Power Laws in Economics: An Introduction*

Xavier Gabaix

August 14, 2014

Abstract

This paper is an elementary expository piece for the Journal of Economic Perspectives. It presents a survey of various power laws for firms, cities, trade and finance, as well as explanations that have been proposed for them. Then, it discusses to what extent power laws may explain aggregate fluctuations. It also given pointers to power laws outside of economics, and to open questions.

Samuelson (1969) was asked by a physicist for a law in economics that was both non-trivial and true. This is a difficult challenge, as many (roughly) true results are in the end rather trivial (e.g. demand curves slope downward), while many non-trivial results in economics in fact require too much sophistication of the agents to be really true. Samuelson answered “the law of comparative advantage”. The story does not say whether the physicist was satisfied, as the law of comparative advantage is a qualitative law, not a quantitative one. Since 1969, economics has found many qualitative insights, but many fewer reliable quantitative laws.

This article will make the case that, if asked now, Samuelson might mention various power laws as non-trivial and true laws in economics. I start by providing several illustrations of power laws, for cities, firms, and the stock market. I summarize some of the explanations that have been proposed. I then go on to suggest that power laws may explain much, including aggregate fluctuations. I conclude with comments on what is so special about power laws.

∗Prepared for the Journal of Economic Perspective. Comments most welcome: please email me if you find that a major mechanism, power law, or reference is missing — subject to space constraints. I thank Jerome Williams for excellent research assistance.

1See Gabaix (2014) for an exploration of this in basic microeconomics: many non-trivial predictions of micro (e.g., Slutsky symmetry) fail when agents are boundedly rational — they rely on too subtle an optimization by the agents.

2Other quantitative (approximate) laws are (i) the Black-Scholes formula, and (ii) some sophisticated mechanisms in implementation theory.
1 Some Empirical Power Laws

Let us look at some pictures of power laws.

1.1 City sizes

Let us look at the data on US cities of size (i.e., population) 250,000 or greater.\(^3\) We rank cities by size, number 1 being the top city: #1 is New York, #2 Los Angeles, etc. We regress log rank on log size, and find the following:

\[
\ln \text{Rank} = 7.88 - 1.03 \ln \text{Size}. \tag{1}
\]

This is close to a straight line \((R^2 = 0.98)\), and the slope very close to 1 (the standard deviation of the slope is 0.01).\(^4\) This means that the rank of a city is proportional to the inverse of its size. This value of \(\sim 1\) has been found repeatedly across many cities and countries, at least after the Middle Ages (Dittmar 2012). There is no artifactual reason to find a power law, and even less for the slope to be 1.

To think about this type of regularity, it is useful to be a bit more abstract, and see the

---

\(^3\)This comprises all “Metropolitan Statistical Areas” (i.e., agglomerations) provided in the *Statistical Abstract of the US* (2012).

\(^4\)Actually, the standard error returned by OLS is incorrect. The correct standard error is \(|\text{slope}| \times \sqrt{2/N} = 0.11\), where \(N = 184\) is the number of cities in the sample (Gabaix and Ibragimov 2011).
cities as coming from a underlying distribution: the probability that the size of a randomly
drawn city is greater than $x$ is proportional to $1/x^\zeta$ with $\zeta \simeq 1$. More generally

$$P(\text{Size} > x) = \frac{a}{x^\zeta}$$

at least for $x$ above a cutoff (here, the 250,000 inhabitants cutoff given by the Statistical
Abstract of the US). A regularity of the type (2) is called a power law. In a given finitely-sized
sample, it generates an approximate relation of type (1): $\ln \text{Rank} = c - \zeta \ln \text{Size}$.

The interesting part is the coefficient $\zeta$, which is called the power law exponent of the
distribution (or Pareto exponent). A “Zipf’s law” is a power law with an exponent of 1.

A lower $\zeta$ means a higher degree of inequality in the distribution: it means a greater
probability of finding very large cities or (in another context) very high incomes. Indeed, the
moments of order greater than $\zeta$ are infinite, while moments of order less than $\zeta$ are finite.
For example, if $\zeta = 1.03$, the expected size is finite, but the variance is formally infinite.

In addition, the exponent is independent of the units (inhabitants or thousands of in-
habitants, say). This makes it at least conceivable that we might find an a priori constant
simple value (such as an integer) in various datasets.

What if we look at cities with size less than 250,000? Does Zipf’s law still hold? Figure
2 shows the distribution of city sizes for the UK (where the data is particularly good). Here
we see the appearance of a straight line for cities of about size 500 and above. Zipf’s law
holds very well too.

Why do we care? As Krugman (1996) put it, after having written his work on economic
geography: “The failure of existing models to explain a striking empirical regularity (one
of the most overwhelming empirical regularities in economics!) indicates that despite con-
siderable recent progress in the modeling of urban systems, we are still missing something
extremely important. Suggestions are welcome.” We shall see that since then, we have im-
proved our understanding of the origin of the Zipf’s law, which has forced a great rethinking
about the origins of cities – and firms, as we shall now see.

1.2 Firm sizes

We now look at the firm size distribution. Axtell (2001) presents it, from the US census data.
He bins firms according to their size (no. of employees), and plots the log of the number of

---

5Because it is delicate to construct agglomerations (rather than fairly arbitrary legal entities), Rozenfeld
et al. (2011) use a new algorithm constructing cities from fine-grained geographical data.
Figure 2: Cumulative distribution of agglomerations in the UK. We see a pretty good power law fit starting at about 500 inhabitants. The exponent is actually statistically non-different from 1 for size $S > 12,000$ inhabitants. Source: Rozenfeld et al. (2011).

firms within a bin. We see a straight line: this is a power law. Here we can even run the regression in “density”, i.e. plot the number of cities of size approximately equal to $x$. If (2) holds, then the density of the firm size distribution is $f(x) = b/x^\zeta$, so the slope in a log-log plot should be $- (\zeta + 1)$ (as $\ln f(x) = - (\zeta + 1) \ln x + \text{constant}$). Impressively, the exponent that Axtell finds is $\zeta = 1.059$. This demonstrates a “Zipf’s law” for firms.

Again, this has forced a rethinking of firms: most static theories (e.g. based on elasticity of demand, fixed cost, economies of scope, etc.) would not predict a Zipf’s law. Some other type of theory is needed, as we shall soon discuss.

1.3 Stock market movements

It is well-known that stock market returns are fat-tailed – i.e. the probability of finding extreme values is larger than for a Gaussian distribution of the same mean and standard deviation. An energetic movement of physicists (“econophysicists”, a term coined after “geophysicists” and “biophysicists”) has quantified a host of power laws in the stock market. For instance, the daily stock market movements are represented in Figure 4. They are consistent with: $P(|r| > x) = a/x^\zeta$ with $\zeta = 3$, the “cubic” law of stock market returns. The left panel of Figure 4 plots the distribution for four different sizes of stocks. The right panel plots the distribution of normalized stock returns, i.e. of stock returns divided by
Figure 3: Log frequency $\ln f(S)$ vs log size $\ln S$ of U.S. firm sizes (by number of employees) for 1997. OLS fit gives a slope of 2.06 (s.e. = 0.054; $R^2 = 0.99$). This corresponds to a frequency $f(S) \sim S^{-2.059}$, i.e. a power law distribution with exponent $\zeta = 1.059$. This is very close to Zipf’s law, which says that $\zeta = 1$. Source: Axtell (2001).

Likewise, lots of other quantities are power law distributed (Plerou et al. 2005, Kyle and Obizhaeva 2013, Bouchaud, Farmer and Lillo 2009, Geerolf 2014). For instance, the number of trades per day is power law distributed with exponent of 3, while the number of shares traded per time interval has an exponent of 1.5, and the price impact is proportional to the volume to the power of 0.5.

Why do we care? One implication of the cubic law is that there are many more extreme events than would occur if the distribution were Gaussian. More precisely, a 10 standard deviation event and a 20 standard deviation event are, respectively, $5^3 = 125$ and $10^3 = 1000$ times less likely than a 2 standard deviation event, whereas if the distribution of returns was Gaussian, the ratios would be $10^{22}$ and $10^{87}$, respectively. Indeed, in a stock market comprising about 1000 stocks, a 10 standard deviation event happens in practice about every day.

More deeply, the explanation for those regularities force us to rethink the functioning of stock markets – we shall discuss later theories that exactly explain these exponents in a unified way.
Figure 4: Cumulative distribution of daily stock market returns. The left panel shows the distributions for 4 different sizes of stocks. The right panel shows the returns, normalized by volatility. The slopes are close to $-3$, reflecting the “cubic law” of stock market fluctuations: $P(|r| > x) \sim kx^{-3}$. The horizontal axis displays returns as high as 100 standard deviations. Source: Plerou et al. (1999).

1.4 Other power laws

Income and wealth also follow roughly power law distributions, as we have known since Pareto (1896), who first documented power laws – not only in economics, but indeed anywhere (to the best of my knowledge). The distribution of wealth is more unequal than the distribution of income: this makes sense, as differences in growth rate of wealth across individuals (due to differences in returns or frugality) pile up and add an extra source of inequality. Typically, the exponent is around 1.5 for wealth, and between 1.5 and 3 for income. Given the recent interest in income inequality, power laws and random growth processes are a central tool to analyze those (Piketty and Zucman 2014, Atkinson, Piketty and Saez 2011, Benhabib, Bisin and Zhu 2011, Lucas and Moll 2014).

2 What Causes Power Laws?

2.1 Random growth

The basic mechanism for generating power laws is proportional random growth (Champernowne 1953, Simon 1955). Suppose that the we start with an initial distribution of firms, and they grow and shrink randomly with independent shocks, and they satisfy “Gibrat’s law”: all firms have the same expected growth rate and the same standard deviation of growth
rate. This model does not even have a steady state distribution, as the distribution becomes a lognormal with larger and larger variance. However, things change altogether if we add to the model some friction that guarantees the existence of a steady state distribution. Suppose for instance that there is also a lower bound on size, so that a firm size cannot go below a given threshold. Then, the model yields a steady distribution, and it is a power law, with some exponent $\zeta$ that depends on the details of the growth process.

Now, this mechanism generates a power law, but with an exponent that is generally not 1. Why would we have exponent of 1? This was Krugman’s question earlier.

One explanation is given in Gabaix (1999) (see also Gabaix (2009) for a thorough review). Suppose that the size of the “friction” (the lower bound) becomes very small, and that we have a given exogenous population size to allocate in the system (between the different cities or firms). Then, the exponent $\zeta$ becomes 1, rather than any other value.\(^6\) This is why we observe lots of Zipf’s laws: proportional random growth, with a small friction, and some “adding-up constraint” for the total size of the system.

Of course, this is the “mechanical” part of the explanation. An economist would like to know why we have random growth in the first place, or in other words, why Gibrat’s law holds. The simplest microfoundation is that cities and firms are basically constant returns to scale, perhaps with small deviations from that benchmark, and lots of randomness. Indeed, many fully economic models for the random growth of cities and firms have been proposed since the 2000s (e.g., Rossi-Hansberg and Wright (2007) and Luttmer (2007)). Likewise, for the income distribution, the details of the underlying mechanism (say, luck vs thrift, responsive to incentives or not) are very important for a variety of questions, and microfounded models are important. Still, to write them sensibly one needs to keep in mind the core mechanics of these models – here, random growth with power laws.

\subsection{2.2 Matching and economics of superstars}

Another manifestation of power laws is in the extremely high earnings of top earners in areas of arts, sports and business. Rosen (1981) gives a qualitative explanation for this with the “economics of superstars”. Gabaix and Landier (2008) give a tractable, calibratable model of this phenomenon, along the following lines. Suppose that lots of firms, of different sizes, compete to hire the talents of CEOs. The “talent” of a CEO is given by how much

\footnote{One intuition is as follows: the exponent cannot be below 1, because then the distribution would have infinite mean. Indeed, an exponent just above 1 is the smallest consistent with a finite total population. As the “friction” becomes very small, the exponent becomes the fattest consistent with a finite population.}
(in percentages) she is expected to increase the firms’ profits. Competition implements the efficient outcome, which is that the largest firm will be matched with the best CEO in the economy, the second largest firm with the second best CEOs, etc.

One might think it hopeless to derive a quantitative theory, as the distribution of talent is very hard to observe. However, we can draw on “extreme value theory” (which is a branch of probability) to obtain some properties of the tail of the distribution of talent, without knowing the distribution itself. One of them is that, given adjacent CEOs in the ordering of talent, the approximate difference in talent between these two CEOs varies like a power law of their rank. The exponent depends on the distribution, but the power law functional form holds for essentially any reasonable distribution (in a way that can be made precise). Given this, Gabaix and Landier work out the pay of CEO number $n$, who manages a firm of size $S(n)$. We denote by $S(n_*)$ the size of a reference firm, i.e. the size of the median firm in the S&P 500. The pay of CEO number $n$ is:

$$w(n) = D(n_*) S(n_*)^{1-b} S(n)^b$$

where one calibrates $b = 1/3$. This is a “dual scaling relation”, because it has two power laws. We see how matching creates a “dual scaling equation” (38), or double power laws, which has three implications:

(a) **Cross-sectional prediction.** In a given year, the compensation of a CEO is proportional to the size of his firm to the power of $1/3$, $S(n)^{1/3}$.

(b) **Time-series prediction.** When the size of all large firms is multiplied by $\lambda$ (perhaps over a decade), the compensation at all large firms is multiplied by $\lambda$. In particular, the pay at the reference firm is proportional to the average market cap of a large firm.

(c) **Cross-country prediction.** Suppose that CEO labor markets are national rather than integrated. For a given firm size $S$, CEO compensation varies across countries with the market capitalization of the reference firm, $S(n_*)^{2/3}$, using the same rank $n_*$ of the reference firm across countries.

It turns out that all three predictions seem to hold empirically since the 1970s. This explains why CEO pay has increased a lot – firm size has increased.

Here, power laws and extreme value theory are the natural language for the “economics of superstas”. In addition, very small differences in talent give rise to very large differences in pay: in the Gabaix-Landier calibration, talent has finite support, but differences of talent are unboundedly large. This is what happens when very large firms compete to hire the

---

7 Empirically, the exponent tends to be in the $[0.3, 0.4]$ range.
services of CEOs.

The same logic should apply to other superstars market: apartments with a large view on Central Park, but also sport and the price of works of art. When the wealth becomes stronger and more unequal, the same should happen to say the price of works of art. As far as I know, a systematic quantitative exploration of those issues still has to be done.\(^8\)

This line of thinking leads to a fresh way of thinking about pay-performance sensitivity, going beyond the classic result of Jensen and Murphy (1990). They define the Pay-Performance Sensitivity (PPS) as how many dollars the CEO compensation (or wealth) changes for a given dollar change in firm value. They find that the PPS is very small: CEOs earn “only” $3 extra when their firm increases by $1000 in value: they conclude that corporate governance may not work well. However, Edmans, Gabaix and Landier (2009) propose a different way to think of the benchmark incentives, resting on scaling arguments. Suppose that to motivate the CEO, it is percent/percent incentives, not dollar/dollar incentives that matter: namely, for a 1% of increase in firm value, the CEO’s wealth should increase by \(k\%\), where \(k\) does not scale with size (this comes from preferences that are multiplicative in effort and consumption). Then, if (3) holds, the PPS of Jensen-Murphy should decrease as \((\text{Firm size})^{-2/3}\).\(^9\) This is actually true empirically, as illustrated in Figure 5. Hence, thinking in terms of scaling leads to new predictions for pay-performance sensitivity, and actually leads to propose new ones.

### 2.3 Optimization and Transfer of power laws

Optimization gives a good way to obtain power laws. For instance, the Allais-Baumol-Tobin rule for the demand for money (which scales as \(i^{-1/2}\), where \(i\) is the interest rate), is a power law. The first scaling relation in economics (and, not coincidentally, the first non-trivial empirical success in economics) may be Hume’s thought experiment that doubling the money supply should lead (after a while) to a doubling of the price level — a basic theory that has stood the test of time.

Power laws have very good aggregation properties: taking the sum of two (independent) power law distributions gives another power law distribution. Likewise, multiplying two

\(^8\)Behrens, Duranton and Robert-Nicoud (2013) propose a theory of Zipf’s law based on matching.\(^9\)The percent/percent incentive is constant \((d\ln w/dr = k\), where \(w\) is CEO pay, and \(r\) the firm return), so the Jensen-Murphy pay-performance sensitivity measure is,

\[
PPS := \frac{dw}{ds} = \frac{dw}{Sdr} = \frac{w d\ln w}{S^{-1} dr} = \frac{k'S^b}{S} k = k''S^{1-b}
\]

using \(w = k'S^b\) from (3, and for firm-independent constants \(k', k''\).
power laws, taking their max or their min, or a power, etc. gives again a power law distribution. This partly explains the prevalence of power laws: they survive many transformations and the addition of noise.

3 Granularity: Aggregate fluctuations from microeconomic shocks

I now turn to an application of power laws: getting a better sense of the origins of aggregate fluctuations in GDP, exports and the stock market.

3.1 Basic ideas

Where do aggregate fluctuations come from? Gabaix (2011) proposes that idiosyncratic shocks to firms (or narrowly defined industries) can generate aggregate fluctuations. A priori, an economist would say that this is not quantitatively possible: there are millions of firms, so by an analogy similar to the central limit theorem, their total fluctuations should be very small (technically, if there are N firms, the total fluctuations should decay as $1/\sqrt{N}$).
However, when the firm size distribution is fat-tailed, things change completely: the central limit theorem doesn’t apply any more.\(^{10}\) Instead, GDP fluctuation decays in \(1/\ln N\).

Empirically, the distribution of economic activity is indeed very concentrated amongst firms. For instance, di Giovanni and Levchenko (2012) find that “In Korea, the 10 biggest business groups account for 54% of GDP and 51% of total exports. [...] The largest one, Samsung, is responsible for 23% of exports and 14% of GDP.” Hence, it is plausible that idiosyncratic shocks to firms would affect GDP activity.

In that view, economic activity is not made of a smooth continuum of firms, but it is made of incompressible “grains” of activities – the firms – whose fluctuations do not wash out in the aggregate. The plain reason is that firms are big (and initial shocks are intensified by a variety of generic amplification mechanisms, e.g. endogenous changes to hours worked).

Is this granular hypothesis relevant empirically? Gabaix (2011) finds that the idiosyncratic shocks to large firms explain about 1/3 of GDP fluctuations in the US. Di Giovanni, Levchenko and Mejean (2014) find that they explain over half the fluctuations in France. Further support is given in Foerster, Sarte and Watson (2011) for industrial production, and di Giovanni and Levchenko (2012) for exports. The exploration continues.

There are two payoffs from this analysis: first, we may better understand the origins of aggregate fluctuations. Second, these large idiosyncratic shocks may give a series of useful instruments for macro. For instance, Amiti and Weinstein (2013) start form the fact that banking is very concentrated, so idiosyncratic bank shocks may have strong ripple effects in the economy, which allows them to quantify banking channels.\(^{11}\)

Another implication of granularity is the importance of networks (Acemoglu et al. 2012). Those large firm-level shocks propagate through networks, which create an interesting amplification mechanism, and a way to quantitatively see the propagation effects. Networks are a particular case of granularity, rather than an alternative to it. They offer a way to “visualize” the propagation of idiosyncratic firm shocks.

### 3.2 The great moderation: A granular post-mortem

This granular perspective offers a way to understand the time variations in volatility. For instance, Carvalho and Gabaix (call granular or “fundamental” volatility the volatility that

\(^{10}\) A power-law variant holds, the Lévy central limit theorem.

\(^{11}\) Incidentally, they find that idiosyncratic bank shock explain 40% of aggregate loan and investment fluctuations.
would come only from idiosyncratic sectoral or firm-level shocks.\footnote{To operationalize this idea, they consider the “fundamental volatility:” $\sigma_{Ft} = \sqrt{\sum_{i=1}^{n} \left( \frac{S_{it}}{\text{GDP}_{it}} \right)^2 \sigma_i^2}$, where $S_{it}$ is the gross (not net) output of sector $i$, and $\sigma_i$ is the standard deviation of the total factor productivity (TFP) in the sector.} Figure 6 reports their findings. The fundamental volatility is quite correlated with actual volatility.

This gives an extra narrative for the “great moderation” and its undoing. i) The long and large decline of fundamental volatility from the 1960s to the early 1990s is due to the smaller size of a handful of heavy-manufacturing sectors: this created a “great moderation” of volatility – not due to monetary factors and the like, but simply because the economy became more diversified. ii) The increased importance of the energy sector (which itself can be traced to the rise of oil prices) accounts for the burst of volatility in the mid 1970s. iii) The increase in the size of the financial sector is an important determinant of the increase in fundamental volatility – and of actual volatility – in the 2000s. Similar findings hold for the other developed countries considered.

Likewise, moving to explaining firm-level volatility, Kelly, Lustig and van Nieuwerburgh (2013) find that taking account of the network can help us understand firm-level and aggregate level stock market volatility: for instance, firms with few, and volatile, customers or suppliers will have large volatility.

Figure 6: Fundamental Volatility and GDP Volatility. The squared line gives the granular / fundamental volatility ($4.5\sigma_{Ft}$, demeaned), i.e. the volatility that would arise if there were only idiosyncratic sectoral shocks, and no aggregate shock. The solid and circle lines are annualized (and demeaned) estimates of GDP volatility, using respectively a rolling-window estimate and an Hodrick-Prescott trend of instantaneous volatility. Source: Carvalho and Gabaix (2013).
3.3 Origins of stock market crashes?

Gabaix et al. (2003, 2006) develop the hypothesis that stock market crashes are due to large financial institutions selling under pressure in illiquid markets (see also Levy and Solomon 1996 and Solomon and Richmond 2001). This may account not only for large crashes, but also for the whole distribution of mini-crashes described by the power law. Large institutions are roughly Zipf distributed, but trade less than proportionally to their size to moderate price impact; the price impact is itself proportional to the square root of the volume traded – all of these relationships arise endogenously. Still, when large institutions sell under time pressure, they make the market fall and (if they are large enough, and under enough pressure) crash. One can speculate that this type of mechanism might have been at work in a variety of well-known events: the crash of Long Term Capital Management in the summer of 1998, the rapid unwinding of very large stock positions by Société Générale after the Kerviel “rogue trader” scandal (which led stock markets to fall in 21-23 January 2008), and the flash crash of May 2010. See the discussion in Gabaix et al. (2006) and Kyle and Obizhaeva (2013), which also argue that the 1929 crash is an example of that phenomenon. This research potential brings us closer to understanding the origins of stock market movements.13

3.4 Volatility of firms and countries

Granularity can also explain the following fact: firms and countries have identical, non-trivial, scaling of growth rates. Stanley et al. (1996) study how the volatility of the growth rate of firms changes with size. To do this, they calculate the standard deviation $\sigma(S)$ of the growth rate of firms’ sales, $S$, and regress its log against log size. They find an approximately affine relationship, displayed in Figure 7: $\ln \sigma^{\text{firms}}(S) = -\alpha \ln S + \beta$. This means that a firm of size $S$ has volatility proportional to $S^{-\alpha}$ with $\alpha = 0.15$.14

Lee et al. (1998) conduct the same analysis for country growth rates and find that countries with a GDP of size $S$ also have a volatility proportional to $S^{-\alpha'}$, with $\alpha' = 0.15$ – a form of diversification (Koren and Tenreyro 2013). The two graphs are plotted in Figure 7. The slopes are indeed very similar. This may be a type of “universality”. This may be explained by granularity: if aggregate fluctuations come from microeconomic shocks, and firm sizes follow Zipf’s law, then the identical scaling of Figure 7 should hold true.

---

13 Barro and Jin (2011) document a power law distribution of macroeconomic disasters, which may explain many puzzles in finance (Gabaix 2012).

14 Large firms have a smaller proportional standard deviation than small firms. However, this diversification effect is weaker than if a firm of size $S$ were composed of $S$ independent units of size 1, which would predict $\alpha = 1/2$. See Riccaboni et al. (2008).
Figure 7: Standard deviation of the distribution of annual growth rates (log-log axes). Note that $\sigma (S)$ decays with size $S$ with the same exponent for both countries and firms: $\sigma (S) \sim S^{-\alpha}$, with $\alpha \simeq 1/6$. The size is measured in sales for the companies (top axis) and in GDP for the countries (bottom axis). The firm data are taken from the Compustat for the years 1974-1993, the GDP data from Summers and Heston (1991) for the years 1950-1992. Source: Lee et al. (1998).

4 Power laws and Universality outside of economics

Mathematics. The idea of “universality” comes from mathematics. The paradigm is the central theorem: when forming an average of demeaned random variables, the distribution of the average (suitably normalized, here multiplied by $\sqrt{N}$) converges to the Gaussian distribution — a “universal” distribution independent of the details of the original distribution. This phenomenon happens for other things, e.g. the distribution of extremes of distributions, and the properties of matrices with random entries.

Physics. The physics of “critical phenomena” is full of “universal” laws, in the sense that different materials behave in the exact same way, after normalization, around the phase transition. There is a sophisticated machinery of the “renormalization group” to understand this (see Sornette 2004). Lots of other phenomena, e.g. forest fires and rivers, have a power law structure (partially explained by the theory of self-organized criticality). Perhaps surprisingly, these notions of percolation and self-organized criticality haven’t (yet) found widespread use in economics (see, however, Bak et al. 1993, Nirei 2006).

Networks. Networks are full of power laws: for instance (probably because of random growth), the popularity of web sites (number of sites linking to it) is a power law (Barabasi and Albert 1999).

Biology. Biology is replete with intriguing, universal relations of the power law type. For instance, the energy that an animal of mass $M$ requires to function is proportional to $M^{3/4}$ –
rather than \( M \) as a simple “constant return to scale” model would predict (Figure 11). West et al. (1997) have proposed the following explanation: If one wants to design an optimal network system to send nutrients to the animal, one designs a fractal system; the resulting efficiency generates the \( M^{3/4} \) law. There is an interesting lesson: lots of things could a priori matter for energy consumption (e.g. climate, predator or prey status, thickness of the fur), and they probably do matter, a little bit. However, in its essence, an animal is best viewed as a network in which nutrients circulate at maximum efficiency. Understanding the power laws forces the researcher to forget, in the first pass, about the details. Likewise, this research shows similar laws for a host of variables, including life expectancy (which scales at \( M^{1/4} \)).

Here the interpretation is that the animal is constructed optimally, given engineering constraints given by biology. Perhaps surprisingly, this type of mechanism doesn’t seem to have been much studied in economics. For instance, the economy resembles a network with power law distributed firms: does it come from optimality, as opposed to randomness? It would be nice to know. Likewise, Zipf’s law holds for the usage frequency of words. The simplest explanation is via random growth (as the popularity of words follow a random growth process). However, perhaps that might reflect an “optimal” organization of mental categories, perhaps in some tree-like structure? Again, one would like to know.

Figure 8: Metabolic rate (i.e., energy requirement per day) a function of mass across animals. The slope of this log-log graph is \( 3/4 \): the metabolic rate of an animal of mass \( M \) is proportional to \( M^{3/4} \) (Kleiber’s law). Source: West, Brown and Enquist (2000).
4.1 Conclusion

Power laws are everywhere (when datasets contains enough variation in some “size”-like factor, such as income or number of employees) – so all economists should know about power laws, and the basic mechanisms that generate them.

Power laws can guide the researcher to the “essence” of the phenomenon. For instance, take city sizes: a priori, lots of things may be important for cities: specialization, transportation cost, elasticity of congestion vs positive externalities in human capital... The power law approach concludes that while those things may exist, they are not the essence: the essence is random growth with a small friction. Now, to generate the random growth, a judicious mix of the “traditional” ingredients may be useful, but to orient the understanding, one should first think about the essence, and only then about the economic underpinning.

Some open questions about power laws

Now is a good time to answer questions on power laws, as many new data are available.

Networks and granularity. How big is the volatility generated from idiosyncratic shocks, propagated and amplified in networks?

The gravity equation. The trade flows between two countries is proportional to the GDP of the two countries (which is trivial, and come from a simple CRS model), and declines with distance (which is intuitive), as the inverse of distance to the power 1 – the latter scaling being very non-trivial. What explains this? Chaney (2013) proposes an ingenious model linking this distance to the probability of forming a link in a random growth (for power laws in trade, see also Melitz, Helpman and Yeaple (2004) and Eaton, Kortum and Kramarz (2011)). Under Zipf’s law, that generates the coefficient of 1. Similar scaling holds for migration, see Levy (2013).

Why is the aggregate production (roughly) Cobb-Douglas with a capital share about 1/3? Jones (2005) generates the functional form, but not the particular exponent. Perhaps generating it will suggest a deeper understand of the causes of technical progress.

Is random growth the canonical origin of power laws? or is it something else, like the economics of superstars, or optimization? Does Gibrat’s law really hold?15

As “big data” makes lots of new datasets available, it will be important to order them. Scaling questions are a natural way to do that, and have met with great success in the

---

15 Gibrat’s law seems to roughly hold (Ioannides and Overman 2003, Eeckhout 2004), though the issue is not settled as the literature hasn’t fully differentiated between permanent shocks and transitory ones and measurement error.
natural sciences. Given the new availability of great datasets, the future of power laws seem bright.

**Recommendations for future readings** A reader seeking a gentle introduction to power law techniques might start with Gabaix (1999), and then move on to more systematic exposition in Gabaix (2009), which contains many other pointers. Mantegna and Stanley (1999) contains an accessible introduction to the field of econophysics, while Sornette (1999) contains many interesting physics mechanisms generating scaling. For extreme value theory, Embrechts, Kluppelberg and Mikosch (1997) is very pedagogical. For networks, Jackson (2010) and Newman (2010) are now classic references.

5 References


Gabaix, Xavier. “A Sparsity-Based Model of Bounded Rationality.” *Quarterly Journal*
of Economics, forthcoming.


Kyle, Albert, and Anna Obizhaeva. 2013. “Large bets and stock market crashes.”
Working paper.


