Scale-free power-laws as interaction between progress and diffusion: a critical evaluation of fat-tail distributions

Martin Hilbert
United Nations ECLAC
Adolfo Ibáñez University, Complex Systems Engineering
MartinHilbert [at] gmail [dot] com

This is a preprint of an article accepted for publication as
Hilbert, M. (2013), Scale-free power-laws as interaction between progress and diffusion. *Complexity*. doi: 10.1002/cplx.21485; http://onlinelibrary.wiley.com/doi/10.1002/cplx.21485/abstract

ABSTRACT

While scale free power-laws are frequently found in social and technological systems, their authenticity, origin and gained insights are often questioned, and rightfully so. The article presents a newly found rank-frequency power-law that aligns the top-500 supercomputers according to their performance. Pursuing a cautious approach in a systematic way, we check for authenticity, evaluate several potential generative mechanisms, and ask the "so what" question. We evaluate and finally reject the applicability of well-known potential generative mechanisms like preferential attachment, self-organized criticality, optimization, and random observation. Instead, the micro-data suggest that an inverse relationship between exponential technological progress and exponential technology diffusion through social networks results in the identified fat-tail distribution. This newly identified generative mechanism suggests that the supply and demand of technology ("technology push" and "demand pull") align in exponential synchronicity, providing predictive insights into the evolution of highly uncertain technology markets.

Keywords: power-law, technological change, diffusion of innovations, pareto.

Acknowledgements: I'm thankful to Aaron Clauset and Cosma Shalizi who not only were so kind to provide their very useful power-law authenticity code online, but also took the time to patiently answer all my questions about power-laws during diverse Santa Fe Institute events over the recent year. Bill McKelvey was instrumental in raising my interest in power-laws.

The fascination with power-laws

The power-law and its variations (the Pareto distribution and Zipf's law) have become a pervasive signature of complex systems: one out of five papers published in the Journal Complexity refer to "power-law", "scale-free", "zipf" or "pareto" (176 of 876 articles between 1996 and 2013). Since the 1960s we have learned of hundreds of power-laws in physics (e.g. sand pile avalanches and earthquakes), biology (e.g. species extinction and body mass), and the social sciences (e.g. city sizes and income) [1, 2, 22].

The current fascination with the power-law is fuelled by at least three sources, all with their respective drawbacks. First, on the upside, power-laws are quite easy to identify visually as a linear relationship on a log-log plot. On the downside, it has been shown that many assumed power-laws are very unlikely authentic power-laws [3]. Second, over recent decades researches have come up with several fascinating generative models for power-laws, among them preferential attachment, self-organized criticality, random walks and optimization [4, 5]. These provide ready-to-use theoretical explanations to choose from once a power-law has been detected. However, superficial analysis or overreliance on intuition can lead to assuming the wrong generative mechanism, resulting in more confusion than insight. Last but not least, it is quite rare to find clear statistical distributions anywhere in nature, much more so in society. Therefore, it is simply alleviating to detect some kind of large scale pattern like power-laws, which act as a straw to hold onto in a complex world. At the same time, simply identifying a pattern does not help us to navigate this complex world any better. We also have to ask about the lessons learned from such descriptive patterns, an important step that is often neglected in the case of power-laws [6]. Unfortunately, the overly hasty celebration of potential power-laws, and the precarious neglect of the before mentioned vulnerabilities have given power-laws a bad name over recent years [7].

This article presents a newly found power-law in the rank-frequency distribution of differently performing technological devices. The main goal of the article is to walk step-by-step through a more careful analysis of this potential power-law. First, we test the authenticity of the visually identified power-law with a more stringent statistical test. Second, we critically ask about the potential generative mechanisms of this distribution. We consider and evaluate the adequacy of several well-known generative mechanisms of power-laws (such as preferential attachment, self-organized criticality, and highly optimized tolerance) and come to the conclusion that they are unlikely to fit the identified phenomenon. We then present a new kind of power-law generative mechanism that fits the micro-data evidence. Third, we discuss the gained lessons learned, as well as the limitations of possible conclusions.

The different faces of power-laws

Before getting started, it is important to remember that "power-laws", "fat-tail-" and "scale free distributions", as well as "Zipf's law" and "Pareto distributions" represent essentially the same relation and can easily be transformed into one another. A rank-frequency power-law[1] relates a variable (size,

[1] Not all power-laws are of the rank-frequency type. Rank-frequency power-laws relate the quantity of a variable (rank) with the frequency (probability p(x)) of appearance of this very same variable. These are sometimes also called "univariate". On the contrary, "bivariate" power-laws relate two different variables, using a generic function

performance, speed, etc.) with its frequency of occurrence as a probability p(x) through the formula: $p(x) = C*x^{-\alpha}$, whereas α is a constant parameter known as the exponent or scaling parameter. A convenient trait of this formula is that it turns out to be a linear relationship when the logarithm is taken on both sides: $\log p(x) = \log (C*x^{-\alpha}) = \log C + \log x^{-\alpha} = -\alpha \log x + c$ (with $c = \log C$). This leads to the eminent straight line when the resulting graph is presented in log-log form. The power-law is the probability density function (PDF) of the associated cumulative distribution function (CDF) known as the Pareto distribution (for a useful discussion see [10]). There are several advantages of using the CDF, instead of the PDF, including the fact that the CDF does not require the binning of the measured variables into groups and therefore avoids the (somewhat arbitrary) decision of bin-width, while additionally not throwing away any information [5]. The CDF of the Pareto distribution can also be presented as Zipf's law [11], which effectively switches the horizontal and vertical x and y axis and ranks the frequency of the entities from 1st to last, instead of cumulating them from 0 to 1, which of course does not change the nature of the underlying relation [5]. This being said, the following analysis uses the cleaner Pareto CDF format for the authenticity analysis (Figure 1 and Table 1) and the visually useful binning in PDF form for the interpretative analysis (Figures 2 - 4).

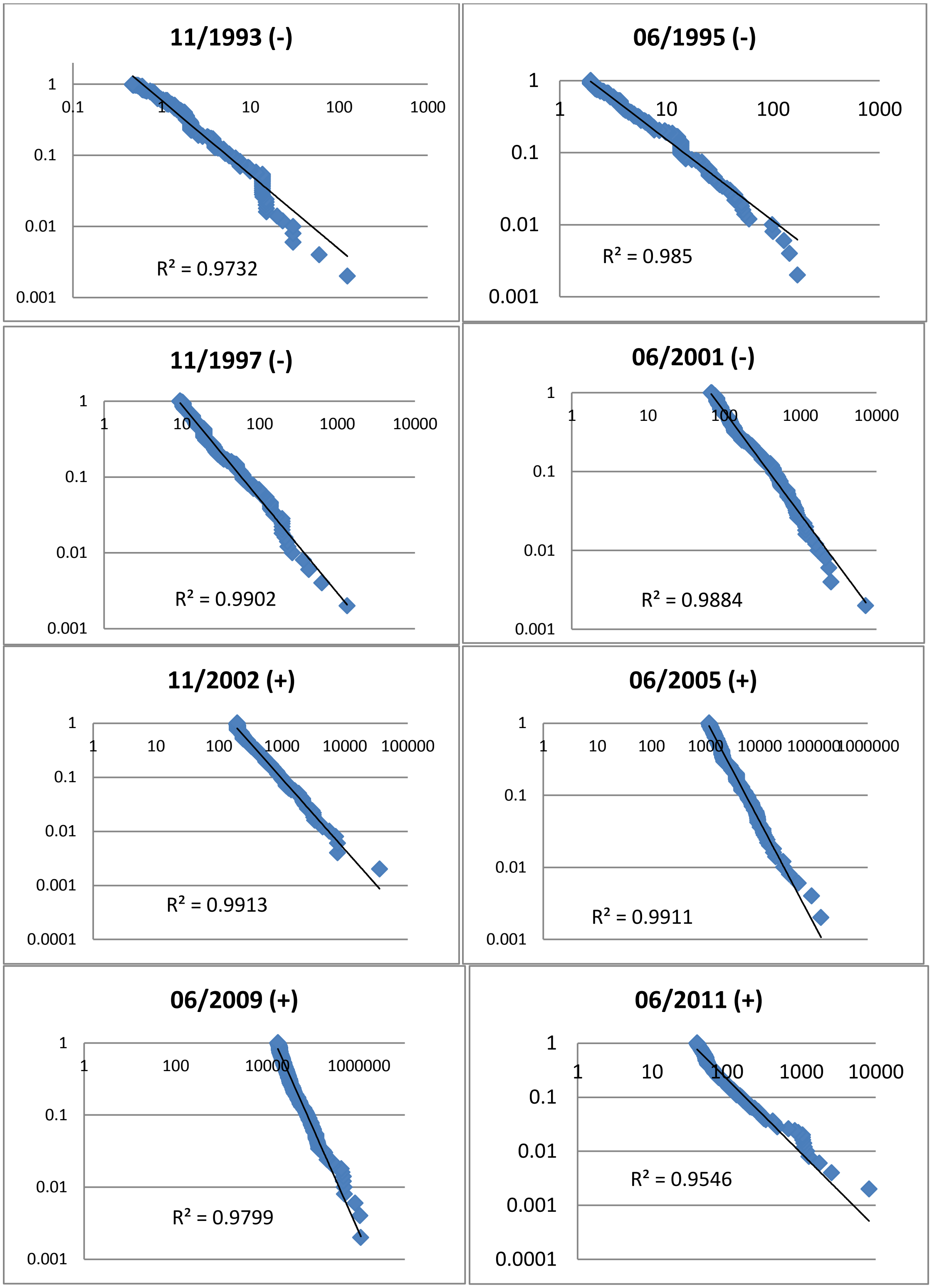
A power-law for supercomputers

Figure 1 relates the cumulative distribution of the world's top performing 500 supercomputers [12] with their respective performance (R_{max} measured in Gflop/s, LINPACK Benchmark [13]).[2] The figures align the computing performance in Gflops on the horizontal x-axis and the cumulative share of supercomputers with certain performance levels on the vertical y-axis. We find an inverse linear log-log relation, with (exponentially) many supercomputers of (exponentially) low performance (the so-called "fat-tail") and (exponentially) few supercomputers of (exponentially) high performance (so-called "large events"). The power-law trend lines fit the measurements very well. Figure 1 only shows eight examples of the 38 available datasets (biannual registries between June 1993 and November 2011 [12]). The good fit also extends to the rest of the data. Over the entire spectrum of 500 supercomputers, almost half of the 38 R^2 values reach a value of 0.99, nine cases with R^2 values of each 0.98 and 0.97, and one case with each 0.96 and 0.95.

f(x) = C*x-α, instead of p(x) = C*x-α. For example, the so-called allometric scaling laws in biology [8] relate body mass and metabolic rate by a power-law (not mass and the probability that a certain mass occurs), while other bivariate social power-laws relate economic production and energy consumption [9], etc.

[2] The performance benchmark is set on solving a dense system of linear equations and measures a computer's floating-point rate of execution [13]. By measuring the actual performance for different problem sizes n, one gets the maximal achieved performance R_{max}.

Figure 1: Samples of CDFs of 500 top supercomputers (y-axis) versus their performance in R_{max} (x-axis). Considering the results from Table 1, (+) and (-) marks likely (authentic) and unlikely (false) power-laws respecively.



Source: author, based on [12].

It has been shown that the visual evaluation or the least-square (R^2) test can be deceptive in the identification of authentic power-laws [14]. Therefore, Clauset, Shalizi and Newman [3] developed a statistical test for the authenticity of rank-frequency power-laws.[1] First they estimate the derived parameters x_{min} and α for the respective power-law model ($f(x) = x^{-\alpha}$).[3] Then they use the method of maximum likelihood to fit a power-law distribution to the observed tail data. With the help of x_{min}, the scaling exponent α can be numerically calculated (not fitted, as in the trendlines in Figure 1).[4] These variables are then used to create a large number of synthetic datasets from a true power-law distribution (1000 in our case). Each of these true power-laws is then compared with the empirical dataset (the Kolmogorov-Smirnov statistic is used) (in line with [14]). This provides a goodness-of-fit test, which generates a p-value that quantifies the plausibility of the power-law hypothesis. The p-value is defined to be the fraction of the synthetic distances that are larger than the empirical dataset, with small p values indicating that it is not a plausible that a true power-law logic created the data (we choose $p>0.05$ for significant* and $p>0.1$ for very significant** results).

Table 1 shows the result of this test for the 38 datasets. Only the second half of the registered period, starting around 2002, makes a quite strong case for the power-law under the stringent test requirements. The test tells us that it is very likely that since 2002 the best performing supercomputers of the tail of the distribution have started to line up in an authentic power-law distribution.[5]

From a theoretical perspective, it is interesting to note the contradictions between the R^2 curve-fitting exercise (Figure 1) and the results of the generative tests [in line with 14, 3]. For example, the R^2 in Figure 1 suggest a better fit of the power-law curve for 11/1997 ($R^2 = 0.9902$) and 06/2001 ($R^2 = 0.9884$) than for 06/2009 ($R^2 = 0.9799$) and 06/2011 ($R^2 = 0.9546$), while the significance test argues in favor of 2009 and 2011 as a power-laws, and rejects 1997 and 2001 as good power-law fits (Table 1). The correlation coefficient between the 38 R^2 (Figure 1) and the p-values (Table 1) is even weakly negative ($\rho_{R2,p} = -0.18$). While a comparative analysis among years has to consider the different sample sizes, which affect the statistical significance test (see Table 1), such comparison between methods underlines the fact that simple trend-line fitting exercises can be deceptive (see (+) and (-) signs in subtitles of Figure 1).

[3] Typically the power-law is most prominent in the "tail" of the distribution (see e.g. [15]), above some minimum value x_{min}. Instead of choosing x_{min} visually, they followed [16] and choose the value of x_{min} that makes the probability distributions of the measured data and the best-fit power-law model as similar as possible above x_{min} (following the Kolmogorov-Smirnos statistic to measure the distance between both probability distributions).

[4] Assuming that the empirical data is drawn from a distribution that follows a power-law exactly for x>xmin, it turns out that the maximum likelihood estimator is $\alpha = 1+ n[\Sigma \ln(x/ xmin)^{-1}$.

[5] It is important to point out that the statistical significance test is sensitive to sample size, which is defined by xmin that defines the length of the power-law tail. Our sample sizes are relatively small (average of 133 valid data points for the years identified as power-laws).

Table 1: Testing empirical datasets against synthetically generated true power-laws

Date	Valid number of samples in tail (out of 500)	α for valid samples of PDFs (± SD) $\alpha_{CDF} = \alpha_{PDF} - 1$	Significance test, with p>0.05 = *; and p>0.1 = **
06/1993	248	2.2101±0.1541	0.0000
11/1993	467	1.9704±0.1206	0.0000
06/1994	500	2.2024±0.3410	0.0000
11/1994	500	2.0228±0.0691	0.0000
06/1995	446	2.0898±0.1062	0.0000
11/1995	499	2.2122±0.0545	0.0000
06/1996	500	2.2228±0.0632	0.0000
11/1996	500	2.2639±0.0642	0.0000
06/1997	274	2.1944±0.1080	0.0010
11/1997	500	2.3262±0.0895	0.0000
06/1998	436	2.2826±0.0865	0.0000
11/1998	500	2.3578±0.1108	0.0000
06/1999	242	2.1940±0.1466	0.0000
11/1999	***199***	***2.1397±0.1217***	***0.1390***
06/2000	206	2.0704±0.1855	0.0090
11/2000	214	2.2302±0.1106	0.0150
06/2001	500	2.3667±0.1497	0.0000
11/2001	384	2.3300±0.1105	0.0230
06/2002	***99***	***2.5944±0.2218***	***0.2780***
11/2002	***159***	***2.3275±0.1308***	***0.8780***
06/2003	***249***	***2.6254±0.1353***	***0.1530***
11/2003	***379***	***2.5102±0.0930***	***0.1820***
06/2004	442	2.4825±0.0810	0.0120
11/2004	500	2.6235±0.1174	0.001
06/2005	***56***	***2.5349±0.1804***	***0.6620***
11/2005	***68***	***2.3624±0.2142***	***0.6740***
06/2006	***90***	***2.3741±0.1902***	***0.7320***
11/2006	***92***	***2.3921±0.1933***	***0.2500***
06/2007	***121***	***2.3408±0.2425***	***0.4090***
11/2007	***154***	***2.4747±0.2663***	***0.0640****
06/2008	***67***	***2.3371±0.2691***	***0.8500***
11/2008	***60***	***2.3421±0.2704***	***0.6880***
06/2009	***136***	***2.2881±0.2009***	***0.7000***
11/2009	***140***	***2.2397±0.2221***	***0.0530****
06/2010	***81***	***2.2757±0.2076***	***0.7280***
11/2010	***123***	***2.1740±0.1163***	***0.5860***
06/2011	***138***	***2.2419±0.1476***	***0.6200***
11/2011	***122***	***2.1727±0.1546***	***0.3970***

Source: based on [12] and the MATLAB implementation of [3] available at: http://tuvalu.santafe.edu/~aaronc/powerlaws.

Fits as good as in Table 1 are exceptionally rare in the social sciences. To use the words of the Nobel-prize winning economist Paul Krugman [17, p. 40]: "we are unused to seeing regularities this exact in economics - it is so exact that I find it spooky".[6] What is the generative mechanism that leads to the fact that the free-wills of a bunch of interdependent agents get "in line" with some general emergent property that subscribes the U.S. Department of Energy to order one supercomputer with performance x, and Los Alamos Laboratory, IBM and a couple of universities additional ones with exactly lesser performance $x^{-\alpha}$, etc.?

Exploring possible generative mechanisms

One of the challenges of power-law science consists in identifying the right generative mechanism of the found distribution. Currently we know of at least 15-20 different scale-free theories that result in power-laws [22, 23]. Some of them are intriguing and it is tempting to declare one of these readily available theories to be the driver of a newly found power-law. Unfortunately, the fit of available theory to existing data is often rushed or done in a qualitative superficial manner, leading to much confusion among students of power-laws. In the following we will critically evaluate the fit of our case to some of the most prominent power-law generative mechanisms.

Do computational tasks follow preferential attachment?

A widely cited generative mechanism for power-laws is preferential attachment [24, 25]. The underlying process consists of a "rich-get-richer" mechanism, whereas some quantity is distributed among a number of objects according to how much they already have (e.g. links or citations). In our case, the quantity is computational capacity. This would imply that the largest supercomputers receive new computational tasks according to how much computational tasks they already carry out. This would imply that the largest problems are assigned to the handful of labs that maintain the most powerful supercomputers ("central nodes"), while the rest takes on an exponentially smaller problems, and therefore require respectively less computational power. While supercomputers are not directly connected to each other, the underlying market place for this dynamic could be the distribution of financial resources for promising research projects, based on the existing workload and reputation of the respective institution. Better computation and respective results attract more financial support, which leads to buying more powerful computers, resulting in a feedback mechanism typical for preferential attachment. In order to verify the validity of this potential explanation we would require micro-data on funding mechanisms and the choice of computation-intensive problems, which we do not have available. However, since the fragmented landscape of supercomputer tasks consists of very diverse and isolated areas (including physics, meteorology, governments, financial markets and retailers), and since financing is surely heavily influenced by the kind of task, the explanation of a far-reaching preferential attachment mechanism seems questionable. In order to make the proposed kind of preferential attachment work, new computational

[6] Krugman makes his often-cited statement with regard to the classical power-law distribution of city-sizes, which often display even weaker relations than the ones presented here: for example, for city-sizes, [18] find R^2 between 0.81 and 0.95; [19] find an R^2-value of 0.92; [20] find an R^2-value of 0.90, and [21] finds an R^2 of 0.986.

tasks would be exclusively guided by the distribution of existing computational power. In reality, however, it is very unlikely that a financial retailor would finance the expansion of the computational power of a meteorology lab, simply because they have the right amount of computing power. In short, while we cannot discard it, pure preferential attachment seems unlikely to be the main generative mechanism of our newly found power-law.

Has the world's leading computational power started to self-organize?

The attention grabbing self-organized criticality [26] holds that non-linear open systems (like earthquakes, solar flares, traffic jams and stock markets) arrange themselves in a critical state away from equilibrium, whereas the emergent structure intermediates some kind of constant tension (or flow) that affects them. An emerging power-law distribution constantly mediates this flow in the open system (such as the typical "avalanches" on a "sand pile"). Adopting this generative mechanism to our case we could hypothesize the arrival of computationally intensive tasks as a constant flow ("grains of sand") and the avalanches as the supercomputer projects that process accumulated flows of computational problems. While it is fascinating to apply this mechanism to our case, the devil in the detail makes this explanation unlikely. In order for self-organized criticality to emerge, the overall system ("grid of sand patches", or the supercomputer landscape) would have to be able to indiscriminately distribute (randomly similar) incoming sub-problems ("incoming grains of sand") between different supercomputers in order to alleviate the arising tension ("sand avalanches"). In contrary to sand-piles with randomly similar sand grains, the supercomputer landscape consists of a myriad of different tasks and sub-tasks. For this theory to apply, one would have to explain how randomly incoming (often highly specialized) computational problems could arbitrarily be shifted between labs of physicists, financial banks, meteorologists, retailers, and national defense authorities, among others. Until the details of such relations cannot be mapped one-to-one to the requirements of the sand-pile model, self-organized criticality has to be rejected as a potential generative mechanism of our power-law.

Is the world's leading computational power designed for optimized tolerance?

While self-organized criticality contains no element of design or planning, another intriguing generative mechanism shows that power-laws can also arise as the result of an engineering process of robustness against unanticipated dangers [27, 28]. Applying so-called highly optimized tolerance to our case, we could hypothesize that the power-law distribution of supercomputers is the result of an optimization process that safeguards against random epidemic failure of the world's most advanced computational capacities, such as caused by viruses or random errors. While such global "super-organizer" that consciously assigns the supercomputers to institutes certainly does not exist, this optimization could be the outcome of natural social selection. This would imply that the regular systemic failure of the blind system naturally shapes the distribution by survival of the fittest (sorting the fittest out through elimination). Notwithstanding this possibility, there is no evidence for such eliminatory forces of natural selection among supercomputers and their institutions.

Do we randomly observe exponentially growing technology?

In spoiling contrast to the rather speculation-inviting previous explanations, two mathematical statisticians came up with a comparatively unspectacular generative mechanism for power-laws [29]. Reed and Hughes explain that the reason "why power-laws are so common in nature" (p. 067103-1) is that time between the observations of different exponentially growing subjects is inversely exponentially distributed. In our case, the performance of the supercomputers of N institutes does grow exponentially, doubling every period Δt, following Moore's law [30]. Now imagine we flip a separate fair coin per institute at every period Δt and randomly assess the performance of the supercomputer of this institute when heads turns up. On average, we will sample half of the institutions on the first round of our assessment (N*0.5) and the performance of these computers will have doubled once (P*2). In the second round we will observe a quarter of the original amount of institutions (N*0.5*0.5) and find computers that already obtained four-fold performance (P*2*2), etc., leading to two inverse exponentials ([0.5,2]; [0.25;4]; [0.125;8]; [0.0625;16]; etc.), representing a power-law. In our case, however, we do have a complete registry of the top-500 supercomputers, not a random sample. Therefore, the pure form of this sampling explanation will not work as a complete explanation.

A new generative mechanism: progress and diffusion

A look at the micro-data of our datasets reveals a new generative mechanism. One of the key realization when thinking about the generation of power-laws consists in searching for two simultaneous exponential processes. Taking the log of both leads to the notorious linear log-log relationship. The first of the two exponential processes in this case is the performance of supercomputers, which increases exponentially through incessant technological progress. As mentioned before, the exponential nature of technological progress can be explained by Moore's law [30], which consists of continuous innovation through learning by scaling ([31]; in this case, a continuous learning process of putting more transistors on a microchip). Empirically, it turns out that the growth rate of the average supercomputer performance (in R_{max}) has been 33 % per semester (doubling every 1.2 years), therefore being considerably faster than the 2 year doubling rate predicted by Moore.

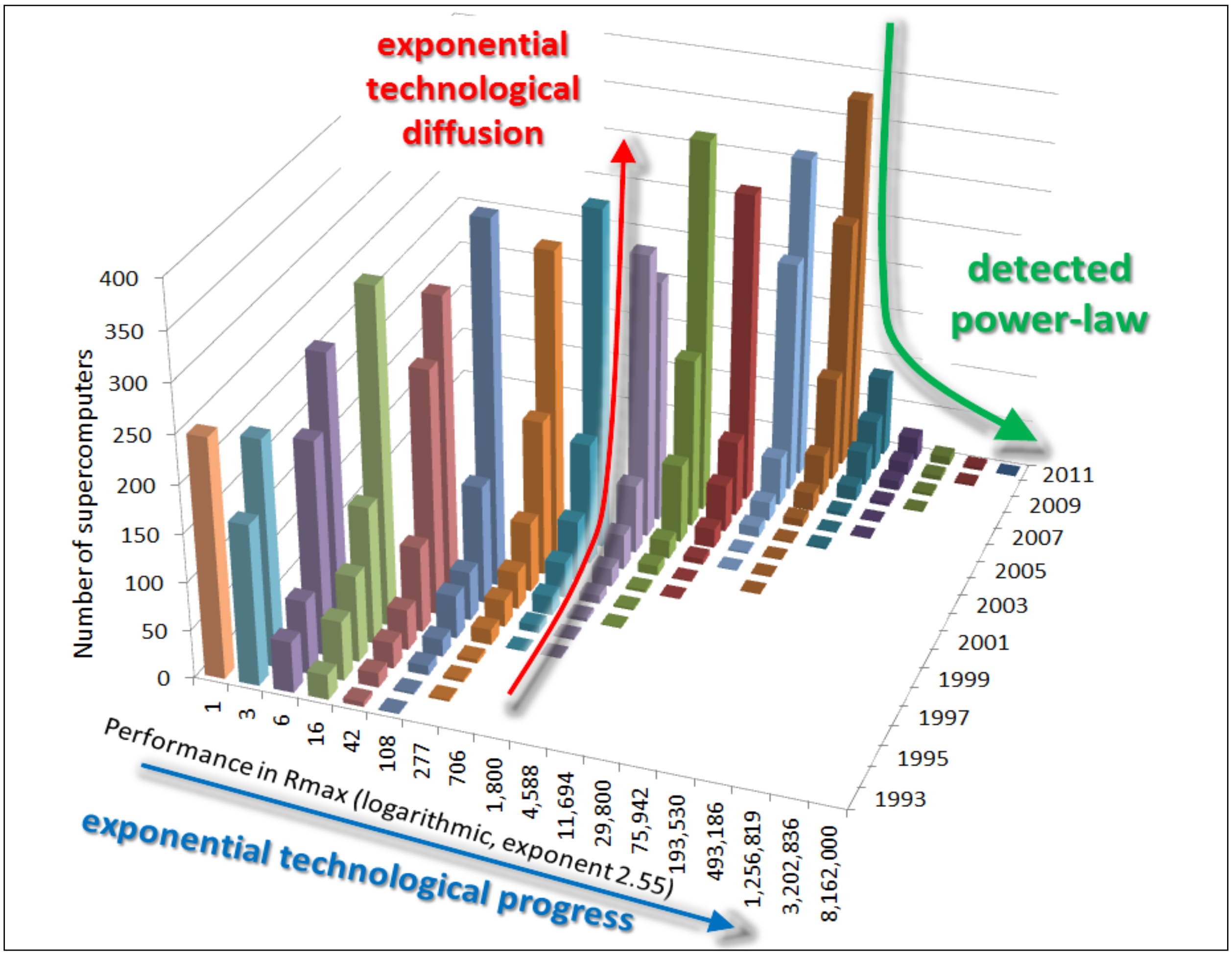
The second exponential process consists in the exponential diffusion of technology through social networks [32-34]. In its simplest form, the diffusion process (be it a disease, meme, rumor or innovation) can be modeled on a homogeneous grid network with a constant rate of contagion. This means that the first adopter "infects" x new adopters, each of whom has a contagious effect on x additional adopters, who will adopt and infect others, etc. The resulting logistic S-shaped diffusion function is exponential before its inflection point. Non-constant contagion rates through more intricate network structures result in a less smooth, but nonetheless exponential diffusion processes.[7]

Figure 2 exemplifies how both exponential processes can align to the detected power-laws for a given year. The Figure uses logarithmic binning of the PDF (not the CDF used in Figure 1), which

[7] After the inflection point, the diffusion S-curve is exponentially decreasing, which should lead to another power-law in the other tail of the distribution. Unfortunately our source only tracks the top-500 supercomputers, which does not allow us to look into the lower part of this distribution.

facilitates a visual interpretation of the dynamic.[8] A generation of supercomputers of a certain performance is introduced in a given year (e.g. the first supercomputer with Rmax between 706 and 1,800 appeared in 1997). In consecutive years, this generation of supercomputers diffused through society. By 2004, 299 of the world's top-500 supercomputers were of this performance level (see arrow "exponential technological diffusion" in Figure 2). One year later, the average computational performance of supercomputers had grown exponentially, and a new generation of supercomputers with exponentially higher performance started to diffuse through society, (see arrow "exponential technological progress"). As a result of the interplay between both exponential processes we find a power-law distribution between the number of supercomputers and their respective performance in any given year (e.g. see arrow "detected power-law" for 2011).

Figure 2: Exponential technological progress (from left to right) and exponential adoption of technology (from bottom up). Binned PDF.



[8] Logarithmic binning of exponent 2.54837 has been used for performance (see lower axis), which is the empirical result of binning the top-500 supercomputers into 5 exponential bins, averaged over the 38 datasets: SUM[(performance$_{min}$ / performance$_{min}$)^(1/5)]/38. As previously mentioned the bin-size is somewhat arbitrary and aggregates information, reason why the more solid mathematical analysis has been carried out with the CDF (Figure 1, Table 1).

Maintining this aligment implies that the rate of contagion (the spread of supercomputers with the same performance; upward in Figure 2) and the rate of technological progress (the average increase in performance; toward the right in Figure 2) need to step in synchrony, more explictly, in exponential synchrony. This logic is visaualized in Figure 3. The straight line of year t is reflected onto the straight line of year t+1 if the slope changes over the entire range. This implies that $\frac{\Delta \mathrm{q}}{\Delta \mathrm{p}} = \frac{\Delta \mathrm{Q}}{\Delta \mathrm{P}}$ over all the chosen bins (see Figure 3):[9]

$$\frac{\log\left({}^{quantity_{t+1}}\!/_{quantity_t} \,\middle|\, for\ given\ performance\ bin \right)}{\log\left({}^{performance_{t+1}}\!/_{performance_t} \,\middle|\, for\ respective\ quantity\ bin \right)} \qquad (1)$$

$$= \frac{\log q_{t+1} - \log q_t}{\log p_{t+1} - \log p_t} = \frac{\log\ \text{ɋ}^{\,y+\Delta Q} - \log\ \text{ɋ}^{\,y}}{\log\ \text{ƥ}^{\,x+\Delta P} - \log\ \text{ƥ}^{\,x}} = \frac{(y+\ \Delta Q) * \log_{\text{ɋ}} \text{ɋ} - y * \log_{\text{ɋ}} \text{ɋ}}{(x+\ \Delta P) * \log_{\text{ƥ}} \text{ƥ} - x * \log_{\text{ƥ}} \text{ƥ}} = \frac{\Delta Q}{\Delta P}$$

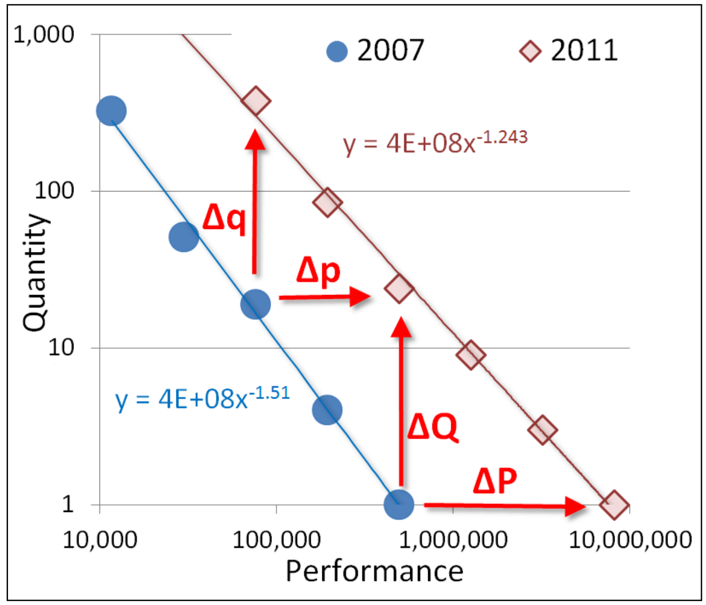
$$= constant\ for\ all\ bins$$

With ɋ and ƥ being the respective bases of exponential growth of quantity and performance. In equation (1), the ratio of quantities is the rate of contagion of the diffusion of a technology with a given level of performance, and the ratio of performances is an indicator of technological progress. In practice, the exactness of this synchrony depends on the goodness of fit of the power-law within specific bins. Using the same binning as in Figure 2 on the supercomputer data,[8] the standard deviation of the yearly average[10] of equation (1) is only half as large for the years in which the test of Clauset, et al. [3] identified an authentic power-law (see Table 1) than for those that fail the test (SD: 0.31 vs. 0.62). This shows that the power-law was maintained more stabiliy after it first got established.

[9] Since the straight line plot is logarithmic, we divide the endpoints of the arrows in Figure 3 (instead of substracting them) and take the logarithm.

[10] Over all bins that fall into the identified power-law tail, see Table 1.

Figure 3: Maintaining the progress-diffusion power-law for supercomputers between 2007 and 2011 requires (exponentially) proportional shifts over the entire distributions (PDFs[8] of tail above x_{min}).

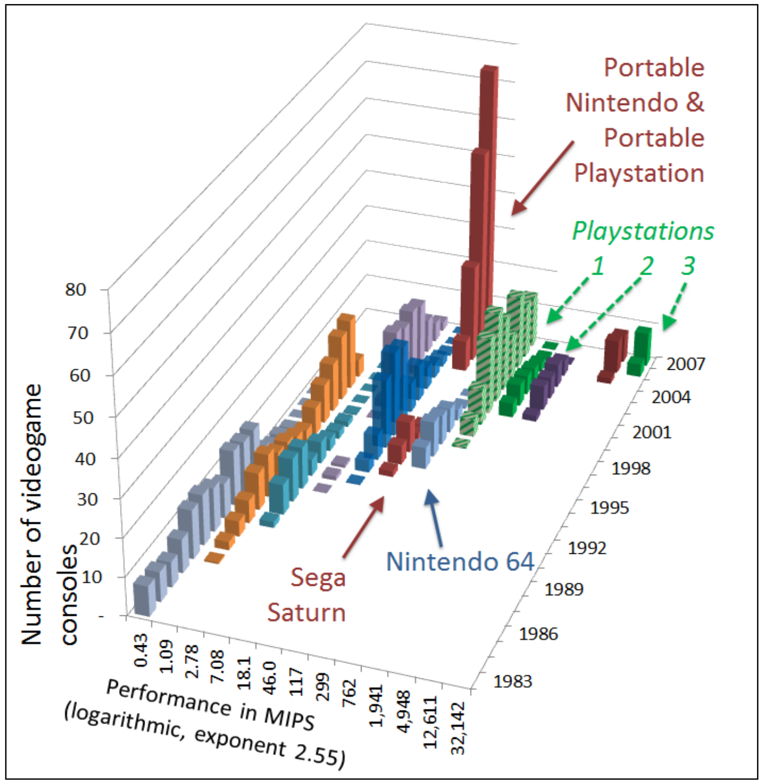


Source: own elaboration, based on [12].

Discussion

It is important to point out that this emerging power-law between exponentially increasing performance and exponential diffusion, stable as it is, is neither automatic, nor the deterministic result of every interplay between exponential progress and social diffusion. We do not find a perfect power-law alignment for earlier years of our dataset (see Table 1). Exploring potential reasons for this, it is helpful to investigate other dynamics of exponential progress and diffusion. Figure 4 shows the equivalent of Figure 2 for videogame consoles. The computational performance of videogame consoles also grows exponentially, and they diffuse throughout society through social networks. However, in this case, it is even visually clear that no power-law emerges between the computational capacity and the number of consoles. The landscape of videgame consoles is too segmented and their performance related to too many other variables than computational power. For example, during the 1990, Sega's Saturn pushed the social diffusion of 50 MIPS consoles. After the Nintendo 64 with 125 MIPS won the race, the 50 MIPS consoles died out, only to be resurected with a vengance in the late 2000s, when Nintendo and Playstation brought their portable solutions to market. At the same time, Playstation 2 could not hold up with the diffusion rate of the much more succesful Playstaion 1, and Playstaion 3 came to market relatively late. In contrary to the supercomputer market, computational power alone might not be the right indicator to measure these developments in the videogame console market. These The particular characteristics of supply markets and demand in the market impede the emergence of the synchronized power-law in this fragmented case. It seems that a progress-diffusion power-law requires (a) having a large and fine-grained enough diffusion phenomenon; (b) finding the right performance indicator (in contrast to videogame consoles, for supercomputers computational power undoubtedly is the main performance indicator), (c) having a mechanism that is little distorted by other mechnisms than the technological progress of the chosen performance indicator.

Figure 4: Number and performance of videogame consoles (PDF, same binning as in Figure 2).



Source: own elaboration, based on Appendix C.13, pp.96-100 of [39].

Looking for reasons why the authenticity test from Table 1 detected more authentic power-laws during the later decade, the data reveal that a larger share of the world's leading supercomputers is replaced instead of maintained in later years, resulting in a smoother alignment between diffusion and progress: during 1993 – 2002, on average on average 42 % of the top 500 supercomputers were new, in contrast to 59 % between 2002-2011. This means that in later years even the majority of models with lower performance are newly aquired, recreating the power-law distriubtion for each year anew, based on new computers (see Figure 2). The yearly replacement of older models with similar performing models of up-to-date technology might just be enough to align the number of supercomputers in a given year with the speed of technological progress even more closely.

Last but not least, let us ask the "so what" question. Why does it matter to discover the kind of power-law we found? The emerging alginment leads to a longstanding and ongoing debate over the origin of innovation as the result of so-called "technology push" (e.g. technology supply through Research & Development) or "demand pull" (technological diffusion demanded by consuption patterns) (e.g. [37, 38]. Despite innumerable decisions of free-willed individuals, we have shown that both supply and demand of technology do align in an almost “spooky” regularity as the result of exponential technological progress and exponential diffusion. This leads to an empirically identifable exponential synchronization of technology push and demand pull (see Figure 3). In cases where this process holds for a specific product (which is not necessarily the case, as shown in Figure 4), economic agents are able to appoximate one through the other for a specific technology. In this sense, the newly found power-law can lead to important insights in the otherwise quite unpredictable market dynamics of today’s fast and violent technological change. This can be useful for producers of technology, who can plan for the magnitude and timing of supply, and for consumers of technology, who can prepare their institutions for upcoming demands, as well as for policy makers and regulators of technology markets.

This all being said, despite the nice fit between the theory and the data and its promising insight, a strict Popperian sense of falsifiability still does not allow us to declare deterministic exclusivity on the relation between the proposed mechanism and the empirical data [35]. The progress-diffusion logic is sufficient, but not necessary to produce the identified power-laws. Once additional data becomes available, we might have to revise the proposed theory. For example, it has long been established that a mixture of individual distributions with different scale parameters can result in power-laws [36]. In the future, it might be revealed that the detected pattern is an intricate combined effect of several different distributions, for example, containing aspects of preferential attachement, self-organized criticiality, and the progress-diffusion dynamic. We cannot find evidence for this in our currently availabe data and while we know that the presented simple logic is sufficient, we can also never reject the possibility of another generative mechanism entirely.

References

1. Andriani P, McKelvey B. Beyond Gaussian averages: redirecting international business and management research toward extreme events and power laws. *J Int Bus Stud*. 2007;38(7):1212–1230.
2. Mandelbrot BB. "New Methods of Statistical Economics," revisited: Short versus long tails and Gaussian versus power-law distributions. *Complexity*. 2009;14(3):55–65.
3. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. *SIAM Review*. 2009;51:661–703.
4. Mitzenmacher M. A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*. 2003;1(2):226–251.
5. Newman M. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*. 2005;46(5):323.
6. Stumpf MPH, Porter MA. Critical Truths About Power Laws. *Science*. 2012;335(6069):665–666.
7. Shalizi, C. R.. So You Think You Have a Power Law — Well Isn't That Special? *Three-Toed Sloth*. 2007, June 15, from http://bactra.org/weblog/491.html
8. West GB, Brown JH, Enquist BJ. A General Model for the Origin of Allometric Scaling Laws in Biology. *Science*. 1997;276(5309):122–126.
9. Brown, J. H., Burnside, W. R., Davidson, A. D., Delong, J. R., Dunn, W. C., Hamilton, M. J., … Zuo, W. Energetic Limits to Economic Growth. BioScience, 2011: 61(1), 19–26.
10. Adamic L. Zipf, Power-law, Pareto - a ranking tutorial. 2000. Available at: http://www.hpl.hp.com/research/idl/papers/ranking/. Accessed May 28, 2010.
11. Zipf GK. *Human behavior and the principle of least effort: an introduction to human ecology*. Addison-Wesley Press; 1949.
12. TOP500. *TOP500 List Releases of supercomputers since June 1993*. TOP500.org; 2012. Available at: http://www.top500.org/lists. Accessed May 28, 2010.
13. Dongarra JJ, Luszczek P, Petitet A. The LINPACK benchmark: Past, present, and future. Concurrency and Computation: Practice and Experience. *Concurrency and Computation: Practice and Experience*. 2003;15:803–820.
14. Goldstein ML, Morris SA, Yen GG. Problems with fitting to the power-law distribution. *Eur Phys J B*. 2004 Sep 1;41(2):255–8.
15. Li W. Analyses of baby name popularity distribution in U.S. for the last 131 years. *Complexity*. 2012;18(1):44–50.
16. Clauset A, Young M, Gleditsch KS. On the Frequency of Severe Terrorist Events. *Journal of Conflict Resolution*. 2007;51(1):58–87.
17. Krugman PR. *The Self-Organizing Economy*. Blackwell Publishers; 1996.
18. Rosen KT, Resnick M. The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics*. 1980;8(2):165–186.
19. Song S, Zhang H. Urbanisation and City Size Distribution in China. *Urban Stud*. 2002;39(12):2317–2327.
20. Shiode N, Batty M. Power Law Distributions in Real and Virtual Worlds. In: *http://www.isoc.org/inet2000/cdproceedings/2a/2a_2.htm.*; 2000. Available at: http://www.isoc.org/inet2000/cdproceedings/2a/2a_2.htm. Accessed May 28, 2010.

21. Gabaix X. Zipf's Law For Cities: An Explanation. *Quarterly Journal of Economics*. 1999;114(3):739–767.
22. Andriani P, McKelvey B. Perspective—From Gaussian to Paretian Thinking: Causes and Implications of Power Laws in Organizations. *Organization Science*. 2009. Available at: http://orgsci.journal.informs.org/content/early/2009/10/07/orsc.1090.0481.
23. Hunt A. Relevance of percolation theory to power-law behavior of dynamic processes including transport in disordered media. *Complexity*. 2009;15(2):13–27.
24. Yule GU. A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S. *Phil Trans R Soc Lond B*. 1925 Jan 1;213(402-410):21–87.
25. Barabasi A-L, Albert R. Emergence of scaling in random networks. *Science*. 1999;286(5439):509–512.
26. Bak P. *How Nature Works: The Science of Self-Organized Criticality*. 1st ed. Springer; 1999.
27. Carlson JM, Doyle J. Highly optimized tolerance: A mechanism for power laws in designed systems. *Physical Review E*. 1999;60(2):1412–1427.
28. Carlson J, Doyle J. Highly optimized tolerance: robustness and design in complex systems. *Physical Review Letters*. 2000;84(11):2529–32.
29. Reed WJ, Hughes BD. From gene families and genera to incomes and internet file sizes: Why power laws are so common in nature. *Phys. Rev. E*. 2002;66(6):067103.
30. Moore GE. Lithography and the future of Moore's law. In: *Proceedings of SPIE (international society for optics and photonics)*. Santa Clara, CA, USA; 1995:2–17. Available at: http://spie.org/x648.html?product_id=209195.
31. Sahal D. Technological guideposts and innovation avenues. *Research Policy*. 1985;14(2):61–82.
32. Rogers EM. *Diffusion of Innovations, 5th Edition*. New York: Free Press; 2003.
33. Valente TW. *Network Models of the Diffusion of Innovations*. Hampton Press (NJ); 1995.
34. Hilbert M. The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making. *Telecommunications Policy*. 2011;35(8):715–736.
35. Popper KR. *The Logic of Scientific Discovery*. Routledge; 2002.
36. Farmer JD, Geanakoplos J. Power laws in economics and elsewhere. Santa Fe: *SFI WP*; 2008. Available from: http://www.santafe.edu/~jdf/papers/powerlaw3.pdf
37. Schmookler J. *Invention and economic growth*. Harvard University Press; 1966.
38. Di Stefano G, Gambardella A, Verona G. Technology push and demand pull perspectives in innovation studies: Current findings and future research directions. *Research Policy*. 2012;41(8):1283–1295.
39. Hilbert M, López P. The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*. 2011;332(6025):60 –65.