

Rich get richer?

Autor: Matheus Oliveira

Trabalho apresentado ao
Departamento de Economia da
Universidade de Brasília (UnB) sob
orientação do Prof. Bernardo Mueller
como requisito parcial a obtenção do
grau de bacharelado em economia

Abstract

O presente trabalho replica o artigo de Di Guilmi *et al* (2003) no objetivo de responder se a distribuição relativa de GDP pode ser dita compatível com um *Power law*. Não obstante, o trabalho busca fundamentação econômica para a suposta regularidade e explicita as principais motivações e intuições por trás do fenômeno tratado. A conclusão é de que os dados utilizados não podem ser ditos não compatíveis com um *Power law* e o principal candidato a mecanismo gerador é uma versão modificada da Lei de Gibrat. O método utilizado para estimar e testar os coeficientes é o proposto na literatura como robusto em contraposição ao método gráfico de validação de *Power laws*.

Palavras chave: *Power laws*, Invariância de escala, Distribuição relativa de GDP, Lei de Gibrat, Crescimento econômico.

Conteúdo

1	Introdução	2
2	<i>Power Laws</i> (PLs)	5
2.1	Forma geral de um PL	5
2.2	Caso particular: <i>Zipf's Law</i> e exemplos em economia	9
2.3	Processo gerador de um PL	12
3	Evidência empírica	15
3.1	Método	15
3.2	Resultados	16
3.3	Teste de compatibilidade do PL	18
4	Possíveis hermenêuticas econômicas do PL	21
5	Conclusões	22

1 Introdução

Em seu trabalho seminal *An inquire into the nature and causes of the wealth of nations*, Adam Smith (1776) argumentou em contraposição a visão mercantilista dominante à época, concluindo que a origem da riqueza de uma nação era dada pela produtividade da força de trabalho e não pela disponibilidade de metais preciosos. A intuição de Smith, apesar de não fornecer uma resposta completa, permitiu o avanço e o desenvolvimento de um campo de pesquisa focado em explicar a dinâmica da riqueza mundial e como essa riqueza evolui.

Seguindo o esforço de Kaldor (1957), diversos pesquisadores apontaram regularidades empíricas sobre crescimento econômico e diversos outros tentaram formular modelos que fossem compatíveis com os chamados fatos estilizados de crescimento¹. Entre os pesquisadores que propuseram regularidades empíricas concernentes à dinâmica da riqueza global, Parente e Prescott (1993) se pautam na análise de GDP per capita no período de 1960 a 1985 para concluir que: i) em todos os anos da análise existe uma grande disparidade de riqueza entre os países, ii) a disparidade de riqueza entre os países permaneceu praticamente a mesma no período de interesse e, iii) a distribuição de riqueza, de fato, deslocou-se para cima, sugerindo a não existência de uma armadilha de pobreza absoluta. Em relação a grande disparidade de renda entre os países, autores como Pritchett (1997) e Ben-David (1994) reportam conclusões semelhantes.

Quatro anos depois, Jones (1997) propõe a utilização do GDP por trabalhador para apontar fatos e regularidades empíricas quanto a distribuição relativa de riqueza entre os países. De acordo com o autor, o uso do GDP per capita subestimaria a real riqueza de países onde a produção fora do mercado fosse significativa². Não obstante, a medida parece ser mais exata para medir riqueza, conforme o que foi sugerido por Smith (1776).

As principais proposições de Jones (1997) apontam, assim como em Parente e Prescott (1993), uma grande disparidade entre os países. O autor também afirma que a distribuição passou a ter “dois picos” em 1988 se opondo à distribuição grosseiramente normal de 1960³. Em termos práticos, esse movimento reflete um deslocamento de massa do centro da distribuição para as extremidades, tornando as caudas mais grossas. Nas palavras do autor, a distribuição de renda para os países analisados é mais ampla em 1988 do que em 1960, tanto na parte correspondente a menores níveis de renda, quanto na parte correspondente a maiores níveis.

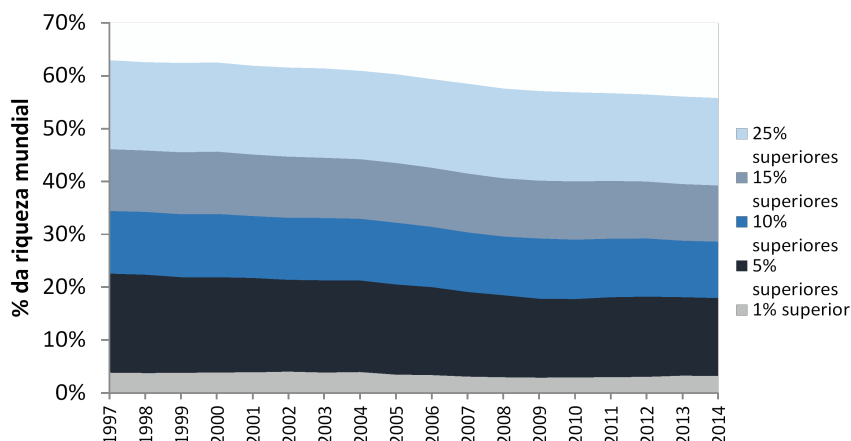
Com relação aos anos mais recentes, os dados de 169 países da Penn World Table (PWT) 9.0 (Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer, 2015) revelam que em 2014 as 25% observações superiores continham cerca de 56% da riqueza mundial medida em termos de GDP por trabalhador. Na figura I.1 é possível observar como a parcela de riqueza correspondente aos 1%, 5%, 10%, 15% e 25% superiores evoluiu no período de 1997 – 2014. Ainda, é possível observar que a disparidade de riqueza entre os países comentada anteriormente persiste no período.

¹Os fatos estilizados de Kaldor (1957) foram extremamente relevantes, por exemplo, para o desenvolvimento do modelo neoclássico de crescimento proposto por Robert M. Solow em 1970. Não obstante a relevância do modelo em compatibilizar uma formalização teórica com a evidência empírica, a proposição de Solow se mostrou relevante também para explicar flutuações em *business cycles* e explicitar consequências de políticas econômicas sobre a riqueza dos países (Parente e Prescott, 1993).

²De acordo com o glossário de termos estatísticos da OECD, a produção fora do mercado seria composta por transações envolvendo serviços ou produtos fornecidos a preços nulos ou não economicamente significantes. Disponível em: <https://stats.oecd.org/glossary/detail.asp?ID=5661>, último acesso em 24/03/2017

³Seguindo o termo *twin-peaks* cunhado por Quah (1993a) para descrever o formato da distribuição de renda entre os países.

Figura I.1 - Evolução da porcentagem de riqueza correspondente às observações superiores



Segundo Clauset, Shalizi e Newman (2009) é comum observar que diversas variáveis de interesse se aglomeram em torno de uma determinada média e, comumente, se situam com porcentagem negligenciável de probabilidade longe do valor típico, tornando, assim, a média um valor representativo da distribuição. Contudo, as considerações de Jones (1997) e Parente e Prescott (1993) sobre a distribuição de GDP entre os países sugerem uma não compatibilidade com esse tipo de modelagem.

Apesar de a adoção da hipótese de normalidade dos dados ser amplamente utilizada, é possível citar Brian Arthur (1994), Kahneman (2003), Brian Arthur (1999), Foley e Doyle Farmer (2009), Nelson e Winter (2002), Dosi e Nelson (1994), entre outros sobre um crescente entendimento das limitações do paradigma econômico atual em explicar uma série de fenômenos econômicos de forma satisfatória. Contudo, é possível abstrair da história da ciência que anomalias e pequenas inconsistências nas teorias *mainstream* não são capazes, por si só, de substituir essas teorias por explicações alternativas. Daí dá-se a grande importância dos estudos realizado em economia sobre o tamanho de cidades (Krugman, 1996) e o tamanho das firmas (Axtell, 2001), onde ambos os fenômenos se mostraram compatíveis com distribuições denominadas de *Power laws* (PL). Mais especificamente, os fenômenos foram caracterizados como um tipo particular de PL: *Zipf's law*.

Em termos práticos, um PL é uma relação onde a variável dependente sofre mudanças relativas proporcionais a mudanças relativas nas demais variáveis. Ou seja, uma variável muda conforme uma potência da outra. Nesse sentido, essa relação é também conhecida como lei de potência.

A distribuição de PL, assim como diversas outras distribuições não normais, não pode ser caracterizada somente pela sua média e desvio padrão, o que torna esses modelos menos práticos, mas essenciais para explicar dinâmicas e processos complexos entre um conjunto de dados. Além disso, segundo Clauset, Shalizi e Newman (p.1, 2009), PLs tem grande potencial explicativo e tem recebido atenção e escrutínio particular:

“Not all distributions fit this pattern (**normal distribution**), however, and while those that do not are often considered problematic or defective for just that reason, they are at the same time some of the most interesting of all scientific observations. The fact that they cannot be characterized as simply as other measurements is often a sign of complex

underlying processes that merit further study. Among such distributions, the power law has attracted particular attention over the years for its mathematical properties, which sometimes lead to surprising physical consequences, and for its appearance in a diverse range of natural and man-made phenomena. The populations of cities, the intensities of earthquakes, and the sizes of power outages, for example, are all thought to have power-law distributions.”

Para o presente trabalho, PLs tem um papel importante porque segundo Di Guilmi *et al* (2003), uma parcela significativa da distribuição mundial de renda é compatível com um PL. De acordo com os autores, no período de 1960 a 1997, a parcela compreendida entre o trigésimo e o octagésimo quinto percentil da distribuição global de renda em termos de GDP per capita segue um PL. Não obstante, os autores trazem que países mais ricos são caracterizados por flutuações de menor amplitude e a diferença de amplitude entre os países ricos e pobres aumentou durante o período de estudo.

Tem-se então a oportunidade de verificar os resultados propostos por Di Guilmi *et al* (2003) para um período de estudo mais atual e para complementar o espaço teórico sobre a dinâmica de distribuição da renda mundial, estudando as implicações e motivações da possível regularidade empírica. O presente trabalho objetiva portanto replicar com alguns ajustes o artigo de Di Guilmi *et al* (2003) e, conseqüentemente, responder se a distribuição de riqueza dos países pode ser dita compatível com um modelo de PL.

O trabalho é dividido da seguinte forma, o capítulo 2 apresentará uma descrição da forma geral de um PL, ilustrará exemplos em economia e os possíveis mecanismos geradores desse tipo de relação. O capítulo 3 apresentará a evidência empírica para o caso concernente a distribuição da renda mundial. O capítulo 4 apresentará a hermenêutica econômica dos resultados e o capítulo 5 apresentará as principais conclusões e limitações da pesquisa.

2 Power Laws (PLs)

2.1 Forma geral de um PL

Um PL é uma relação entre duas variáveis de interesse X e Y que assume a forma da equação (1), onde C é uma constante e α é chamado de coeficiente do PL:

$$Y = CX^{-\alpha}, \quad (1)$$

Uma das características de maior interesse da relação (1) é o fato de ser homogênea. Em termos práticos, um PL afirma que mudanças relativas na variável X causam mudanças relativas proporcionais em Y independentes do tamanho inicial ou da dimensão dessas variáveis⁴. Tal propriedade pode ser denominada de invariância de escala e é particularmente conveniente no esforço de observar regularidade empírica em um conjunto de dados⁵.

O primeiro cientista social a aplicar esse tipo de relação foi o economista Vilfredo Pareto em 1896 na tentativa de explicitar a dinâmica de distribuição de riqueza. Pareto estipulou que entre os indivíduos na parte superior da distribuição, a probabilidade de que houvesse riquezas W superiores a dado valor x era proporcional à $x^{-\zeta}$ para algum número ζ positivo (Gabaix, 2009). Dessa forma, a relação proposta por Pareto pode ser encarada como resultante de uma distribuição de PL no formato da equação (2):

$$P(S > x) = Cx^{-\zeta} \quad (2)$$

Em que a função densidade de probabilidade (fdp) é dada por (3):

$$P(S < x) = 1 - P(S > x) = (x^\zeta - C)x^{-\zeta} \quad (3)$$

É importante perceber que a relação não se sustenta para valores irrestritos⁶ de x e a função de probabilidade não necessariamente existe, divergindo para $\zeta \leq 1$. Contudo, é direto ver que a probabilidade converge para $\zeta > 1$ e $x \geq x_{min}$ tal que:

$$p(x) = \frac{\zeta - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\zeta} \quad (4)$$

e,

$$P(X \geq x) = \int_x^\infty p(x)dx = \left(\frac{x}{x_{min}} \right)^{1-\zeta} \quad (5)$$

Onde $p(x)$ é chamada função de probabilidade e $P(X \geq x)$ a função cumulativa complementar de $p(x)$. Ou seja, a relação proposta por Pareto que posteriormente deu origem a chamada distribuição de pareto exemplifica uma dinâmica fundamentada em dois parâmetros principais: o valor mínimo para o qual o PL se sustenta (x_{min}) e o coeficiente ζ .

⁴Note que: $Y(tX) = Ct^{-\alpha}X^{-\alpha} = t^{-\alpha}Y(X)$

⁵Conforme será visto posteriormente, esse tipo de propriedade provê a possibilidade de estudar PLs mesmo na presença de erros de medição e barulho nos dados.

⁶Note que a relação não vale para $x \rightarrow 0$, de forma que deve haver um valor mínimo x_{min} para o qual a função de probabilidade não diverge

PLs possuem outras propriedades matemáticas interessantes. Segundo Gabaix (2009), multiplicar um PL por variáveis normais, adicionar ruído nas regiões de cauda ou somar variáveis independente e identicamente distribuídas preservam o expoente do PL. Ainda, conforme Jessen e Mikosch (2006), PLs mantêm suas relações funcionais quando submetidos a transformações polinomiais, min e máx, de forma que, em geral, quando combinadas duas variáveis que seguem um PL, a com menor expoente predomina.

Em termos práticos, essas propriedades de agregação fornecem a possibilidade de que PLs possam ser mensurados mesmo na presença de certo ruído nos dados. Apesar de afetar as variâncias, essas circunstâncias não afetam consideravelmente o expoente (responsável por exemplificar a essência do fenômeno de interesse). Nesse sentido, essas propriedades são também referidas como um mecanismo de herança, pois podem ser utilizadas para gerar outros PLs semelhantes.

Com relação aos momentos da relação (2), temos que o primeiro e o segundo momento satisfazem, respectivamente:

$$[x] = E(x) = \int_x^\infty xp(x)dx = \int_x^\infty Cx^{1-\zeta}dx \quad (6)$$

e,

$$E(x^2) = \int_x^\infty Cx^{2-\zeta}dx \quad (7)$$

O que implica que a média é finita para $\zeta > 2$ e a variância, para $\zeta > 3$, de forma que:

$$[x] = \frac{\zeta - 1}{x_{min}^{1-\zeta}} \frac{x^{2-\zeta}}{\zeta - 2} \quad (8)$$

e,

$$E(x^2) = \frac{\zeta - 1}{x_{min}^{1-\zeta}} \frac{x^{3-\zeta}}{\zeta - 3} \quad (9)$$

Ainda, segundo Gabaix (2009) todos os momentos maiores do que ζ divergem. Com relação à implicação dessa divergência, claramente, quando calculada a média de uma amostra finita retirada de uma distribuição PL o resultado obtido não será infinito. O fato da média não ser definida para $\zeta \leq 2$ significa que a média pode variar consideravelmente entre experimentos. Não obstante, conforme Newman (2005), esse fato implica também que conforme o número da amostra se torna grande, a média é dominada pelo maior valor da amostra, que diverge conforme n se torna grande.

Dessa forma, para PLs com $2 < \zeta \leq 3$ temos uma média bem definida, mas uma variância infinita. Conforme Clauset, Shalizi e Newman (2009), grande parte dos fenômenos naturais explicados por PLs possuem ζ nesse intervalo. Isso implica que esses fenômenos podem ser, de certa forma, caracterizados por uma média, mas são compatíveis com uma dinâmica de *black swan*⁷ ou, em outras palavras, caudas grossas da distribuição. Nesse mesmo sentido, o terceiro momento revela que para o intervalo de ζ referido, a distribuição relativa a um PL pode ser classificada como leptocúrtica. Em essência, isso significa que o PL é uma distribuição que apresenta caudas mais grossas do que uma distribuição normal.

⁷O termo é uma referência utilizada por Nassim Taleb (2007) para descrever eventos improváveis de impacto significativo e imprevisíveis *à priori*

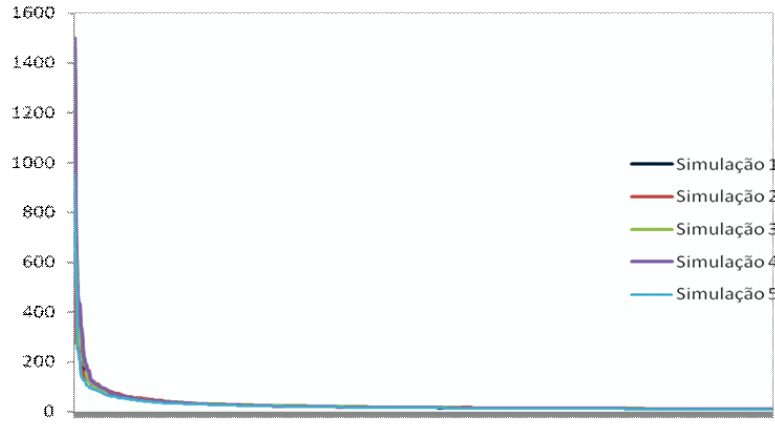
Outra importante característica dos PLs pode ser observada em um gráfico plotando o *rank* das N realizações da variável x_t para determinado período t . Dessa forma, utilizei o método proposto por Newman (equação 12, p. 4, 2005) para gerar $\mathbf{x}_t = (x_{1t}, \dots, x_{Nt})$, onde \mathbf{x}_t segue um PL para $x_{it} \geq x_{min}$, $i = 1, \dots, N$ e para os valores arbitrários $x_{min} = 10$, $\zeta = \frac{5}{2}$, com r distribuído aleatória e uniformemente entre 0 e 1. Desse método obtemos algumas simulações com distribuição no formato (10) e podemos observa-lás na figura 2.1 utilizando $N = 1000$:

$$p(x) = 3 \frac{10^{\frac{3}{2}}}{x^{\frac{5}{2}}} \quad (10)$$

$$x = x_{min}(1 - r)^{-\frac{1}{\zeta-1}} = 10(1 - r)^{-\frac{2}{3}} \quad (11)$$

É válido notar de (10) que $p(x) \rightarrow 0$ se e somente se $x \rightarrow \infty$. Em outras palavras, a dinâmica de *black swan* comentada anteriormente é materializada por meio do fato de que um evento pode ter probabilidade de ocorrência muito pequena, mas efetivamente não nula.

Figura 2.1 - Simulações para o PL arbitrário



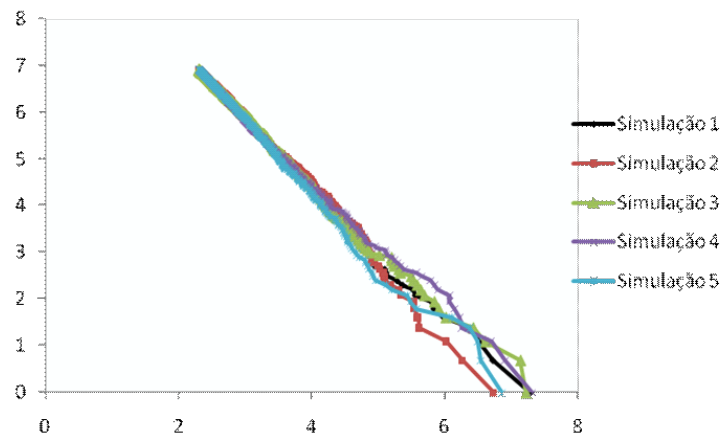
É possível observar o caráter de cauda longa nas simulações e uma região onde a curva decresce rapidamente. Ainda, da relação funcional (1) obtemos para $K = \ln(c)$:

$$\ln(Y) = K - \alpha \ln(x) \quad (12)$$

Seguindo Gabaix (2009) e tendo em vista a equação acima, é esperado que um gráfico plotando $\ln(Rank)$ vs $\ln(x)$ retorne uma reta negativamente inclinada com coeficiente angular α , onde o *Rank* é construído atribuindo $1, \dots, n$, respectivamente, da maior até a n-ésima maior variável x_{Nt} . Essa representação gráfica é referida como *Zipf plot* (Stanley *et al*, 1995). É possível rodar uma regressão OLS para retornar uma estimativa⁸ para α . No gráfico 1.2 é possível observar que as simulações realizadas compartilham uma região onde os dados estão arranjados ao longo de uma reta:

⁸De acordo com Newman (2005), Clauset, Shalizi e Newman (2009) o estimador da regressão OLS é viesado e significativamente distante de α , sendo necessário a utilização de outros métodos para retornar uma estimativa robusta, conforme será visto com mais profundidade no capítulo 2.

Figura 2.2 - Zipf plot para as simulações apresentadas na Figura 1.1



Essa característica é frequentemente usada como condição suficiente para verificar a compatibilidade dos dados com um modelo de PL. Contudo, a relação linear mostrada no gráfico 1.2 é somente condição necessária para tal conclusão (Clauset, Shalizi e Newman, 2009). Não obstante, o gráfico 1.2 também revela outra particularidade relevante do PL. A inclinação da reta pode ser encarada, intuitivamente, como uma medida de quão desigual é a distribuição em sua cauda superior. Quanto menor a inclinação (menor α), maior a disparidade entre as observações superiores. Nas figuras 2.3 e 2.4 uma comparação entre diferentes coeficientes a partir do exemplo anterior é realizada.

Figura 2.3 - Comparação entre PLs com diferentes coeficientes

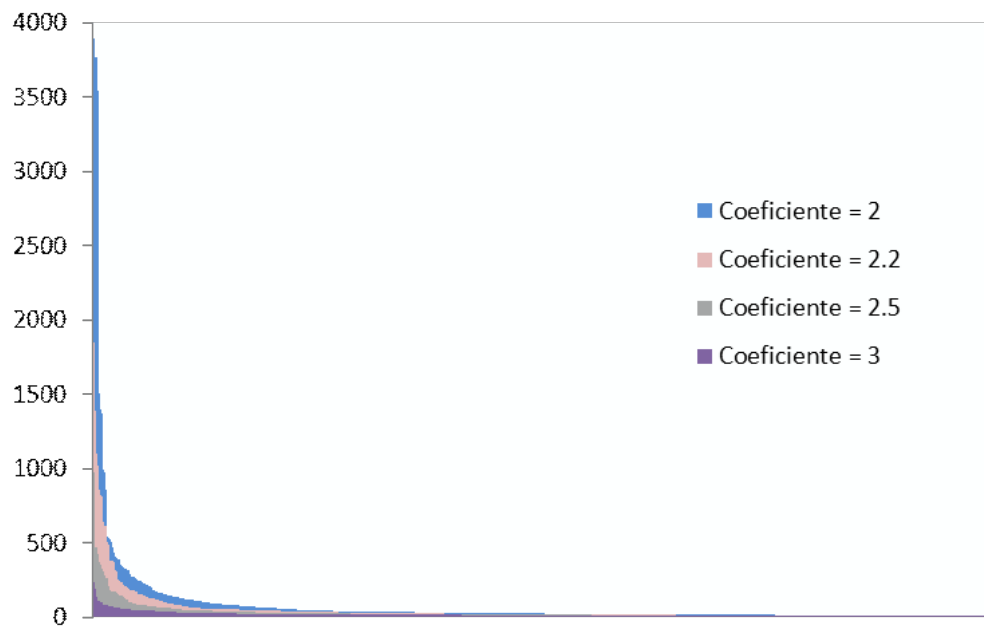
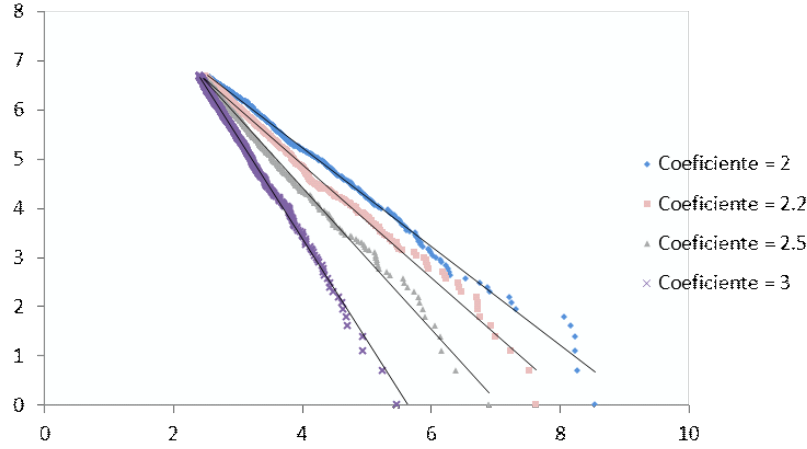


Figura 2.4 – Comparação entre diferentes coeficientes (*Zipf Plot*)



2.2 Caso particular: *Zipf's Law* e exemplos em economia

Um tipo particular de PL denominado de *Zipf's Law* foi popularizado pelo linguista norte americano George Kingsley Zipf (1902-1950). Zipf foi responsável por formular a lei empírica e propor que dado algum corpo linguístico, a frequência com que determinada palavra é utilizada é inversamente proporcional ao seu *Rank* em uma tabela de frequência (Zipf 1935; 1949). Em termos práticos, a lei proposta por Zipf é do tipo (1) com $\alpha \approx 1$:

$$Y = \frac{C}{X} \quad (13)$$

Surpreendentemente, Fagan e Gençay (2010) apontam que, em conformidade com a lei proposta por Zipf, no *Brown corpus* de inglês americano a palavra mais frequente ocorre cerca de 7% das vezes, enquanto a segunda 3,5% das vezes, assim por diante.

Conforme comentado na seção introdutória, tanto o tamanho de firmas (Axtell, 2001; Luttmer, 2007), quanto o tamanho de cidades (Krugman 1996; Gabaix e Ioannides, 2004) pode ser explicado pela Lei de Zipf. No primeiro caso, Axtell (2001) conclui que a relação segue um PL com $\alpha = 1,059$ e, no segundo, Krugman (1996) reporta que a regularidade satisfaz:

$$\ln(Rank) = 10,53 - 1,005\ln(Size) \quad (14)$$

Onde α seria, portanto, igual a 1,005. Ambos os casos satisfazem a Lei de Zipf com bastante proximidade. Vale ainda apontar, conforme Gabaix (2009), que podemos encarar o x_{min} como uma força ou fricção econômica. No caso das cidades, o x_{min} impede que as cidades ou aglomerações populacionais fiquem muito pequenas e pode ser representado por uma probabilidade positiva de morte, custos de transporte, custos fixos, entre outros. No caso das firmas, o x_{min} pode representar um tamanho mínimo para que a empresa seja economicamente viável.

Apesar da validade da lei para diversos casos⁹ incluindo línguas artificiais como “Esperanto” (Bill

⁹Piantadosi (2014) investiga a validade da lei para diversos idiomas e provê uma pesquisa mais abrangente sobre os mecanismos específicos por trás da regularidade nesse caso.

Manaris *et al.* 2006), a regularidade não parece ter explicação única. Zipf argumentou que nem falantes e nem ouvintes de um idioma desejam trabalhar mais do que o necessário para entender o que se está comunicando. Dessa forma, um princípio de menor esforço geraria poucas palavras muito utilizadas e muitas palavras pouco utilizadas a fim de possibilitar uma comunicação eficiente. Contudo, não existe uma tautologia para derivar *a priori* o coeficiente $\alpha \approx 1$ ou a relação funcional (1).

De fato, é surpreendente que algo complexo como a linguagem humana possa ser descrito por uma relação tão simples como um PL, conforme comentado por Piantadosi (2014, p.2):

“But given that words do vary in frequency, it is unclear why words should follow such a precise mathematical rule—in particular one that does not reference any aspect of each word’s meaning. Speakers generate speech by needing to communicate a meaning in a given world or social context; their utterances obey much more complex systems of syntactic, lexical, and semantic regularity. How could it be that the intricate processes of normal human language production conspire to result in a frequency distribution that is so mathematically simple—perhaps “unreasonably” so?”

Köhler (2002) afirma que dentre os críticos da Lei de Zipf o argumento mais comum é que uma distribuição do tipo Zipf pode ser gerada como simples consequência de se digitar aleatoriamente em um teclado¹⁰. Intuitivamente, se cada tecla tem a mesma probabilidade de ser pressionada (sendo que cada tecla representa uma de 26 letras e uma adicional representa a “barra de espaço”), então existirão exponencialmente mais palavras diferentes com mais letras do que com menos letras. Por exemplo, o alfabeto pode ser usado para fazer 26 palavras de uma sílaba, 26^2 palavras de duas sílabas e assim por diante. Semelhantemente, como a barra de espaço representaria o final de uma palavra e a sua probabilidade de ocorrência é positiva, palavras com muitas letras são exponencialmente menos prováveis do que palavras menores.

O argumento acima é demonstrado por Newman (p. 13, 2005) supondo uma probabilidade p_l não nula e equivalente para todas as m letras possíveis e uma probabilidade $p_e = \frac{(1-p_l)}{m}$ de se pressionar a barra de espaço (consequentemente formando uma palavra se ocorrer após uma letra). Dessa forma, a frequência x com que uma palavra de tamanho y ocorre é:

$$x = \frac{(1 - p_l)^y}{m} p_l = e^{ya} \quad (15)$$

em que $a = \ln(\frac{(1-p_l)}{m}) + \ln(p_l)$

Assim, assumindo que $p(y) = e^{yb}$ para uma constante não nula qualquer b , têm-se:

$$p(x) = \frac{p(y)dy}{dx} = \frac{e^{yb}}{ax} = \frac{x^{\frac{b}{a}-1}}{a} \quad (16)$$

que é um PL com coeficiente $(1 - \frac{b}{a})$

Apesar da validade do argumento, Newman (p.13, 2005) prossegue comentando sobre a irrealidade da hipótese de aleatoriedade da linguagem humana:

“This is a reasonable theory as far as it goes, but real text is not made up of random letters. Most combinations of letters don’t occur in natural languages; most are

¹⁰Esse argumento é conhecido como “monkeys with typewriters” (Newman, 2005)

not even pronounceable. We might imagine that some constant fraction of possible letter sequences of a given length would correspond to real words and the argument above would then work just fine when applied to that fraction, but upon reflection this suggestion is obviously bogus.”

Ainda segundo Köhler (2002), também é argumentado que a ubiquidade das distribuições PL torna contraditória a interpretação de que a lei teria, de fato, relevância como fato linguístico. Contudo, é visão predominante nas ciências naturais que o fato de PLs explicarem fenômenos tão diversos como a frequência de palavras em um texto, tamanho de cidades, tamanho de firmas, intensidade de terremotos e de erupções solares, entre outros¹¹, permite novas perspectivas sobre os fenômenos e não implicaria em menor relevância ou validade da lei (Schroeder, 2012).

Semelhante ao princípio de menor esforço, outra ilustração pode ser obtida com base no conceito de *preferential attachment*. Isto é, as palavras formadas com mais facilidade se tornariam mais usadas e, por isso, ainda mais usadas, ocasionando uma distribuição com poucas palavras muito frequentes e muitas possíveis palavras pouco utilizadas. Esse tipo de processo foi estudado por Yule (1925) e se relacionava com o número de espécies por gênero biológico. Em termos práticos, o argumento de *preferential attachment* está bastante relacionado com a chamada Lei de Gibrat¹² e representa algum processo estocástico em que a distribuição de uma quantidade depende de quanto já se tem dessa quantidade. Com isso, esses mecanismos geram distribuições de cauda longa e resultam em um efeito denominado de *Rich get richer*. Isto é, como esses processos estocásticos distribuem uma quantidade de acordo com a quantidade já existente a distribuição resultante seria composta por poucos indivíduos muito significantes e muitos indivíduos pouco significantes em termos de massa de probabilidade.

Retornando ao fenômeno tratado na literatura como exemplo de PL e comentado no início do capítulo, a distribuição de Pareto é, na verdade, o caso contínuo da distribuição de Zipf¹³. Ainda, as duas distribuições diferem na forma como são plotadas. Segundo Newman (2005), Zipf utilizou x no eixo horizontal e $P(x)$ no eixo vertical, tendo sido realizado o inverso para a distribuição de Pareto¹⁴.

Com relação ao coeficiente unitário no caso estudado por Pareto, isso implicaria que o indivíduo mais rico da distribuição possuiria o dobro do segundo indivíduo, o triplo do terceiro e assim por diante. Esse resultado é surpreendente pois provê uma projeção agregada que segue uma regra *ad hoc* em princípio. Em outras palavras, não há um motivo aparente para que as decisões individuais dos agentes resultem em uma distribuição de riqueza no formato comentado. O resultado é, dessa forma, importante para revelar aspectos do mecanismo por trás da alocação de riqueza e também para elucidar novas questões sobre a ligação entre as inúmeras decisões no âmbito microeconômico e seu resultado agregado.

Em outro exemplo de PL, Gabaix (2009) e diversos outros estudos empíricos documentaram que a compensação salarial (w) de um CEO aumenta como potência do tamanho da firma (S) que trabalha (Barro e Barro, 1990; Kostiuk, 1990; Frydman e Saks, 2007) de forma que $w \sim S^k$ e, tipicamente, $0,2 \leq k \leq 0,4$ (Gabaix, 2009)

¹¹Newman (2005) e Clauset, Shalizi e Newman (2009) realizam uma pesquisa abrangente apresentando diversos exemplos de PL.

¹²Esse mecanismo será abordado em maior profundidade na próxima seção. Contudo, a Lei de Gibrat é a proposição de que a taxa de crescimento de determinada quantidade é independente do seu tamanho ou dimensão

¹³Como a regularidade proposta por Zipf se refere a frequência com que palavras ocorrem em um corpo linguístico, trata-se de um ferramental discreto.

¹⁴Apesar da possível confusão na literatura, os resultados são os mesmos

Mais recentemente, Di Guilmi *et al* (2003) reportam que durante o período de 1960 a 1997, os países compreendidos entre o 30º e o 85º percentil da distribuição de GDP per capita se distribuem conforme um PL. Não obstante, os autores concluem também que o expoente apresentou tendência de queda no período e sugerem que o processo de crescimento desses países é invariável em suas escalas.

Os resultados de Di Guilmi *et al* (2003) são interessantes pois tem consequências relevantes para as teorias de desenvolvimento e dinâmica da renda mundial. Caso a regularidade seja confirmada também por pesquisas posteriores, é de suma importância que os modelos de desenvolvimento incorporem a evidência empírica nas suas explicações. Não obstante, é válida também a citação proveniente de uma discussão em Schumpeter (1949, p. 155) sobre a observação de Pareto, em que o autor aponta o potencial da relação para sustentar um tipo novo de teoria econômica:

“Few if any economists seem to have realized the possibilities that such invariants hold for the future of our science. In particular, nobody seems to have realized that the hunt for, and the interpretation of, invariants of this type might lay the foundations for an entirely novel type of theory.”

Por fim, Gabaix (2009) aponta que regularidades do tipo PL estão sendo propostas para explicar não somente o movimento de preços de ativos no mercado financeiro, mas também a própria distribuição de crises financeiras e flutuações de grande escala.

2.3 Processo gerador de um PL

Um candidato importante a processo gerador de um PL é o chamado mecanismo de crescimento aleatório proporcional (Gabaix, 2009). Conforme comentado na primeira seção do capítulo, PLs são invariantes em escala. Isto é, a taxa de crescimento da variável de interesse é independente da dimensão dessa variável.

A intuição acima é uma das hipóteses do modelo desenvolvido por Gibrat (1931) para explicar a trajetória de crescimento e a estrutura do mercado das firmas (Sutton, 1996). A intuição do raciocínio de Gibrat pode ser ilustrada denotando o tamanho de uma firma no período t por x_t e o crescimento do período $(t - 1)$ para o período t pela variável aleatória ζ_t . Dessa forma:

$$x_t - x_{(t-1)} = \zeta_t x_{(t-1)} \quad (17)$$

em que $x_t = (1 + \zeta_t)x_{(t-1)} = x_0(1 + \zeta_1)(1 + \zeta_2)\dots(1 + \zeta_t)$

Supondo t pequeno o suficiente, é possível aproximar $\log(1 + \zeta_t)$ por ζ_t . De forma que, aplicando o logaritmo em (13), temos:

$$Z = \log(x_t) = \log(x_0) + \sum_{i=1}^t \zeta_i \quad (18)$$

Assumindo que os termos em $\sum_{i=1}^t \zeta_i$ são variáveis independentemente distribuídas e normais com média μ e desvio padrão θ , temos pelo teorema do limite central que, conforme $t \rightarrow \infty$, o termo $\log(x_0)$ se torna pequeno e a variável Z é aproximada por uma normal com média $t\mu$ e desvio padrão $t\theta$. Ou seja, a variável x_t , de acordo com o desenvolvimento de Gibrat, se distribuiria conforme uma lognormal.

Segundo Sutton (1996), Gibrat primeiro aplicou o modelo descrito para distribuições de renda e, depois, para tamanho das firmas no setor de manufatura, obtendo resultados coerentes com o proposto.

É importante notar que esse modelo de Gibrat não permite a entrada ou saída de firmas de forma que a variância cresce ilimitadamente¹⁵ com t , um aspecto pouco desejável no modelo. Não obstante, o processo gera uma distribuição que pode compartilhar, por exemplo, de características gráficas de um PL (Clauset, Shalizi e Newman, 2009).

Nesse sentido, é válido apontar que se o processo segue o modelo irrestrito de Gibrat, então a distribuição resultante pode parecer um PL, mas não o sê-lo assintoticamente.

A observação feita por Gabaix (2009) com base no teorema de Kesten (1973) em relação a Lei de Gibrat é de que se o processo descrito sofre de fricção¹⁶ (um custo fixo que inviabiliza a existência de firmas pequenas, por exemplo) a distribuição no *steady state*, caso exista, será um PL com coeficiente ζ . Essa observação é extremamente relevante para o desenvolvimento do trabalho pois garante que em determinadas condições, o mecanismo de crescimento proporcional randômico resulta em uma distribuição PL com coeficiente ζ e não em uma distribuição lognormal.

Com isso, uma intuição que podemos retirar desse mecanismo é que diferenças passadas nas taxas de crescimento podem ser responsáveis por diferenciar significativamente o GPD por trabalhador de países no presente. Isto é, erros de política econômica, choques de curto prazo, entre outros fatores podem ter afetado o GDP de alguns países em detrimento de outros e isso, supondo um mecanismo como o proposto acima, acarretaria em uma distribuição com poucos países muito ricos e outros com significativamente menos riqueza.

Em relação a taxa média de crescimento entre as observações de um PL¹⁷, esta pode ser encarada também como uma consequência de auto organização em um sistema. Isto é, emerge de forma espontânea das decisões individuais agregadas um padrão: uma taxa de crescimento média comum. De uma perspectiva biológica, diversos padrões e estruturas formadas por auto organização ou organização espontânea possuem funcionalidade¹⁸. Em suma, um questionamento relevante trago pelo presente trabalho é se essa característica não seria resultado de uma auto organização no sistema internacional de países e se essa auto organização não seria responsável por uma forma eficiente de distribuir crescimento entre os países. Em outras palavras, é possível que, semelhante a sistemas naturais onde emergem relações do tipo PL, o caso estudado reflita uma forma eficiente de alocação de riqueza.

Resumindo as considerações acima temos: **1 - A distribuição resultante de um mecanismo como o de crescimento aleatório proporcional incorpora uma dinâmica de *Rich get richer* resultando em poucos países muito ricos e vários significativamente mais pobres** e **2 - Isso pode sugerir alguma forma de auto organização entre os países de forma que**

¹⁵Note que $Var(Z) = Var(\log(x_0)) + \sum_{i=1}^t t Var(\zeta_i)$, que cresce em função de t

¹⁶Segundo o autor, essa fricção pode ser representada por uma constante positiva na equação (12)

¹⁷Cordoba (2001) demonstra por meio da lei de movimento que para uma distribuição $F(x)$ ser um Power Law com coeficiente α este deve obrigatoriamente ter uma taxa média ϕ de crescimento entre as observações.

¹⁸Conforme Gabaix (2016, p. 200): “Biology is replete with intriguing and seemingly universal relations of the power law type. For instance, the energy that an animal of mass M requires to function is proportional to $M^{\frac{3}{4}}$ — rather than M as a simple “constant return to scale” model would predict, as illustrated. 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 $M^{\frac{3}{4}}$ law. There is an interesting lesson: a priori, lots of things could matter for energy consumption—for example, climate, predator or prey status, thickness of the fur—and they probably do matter to a limited extent. However, in its essence, an animal is best viewed as a network in which nutrients circulate at maximum efficiency”

o crescimento estaria sendo distribuído de forma mais ou menos normal. Com isso, os países teriam uma taxa média compartilhada.

3 Evidência empírica

3.1 Método

Conforme comentado nas seções anteriores, Di Guilmi *et al* (2003) concluem que no período de 1960 a 1997 os países compreendidos entre o 30º e o 85º percentil da distribuição são compatíveis com um modelo de PL. Os autores prosseguem afirmando que a emergência dessa regularidade na distribuição *steady state* de GDP per capita pode ser explicada por meio de processos de crescimento aleatório específicos de cada país, obedecendo a uma versão modificada da Lei de crescimento proporcional de Gibrat. Em outras palavras, o crescimento desses países seria invariável.

Os autores utilizam dados da Penn World Table (PWT) 6.1 (Summers, Heston and Ater, 2002). Infelizmente, a base contém no máximo cerca de 50 países válidos como amostra durante o período relatado. Não obstante, a metodologia aplicada pelos autores é baseada somente na identificação visual de uma reta em um gráfico *log-log* e o uso de uma regressão OLS, conforme citado a seguir:

“The hypothesis that the central part of the world income distribution follows a power law seems to be corroborated by the extremely good fit of linear regressions, as one can appreciate by noting that the value of the OLS R^2 is never below 0.978. Furthermore, note that ζ shows a clear tendency to decrease over time.”

Nesse ponto é muito importante considerar os argumentos de Clauset, Shalizi e Newman (2009) sobre o quão viesado pode ser o estimador de ζ calculado com a metodologia descrita na citação acima. É possível exemplificar o caso utilizando as simulações apresentadas no Capítulo 1 para os parâmetros $x_{min} = 10$ e $\zeta = 2, 5$. Aplicando o método OLS, têm-se uma estimativa de 1,56 para ζ com $R^2 = 0.998$ e erro padrão de aproximadamente 7%. É válido notar que a estimativa é razoavelmente distante do verdadeiro coeficiente.

Além disso, Köhler (p. 52, 2002) afirma que uma grande fonte de ceticismo quanto a validação de um PL é o argumento de que a própria fabricação do *Rank* e o rearranjo dos dados no *Zipf Plot* poderia criar artificialmente o formato referido, tornando o método de visualização gráfica insuficiente, por si só, para afirmar a validade do modelo:

“In most cases, however, where linguists or mathematicians publish their scepticism about the validity of Zipf’s Law the main argument is that the rank-frequency distribution with its “typical” shape be the “artificial” result of re-arranging the data according to their frequency. This criticism may be justified with respect to some recent publications in physics, where sometimes power law curves are considered as identical even if one of them is approximately a straight line in log-log coordinates whereas the other one appears as near to linear in linear log coordinates”

Assim, é importante aplicar uma metodologia diferente da utilizada por Di Guilmi *et al* (2003) com o objetivo de obter um estimador não viesado e consistente para ζ . Conforme proposto por Clauset, Shalizi e Newman (2009), o método de *Maximum Likelihood Estimators* (MLEs) é mais adequado e os autores derivam um estimador equivalente ao chamado estimador de Hill:

$$\hat{\alpha} = 1 + n \left(\sum_{i=1}^n \frac{\ln(x_i)}{\ln(x_{min})} \right)^{-1} \quad (19)$$

em que x_i , $i = 1, 2, 3, \dots, n$ são os valores observados e supostamente retirados de uma distribuição de PL para $x_i \geq x_{min}$ e o estimador proposto é assintoticamente normal e consistente (Clauset, Shalizi e Newman, 2009).

Um estimador para o erro padrão (θ) de ($\hat{\alpha}$) é derivado por Newman (2005):

$$\hat{\theta} = n^{0.5} \left(\sum_{i=1}^n \frac{\ln(x_i)}{\ln(x_{min})} \right)^{-1} = \frac{(\hat{\alpha} - 1)}{n^{0.5}} \quad (20)$$

Dessa forma, assumindo que o modelo de PL é adequado aos dados, os estimadores acima restringem as limitações impostas pelo método de regressão linear simples¹⁹.

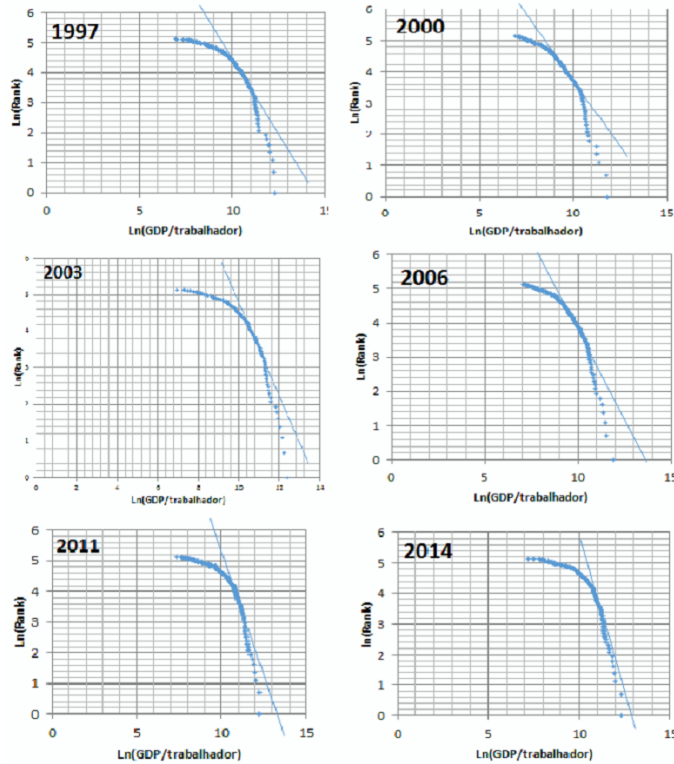
Com relação ao cálculo do x_{min} , Clauset, Shalizi e Newman (2009) propõe o uso do teste não paramétrico de Kolmogorov-Smirnov ou estatística de KS. Essa estatística representa a distância máxima entre a função cumulativa de probabilidade dos dados e do modelo que se deseja testar. Por enquanto, por motivos expositivos, supõe-se que o modelo de PL é adequado aos dados de interesse e que o x_{min} pode ser obtido por meio do método de visualização gráfica de forma satisfatória.

3.2 Resultados

Com esse ferramental, é possível iniciar a análise do caso de interesse do trabalho. Na figura 3.1, é possível observar o *Zipf plot* das distribuições de GDP por trabalhador e respectivos *Rankings* obtidas da Penn World Table (PWT) 9.0 (Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer, 2015) de 169 países em alguns cortes temporais no período de 1997 a 2014:

Figura 3.1 - Zipf plot para os dados de GDP por trabalhador

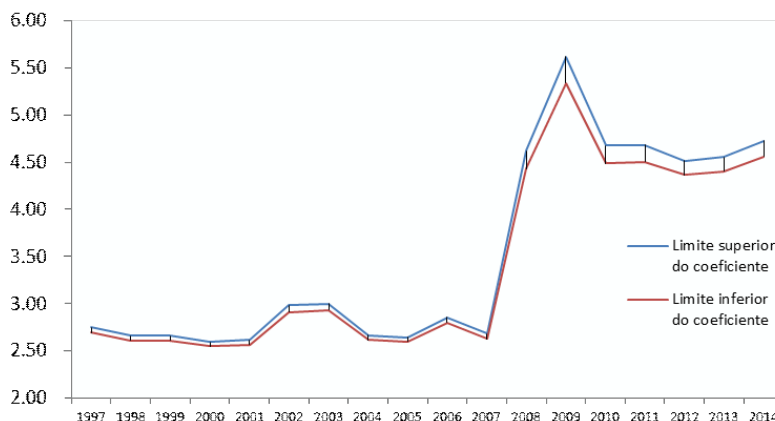
¹⁹É importante comentar que o caso de interesse possui também um limite superior para o PL. A ideia é que os estimadores acima podem ser utilizados também nesse caso pois o interesse é medir a inclinação α . Em termos práticos, a forma funcional para testar o modelo com limite superior deveria incluir outra constante $e^{a\alpha}$ para alguma constante a a fim de obter a normalização da função de probabilidade.



É possível notar da Figura 3.1 uma região onde os dados estão arranjados ao longo de uma reta nos períodos descritos²⁰. Têm-se então a suspeita de que esses dados podem ser compatíveis com um modelo de PL.

Utilizando o ferramental descrito anteriormente, é possível observar na Figura 2.2 a evolução das estimativas de α conjuntamente com o erro $\hat{\theta}$, ambos obtidos por meio dos estimadores derivados por Clauset, Shalizi e Newman (2009) e Newman (2005), respectivamente. Na figura, o limite superior do coeficiente representa $(\hat{\alpha} + \hat{\theta})$ e o limite inferior $(\hat{\alpha} - \hat{\theta})$.

Figura 3.2 – Evolução das estimativas do coeficiente do PL



Para melhor analisar os resultados é válido separar a amostra em dois momentos. No primeiro momento (1997 - 2007) O coeficiente ($\hat{\alpha}$) é relativamente constante em torno de 2,65 com erros de 3% em média e desvio padrão de 0,19. Essa consideração é importante porque, conforme exposto no Capítulo 2, um coeficiente com tendência de queda significaria uma diferença crescente ao longo do período entre as observações superiores. Por outro lado, uma tendência crescente de ($\hat{\alpha}$) sugeriria uma menor concentração de riqueza entre as observações.

No segundo momento (2008 - 2014), o coeficiente aumenta significativamente ($\hat{\alpha} = 4,58 \pm 0,09$). Esse aumento pode ter duas explicações: i) a crise financeira internacional ocorrida no período foi um choque na distribuição que afetou mais as observações superiores, ii) O aumento foi resultado da diminuição da amostra (possivelmente pelo fato da distribuição não estar em *steady state* dado o choque). Cerca de 37% dos países ($n = 65$) e 41% da renda mundial medida em termos de GDP por trabalhador se encontram na região do PL de forma relativamente constante de 1997 a 2007. A partir de 2008, cerca de 24% dos países ($n = 41$) e 37% da riqueza mundial estão contidas na região do suposto PL.

É válido notar que em 2002 e 2003 há um aumento temporário de cerca de 14% no coeficiente do PL. A amostra utilizada nesse período caiu cerca de 17% e não há motivo aparente para pensar que o movimento reflete uma real mudança na essência do fenômeno estudado, mas sim um reflexo de uma flutuação estatística.

²⁰Resultados semelhantes são encontrados para os demais cortes temporais no período de 1997 a 2014

3.3 Teste de compatibilidade do PL

É importante perceber que as análises e o ferramental utilizado até o momento no trabalho permitem estimar os parâmetros e estudar o comportamento de um PL, mas não dizem se, de fato, o modelo utilizado é uma boa hipótese. Em outras palavras, quando consideramos que os dados são compatíveis com um PL, o desafio de estimar com robustez os parâmetros deve vir acompanhado de suficiente evidência de que a distribuição de PL é, em termos práticos, uma hipótese razoável.

Gabaix e Ibragimov (2008b) sugerem um método baseado em OLS para aferir se a hipótese dos dados proverem de um PL pode ser rejeitada. Já Clauset, Shalizi e Newman (2009) propõe um teste baseado na estatística de Kolmogorov-Smirnov (KS). Como a estatística de KS é obtida por meio de um teste não paramétrico pautado em estimação por MLE e não sofre das restrições comentadas anteriormente para o método OLS, o método adotado no trabalho é o sugerido por Clauset, Shalizi e Newman (2009).

O teste de Kolmogorov-Smirnov é um teste não paramétrico que quantifica a distância entre a função cumulativa obtida dos dados e a função cumulativa hipotética. Nesse caso, queremos testar se os dados observados podem ter sido realmente obtidos a partir de uma distribuição PL do tipo (3). Essa distância é representada pela estatística de KS e é menor para distribuições que são mais compatíveis ou melhor explicativas da distribuição real dos dados. Formalmente, a estatística de KS é representada por D , onde $S(x)$ é a função cumulativa observada dos dados e $P(x)$ a função cumulativa de um PL com o coeficiente estimado na seção anterior ($\hat{\alpha}$).

$$D = \max_{x \geq x_{min}} |S(x) - P(x)| \quad (21)$$

Em termos práticos, apesar de o método proposto oferecer resultados satisfatórios (Clauset, Shalizi e Newman, 2009), o teste possui duas limitações principais. Primeiro, o fato de não rejeitar a hipótese nula de que os dados seguem um modelo de PL não implica que o modelo é, de fato, a verdadeira distribuição da população estudada, apenas que falha-se em rejeitar essa hipótese. Segundo, a estatística de KS tem flutuações decrescentes em n , mas deve ser analisada com cuidado para amostras menores. Além disso, é realmente mais difícil rejeitar a hipótese de que os dados seguem um PL quando se tem uma amostra pequena, não somente pelas características do teste. A ideia é que torna-se mais provável obter um resultado coerente com o modelo simplesmente por acaso quando n é pequeno.

Os valores críticos foram obtidos por meio da aproximação (22) obtida por Zar (2010) para o teste bilateral:

$$KS_{critico} = D_{\alpha(2),n} = \sqrt[2]{\frac{-\ln(\frac{\alpha}{2})}{2n}} \quad (22)$$

Que é uma aproximação válida para $n > 35$ e foi aplicada com nível de significância $\alpha = 5\%$. Na Tabela 1 são apresentados os resultados obtidos. Conforme comentado na seção anterior, o fato da estatística de KS estimada ser inferior a estatística crítica não implica que o modelo é de fato um PL, mas que falhamos em rejeitar essa hipótese dada a aproximação utilizada e um nível de significância de 5%. É importante comentar que a proximidade das duas estatísticas em alguns anos sugere que o p -valor não é muito superior do que 5%. Contudo, ao nível de significância escolhido a conclusão é de que falha-se em rejeitar que os dados foram retirados de uma distribuição PL com coeficiente α .

Table 1: Estatísticas $KS_{critico}$ e $KS_{estimadas}$ para o periodo de 1997 - 2014

Ano	$KS_{critico}$	$KS_{estimado}$
1997	0.18	0.15
1998	0.17	0.13
1999	0.17	0.11
2000	0.16	0.12
2001	0.17	0.13
2002	0.18	0.16
2003	0.18	0.16
2004	0.17	0.15
2005	0.16	0.14
2006	0.17	0.14
2007	0.17	0.15
2008	0.22	0.13
2009	0.24	0.15
2010	0.22	0.15
2011	0.21	0.15
2012	0.20	0.14
2013	0.20	0.12
2014	0.20	0.14

Ou seja, em 95% dos casos onde um PL com o coeficiente estimado ($\hat{\alpha}_t$ para $t = 1997, \dots, 2014$) e amostra n_t obtemos estatísticas de KS iguais ou menores do que as estatísticas críticas apresentada na Tabela 1.

É válido acompanhar os resultados da Tabela 1 conjuntamente com a Figura 3.3, que apresenta a amostra (n_t) utilizada no período de estudo. Nota-se que em 2002 e 2003 há uma queda temporária em n e após 2008 a amostra cai em cerca de 30 países. Isso explica o aumento significativo na estatística crítica após 2008, refletindo a maior dificuldade de se rejeitar o modelo conforme n decresce.

É importante comentar também que em 2009 $n < 35$, de forma que a aproximação não é mais válida e a estatística crítica foi, nesse caso, retirada da tabela em Massey (1951, p. 70) que está reproduzida na Figura 3.4. Em termos práticos, os resultados da aproximação e da tabela são similares quando $n > 35$. Por exemplo, a estatística crítica é 0.15801 pela regra da tabela no caso de $n = 60$ e igual a 0.15800 pela aproximação.

Figura 3.3 - Amostra utilizada no teste entre 1997 e 2014

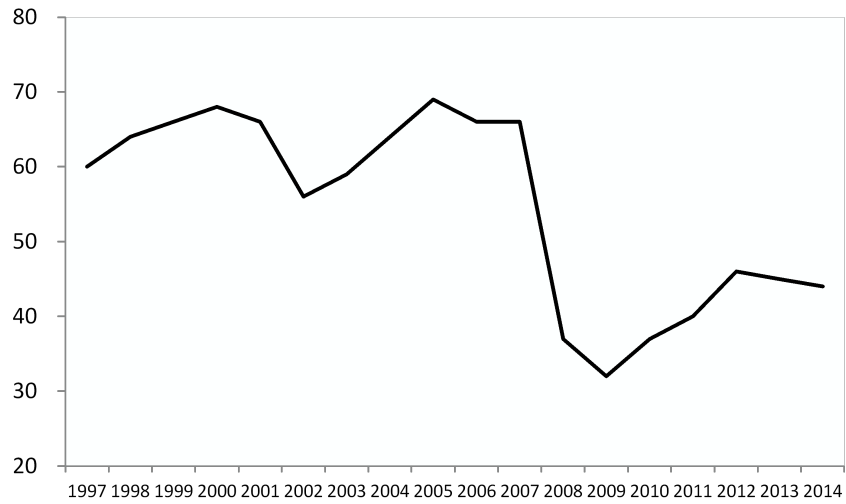


Figura 3.4 - Tabela retirada de Massey (1951, p. 70)

TABLE 1. Critical values, $d_\alpha(N)$, of the Maximum Absolute Difference between Sample and Population Cumulative Distributions.
Values of $d_\alpha(N)$ such that $Pr[\max |S_N(x) - F_0(x)| > d_\alpha(N)] = \alpha$, where $F_0(x)$ is the theoretical cumulative distribution and $S_N(x)$ is an observed cumulative distribution for a sample of N .

Sample size (N)	Level of significance (α)				
	0.20	0.15	0.10	0.05	0.01
1	0.900	0.925	0.950	0.975	0.995
2	0.684	0.726	0.776	0.842	0.929
3	0.565	0.597	0.642	0.708	0.828
4	0.494	0.525	0.564	0.624	0.733
5	0.446	0.474	0.510	0.565	0.669
6	0.410	0.436	0.470	0.521	0.618
7	0.381	0.405	0.438	0.486	0.577
8	0.358	0.381	0.411	0.457	0.543
9	0.339	0.360	0.388	0.432	0.514
10	0.322	0.342	0.368	0.410	0.490
11	0.307	0.326	0.352	0.391	0.468
12	0.295	0.313	0.338	0.375	0.450
13	0.284	0.302	0.325	0.361	0.433
14	0.274	0.292	0.314	0.349	0.418
15	0.266	0.283	0.304	0.338	0.404
16	0.258	0.274	0.295	0.328	0.392
17	0.250	0.266	0.286	0.318	0.381
18	0.244	0.259	0.278	0.309	0.371
19	0.237	0.252	0.272	0.301	0.363
20	0.231	0.246	0.264	0.294	0.356
25	0.21	0.22	0.24	0.27	0.32
30	0.19	0.20	0.22	0.24	0.29
35	0.18	0.19	0.21	0.23	0.27
over 35	1.07	1.14	1.22	1.36	1.63
	\sqrt{N}	\sqrt{N}	\sqrt{N}	\sqrt{N}	\sqrt{N}

4 Possíveis hermenêuticas econômicas do PL

Uma das principais consequências dos resultados apresentados anteriormente é o fato de que os dados compatíveis com uma distribuição de PL compartilhariam uma taxa média de crescimento. Em outras palavras, os países que se encontram na região do PL teriam uma taxa comum de crescimento invariável em seus GDPs por trabalhador. Não obstante, a emergência da dinâmica de invariância de escala sugere que o crescimento desses países pode estar sendo determinado por meio de um mecanismo como o de crescimento proporcional randômico descrito no Capítulo 1.

Dessa forma, a grande disparidade entre os países comentada na introdução poderia ter origem em fatores aleatórios que impactaram as taxas de crescimento no passado causando um efeito acumulado mais significativo no presente. Por exemplo, no caso de firmas, esse raciocínio pode ser ilustrado por meio do resultado final de diversos fatores econômicos que incluem componentes aleatórios como a efetividade de uma campanha publicitária, o tempo de absorção de um novo produto, greves, etc. Nesse cenário, pequenas diferenças de crescimento no passado se tornariam, eventualmente, responsáveis por grandes disparidades no tamanho das firmas e gerariam o formato exponencial característico de um PL.

Outra intuição importante que pode ser retirada é a sugestão de que o processo de crescimento comentado estaria associado a algum tipo de auto organização dos países. Um exemplo físico é fornecido por Dahr, Sadhu, Chandra (2009) em um modelo de formação de padrões em montes de areia. Os autores utilizam o mecanismo de crescimento proporcional para gerar padrões complexos que remetem a uma situação de auto organização. Esse modelo concretiza o fato de que o mecanismo de crescimento aleatório proporcional e a emergência de relações do tipo PL podem ser sinais de auto organização em determinado sistema. Por fim, a característica de compartilhar uma taxa média de crescimento pode ter funcionalidade para esses países. Intuitivamente, duas sugestões principais podem ser comentadas, i) essa forma de organização dos países no PL permitiria a difusão de capital tecnológico e humano proporcionando o que proporcionaria a taxa comum de crescimento entre as observações, ii) Esse tipo de organização poderia ser também resultado de acordos e blocos comerciais permitindo a troca de bens e serviços novamente suportando a ideia de uma taxa média de crescimento entre os países.

5 Conclusões

As principais conclusões do presente trabalho são: i) No período de 1997 a 2007 a parcela correspondente a cerca de 37% da amostra total de 169 países disponíveis na Penn World Table (PWT) (Feenstra, Robert C., Tobert Inklaar e Marcel P. Timmer, 2015) não pode ser dita não compatível com um PL com distribuição no formato (2) e no período de 2008 a 2014 a conclusão se mantém para cerca de 24% da amostra total; ii) Essa regularidade sugere que o processo por trás da alocação de riqueza relativa entre esses países possa ser o mecanismo de crescimento proporcional aleatório. Essa sugestão implicaria que os países compartilhariam uma taxa média de crescimento e estariam sujeitos a uma dinâmica de *Rich get richer*. Ou seja, a distribuição resultante é composta por poucos indivíduos muito ricos e vários significativamente menos ricos; iii) Em uma outra perspectiva, o trabalho também aborda o papel do mecanismo de crescimento proporcional em situações de auto organização. Dentro dessa perspectiva, o questionamento que surge é se, semelhante a outros sistemas em que o mecanismo de crescimento aleatório proporcional está presente e emerge uma relação do tipo PL, o caso estudado aqui representaria uma forma de auto organização eficiente ou que estaria alocando eficientemente a riqueza em termos de GDP por trabalhador.

Com relação as principais limitações do presente trabalho é importante apontar que conforme sugerido por Gabaix (2009) o valor mínimo para o qual o PL se sustenta é reflexo de uma força de natureza econômica. O trabalho não explora esse tema, sendo este uma oportunidade futura de pesquisa. Não obstante, permanece o questionamento importante sobre a relação entre as decisões individuais dos agentes e o resultado agregado caracterizado pelo PL. Em essência, essa questão põe em voga a possível relação entre as decisões maximizadoras de lucro das firmas, por exemplo, e o padrão agregado observado pela literatura (poucas firmas muito grandes de forma que a segunda maior tem aproximadamente metade do tamanho da maior, a terceira um terço e assim por diante). Por fim, um ponto deixado em aberto para futuras pesquisas é a formalização da ideia da possível auto organização entre os países e sua relação com a regularidade empírica estudada aqui.

Referências

- [1] DI GUILMI, C., *et al.* “Power Law Scaling in the World Income Distribution”. *Economics Bulletin*, Vol. 15, No. 6, p. 1-7, ago. 2003
- [2] SMITH, Adam. “Wealth of Nations”. University of Chicago Bookstore, 2005
- [3] KALDOR, N. “A model of economic growth”. *The economic journal*, Vol. 67, No. 268, p. 591-624, dez. 1957
- [4] PARENTE, Stephen L., e PRESCOTT, E. “Changes in the Wealth of Nations”. Federal Reserve Bank of Minneapolis. *Quarterly Review-Federal Reserve Bank of Minneapolis*, Vol. 17, No. 2, p. 3, mar. 1993
- [5] PRICHETT, L. “Divergence, big time”. *The Journal of Economic Perspectives*, Vol. 11, No. 3, p. 3-17, jul. 1997
- [6] BEN-DAVID, D. “Convergence clubs and diverging economies”. No. 95. Centre for Economic Policy Research, fev. 1994
- [7] JONES, C. “On the evolution of the World Income Distribution”. *The Journal of Economic Perspectives*, Vol. 11, No. 3, p.19-36, 1997
- [8] FEENSTRA, Robert C., INKLAAR, R. e TIMMER, M. “The next generation of the Penn World Table.” *The American Economic Review*, Vol. 105, No. 10, p. 3150-3182, out. 2015
- [9] CLAUSET, A., SHALIZI, C., e NEWMAN, E. J. “Power-law distributions in empirical data.” *SIAM review*, Vol. 51, No. 4, p. 661-703, nov. 2009
- [10] BRIAN ARTHUR, W. *et al* “Complexity and the Economy”. *SCIENCE*, Vol. 284, p. 107-109, abr. 1999
- [11] FOLEY, D.; DOYLE FARMER, J. “The economy needs agent-based modelling”. *NATURE*, Vol 460, p. 685-686, ago. 2009
- [12] DOSI, G; NELSON, R. “An Introduction to evolutionary theories in economics”, *Journal of Evolutionary Economics*, v. 4, n. 3, p. 153-172, set. 1994
- [13] NELSON, R.; WINTER, S. “Evolutionary Theorizing in Economics”, *The Journal of Economic Perspectives*, v. 16, n. 2, p. 23 - 46, jan. 2002
- [14] BRIAN ARTHUR, W. ”Inductive reasoning and bounded rationality (The El Farol Problem).” *American Economic Review*, vol. 84, issue 2, p. 406-11, 1994
- [15] KAHNEMAN, D. “Maps of Bounded Rationality: Psychology for Behavioral Economics”. *The American Economic Review*, Volume 93, No 5, p. 1449-1475, dec. 2003
- [16] KRUGMAN, P. “Confronting the mystery of urban hierarchy.” *Journal of the Japanese and International economies*, Vol. 10, No. 4, p.399-418, dez. 1996
- [17] AXTELL, Robert L. “Zipf distribution of US firm sizes.” *Science*, Vol. 293, No. 5536, p. 1818-1820, set. 2001

- [18] GABAIX, Xavier. "Power laws in economics and finance." *Annu. Rev. Econ.* Vol. 1, No. 1, p. 255-294, set. 2009
- [19] JESSEN, *et al.* "Regularly varying functions". University of Copenhagen, laboratory of Actuarial Mathematics, jul. 2006
- [20] NEWMAN, E. J. "Power laws, Pareto distributions and Zipf's law." *Contemporary physics*, Vol. 46, No. 5, p. 323-351, set. 2005
- [21] TALEB, N. "Black swans and the domains of statistics." *The American Statistician*, Vol. 61, No. 3, p. 198-200, ago. 2007
- [22] STANLEY, M., *et al.* "Zipf plots and the size distribution of firms." *Economics letters*, Vol. 49, No. 4, p. 453-457, out. 1995
- [23] ZIPF, G. "The psycho-biology of language." 1935
- [24] ZIPF, G. "Human behavior and the principle of least effort: an introduction to human ecology". Addison-Wesley Press, 1949.
- [25] FAGAN, S., e GENÇAY, R. "An introduction to textual econometrics." *Handbook of empirical economics and finance*, P.133-154, 2011
- [26] LUTTMER, E. "Selection, growth, and the size distribution of firms." *The Quarterly Journal of Economics*, Vol. 122, No. 3, p. 1103-1144, ago. 2007
- [27] GABAIX, X., e IOANNIDES, Y. "The evolution of city size distributions." *Handbook of regional and urban economics*, Vol. 4, p. 2341-2378, dez. 2004
- [28] MANARIS, B., *et al.* "Investigating Esperanto's Statistical Proportions Relative to other Languages using Neural Networks and Zipf's Law." *Artificial Intelligence and Applications*, p. 102-108, fev. 2006.
- [29] PIANTADOSI, S. "Zipf's word frequency law in natural language: A critical review and future directions." *Psychonomic bulletin e review*, Vol. 21, No. 5, p. 1112-1130, out. 2014
- [30] KOHLER, R. "Power Law Models in Linguistics: Hungarian." *Glottometrics*, Vol. 5, p. 51-61, jan. 2002
- [31] SCHROEDER, M. "Fractals, chaos, power laws: Minutes from an infinite paradise". Courier Corporation, mar. 2012.
- [32] YULE, G. "A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FRS." *Philosophical transactions of the Royal Society of London*, Vol. 213, p. 21-87, jan. 1925
- [33] BARRO, J., e BARRO, Robert J. "Pay, performance, and turnover of bank CEOs." *Journal of Labor Economics*, Vol. 8, No. 4, p. 448-481, out. 1990
- [34] KOSTIUK, P. "Firm size and executive compensation." *Journal of human Resources*, p. 90-105, jan. 1990
- [35] FRYDMAN, C., e SAKS, R. "Historical trends in executive compensation, 1936-2003." *Journal of Economic History*. Vol. 67. No. 2, p. 520-521, jun. 2007.

- [36] SCHUMPETER, J. "Vilfredo Pareto (1848-1923)." The Quarterly Journal of Economics, Vol. 1, p.147-173, maio 1949
- [37] GIBRAT, R. "Les inégalités économiques". Recueil Sirey, 1931
- [38] KESTEN, H. "Random difference equations and renewal theory for products of random matrices." Acta Mathematica, Vol. 131, No. 1, p. 2017-248, dez. 1973
- [39] CORDOBA, Juan-Carlos. "Zipf's law: a case against scale economies." Processed, University of Rochester, 2001
- [40] GABAIX, X. "Power laws in economics: An introduction." The Journal of Economic Perspectives, Vol. 30, No. 1, p. 185-205, jan. 2016
- [41] HESTON, A., SUUMMERS, R., e ATEN, B. "Penn world table version 6.1." Center for International Comparisons at the University of Pennsylvania (CICUP), Vol. 18, out. 2002
- [42] GABAIX, X., e IBRAGIMOV, R. "Rank - $1/2$: a simple way to improve the OLS estimation of tail exponents." Journal of Business and Economic Statistics, Vol. 29, No. 1, p. 24-39, jan. 2011
- [43] MASSEY, Jr., F. J. "The Kolmogorov-Smirnov test for goodness of fit." Journal of the American statistical Association, Vol. 46, No. 253, p. 68-78, mar. 1951
- [44] DHAR, D. *et al.* "Pattern formation in growing sandpiles." EPL (Europhysics Letters), Vol. 85, No. 4, p. 48002, mar. 2009