where p is extremely small, which explains the visual and numeric differences between the left panels in Figure 12.

This structural difference notwithstanding, one is led to believe that both processes will lead the upper segment of the resulting distribution to follow a power-law. While this relationship is quite robust for the Yule model it is less so for the Polya model, which, due to its slow overall growth rate, requires more time to eventually arrive at such an outcome for a larger number of players. Hence, the result is sensitive to the exact choices for n and t in the Polya model, but much less so in the Yule model.

As in the previous chapters we tried to obtain or employ formulations that express the core intuition of the underlying mechanism as clean and simply as possible. By suggesting a simple scenario for a large-N application of the Polya and the Yule model, we show how the similarities between these setups and the Gibrat model map unto the joint feature of producing power law tails. At the same time these models indicate how small changes in the exact growth mechanism allow for explaining different constellations potentially emerging from such dynamics in the real world, i.e., they allow for distinguishing explosive and stable, slow and quick, persistent as well as flexible dynamics. A common foundation of these approaches is thereby found in the principle of random multiplication and the associated emergence of power law tails in the resulting distributions. We argue that that mixture of unity and diversity is often overlooked, but makes a valid argument to consider random multiplication as key analytical approach in heterodox economics and political economy.

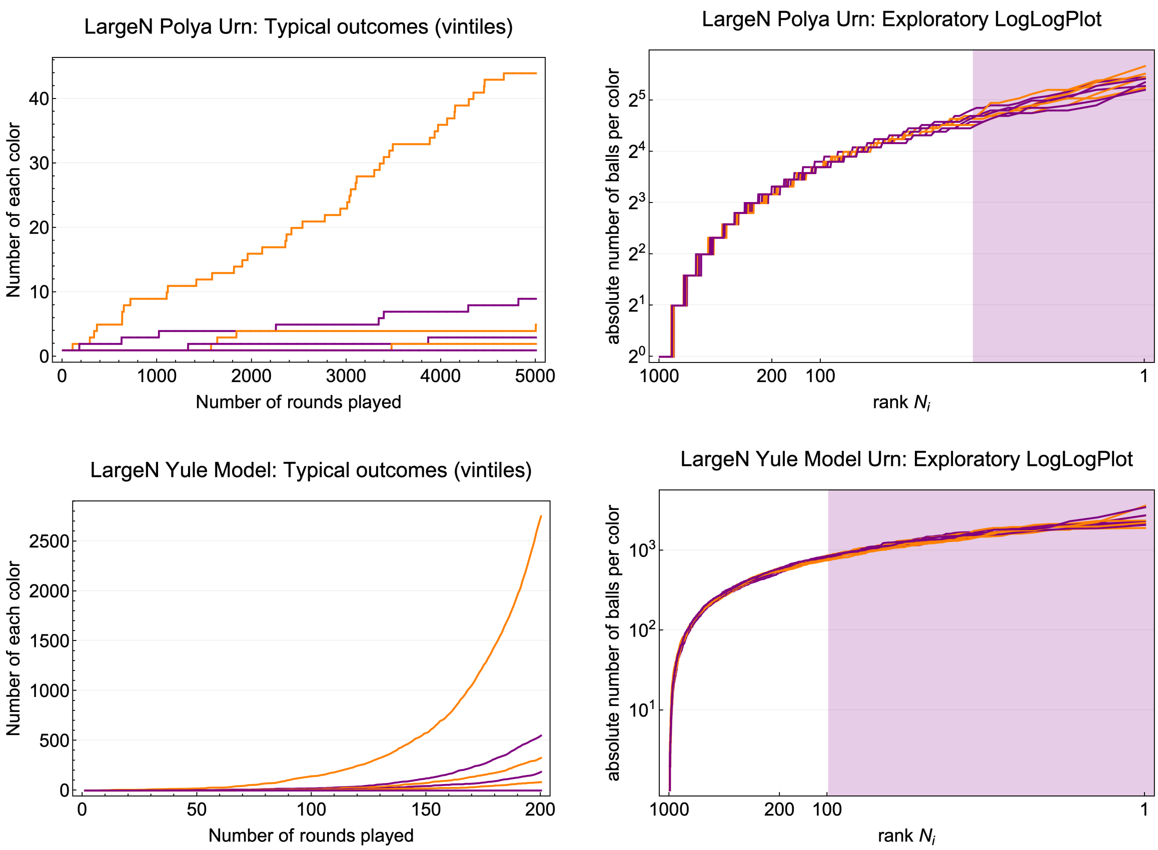


Figure 12: The upper panels show exemplary model runs (left) for a largeN Polya model (with number of colors $n = 1000$) and the final distribution after 5000 rounds for ten different runs (right) in a log-log plot. The lower panels show the same for a largeN Yule model with $n = 1000$ and reproduction probability of $p = 0.03$ for each entity in each period. The left panel provides exemplary model runs, the right panel showing the final distribution after 200 rounds again).

5 Concluding thoughts

In this paper we aimed to show that two generic mechanisms associated with the emergence of power law distribution stand out from a political economy perspective: cumulative advantage on the one and randomness in multiplicative dynamics on the other hand. Both of these mechanisms have a wide area of application and can help to better rationalize diverse real-world patterns in economics and beyond – the repeated occurrence of power law tails in empirical distributions being the key example. Nonetheless, in this paper we show that these mechanisms do not only relate to statistical properties in the more narrow sense, but also allow for illuminating related phenomena, like the existence of social classes, structural power asymmetries and persistent inequalities, the emergence of path-dependence in technological trajectories, attention patterns and economic development or the distinction between explosive and stable socio-economic dynamics. The fact that many different analytical applications in political economy show some explicit or implicit connection to these two forces shaping economic dynamics thereby underscores the prominence and, hopefully, relevance of these generic mechanisms.

Starting from this assessment a key pedagogical question is how to better exploit this shared analytical core in terms of applied analysis, shared language and pedagogical innovations. In terms of conceptual advancement it seems natural to ask for and explore different ways to combine these two mechanisms to allow for a genuine interaction between random multiplication and cumulative advantage. Pursuing this line of thought can build on the simple models presented in this paper and introduce modifications to look for plausible and empirically suitable intermediate formulations that might foster a better understanding of how these two forces interact in practice. Assuming that multi-causality pervades the subject matter of political economy (and the broader social and economic sciences) and seems plausible to assume that some form of such an interaction will take place in most practically relevant contexts. A potentially exemplary line of research that explicitly introduces a continuum of hybrid models is found in the realm of preferential attachment, where the distinction between linear and super-linear models of preferential attachment allows for such an explicit integration (Kunegis, Blattner, and Moser, 2013). Here, the former variant more closely corresponds to the Gibrat model, while the latter allows for introducing notions of cumulative advantage leading to an increasing variance in growth rates over time as observed in the models discussed in 3. Unsurprisingly, such super-linear preferential attachment models are characterized by power law distribution in the pre-asymptotic stage, while in the limit these models typically show a 'winner-takes-it-all' pattern.

Eventually, such a line of research might also directly engage and confront mainstream claims about the evolution of inequality and related question on the emergence of power law tails. While random multiplication is also an acceptable null-model for many mainstream researchers, which seems like a first cornerstone for constructive engagement (Benhabib and Bisin, 2018; Gabaix, Lasry, et al., 2016), adaptions of these models to better align them with empirical often rely on narratives building on unobservables, like ability or talent, to induce relevant changes in the formation of the underlying distribution. Such an approach inherently frames existing differences as a natural, meritocratic result emerging from the heterogeneity of agents to reaffirm the common conventions associated with marginal productivity theory.[22] The challenge of political economy to show, why such additions

[22]Such a framing is even applied in cases where, actually, a mechanism of cumulative advantage is employed. An example is given by the notion of *scale dependence* in Gabaix, Lasry, et al. (2016), where cumulative advantage is reframed (and, thereby, mystified) as an exogenous shock that overproportionally affects high incomes.

are eventually ad-hoc, superfluous and potentially misleading, as a description based on cumulative advantage and random multiplication understood as core dynamic forces inherent to capitalist systems should be eventually sufficient to derive the rough characteristics of the empirically observed distributions.

References

Akhundjanov, S. B., S. Devadoss, and J. Luckstead (2017). "Size distribution of national CO2 emissions". In: *Energy Economics* 66, pp. 182–193. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2017.06.012.

Albers, T. N. H., C. Bartels, and M. Schularick (2022). *Wealth and Its Distribution in Germany, 1895–2018*.

Albert, R. and A.-L. Barabási (2002). "Statistical mechanics of complex networks". In: *Reviews of Modern Physics* 74.1, pp. 47–97. ISSN: 0034-6861. DOI: 10.1103/revmodphys.74.47. eprint: cond-mat/0106096.

Alfani, G. (2021). "Economic Inequality in Preindustrial Times: Europe and Beyond". In: *Journal of Economic Literature* 59.1, pp. 3–44. ISSN: 0022-0515. DOI: 10.1257/jel.20191449.

Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*. New Approaches to Economic and Social History. Cambridge University Press.

Arthur, W. B. (1994). *Increasing Returns and Path Dependence in the Economy*. Book collections on Project MUSE. University of Michigan Press. ISBN: 9780472022403. URL: https://books.google.at/books?id=94MTEAAAQBAJ.

Arthur, W. B., Y. Ermoliev, and Y. Kaniovski (1987). "Path-dependent processes and the emergence of macro-structure". In: *European Journal of Operational Research* 30.3, pp. 294–303. ISSN: 0377-2217. DOI: 10.1016/0377-2217(87)90074-9.

Axtell, R. (2001). "Zipf Distribution of U.S. Firm Sizes". In: *Science* 293, pp. 1818–1820.

Balboni, C. et al. (2021). "Why Do People Stay Poor?" In: *The Quarterly Journal of Economics* 137.2, qjab045–. ISSN: 0033-5533. DOI: 10.1093/qje/qjab045.

Bavel, B. van (2016). *The Invisible Hand?: How Market Economies Have Emerged and Declined Since AD 500*. UPSO - Oxford University Press E-Books. Oxford University Press. ISBN: 9780199608133. URL: https://books.google.at/books?id=lElRDAAAQBAJ.

Benhabib, J. and A. Bisin (2018). "Skewed Wealth Distributions: Theory and Empirics". English. In: *Journal of Economic Literature* 56.4, pp. 1261–1291. DOI: 10.1257/jel.20161390.

Berman, Y., O. Peters, and A. Adamou (2021). “Wealth Inequality and the Ergodic Hypothesis: Evidence from the United States”. In: *Journal of Income Distribution®*. ISSN: 0926-6437. DOI: 10.25071/1874-6322.40455.

Bowles, S. (2023). “The Dawn of Everything: A New History of Humanity”. In: *Journal of Economic Literature* 61.3, pp. 1190–1198. ISSN: 0022-0515. DOI: 10.1257/jel.61.3.1188.r2.

Bowles, S. and M. Fochesato (2024). “The Origins of Enduring Economic Inequality”. In: *Journal of Economic Literature* 62.4, pp. 1475–1537. ISSN: 0022-0515. DOI: 10.1257/jel.20241718.

Champernowne, D. G. (June 1953). “A Model of Income Distribution”. In: *The Economic Journal* 63.250, pp. 318–351. ISSN: 0013-0133. DOI: 10.2307/2227127. eprint: https://academic.oup.com/ej/article-pdf/63/250/318/27395553/ej0318.pdf. URL: https://doi.org/10.2307/2227127.

Chetty, R. et al. (Apr. 2017). “The fading American dream: Trends in absolute income mobility since 1940”. English. In: *Science* 356.6336, pp. 398–406. DOI: 10.1126/science.aal4617.

David, P. A. (2007). “Path dependence: a foundational concept for historical social science”. In: *Cliometrica* 1.2, pp. 91–114. ISSN: 1863-2505. DOI: 10.1007/s11698-006-0005-x.

Dosi, G., Y. Ermoliev, and Y. Kaniovski (1994). “Generalized urn schemes and technological dynamics”. In: *Journal of Mathematical Economics* 23.1, pp. 1–19. ISSN: 0304-4068. DOI: 10.1016/0304-4068(94)90032-9.

Ederer, S. and M. Rehm (2018). *Making Sense of Piketty's ‘Fundamental Laws’ in a Post-Keynesian Framework: The Transitional Dynamics of Wealth Inequality.*

Eggenberger, F. and G. Pólya (1923). “Über die statistik verketteter vorgänge”. In: *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 3.4, pp. 279–289.

Fix, B. (2019). “Personal Income and Hierarchical Power”. In: *Journal of Economic Issues* 53.4, pp. 928–945. ISSN: 0021-3624. DOI: 10.1080/00213624.2019.1657746.

Gabaix, X. (1999). “Zipf's Law for Cities: An Explanation”. English. In: *The Quarterly Journal of Economics* 114.3, pp. 739–767. DOI: 10.1162/003355399556133.

— (2009). “Power laws in economics and finance”. English. In: *Annual Review of Economics* 1.1, pp. 255–294. DOI: 10.1146/annurev.economics.050708.142940.

Gabaix, X., P. Gopikrishnan, et al. (2003). “A theory of power-law distributions in financial market fluctuations”. In: *Nature* 423.6937, pp. 267–270. ISSN: 0028-0836. DOI: 10.1038/nature01624.

Gabaix, X., J.-M. Lasry, et al. (2016). “The Dynamics of Inequality”. In: *Econometrica* 84.6, pp. 2071–2111. ISSN: 1468-0262. DOI: 10.3982/ecta13569.

Georgescu-Roegen, N. (1971). *The Entropy Law and the Economic Process*. A Harvard paperback: Economics. Harvard University Press. ISBN: 9780674257801. URL: https://books.google.at/books?id=WYQxAAAAMAAJ.

Gibrat, R. (1931). *Les inégalités économiques*. Sirey. URL: https://books.google.at/books?id=m9fuoAEACAAJ.

Goldstein, L. and G. Reinert (2013). "Stein's method for the Beta distribution and the Pólya-Eggenberger urn". In: *Journal of Applied Probability* 50.4, pp. 1187–1205.

Graeber, D. (2011). *Debt: The First 5,000 Years*. Melville House. ISBN: 9781612190983. URL: https://books.google.at/books?id=GYhajCQU8XIC.

Jones, C. I. (2015). "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality". English. In: *Journal of Economic Perspectives* 29.1, pp. 29–46. DOI: 10.1257/jep.29.1.29. URL: http://web.stanford.edu/%5Ctextbackslashtextasciitildechadj/papers.html.

Kahn, R. F. (1959). "Exercises in the Analysis of Growth". In: *Oxford Economic Papers* 11.2, pp. 143–156. ISSN: 0030-7653. DOI: 10.1093/oxfordjournals.oep.a040820.

Kaldor, N. (1955). "Alternative Theories of Distribution". In: *The Review of Economic Studies* 23.2, pp. 83–100. ISSN: 0034-6527. DOI: 10.2307/2296292.

Kamiyama, N. and M. Murata (2019). "Reproducing Popularity Distribution of YouTube Videos". In: *IEEE Transactions on Network and Service Management* 16.3, pp. 1100–1112. ISSN: 1932-4537. DOI: 10.1109/tnsm.2019.2914222.

Kirstein, M. (2019). "The Ergodicity Problem in Economics". PhD thesis. Technical University Dresden.

Kunegis, J., M. Blattner, and C. Moser (2013). "Preferential Attachment in Online Networks: Measurement and Explanations". In: *arXiv*. DOI: 10.48550/arxiv.1303.6271. eprint: 1303.6271.

Locke, J. and P. Laslett (1965). *Two Treatises of Government*. A Mentor book ; ME1621. New American Library. ISBN: 9780451622037. URL: https://books.google.at/books?id=ys0jAQAAIAAJ.

Lux, T. and S. Alfarano (2016). "Financial power laws: Empirical evidence, models, and mechanisms". In: *Chaos, Solitons & Fractals* 88, pp. 3–18. ISSN: 0960-0779. DOI: 10.1016/j.chaos.2016.01.020.

Mahmoud, H. (2008). *Pólya urn models*. Chapman and Hall/CRC.

Mani, A. et al. (2013). "Poverty Impedes Cognitive Function". In: *Science* 341.6149, pp. 976–980. ISSN: 0036-8075. DOI: 10.1126/science.1238041.

Merton, R. K. (1968). "The Matthew Effect in Science". In: *Science* 159, pp. 56–63.

Metcalfe, S. (Jan. 1994). "Competition, Fisher's principle and increasing returns in the selection process". In: *Journal of Evolutionary Economics* 4, pp. 327–346.

Mitzenmacher, M. (Jan. 2004). "A Brief History of Generative Models for Power Law and Lognormal Distributions". English. In: *Internet Mathematics* 1.2, pp. 1385–251. DOI: 10.1080/15427951.2004.10129088.

Mulder, M. B. et al. (Oct. 2009). "Intergenerational wealth transmission and the dynamics of inequality in small-scale societies." English. In: *Science* 326.5953, pp. 682–688. ISSN: 1095-9203. DOI: 10.1126/science.1178336. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2792081%5C&tool=pmcentrez%5C&rendertype=abstract.

Nair, J., A. Wierman, and B. Zwart (2022). *The fundamentals of heavy-tails.* Cambridge University Press. DOI: 10.1145/2494232.2466587.

Newman, M. (2005). "Power laws, Pareto distributions and Zipf's law". English. In: *Contemporary Physics* 46.5, pp. 323–351. DOI: 10.1080/00107510500052444.

Oswald, Y., A. Owen, and J. K. Steinberger (2020). "Large inequality in international and intranational energy footprints between income groups and across consumption categories". In: *Nature Energy* 5.3, pp. 231–239. DOI: 10.1038/s41560-020-0579-8.

Page, S. E. (2006). "Path Dependence". In: *Quarterly Journal of Political Science* 1.1, pp. 87–115. DOI: 10.1561/100.00000006.

Pareto, V. (2014). *Manual of Political Economy.* Oxford University Press.

Pasinetti, L. (1962). "Rate of profit and income distribution in relation to the rate of economic growth". In: *The Review of Economic Studies* 29.4, p. 267. DOI: 10.2307/2296303.

— (1974). *Growth and income distribution.* CAMBRIDGE University Press. URL: https://books.google.at/books?id=HpBrzgEACAAJ.

— (1983). "Conditions of Existence of a Two Class Economy in the Kaldor and More General Models of Growth and Income Distribution". In: *Kyklos* 36.1, pp. 91–102. ISSN: 0023-5962. DOI: 10.1111/j.1467-6435.1983.tb02662.x.

Piketty, T. (2014). *Capital in the Twenty-First Century.* Harvard University Press. ISBN: 9780674369559. URL: https://books.google.at/books?id=J222AgAAQBAJ.

Pistor, K. (2020). *The Code of Capital: How the Law Creates Wealth and Inequality.* Business book summary. Princeton University Press. ISBN: 9780691208602. URL: https://books.google.at/books?id=wkjWDwAAQBAJ.

Polanyi, K. (1957). *The Great Transformation.* Beacon paperback no. 45. Beacon Press. ISBN: 9780807056790. URL: https://books.google.at/books?id=848GAQAAIAAJ.

Price, D. J. d. S. (1965). “Networks of Scientific Papers”. In: *Science* 149.3683, pp. 510–515. ISSN: 0036-8075. DOI: 10.1126/science.149.3683.510.

Quelle, D. and A. Bovet (2025). “Bluesky: Network topology, polarization, and algorithmic curation”. In: *PLOS ONE* 20.2, e0318034. DOI: 10.1371/journal.pone.0318034. eprint: 2405.17571.

Rigney, D. (2010). *The Matthew effect: How advantage begets further advantage.* New York: Columbia University Press.

Rothschild, K. W. (1947). “Price Theory and Oligopoly”. In: *The Economic Journal* 57.227, p. 299. ISSN: 0013-0133. DOI: 10.2307/2225674. URL: http://www.jstor.org/stable/2225674.

Scheidel, W. (2018). *The Great Leveler: Violence and the History of Inequality from the Stone Age to the Twenty-First Century.* The Princeton Economic History of the Western World. Princeton University Press. ISBN: 9780691184319. URL: https://books.google.at/books?id=CD1hDwAAQBAJ.

Scherer (1998). “The Size Distribution of Profits from Innovation”. In: *Annales d’Économie et de Statistique* 49/50, p. 495. ISSN: 0769-489X. DOI: 10.2307/20076127.

Schulz, J. and M. Milaković (2021). “How Wealthy are the Rich?” In: *Review of Income and Wealth.* ISSN: 0034-6586. DOI: 10.1111/roiw.12550.

Shaikh, A. (2016). *Capitalism – Competition, Conflict, Crises.* Oxford University Press.

Shaikh, A. and A. Ragab (2023). “Some universal patterns in income distribution: An econophysics approach”. In: *Metroeconomica* 74.1, pp. 248–264. ISSN: 0026-1386. DOI: 10.1111/meca.12412.

Silva, A. C. and V. M. Yakovenko (2004). “Temporal evolution of the ”thermal” and ”superthermal” income classes in the USA during 1983-2001”. In: *Europhysics Letters* 69.2. DOI: 10.1209/epl/i2004-10330-3. eprint: cond-mat/0406385.

Simon, H. A. (1955). “On a Class of Skew Distribution Functions”. In: *Biometrika* 42.3/4, p. 425. ISSN: 0006-3444. DOI: 10.2307/2333389.

— (1960). “Some further notes on a class of skew distribution functions”. In: *Information and Control* 3.1, pp. 80–88. ISSN: 0019-9958. DOI: 10.1016/s0019-9958(60)90302-8.

Smith, E. A. and B. F. Codding (2021). “Ecological variation and institutionalized inequality in hunter-gatherer societies”. In: *Proceedings of the National Academy of Sciences* 118.13, e2016134118. ISSN: 0027-8424. DOI: 10.1073/pnas.2016134118.

Stanley Metcalfe, J. (1998). *Evolutionary Economics and Creative Destruction.* London: Routledge.

Steindl, J. (1965). *Random Processes and the Growth of Firms: A Study of the Pareto Law.* Economic theory and applied statistics. Hafner Publishing Company. ISBN: 9780028529509. URL: https://books.google.de/books?id=wYZRAAAAMAAJ.

Wildauer, R. and J. Kapeller (2021). “Fitting Pareto Tails to Wealth Survey Data: A Practioners’ Guide”. In: *Journal of Income Distribution®*. ISSN: 0926-6437. DOI: 10.25071/1874-6322.40447.

— (2022). “Tracing the invisible rich: A new approach to modelling Pareto tails in survey data”. In: *Labour Economics* 75, p. 102145. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2022.102145.

Wold, H. O. A. and P. Whittle (1957). “A Model Explaining the Pareto Distribution of Wealth”. In: *Econometrica* 25.4, p. 591. ISSN: 0012-9682. DOI: 10.2307/1905385.

Yule, G. U. (1925). “II.—A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FR S”. In: *Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character* 213.402-410, pp. 21–87.

Zipf, G. K. (1932). *Selected Studies on the Principle of Relative Frequency in Language*. Harvard MA: Harvard University Press.

ifso working paper

ifso working papers are preliminary scholarly papers emerging from research at and around the Institute for Socio-Economics at the University of Duisburg-Essen.

All **ifso working papers** at uni-due.de/soziooekonomie/wp

ISSN 2699-7207

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

ifso institute for socio-economics

Institute for Socio-Economics
University of Duisburg-Essen

Lotharstr. 65
47057 Duisburg
Germany

uni-due.de/soziooekonomie
wp.ifso@uni-due.de