Lecture 8 - Firm Size Distributions

UCLA - Econ 221 - Fall 2018

François Geerolf

UCLA

November 21, 2018

Outline

1. Introduction: Pareto Distributions
   - Zipf's Law for Cities
   - Zipf's Law for Firm Sizes
   - Pareto Distributions for Labor Incomes
   - Pareto Distributions for Wealth
   - Other: Leverage Ratio Distributions, etc.

2. Mathematical Preliminaries
   - Kolmogorov Forward Equation for Diffusion
   - Intuition for Wiener Process

3. First Random Growth Model: Gabaix (1999)

Basics on Pareto Distributions

- Denote by $f(x)$ the density function for the population x of cities, and $F(x)$ the cumulative distribution.
- Then the survivor function for a standard Pareto distribution with:
  - A **scale parameter $\mathbf{x_m}$** (minimum value that x can take).
  - A **shape parameter a**, governing the fattailedness. Also named **Pareto parameter**, or **tail coefficient**.

  is given by (probability that population is greater than x):

$$\mathbb{P}[x' > x] = \boxed{1 - F(x) = \left(\frac{x_m}{x}\right)^a}.$$

- The density is then given by:

$$f(x) = a\frac{x_m^a}{x^{a+1}}.$$

- Pareto parameter is obtained as linear regression of log (survivor) on log(x):

$$\boxed{\log\left(1 - F(x)\right) = -a\log(x) + a\log(x_m).}$$

Intuition

- Fractal inequality in the case of power law distributions.
- Many way to say the same thing:
  - Average population size above a certain threshold is proportional to this threshold:

$$\mathbb{E}\left[x' | x' \geq x\right] = \frac{a}{a-1} x.$$

    Note: $b = a/(a-1)$ is called the inverted Pareto-Lorenz coefficient (cf. next slide).

  - Top 0.01% cities have $10^{1/a}$ more inhabitants on average than top 0.1% cities.
  - Share of inhabitants living in the top p percent of cities is given by:

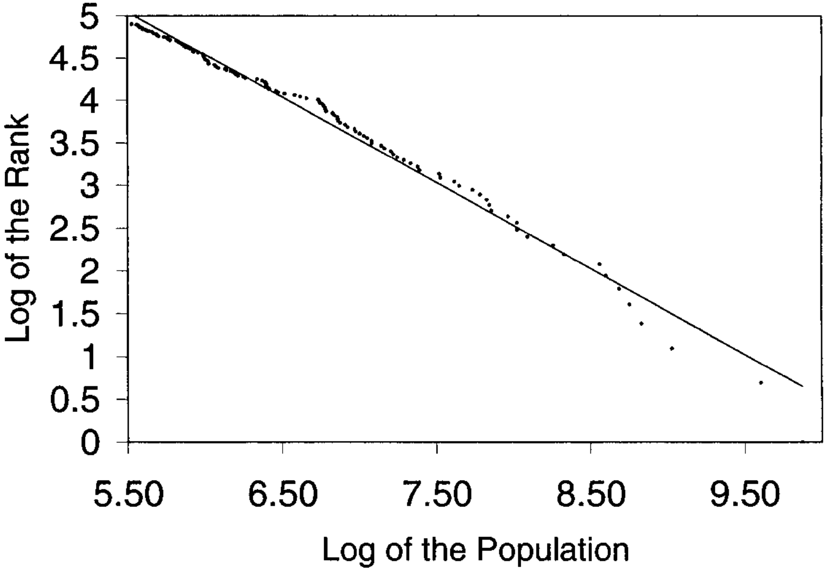
$$S(p) = \left(\frac{100}{p}\right)^{1/a-1}.$$

Pareto VS Inverted Pareto Coefficients

TABLE 3
PARETO-LORENZ α COEFFICIENTS VERSUS INVERTED-PARETO-LORENZ β COEFFICIENTS

α	$\beta = \alpha/(\alpha - 1)$	β	$\alpha = \beta/(\beta - 1)$
1.10	11.00	1.50	3.00
1.30	4.33	1.60	2.67
1.50	3.00	1.70	2.43
1.70	2.43	1.80	2.25
1.90	2.11	1.90	2.11
2.00	2.00	2.00	2.00
2.10	1.91	2.10	1.91
2.30	1.77	2.20	1.83
2.50	1.67	2.30	1.77
3.00	1.50	2.40	1.71
4.00	1.33	2.50	1.67
5.00	1.25	3.00	1.50
10.00	1.11	3.50	1.40

Gabaix (1999)



Log of the Rank
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
5.50
6.50
7.50
8.50
9.50
Log of the Population

Zipf's Law for cities

- Gabaix (1999): "Zipf's law for cities is one of the most conspicuous empirical facts in economics or in the social sciences generally."
- Example of the United States: order cities by population. And lot log(rank) as a function of log (Population). One obtains a linear relationship, with a slope of -1.
- Very concretely, this means that the second largest city is twice as large as the first, the third is three times as large as the third, etc. Examples (with 2013 numbers from Census via Wikipedia):

  1. New York: 8,405,837.
  2. Los Angeles: 3,884,307.
  3. Chicago: 2,718,782.

Taken from Wikipedia (Feb 28, 2015)

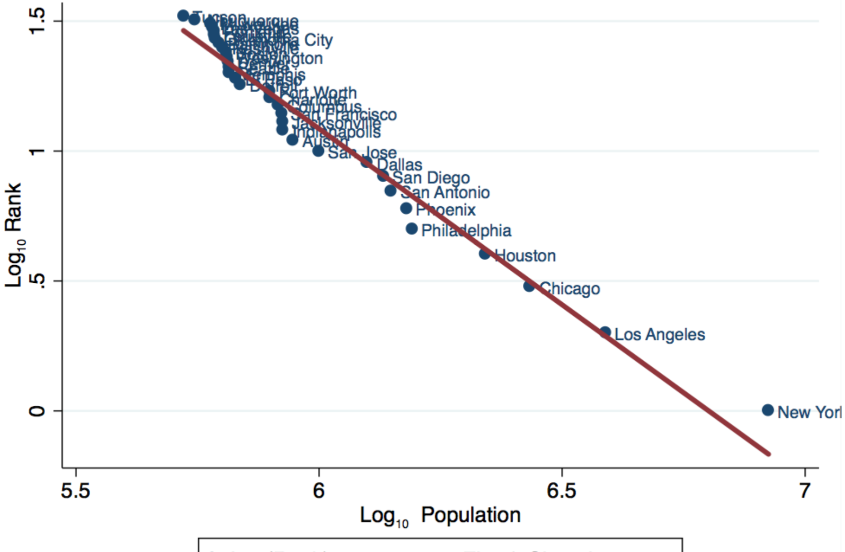
2013 rank	City	State[5]	2013 estimate
1	*New York*[6]	New York	8,405,837
2	*Los Angeles*	California	3,884,307
3	*Chicago*	Illinois	2,718,782
4	*Houston*[7]	Texas	2,195,914
5	*Philadelphia*[8]	Pennsylvania	1,553,165
6	***Phoenix***	Arizona	1,513,367
7	San Antonio	Texas	1,409,019
8	San Diego	California	1,355,896
9	Dallas	Texas	1,257,676
10	San Jose	California	998,537
11	**Austin**	Texas	885,400
12	***Indianapolis***[9]	Indiana	843,393
13	*Jacksonville*[10]	Florida	842,583
14	San Francisco[11]	California	837,442
15	***Columbus***	Ohio	822,553
16	*Charlotte*	North Carolina	792,862
17	Fort Worth	Texas	792,727
18	*Detroit*	Michigan	688,701
19	El Paso	Texas	674,433
20	*Memphis*	Tennessee	653,450
21	*Seattle*	Washington	652,405
22	***Denver***[12]	Colorado	649,495

Taken from Wikipedia (Feb 28, 2015)

Rank	City	State	2013 estimate	Log(Rank)	Log(Population)
1	New York	New York	8,405,837	0	6.924580964
2	Los Angeles	California	3,884,307	0.301029996	6.589313547
3	Chicago	Illinois	2,718,782	0.477121255	6.434374386
4	Houston[7]	Texas	2,195,914	0.602059991	6.341615328
5	Philadelphia[8]	Pennsylvania	1,553,165	0.698970004	6.191217595
6	Phoenix	Arizona	1,513,367	0.77815125	6.17994426
7	San Antonio	Texas	1,409,019	0.84509804	6.148916849
8	San Diego	California	1,355,896	0.903089987	6.13222638
9	Dallas	Texas	1,257,676	0.954242509	6.099568773
10	San Jose	California	998,537	1	5.999364162
11	Austin	Texas	885,400	1.041392685	5.947139518
12	Indianapolis[9]	Indiana	843,393	1.079181246	5.926029992
13	Jacksonville[10]	Florida	842,583	1.113943352	5.925612693
14	San Francisco[	California	837,442	1.146128036	5.922954738
15	Columbus	Ohio	822,553	1.176091259	5.915163891
16	Charlotte	North Carolina	792,862	1.204119983	5.899197604
17	Fort Worth	Texas	792,727	1.230448921	5.89912365
18	Detroit	Michigan	688,701	1.255272505	5.838030714
19	El Paso	Texas	674,433	1.278753601	5.828938812
20	Memphis	Tennessee	653,450	1.301029996	5.815212362
21	Seattle	Washington	652,405	1.322219295	5.814517281
22	Denver[12]	Colorado	649,495	1.342422681	5.812575812
23	Washington[13]	District of Columbia	646,449	1.361727836	5.810534268
24	Boston	Massachusetts	645,966	1.380211242	5.81020966
25	Nashville[14]	Tennessee	634,464	1.397940009	5.802406985
26	Baltimore[15]	Maryland	622,104	1.414973348	5.793862994
27	Oklahoma City	Oklahoma	610,613	1.431363764	5.785766046
28	Louisville[16]	Kentucky	609,893	1.447158031	5.785253649
29	Portland	Oregon	609,456	1.462397998	5.784942357
30	Las Vegas	Nevada	603,488	1.477121255	5.780668639

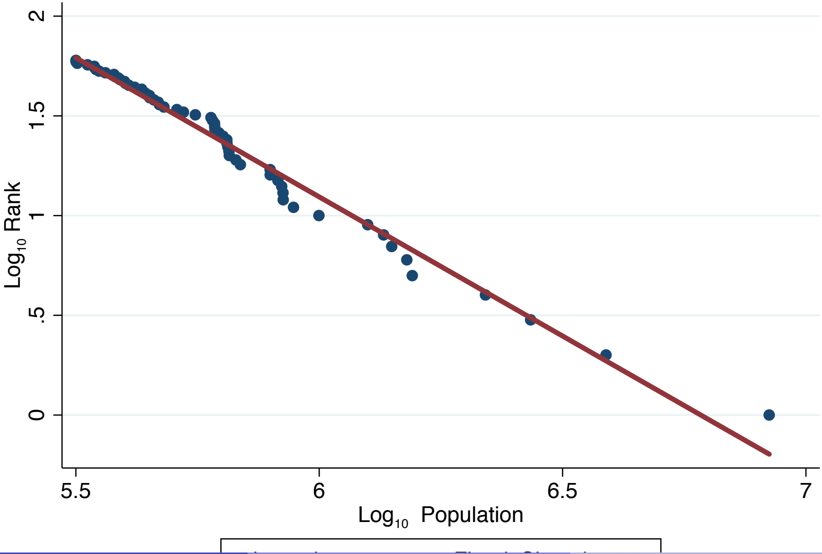
Plotting with Stata

- Beware: compared with Gabaix (1999), this is log10(.).

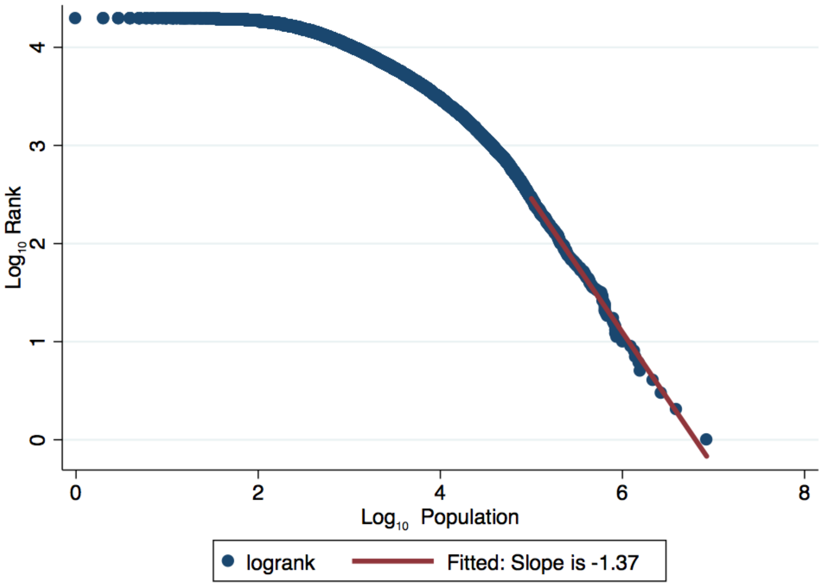


More cities: Entire Distribution

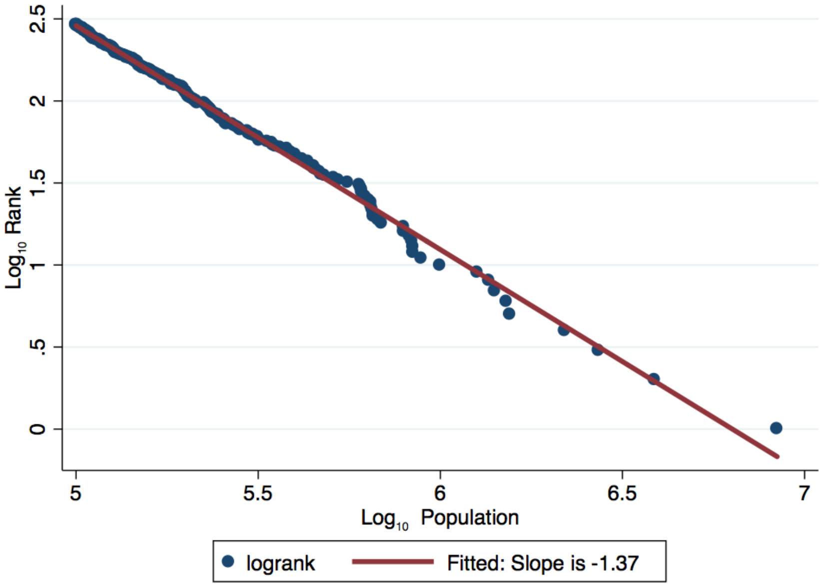
- Beware: compared with Gabaix (1999), this is log10(.).



Needed Cutoff for Zipf ? $log10 > 3.75$



Needed Cutoff for Zipf ? $log10 > 3.75$

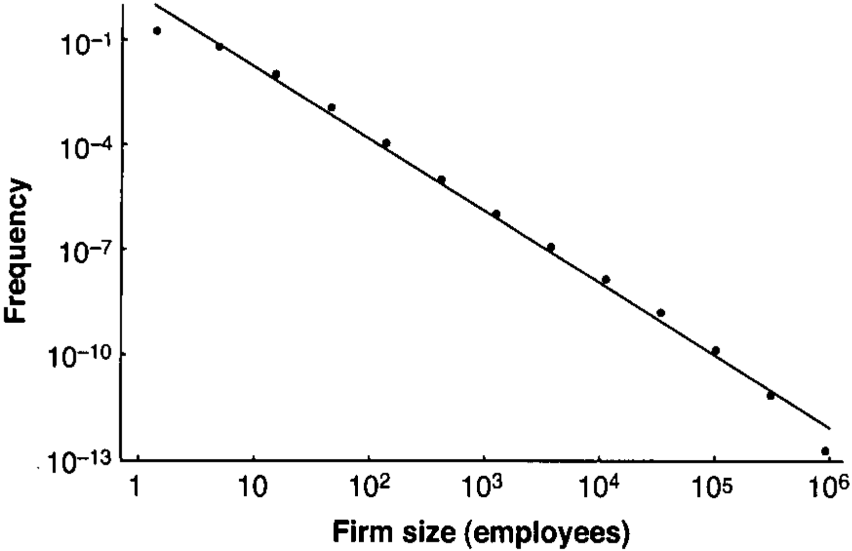


1. Introduction: Pareto Distributions
   - Zipf's Law for Cities
   - Zipf's Law for Firm Sizes
   - Pareto Distributions for Labor Incomes
   - Pareto Distributions for Wealth
   - Other: Leverage Ratio Distributions, etc.

2. Mathematical Preliminaries
   - Kolmogorov Forward Equation for Diffusion
   - Intuition for Wiener Process

3. First Random Growth Model: Gabaix (1999)

Axtell (2001)



Frequency
10^{-1}
10^{-4}
10^{-7}
10^{-10}
10^{-13}
1
10
10^2
10^3
10^4
10^5
10^6
Firm size (employees)

Not just the US: Evidence in di Giovanni and Levchenko (2013)

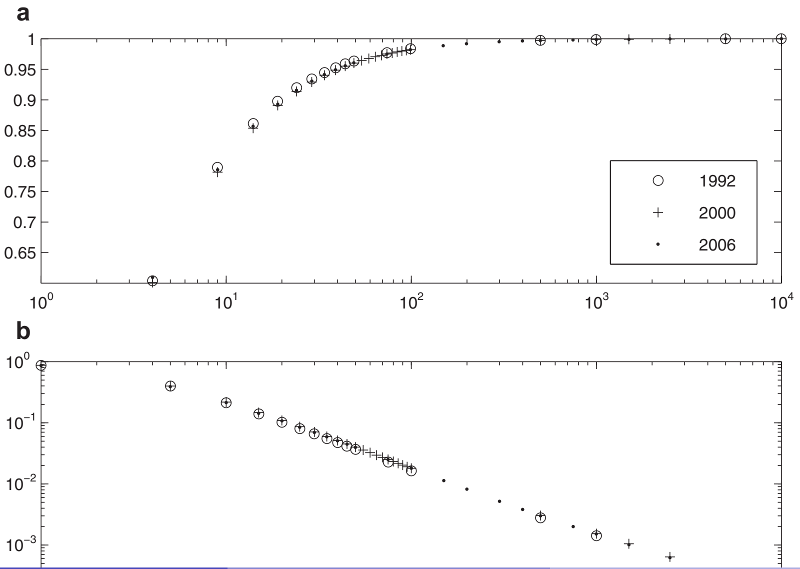
Table A2
Country-by-country estimates of power laws in firm size.

Country	CDF estimation			PDF estimation		
	PL coef.	R^2	*p*-Value	PL coef.	R^2	*p*-Value
Argentina	−1.046**	0.988	0.243	−2.039**	0.994	0.466
Australia	−0.992**	0.986	0.838	−1.905**	0.994	0.076
Austria	−0.695**	0.963	0.000	−1.677**	0.989	0.000
Belgium	−0.972**	0.999	0.011	−1.956**	0.998	0.150
Bosnia and Herzegovina	−1.022**	0.990	0.508	−2.036**	0.992	0.550
Brazil	−0.918**	0.966	0.162	−1.892**	0.991	0.096
Bulgaria	−0.981**	0.979	0.686	−2.007**	0.992	0.908
Canada	−0.888**	0.989	0.004	−1.913**	0.995	0.069
China	−1.117**	0.976	0.060	−2.091**	0.996	0.061
Croatia	−1.094**	0.988	0.034	−2.120**	0.992	0.074
Czech Republic	−1.083**	0.992	0.020	−2.072**	0.998	0.031
Denmark	−0.776**	0.950	0.003	−1.684**	0.987	0.001
Estonia	−1.017**	0.986	0.674	−2.067**	0.987	0.389
Finland	−0.869**	0.989	0.001	−1.879**	0.997	0.006
France	−0.886**	0.999	0.000	−1.894**	1.000	0.000
Germany	−0.853**	0.999	0.000	−1.960**	0.981	0.653
Greece	−0.992**	0.997	0.620	−1.951**	0.998	0.089
Hungary	−0.953**	0.995	0.050	−1.987**	0.996	0.741
India	−0.975**	0.988	0.476	−1.954**	0.995	0.319
Ireland	−0.761**	0.998	0.000	−1.718**	0.999	0.000
Italy	−1.030**	0.996	0.172	−2.037**	0.999	0.093
Japan	−0.955**	0.990	0.177	−1.985**	0.996	0.716
Latvia	−1.118**	0.989	0.011	−2.054**	0.995	0.281
Lithuania	−1.153**	0.992	0.001	−2.151**	0.996	0.009
Macedonia	−1.109**	0.999	0.000	−2.095**	0.990	0.176
Netherlands	−0.906**	0.994	0.002	−1.917**	0.995	0.082
Norway	−1.045**	0.970	0.454	−1.975**	0.997	0.516
Poland	−1.086**	0.987	0.051	−2.125**	0.995	0.028
Portugal	−0.919**	0.996	0.001	−1.924**	0.999	0.001
Korea	−0.880**	0.999	0.000	−1.860**	1.000	0.000
Romania	−1.002**	0.990	0.956	−2.047**	0.995	0.349
Russia	−1.039**	0.996	0.086	−2.027**	0.998	0.384
Serbia	−1.181**	0.989	0.001	−2.163**	0.996	0.004
Singapore	−0.888**	0.979	0.021	−1.825**	0.995	0.002
Slovakia	−1.139**	0.990	0.003	−2.124**	0.996	0.018
Slovenia	−0.993**	0.986	0.846	−1.998**	0.989	0.981
Spain	−0.978**	1.000	0.005	−2.011**	0.997	0.769
Sweden	−0.884**	0.997	0.000	−1.895**	0.998	0.002
Switzerland	−0.791**	0.990	0.000	−1.760**	0.996	0.000
Taiwan POC	−0.889**	0.989	0.003	−1.863**	0.991	0.031
Thailand	−0.956**	0.976	0.381	−1.953**	0.994	0.358
Ukraine	−1.058**	0.991	0.102	−2.007**	0.999	0.802
United Kingdom	−1.010**	0.975	0.856	−2.017**	0.992	0.775

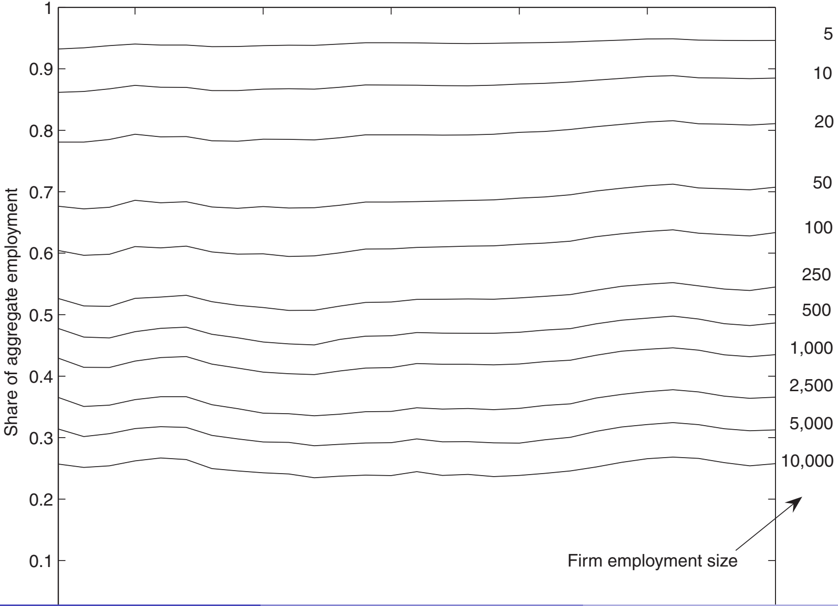
Notes: This table reports the estimated of power laws in firm size across countries. Column "PL coef." reports the coefficient on the power law for each country, the second column reports the R^2, the third column reports the *p*-value of the test that the power law coefficient is statistically different from -1 (-2 in the right panel). The estimates are based on firm-level sales data from ORBIS. Variable definitions, sources, and estimation techniques are described in detail in the text.

** Significant at the 1% level.

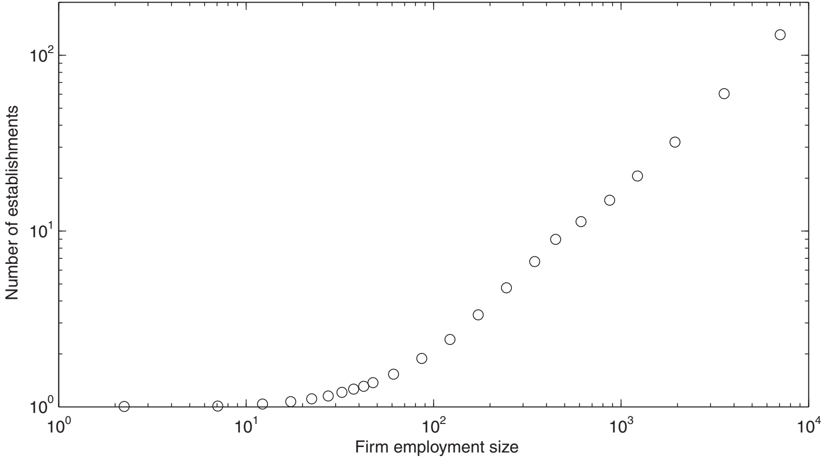
Firm Size Distribution reported by SBA for 1992, 2000, and 2006 - implied $a \approx 1.05$. Source: Luttmer (2010)



Repeated Cross-Section - Source: Luttmer (2010)



Average Number of Establishments Per Firms - Source: Luttmer (2010)



Known since Pareto (1896)

SAXE

Années	Revenus I	Revenus II
1879	959	327
1880	982	330
1882	1.059	346
1884	1.141	362
1886	1.237	386
1888	1.338	407

I = total en millions de marks. II = moy^nne par habitant en marks.

957. Répartition de la richesse. La répartition de la richesse peut dépendre de la nature des hommes dont se compose la société, de l'organisation de celle-ci, et aussi, en partie, du *hasard* (les *conjonctures* de Lassalle), c'est-à-dire de cet ensemble de causes inconnues, agissant tantôt dans un sens, tantôt dans un autre, auxquelles, dans notre ignorance de leur vraie nature, nous donnons le nom de hasard.

C'est à l'observation de nous renseigner sur la part qu'ont réellement ces causes dans la répartition de la richesse. Si nous trouvons que la répartition de la richesse varie considérablement et d'une manière irrégulière, nous en conclurons que « le hasard » a une part considérable dans la production de ce phénomène. Si les variations de la répartition de la richesse suivent les variations de l'organisation économique, c'est à cette organisation que nous devrons attribuer une part prépondérante. Enfin, si la répartition de la richesse varie peu pour des contrées, des époques, des organisations différentes, il nous faudra conclure que, sans vouloir négliger les autres causes, nous devons chercher dans la nature de l'homme la cause principale qui détermine le phénomène.

958. Malgré les incertitudes que comportent les déclarations des contribuables pour l'impôt sur le revenu, c'est encore la base la plus sûre que nous ayons pour connaître, au moins d'une manière approchée, comment se répartit la richesse.

Dans ce qui suit, nous indiquerons par x un certain revenu, et par N le nombre de contribuables ayant un revenu supérieur à x.

En Angleterre, c'est seulement pour la schedule D : Commerce et professions, que nous avons une classification étendue des contribuables suivant l'importance des revenus. Mais, en compensation, il y a l'avantage d'avoir ces résultats pour des époques assez éloignées et pour des organisations économiques aussi différentes que le sont celles de l'Angleterre proprement dite et de l'Irlande.

Traçons deux axes AB et AC. Sur AB portons les logarithmes de x, sur AC les logarithmes de N.

Nous sommes tout de suite frappé du fait que les points ainsi déterminés, ont une tendance très marquée à se dis-

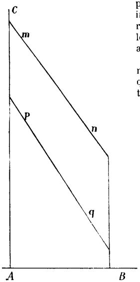


Fig. 47.

poser en ligne droite[1]. Disons immédiatement que nous allons retrouver cette tendance dans les nombreux exemples que nous aurons encore à examiner.

Un autre fait, tout aussi, et même plus remarquable, c'est que les courbes de la répartition des revenus, en Angleterre

Schedule D — Année 1893-94.

x £	N: Great Britain	N: Ireland
150	400 648	17 717
200	234 185	9 365
300	121 996	4 592
400	74 041	2 684
500	54 419	1 898
600	42 072	1 428
700	34 269	1 104
800	29 311	940
900	25 033	771
1000	22 896	684
2000	9 880	271
3000	6 069	142
4000	4 161	88
5000	3 081	68
10000	1 104	22

et en Irlande, présentent un parallélisme à peu près complet. Ce fait est à rapprocher d'un autre, que nous allons bientôt constater : les inclinaisons des lignes mn, pq obtenues pour dif-

(**958**) [1] C'est-à-dire que la courbe réelle est interpolée par une droite dont l'équation est

$$\log N = \log A - \alpha \log x. \tag{1}$$

L'équation générale de la courbe est peut-être

$$\log N = \log A - \alpha \log (a + x) - \beta x \,; \tag{2}$$

mais ce n'est que dans un seul cas (Oldenbourg) que nous avons trouvé une valeur appréciable pour β. Il est donc fort probable que β est, en général, négligeable, et qu'on a simplement

$$\log N = \log A = \alpha \log (a + x). \tag{3}$$

Quand il s'agit du revenu total, a est aussi, en général, fort petit et le plus souvent, de l'ordre des erreurs d'observation. Nous sommes donc ainsi ramené à l'équation (1).

Pays	Inclinaison α	Pays	Inclinaison α
Angleterre, 1843	1,50	Pérouse, campagne	1,37
» 1879-80	1,35	Ancône, Arezzo, Parme et Pise (ensemble)	1,32
Prusse, 1852	1,89	Villes italiennes (ensemble), (958 [2])	1,45
» 1876	1,72	Bâle, 1887	1,24
» 1881	1,73	Paris (loyers)	1,57
» 1886	1,68	Augsburg, en 1471	1,43 (*i*)
» 1890	1,60	» en 1498	1,47 (*i*)
» 1894	1,60	» en 1512	1,26 (*i*)
Saxe, 1880	1,58	» en 1526	1,13 (*i*)
» 1886	1,51	Perou (fin du 18e siècl.)	1,79 (*i*)
Florence	1,41		
Pérouse, ville	1,69		

Nous verrons plus loin (965 [1]) qu'une diminution de l'inclinaison α, indique une moindre inégalité des revenus.

960. Ces résultats sont très remarquables. Il est absolument impossible d'admettre qu'ils sont dus seulement au hasard. Il y a bien certainement une *cause* qui produit la tendance des revenus à se disposer suivant une certaine courbe. La forme de cette courbe parait ne dépendre que faiblement des différentes conditions économiques des pays considérés, puisque les effets sont à peu près les mêmes pour des pays dont les conditions économiques sont aussi différentes que celles de l'Angleterre, de l'Irlande, de l'Allemagne, des villes italiennes, et même du Pérou[1].

Certes, lorsqu'il s'agit de lois purement empiriques, on ne saurait être trop prudent. En tous cas, les conséquences que nous allons tirer de cette loi seront toujours valables, au moins, pour les peuples pour lesquels nous avons vu qu'elle se vérifie.

961. Si nous repassons des logarithmes aux nombres, nous aurons la courbe de la répartition des revenus[1]. C'est-

(**960**) [1] Denys d'Halic., *Ant., Rom. VII,* 59, dit qu'à Rome, les plus pauvres citoyens n'étaient pas moins nombreux que tous les autres, pris ensemble : *Οἱ δὲ ἀπορώτατοι τῶν πολιτῶν οὐχ ἐλάττους τῶν ἄλλων ἁπάντων ὄντες…* Sans attacher trop d'importance à ce rapprochement, on peut observer qu'en prenant, par exemple, la statistique des revenus en Saxe, le nombre des citoyens ayant un revenu de 500 à 800 marks est à peu près égal au nombre des citoyens ayant un revenu supérieur à 800 marks. Les revenus actuels de 500 à 800 marks peuvent correspondre à ce qu'étaient autrefois les revenus des citoyens les plus pauvres. Les esclaves représentent la partie de la population dont, actuellement, les revenus sont au-dessous de 500 marks.

(**961**) [1] La courbe $n\,t\,s$ de la Fig. 48 est celle qui correspond aux équations (4) et (5) de **158**) [1]. La surface $m\,n\,s\,x$ représente le nombre total des revenus.

Considérons l'équation

$$N_x = \frac{A}{(x+a)^{\alpha}}, \tag{1}$$

Known since Pareto (1896)

La question de savoir quelle est la forme de la partie $s\ t\ v$ n'est pas de simple curiosité. Des conséquences importantes découlent du fait que cette forme se rapproche de celle qui est indiquée par la Fig. 51.

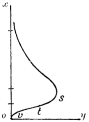


Fig. 50.

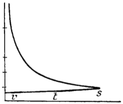


Fig. 51.

Il faut observer qu'en recherchant la répartition des revenus, nous ne nous occupons pas de leur provenance. L'homme, même le plus pauvre, doit être considéré comme ayant pour revenu la somme qui le fait vivre. Il importe peu que cette somme soit le fruit de son travail, ou qu'elle lui soit donnée par charité ou, enfin, qu'elle lui parvienne d'une manière quelconque, licite ou illicite.

962. La répartition des revenus n'est pas l'effet du hasard. A première vue, la courbe de la répartition des revenus ressemble à la courbe des probabilités, bien connue sous le nom de « courbe des erreurs ». On pourrait donc supposer que la répartition des revenus est simplement l'effet du hasard (les *conjonctures* de Lassalle). Les riches auraient eu les gros lots.

Il n'en est rien. Le profil qui résulterait de la loi des probabilités est beaucoup plus creusé que ne l'est celui de la Fig. 48. En d'autres termes, la courbe des probabilités se rapproche des axes beaucoup plus que la courbe de la Fig. 48.

L'importance de cette proposition nous a engagé à faire plusieurs essais pour tâcher de trouver une démonstration sans recourir aux mathématiques. Malheureusement, ces essais sont demeurés infructueux [1].

(**962**) [1] Plusieurs personnes qui manquent des connaissances scientifiques nécessaires pour bien comprendre les nouvelles théories, affirment que l'usage des mathématiques n'ajoute rien à nos connaissances en Economie politique, et elles croient le prouver en citant Cairnes. La seule preuve vraiment efficace serait de faire voir que l'on peut, sans re-

963. La base $v\ t\ s$ (Fig. 51) de la « pyramide sociale » étant fort écrasée, on peut, au moins pour une première approximation, la supposer plane. Alors la représentation de la répartition des revenus prend la forme $m\ n\ b\ s\ x$, Fig. 48.

Pouvons-nous étendre à toute la population la courbe des revenus ? Il paraît bien que oui, au moins approximativement.

Prenons comme exemple le royaume de Saxe. Nous avons vu (956) que l'impôt frappe toute personne, homme ou

courir aux mathématiques, démontrer le théorème dont nous venons de parler et bien d'autres encore.

A peine nos savants critiques auront daigné donner de telles démonstrations, nous ne manquerons pas de les substituer aux nôtres. En attendant, ils voudront bien nous permettre de donner ces démonstrations de la seule manière actuellement connue.

Si la répartition des revenus était seulement l'effet du hasard, la courbe $v\ t\ s$, Fig. 50, serait la courbe des probabilités (pour retrouver la forme qu'on a l'habitude de donner à la courbe des erreurs, on doit regarder la figure en disposant verticalement l'axe $o\ y$). L'évènement qui correspond à s serait le plus probable. La forme donnée par la Fig. 51 indique qu'il s'agit de la répétition d'un événement ayant une assez faible probabilité. Jusqu'ici, il n'y a pas de désaccord entre les faits et notre hypothèse ; car, en effet, la probabilité de s'enrichir est, partout, assez faible.

Soient, comme d'habitude, μ le nombre total des épreuves, m le nombre des épreuves favorables, n celui des épreuves contraires, p la probabilité de l'évènement favorable (le gain d'une certaine somme) et

$$q = 1 - p.$$

La probabilité d'avoir un revenu proportionnel à m sera

$$U = \frac{1.\ 2 \ldots \mu}{1.\ 2 \ldots m.\ 1.\ 2 \ldots n}\, p^m q^n . \tag{1}$$

On sait que l'évènement le plus probable est celui pour lequel m est égal à $p\,\mu$. Le plus grand nombre d'individus aurait donc le revenu $p\,\mu$. Les revenus inférieurs, ou supérieurs à $p\,\mu$ seraient ceux qui appartiendraient à des nombres moindres d'individus. Posons donc, en général

$$m' = p\mu\,, \quad m = m' + t\,,$$
$$n' = q\mu\,, \quad n = n' - t\,,$$

$$P_o = \frac{1.\ 2 \ldots \mu}{1.\ 2 \ldots m'.\ 1.\ 2 \ldots n'} \left(\frac{m'}{\mu}\right)^{m'} \left(\frac{n'}{\mu}\right)^{n'},$$

$$P = \frac{1.\ 2 \ldots m'.\ 1.\ 2 \ldots n'}{1.\ 2 \ldots (m' + t).\ 1.\ 2 \ldots (n' - t)} \left(\frac{m'}{\mu}\right)^{t} \left(\frac{n'}{\mu}\right)^{-t};$$

nous aurons

$$U = P_o\, P. \tag{2}$$

des faits, une grande importance aux mots [1]. On se dispute donc pour savoir auquel des deux phénomènes indiqué doit être réservée la dénomination de *moindre inégalité des fortunes*.

Les opinions qui ont cours à ce sujet sont fort bien expliquées par M^r Leroy Beaulieu, *Essai sur la répartition des richesses*, p. 45 et suiv. Il commence par rappeler ce que dit Lassalle : « Toute souffrance et toute privation humaine, de même que toute satisfaction humaine, par conséquent aussi la situation de chaque partie de l'humanité, ne peuvent se mesurer que par comparaison avec la situation dans laquelle se trouvent d'autres hommes du même temps relativement à la moyenne habituelle des besoins. La situation de chaque classe a toujours pour unique mesure la situation des autres classes dans le même temps. » Là dessus, M^r Leroy Beaulieu observe que, suivant Lassalle, « ce n'est pas la situation absolue de la population ouvrière qui importe, c'est la situation relative. Que les ouvriers soient bien nourris, bien logés, bien meublés, bien vêtus, qu'ils aient des loisirs, qu'ils jouissent de la sécurité du lendemain et du repos de la vieillesse, tout cela n'a pas d'importance... si d'autres hommes ont une table plus raffinée, des palais plus amples, des meubles plus agréables. *Sans doute, Lassalle aimerait mieux que la classe ouvrière fût plus misérable,* mais qu'il

(**964**) [1] On n'a pas tort lorsque l'on veut composer, non une œuvre scientifique, mais un plaidoyer. La première condition pour persuader les gens, c'est d'en être compris ; il est donc évident qu'un plaidoyer qui s'adresse au gros du public, ne doit employer que des termes que chacun comprend. Un tel plaidoyer, *s'il est réellement efficace*, peut être infiniment plus utile que l'œuvre scientifique la plus profonde.

Il faut, malheureusement, observer que les économistes littéraires, bien qu'ils aient composé des œuvres d'une réelle valeur, n'ont pas réussi jusqu'à présent, à persuader le gros du public et que loin de gagner du terrain, ils en perdent de jour en jour. Sauf l'Angleterre, où règne le libre échange, principalement parce qu'il est favorable aux intérêts de certains entrepreneurs, le reste des pays civilisés verse de plus en plus dans le protectionnisme. Le socialisme d'Etat et le socialisme tout court font chaque jour des progrès. Il se peut que la science économique soit, pratiquement, tout aussi inutile que l'économie politique littéraire, elle ne saurait, vraiment, l'être plus, et elle a au moins le mérite d'arriver à la connaissance des vraies causes des phénomènes.

D'ailleurs, toute œuvre scientifique a et doit avoir, de par sa nature même, moins de lecteurs qu'une œuvre littéraire. Le nombre d'exemplaires auquel ont été tirés les *Principia* de Newton, n'est rien en comparaison du nombre d'exemplaires qu'on a tiré de l'*Assommoir* de Zola.

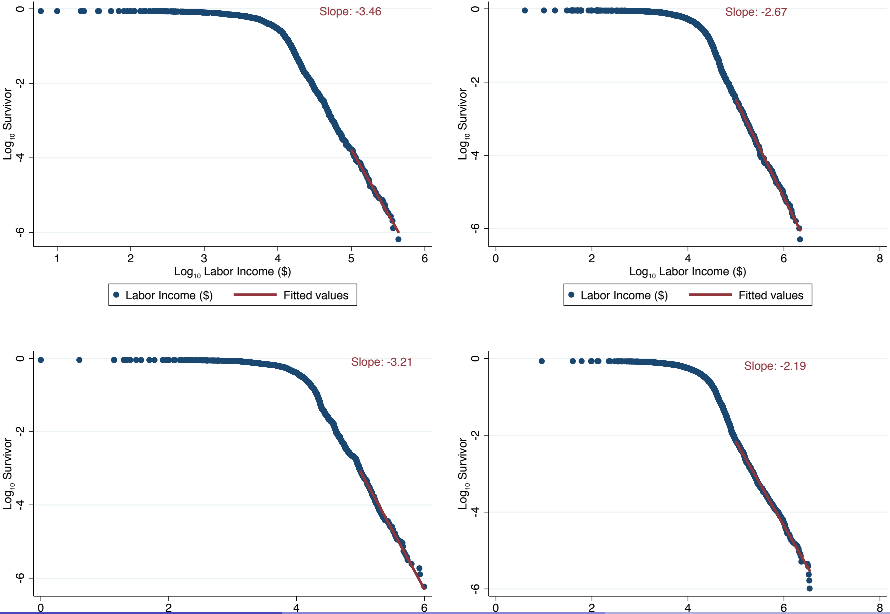
More contemporaneous evidence - Feenberg and Poterba (1993)

TABLE A-1.
Estimated Pareto Distribution Parameters, 1951–1990.

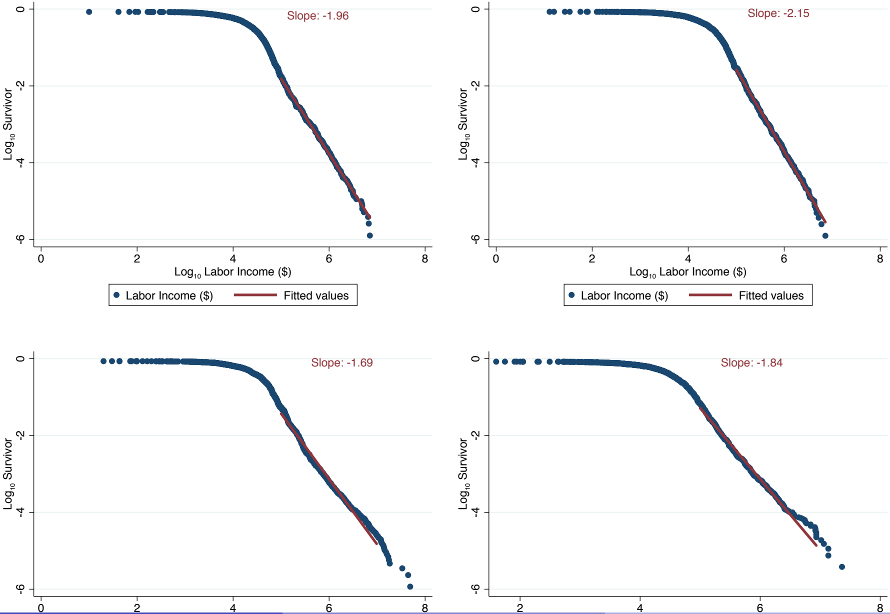
Year	$\hat{\alpha}$	$\hat{k}$
1951	1.83	1061
1952	1.79	967
1053	1.89	1159
1954	1.90	1205
1955	2.08	1720
1956	2.03	1661
1957	2.06	1731
1958	2.08	1782
1959	1.98	1685
1960	2.17	2124
1961	2.18	2240
1962	2.20	2366
1963	2.20	2503
1964	2.15	2454
1965	2.11	2505
1966	2.13	2713
1967	2.12	2919
1968	2.22	3558
1969	2.32	4006
1970	2.46	4725
1971	2.44	4892
1972	2.38	4959
1973	2.43	5587
1974	2.38	5674
1975	2.38	5891
1976	2.37	6342
1977	2.35	6621
1978	2.36	7445
1979	2.27	7324
1980	2.26	7904
1981	2.24	8293
1982	2.13	7614
1983	2.04	7174
1984	2.04	7876
1985	1.99	8036
1986	1.96	8711
1987	1.73	6830
1988	1.54	5390
1989	1.62	6845
1990	1.59	6698

Source: Authors' estimates using the method described in the text.

Public Tax Use - 1970, 1975 (top), 1980, 1985 (bottom)



Still Works - 1990, 1995 (top), 2000, 2005 (bottom)



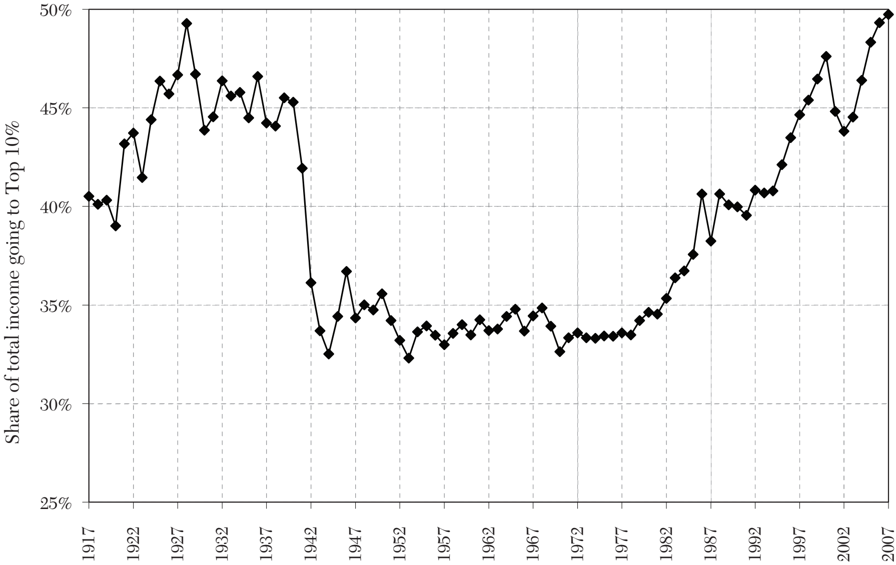
Remarks

- Contrary to the city size distribution, or the firm size distribution: the exponent varies quite a bit.
- In particular, Piketty and Saez express some of their results in terms of the top income share, as well as directly as a function of the Pareto coefficient. Recall that:

$$S(p) = \left(\frac{100}{p}\right)^{1/a-1}.$$

- Ideally, we want our theory to be able to explain the variations in these coefficients as well.

Top Income Shares



Top Income Shares

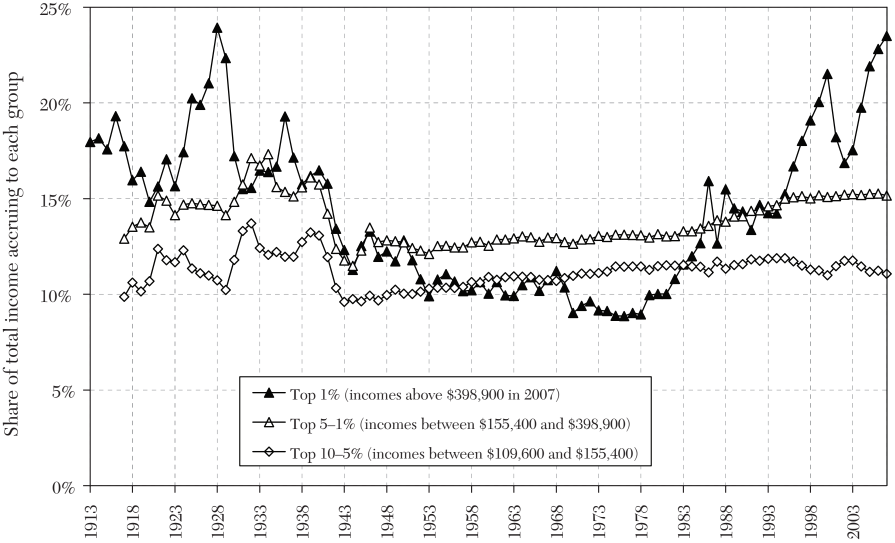
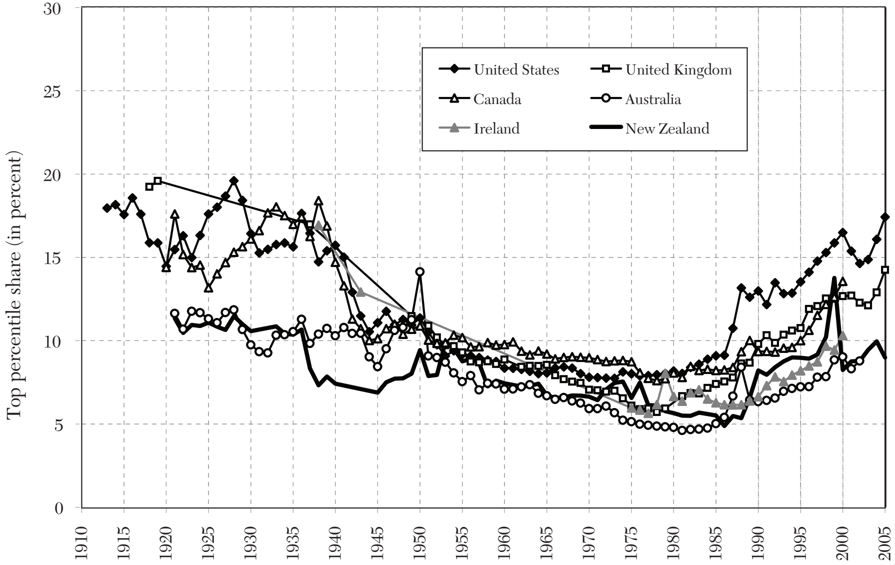
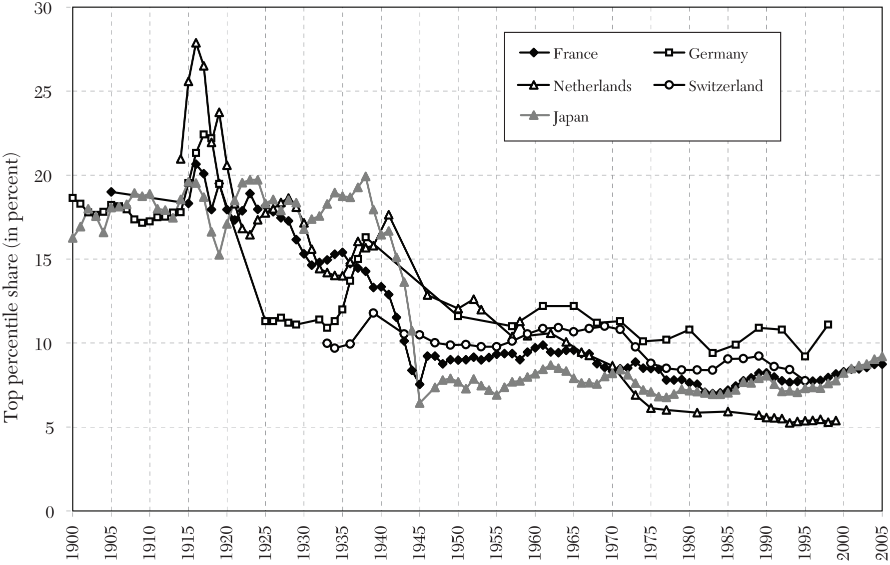


Figure 2. Decomposing the Top Decile US Income Share into three Groups, 1913–2007

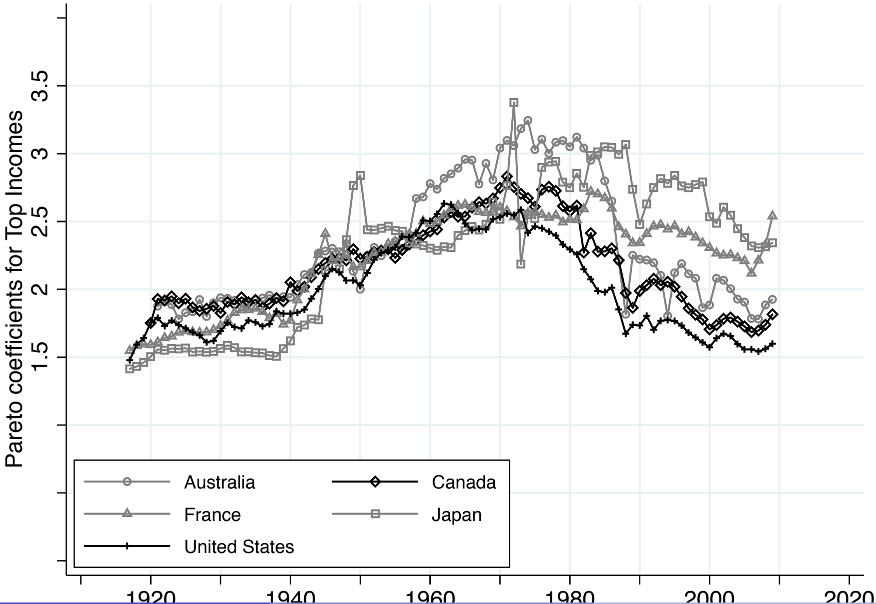
Top Income Shares



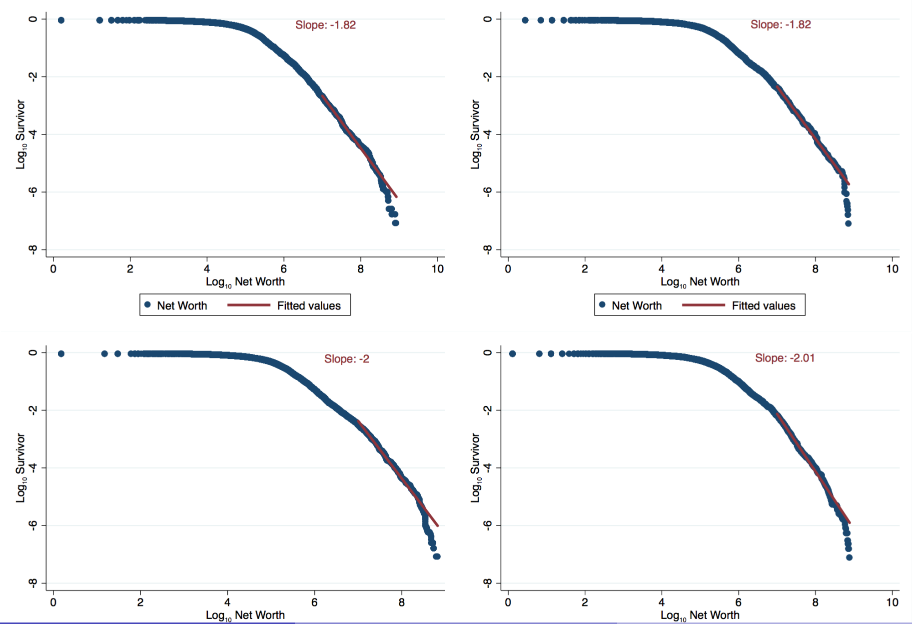
Top Income Shares



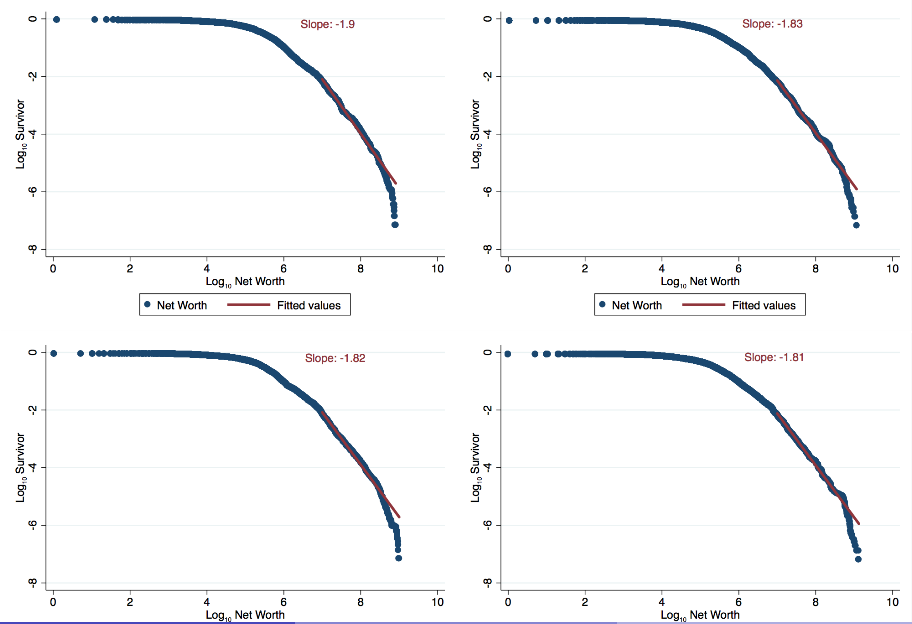
Pareto Coefficients Directly



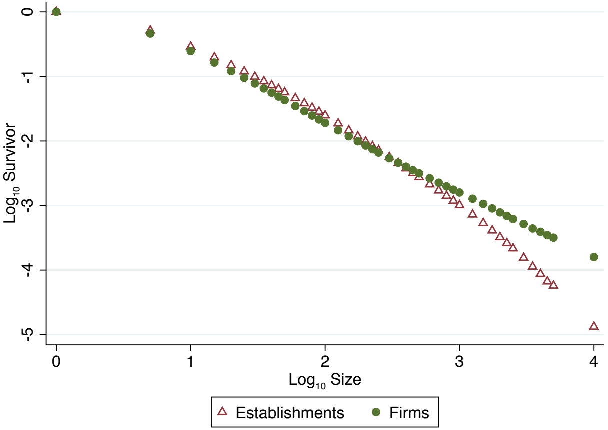
Source: SCF. Years: 1992, 1995, 1998, 2001



Source: SCF. Years: 2004, 2007, 2010, 2013

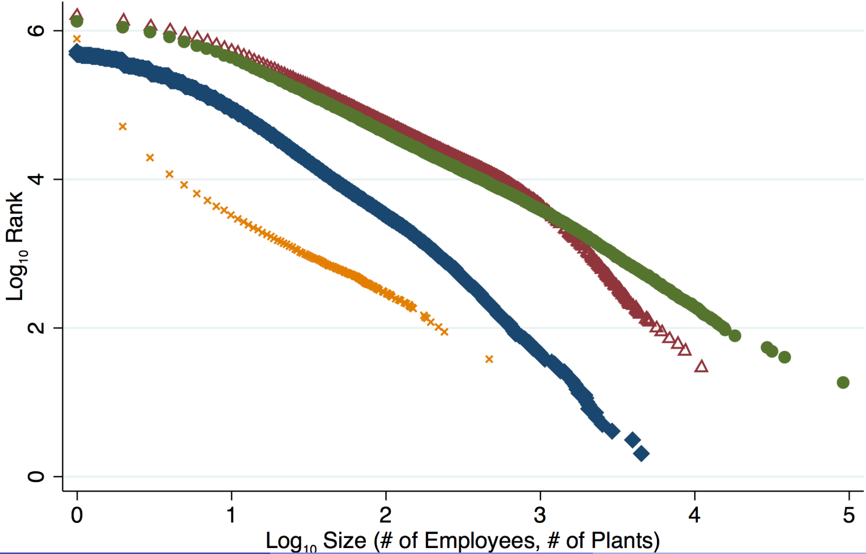


- More disaggregated levels of hierarchical organization are also Pareto distributed.

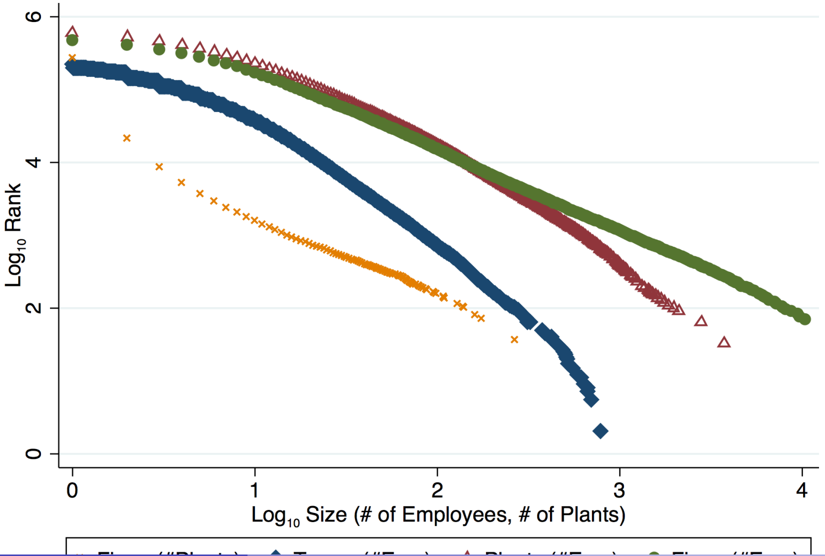


- One sees that on the US data (cf publicly available Census Data, here 1990).
- But even more clearly on French data (next 3 slides).

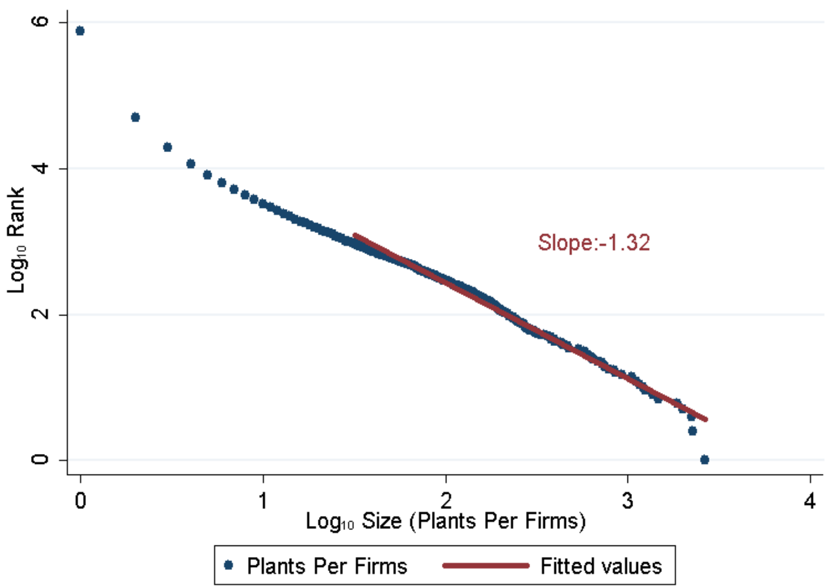
Source: French Matched Employer-Employee Data, 2007, All Sectors



Source: French Matched Employer-Employee Data, 2007, Wholesale Trade



Source: French Matched Employer-Employee Data, 2007



Kolmogorov Forward Equation (will be useful for Luttmer...)

- Assume process x follows a diffusion (of course, implicitly x is actually x_t, and so on):

$$dx = \mu(x)dt + \sigma(x)dw.$$

- Denote by $f(x,t)$ the cross-sectional density of x at time t.
- Time evolution of $f(x,t)$ given by a PDE known as the **Kolmogorov Forward Equation**:

$$\boxed{\frac{\partial f(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[\mu(x)f(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma(x)^2 f(x,t)\right]},$$

  with initial condition $f_0(x) := f(x,0)$
- If stationary distribution $f(x)$ exists, then it satisfies:

$$-\frac{d}{dx}\left[\mu(x)f(x)\right] + \frac{1}{2}\frac{d^2}{dx^2}\left[\sigma(x)^2 f(x,t)\right] = 0,$$

- Where does the Kolmogorov Forward Equation come from?

Kolmogorov Forward Equation

- To get intuition, let me simplify what we are looking at, and look at a standard Brownian motion (or Wiener Process):

$$dx = \mu dt + \sigma dw.$$

- Remember that heuristically:

$$dw = w(t+\Delta t) - w(t) = \epsilon\sqrt{\Delta t}, \quad \text{with} \quad \epsilon \sim IID\mathcal{N}(0,1)$$
$$\mathbb{E}[dw] = \mathbb{E}[\epsilon_t\sqrt{dt}] = 0 \qquad \mathbb{V}ar[dw] = \mathbb{E}[\epsilon_t^2 dt] = dt.$$

- Note that integrating we get:

$$x(t) = x(0) + \mu t + \sigma w(t).$$

That is, $x(t) - x(0)$ is normal with mean and variance linear in t:

$$x(t) - x(0) \sim \mathcal{N}(\mu t, \sigma^2 t).$$

Approximation by random walk 1/2

- Consider a discrete random walk with period length Δt.
- Over period of length Δt process $\Delta x = \{+h, -h\}$ with probabilities $\{p, 1-p\}$. Then we have that:

$$\mathbb{E}[\Delta x] = ph - (1-p)h = (2p-1)h.$$

and

$$\mathbb{V}ar[\Delta x] = \mathbb{E}[(\Delta x)^2] - \mathbb{E}[(\Delta x)]^2 = \left[1 - (2p-1)^2\right] h^2.$$

- Brownian motion is obtained as continuous time limit of this process when in particular h and Δt go to zero together in a particular way. That is, because we want nowhere differentiable process, we can show that if:

$$h = \sigma\sqrt{\Delta t}.$$

and then $\Delta t \to 0$ in this way, then standard Brownian motion.

Approximation by random walk 2/2

- Matching moments:

$$\mathbb{E}[\Delta x] = (2p-1)h = \mu \Delta t$$
$$\mathbb{V}ar[\Delta x] = \left[1 - (2p-1)^2\right] h^2 = \sigma^2 \Delta t.$$

- Eliminating h and solving for p gives quadratic equation:

$$p = \frac{1}{2}\left(1 + \sqrt{1 - \frac{1}{1 + (\mu/\sigma)^2 \Delta t}}\right) \approx \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{\Delta t}\right).$$

- Plugging back gives the intuitive (cf case with zero drift !):

$$h = \frac{\mu \Delta t}{2p-1} = \sigma\sqrt{\Delta t}.$$

Aside

- Look at Stokey (2009) to know more about diffusion processes more generally, that is processes of the form:

$$dx = \mu(x)dt + \sigma(x)dw.$$

- For our purposes of random growth models, we will need the geometric Brownian motion:

$$\mu(x) = \bar{\mu}x, \qquad \sigma(x) = \bar{\sigma}x.$$

(ie, dx/x is a standard Brownian motion).

- Orstein-Uhlenbeck process:

$$\mu(x) = -\alpha(x - \bar{x}), \qquad \sigma(x) = \bar{\sigma}, \quad \alpha > 0.$$

(ie, mean reversion to $\bar{x}$).

Kolmogorov Forward Equation Intuition (finally) 1/2

- Consider Brownian motion approximated by random walk as previously. We want to calculate how the cross-sectional distribution of an economic quantity following the Brownian motion evolves over time: $\frac{\partial f(x,t)}{\partial t}$??? We write it very naturally:

$$\frac{\partial f(x,t)}{\partial t} = \frac{f(x,t+\Delta t) - f(x,t)}{\Delta t}.$$

- Position at x might come from up or down move:

$$f(x,t+\Delta t) = pf(x-h,t) + (1-p)f(x+h,t).$$

- Expand this at the second order (like for the heuristics of Ito's lemma !):

$$f(x-h,t) \approx f(x,t) - \frac{\partial f(x,t)}{\partial x}h + \frac{1}{2}\frac{\partial^2 f(x,t)}{\partial x^2}h^2$$
$$f(x+h,t) \approx f(x,t) + \frac{\partial f(x,t)}{\partial x}h + \frac{1}{2}\frac{\partial^2 f(x,t)}{\partial x^2}h^2.$$

Kolmogorov Forward Equation Intuition (finally) 2/2

- Replacing, we therefore have:

$$f(x, t + \Delta t) = f(x, t) + \frac{1}{2}\frac{\partial^2 f(x, t)}{\partial x^2} h^2 - (2p - 1)\frac{\partial f(x, t)}{\partial x} h.$$

- Using the approximation of the Brownian motion through $h = \sigma\sqrt{\Delta t}$ and $(2p - 1)h = \mu\Delta t$, this becomes:

$$f(x, t + \Delta t) \approx f(x, t) + \frac{1}{2}\frac{\partial^2 f(x, t)}{\partial x^2}\sigma^2 \Delta t - \frac{\partial f(x, t)}{\partial x}\mu\Delta t.$$

- We therefore know how to calculate the transformation of this distribution over time:

$$\boxed{\frac{\partial f(x, t)}{\partial t} = \frac{1}{2}\frac{\partial^2 f(x, t)}{\partial x^2}\sigma^2 - \frac{\partial f(x, t)}{\partial x}\mu}.$$

Gibrat's Law

- **Dynamic** models generating Pareto distributions all at some level rely on a dynamic random growth process with statistical frictions:
- Emphasis on Dynamic.
  - Lucas span of control model gives that, for example.
  - Hierarchies model? (next week)
- There are many variations, through.
- In particular, for the source of statistical frictions.
- But all rely at some level on Gibrat's law.
  - Example on city sizes.
  - Discussion of Gabaix (1999).
- "Rome was not built in a day." Large cities from a number of lucky draws (for example: amenity shock, etc.). Similarly, large firms from a number of lucky draws.

Taken from Luttmer (2011)

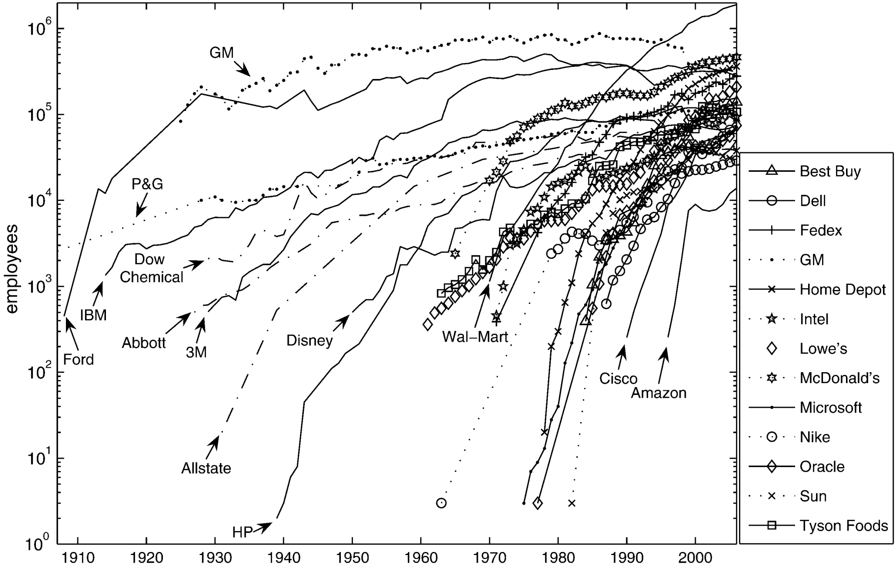


FIGURE 1

Bibliography I

Axtell, Robert L., "Zipf Distribution of U.S. Firm Sizes," *Science*, September 2001, *293* (5536), 1818–1820.

di Giovanni, Julian and Andrei a. Levchenko, "Firm entry, trade, and welfare in Zipf's world," *Journal of International Economics*, March 2013, *89* (2), 283–296.

Feenberg, Daniel R. and James M. Poterba, "Income inequality and the incomes of very high-income taxpayers: evidence from tax returns," *Tax Policy and the Economy, Volume 7*, 1993, *7* (January), 145–177.

Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *The Quarterly Journal of Economics*, 1999, *114* (3), 739–767.

Luttmer, Erzo G. J., "Models of Growth and Firm Heterogeneity," *Annual Review of Economics*, 2010, *2* (1), 547–576.

Luttmer, Erzo G.J., "On the Mechanics of Firm Growth," *Review of Economic Studies*, 2011, *78* (February), 1042–1068.

Pareto, Vilfredo, "Cours d'Economie Politique, vol. 1," 1896.