The Forbes 400, the Pareto Wealth Distribution and Efficient Markets

Oren S. Klass*, Ofer Biham*, Moshe Levy**, Ofer Malcai*, and Sorin Solomon*

May 2004

Abstract

Statistical regularities at the top end of the wealth distribution are examined using the Forbes 400 lists. We find that wealth is distributed according to a Pareto (power-law) distribution. This result is explained by a simple model based on stochastic returns and market efficiency.

Keywords: *wealth distribution, Pareto, market efficiency.*
JEL Classification: D31, 016.

*Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel
**Corresponding author. School of Business Administration, The Hebrew University, Jerusalem 91905, Israel . Tel: 972 2 588 3219; Fax: 972 2 5881341; mslm@mscc.huji.ac.il.
We thank Chani Charush for much help in processing the Forbes 400 data. This work was supported by a grant from the Israel Science Foundation and by the Zagagi Fund.

Introduction

Every year, in September, Forbes magazine publishes the list of 400 richest people in the United States (see www.forbes.com/lists). The list includes the net worth of each individual as well as background information about the businesses that have lead to this prosperity. The list is amazingly diverse including individuals involved in all sectors of the economy, such as computer software (Bill Gates, Paul Allen and Larry Ellison), financial investments (Warren Buffet), retailing (the Walton family), computer hardware (Michael Dell) as well as media, entertainment, communication, real estate and many other sectors.

Although the people included in the Forbes 400 list made their fortunes in various different ways, the distribution of their wealth exhibits a striking statistical regularity. This regularity describes not only the wealth distribution at the very top of the wealth range, but also provides insight about the distribution of wealth at wealth levels many orders of magnitude below this range. Moreover, the empirical findings shed light on the nature of the wealth accumulation process, and on its relation to the central issues of market efficiency and market fluctuations.

Empirical Findings

The wealth distribution is typically studied by employing data about the number of individuals within each wealth range (e.g. the number of individuals with wealth between $100,000 and $150,000). Here we employ a different approach. We use the Forbes 400 lists to study the distribution of wealth at the top wealth range. This approach has the advantage of employing very specific data on the wealth of each of the individuals on the list, and avoiding the aggregation problems of the standard approach. The disadvantage, of course, is that this approach is limited to the very top of the wealth distribution. However, as we argue below, the results obtained with this approach shed light on a much wider range of the wealth scale and on the capital investment process in general.

A Pareto (power-law) distribution of wealth implies the following relationship between the rank of an individual in the wealth hierarchy and her wealth:

$$w_r = Ar^{-\beta}, \tag{1}$$

where r is the rank (r=1 being the wealthiest person, etc.), w_r is the wealth of the individual with rank r, and A and the exponent β are constants (for a derivation this relationship, see for example, Takayasu [1990], Levy and Solomon [1997], and Levy [2004]). This implies that when wealth is plotted against rank on a log-log scale a linear relation is to be observed for a Pareto distribution.

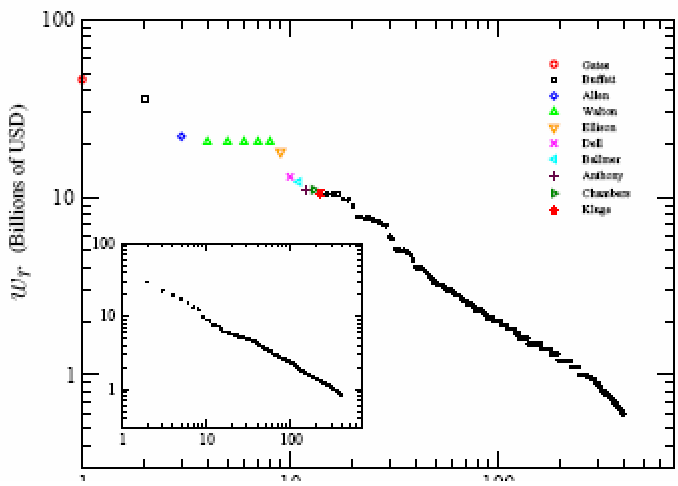

Figure 1. Zipf plot of the wealth w_r of the investors in the Forbes 400 list of 2003 vs. their ranks r. The power-law fit, with $\beta = 0.71$ and R^2=0.994 (see eq.(1)), was obtained in the central range $50 \le r \le 300$. The corresponding simulation results are shown in the inset.

Figure 1 depicts the wealth, w_r, of each of the 400 richest individual in the US in 2003, as a function of their rank, r, on a log-log scale. The data on this graph, known as the Zipf plot [1949], are fitted very closely by a straight line with $\beta = 0.71$ (R^2=0.994). This indicates an excellent agreement of the empirical data with a Pareto wealth distribution. Note, however, that due to the relatively small number of data points the distribution turns out to be rather noisy. To obtain more reliable results we perform a multi-year analysis by combining the data of the Forbes 400 lists from a sixteen-year period (1988-2003) to construct a much larger data set $w_r(t)$, where t is the year. The multi-year analysis requires using the normalized wealth variables $x_r(t) \equiv w_r(t)/\overline{w}(t)$, where $\overline{w}(t) = \sum_{r=1}^{400} w_r(t)/400$ is the average wealth of the 400 investors at year t. The time dependence of $\overline{w}(t)$ for the Forbes 400 investors between 1988 and 2003, shown in the inset of Fig. 2 (circles), reflects the stock prices after the 1987 crash, through the recession of the early 1990's, the bubble economy of the late 1990's and its aftermath, and a recovery in 2003.

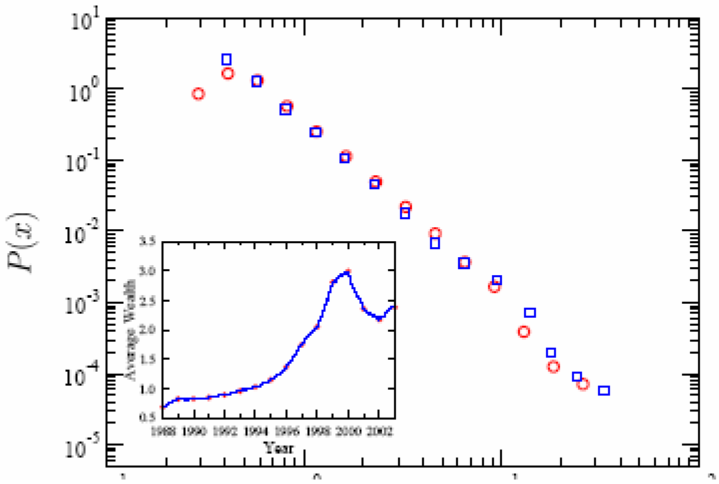

Figure 2. The distribution P(x) of the normalized wealth of the Forbes 400 investors (circles, $\alpha = 1.49$ R^2=0.992) and the corresponding simulations results (squares, $\alpha = 1.46$ R^2=0.996). The inset shows the average wealth $\overline{w}_t$ vs. time (circles), and simulation results with finer temporal resolution (rough solid line).

Figure 2 shows the probability density P(x) as a function of x (circles). In spite of the dramatic variations in the economic climate during the past sixteen years, the wealth distribution is in excellent agreement with the Pareto power-law distribution given by

$$P(x) = k\, x^{-(1+\alpha)} \tag{2}$$

with an exponent of $\alpha = 1.49$. The R^2 of this fit is 0.992. Note that in general the connection between the Pareto exponent α and the Zipf exponent β is $\alpha = 1/\beta$ (see Takayasu [1990]). Thus, the value of $\alpha = 1.49$ obtained with the multi-year analysis and the probability density approach is consistent with the value of $\beta = 0.71$ obtained as the Zipf exponent for 2003 (1/0.71 = 1.41).

This striking regularity of the wealth distribution confirms the hypothesis made over a century ago by the famous economist Vilfredo Pareto [1897]. During the century since Pareto's work, empirical evidence has been accumulated in support of his hypothesis (see, for example, Steindl [1965], Atkinson and Harrison [1978], and Persky [1992]). This is a remarkable result because power-law distributions exhibit the special property that they have no characteristic scale. It may indicate that the same dynamical rules of gains/losses apply across the entire economy independently of the particular sector or the wealth and sophistication of different investors (Anderson [1997]). Thus, the Forbes 400 data may provide useful information not only about the richest individuals, but also about the wealth of people in wealth percentiles far away from the top 400. These findings raise the broader, puzzling question about the origins of the nature of this wealth distribution. While the physical properties of humans (such as height) as well as their mental and social abilities follow Gaussian distributions that tend to be rather narrow, their wealths are very widely distributed and span over seven orders of magnitude (which, in terms of other human properties such as IQ or height, would correspond to observing an individual with an IQ of 10^9, or an individual who is 10,000 miles tall). What is the underlying reason for this power-law distribution of wealth? Below we argue that the Pareto wealth distribution is a consequence of the fundamental nature of the capital market.

The Model

The fundamental property of financial markets that enables some people to get so rich is the multiplicative nature of capital investments. To exemplify this, consider an investor who has a capital of $1,000. What does it take to make a million dollars from this initial investment? In an *additive* investment process that yields a fixed amount of, say, $1,000 per year, it would take 999 years. However, a *multiplicative* process that doubles the investment each year would require only 10 years. Indeed, analysis of stochastic models with multiplicative dynamics has provided much insight about the origin of the power-law distributions in economic systems[1].

Here we show that the wealth distribution obtained from the Forbes data can be reproduced by a very simple stochastic-multiplicative model based on the following three assumptions:

(I) Returns are stochastic;

(II) Markets are efficient, namely no investor can consistently "beat the market" reaping abnormal returns. In our model this feature is incorporated by drawing the random returns for all investors from the same probability distribution;

(III) There is a minimal amount of wealth, w_{min}, which is required in order to participate in the capital investment process. This can be thought of as a minimal wealth level needed for basic existence. We assume that w_{min} is equal to some fraction c of the average wealth $\overline{w}$.

[1] See Champernowne [1953], Wold and Whittle [1957], Simon [1955], Simon and Bonini [1958], Kesten [1973], Levy and Solomon [1996], Sornette and Cont [1997], Takayasu, Sato, and Takayasu [1997], Marsili, Maslov, and Zhang [1998], Malcai, Biham, and Solomon [1999], Bouchaud and Mézard [2000], and Levy and Levy [2003].

Our model consists of N investors whose wealths at time t are given by $w_n(t)$, n=1, …,N. The wealths are updated asynchronously such that the average time between successive updates of each of the w_n's is Δt . At each update, the randomly chosen wealth w_n is multiplied by a stochastic factor λ drawn from a given distribution $p(\lambda)$. As a result

$$w_n \to \lambda w_n \tag{3}$$

while all the other w_m's $(m \neq n)$ remain unchanged. The threshold wealth $w_{min}(t)$ required for entering and staying in the market at time t is given by $w_{min}(t)=c\,\overline{w}(t)$, where c is a parameter and $\overline{w}(t)$ is the average wealth at time t. If $w_n(t)$ is reduced below $w_{min}(t)$ (by the realization of a low return) investor n is dropped from the list and a "new" investor then enters, taking over the index n, with an initial wealth $w_n(t)$ = $w_{min}(t)$ (the results are robust to the relaxation of this assumption regarding the initial wealth of the new investor). Note that as it is impossible for all investors to have wealth exceeding the average wealth, we have c<1.

Numerical Simulations

We have performed computer simulations of the model for N=400. The value of the parameter c was obtained directly from the Forbes data, as the ratio $w_{min} / \overline{w}$, averaged over the 16 years, where w_{min} is the wealth of the least wealthy investor on the list. We find a value of c=0.337. Starting from a homogeneous distribution of the wealth, the wealth distribution spontaneously evolved towards a power-law distribution [shown in Fig. 2 (squares)], with an exponent $\alpha = 1.46$ (R^2=0.996), which is in excellent agreement with the empirical data ($\alpha = 1.49$, circles in Fig. 2).

The value of the exponent α depends on the value of the lower wealth bound. Indeed, empirical studies show that the value of α changes across different countries, and is typically in the range $1<\alpha<2$ (see, for example, Steindl [1965], Atkinson and Harrison [1978], and Levy [2004]). Theoretical analysis shows that for N not too small, the exponent α is determined by the parameter c, according to:

$$\alpha = \frac{1}{1-c} \tag{4}$$

(see Levy and Solomon [1996], and Malcai, Biham, and Solomon [1999]). This result provides a strong connection between the lower-cutoff, and the exponent α that affects the wealth distribution of the entire population, including those at the top. The dependence of α on c given by Eq. (4) indeed confirms that $\alpha>1$ for c<1, in agreement with the empirically observed values of α. The relationship (4) holds almost precisely for the Forbes 400 lists (and for the simulation)- the value of c=0.337 corresponds to a value of $\alpha = 1/(1-0.337) = 1.508$, which is very close to $\alpha = 1.49$ which is measured directly.

The model predicts that the power-law distribution of wealth extends well beyond the top 400 investors. To examine this feature we performed simulations with N=10,000 investors and the same value of c. We obtained a power-law distribution of wealth with the same value of α (α=1.46 R^2=0.998).

Discussion

As mentioned above, the power-law distribution of wealth is insensitive to the properties of the return distribution $p(\lambda)$. This may explain the robustness of the power-law behavior in the Forbes data over the sixteen-year period. However, a crucial assumption in the model is that the same distribution $p(\lambda)$ is used for all the investors. It was found that simulations in which this assumption is violated do not give rise to a homogeneous power-law distribution of the wealth (Levy [2003]). Thus, according to the model, the power-law wealth distribution in the Forbes data indicates that the *ex-ante* return distribution is similar for all the investors. This conclusion may look like a paradox. While it is clear that to be in the top of the list one has to do 'something right' in a big and consistent way, the statistical regularities of the wealth distribution can be captured by a model that exhibits completely random dynamics. The resolution of this paradox is that although the *ex-ante* distribution $p(\lambda)$ is similar for all the investors, the realized returns that each investor draws are different. The multiplicative dynamics greatly magnifies the differences between more successful and less successful investors, and builds up the power-law distribution of wealth. The assumption that $p(\lambda)$ is the same for all the investors is an implementation of the efficient market hypothesis, which states that no investor can consistently obtain a return distribution better (adjusted for risk) than the return distributions of the other investors. Thus, our model provides a connection between the efficient market hypothesis and the Pareto distribution of wealth.

The model also illuminates the relationship between the distribution of wealth and the distribution of stock market returns. It was observed long ago that fluctuations in financial markets exhibit non-Gaussian "fat tailed" distributions. Already in the 1960's Mandelbrot studied the fluctuations in cotton prices and discovered that they can be expressed by a Lévy-stable distribution (Mandelbrot [1963]). Further evidence supporting the Lévy-stable distribution was provided by Fama [1965], Roll [1968], Teichmoeller [1971], and Officer [1972]. Recent empirical studies of the short-term fluctuations of the S&P500 index provided apparently conflicting results. The central peak of the distribution of returns has been found to exhibit scaling behavior which is consistent with the Lévy distribution (see Mantegna and Stanley [1995]). However, the tails of the distribution of returns have been found to decay as a power-law with an exponent of roughly 3, namely beyond the Lévy-stable range, (see Gopikrishnan et al. [1999]).[2] It is interesting that our very simple model reproduces both of these empirical findings, as explained below.

We define the market log-return over horizon τ as :

$$G(\tau) = \ln(\overline{w}(t+\tau)) - \ln(\overline{w}(t)), \tag{5}$$

where $\overline{w}$ is the average wealth. The return distribution P(G) depends, of course, on the horizon τ: the longer the horizon the more spread out the distribution, and the lower its peak. For a symmetric Lévy distribution the height of the peak of the distribution , P(0), depends on τ as follows:

$$P_\tau(0) = D\tau^{-1/\alpha_L}, \tag{6}$$

where D is a constant and α_L is the Lévy exponent (Mantegna and Stanley [1995]).

Figure 3 shows the relationship we find in the simulation of our model between $P_\tau(0)$ and τ on a log-log scale. The relationship is very close to linear over three orders of magnitude, with a slope –0.67 and R^2=0.99. This is in excellent agreement with the Lévy distribution in the central range.

[2] Two alternative modeling approaches to explain the shape of the return distribution are the mixture of Gaussians approach developed by Hsu, Miller and Wichern [1974], Pedrosa and Roll [1998], and the ARCH/GARCH approach of Engle [1982], Bollerslev, Chou and Kroner, [1990].

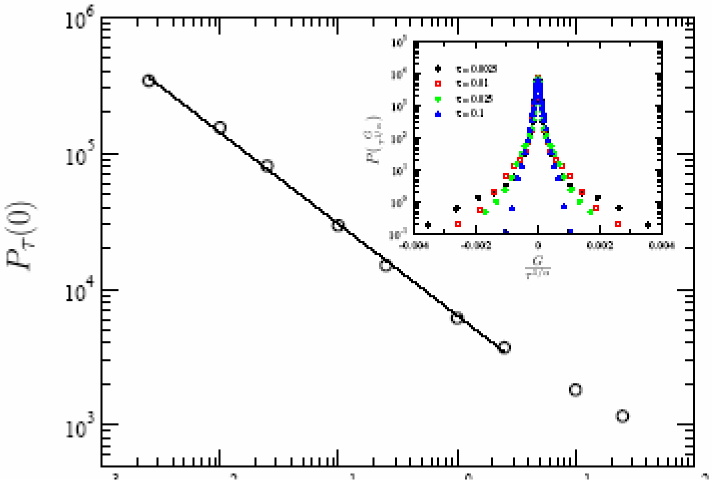

Figure 3. The height of the central peak $P_{\tau}(0)$ vs. τ on a log-log scale. The inset shows a scaling plot of $P_{\tau}(G/\tau^{1/\alpha})$ for $\tau = 1,2,4,8$. The central peaks coincide confirming the scaling in the main Figure, but the tails deviate from the Lévy-stable form.

Levy and Solomon [1997] and Levy [2003] argue that the source of the Lévy distribution of stock returns is rooted in the Pareto distribution of wealth. Their argument is based on the Doblin-Gnedenko Theorem (see Feller [1971]), and it leads to the prediction that the Lévy exponent, α_L, should be equal to the Pareto exponent α. Our simulation offers a test of this prediction. The slope of –0.67 in Fig. 3 corresponds to a value of $\alpha_L = 1.49$ (see eq.(6)). This is very close to the Pareto wealth exponent in the simulation (α=1.46). Moreover, the value α=1.49 which is empirically obtained from the Forbes data is similar to the value $\alpha_L = 1.40$ empirically measured by Mantegna and Stanley [1995] for the S&P index.

Another test for the agreement of the return distribution with the Lévy distribution is suggested by the scaling properties of the Lévy distribution. A Levy distribution $P_\tau(G)$ scales as $G/\tau^{1/\alpha}$. In other words, with this scaling the distributions $P_\tau(G)$ with different horizons τ all fall one on top of the other (see Mantegna and Stanley [1995]). The inset of Figure 3 shows this scaling. In the central part of the distribution we find an excellent agreement with the predictions of the Lévy distribution – in this range the distributions all fall on top of each other. However, at the far tails of the distributions this scaling breaks down. These results are completely consistent with the empirical findings regarding the return distributions of market indices.

In conclusion, the Forbes 400 lists reveal a striking agreement with the Pareto power-law distribution of wealth. A simple model that incorporates stochastic returns and a lower-bound on the wealth of investors, provides an interesting connection between market efficiency, the Pareto distribution of wealth and the distribution of stock market returns.

References

P. W. Anderson, Some Thoughts about Distributions in Economics, in *The Economy as an Evolving Complex System II*, Eds. W.B. Arthur, S.N. Durlauf, and D.A. Lane, SFI Studies in the Sciences of Complexity , Addison Wesley Longman, (1997).

A.B. Atkinson and A.J. Harrison, “Distribution of Total Wealth in Britain,” Cambridge University Press, Cambridge, (1978).

O. Biham, ZF Huang, O. Malcai, S. Solomon. “Long-time fluctuations in a dynamical model of stock market indices”. *Phys Rev E* 64, 2, (2001), 6101-5.

Bollerslev, Chou, and Kroner, “ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence.” *Journal of Econometrics*, 52, (1992), 5-59.

J.P. Bouchaud, Mézard M. “Wealth condensation in a simple model of economy” *Physica A* 282, (2000), 536-545.

D.G. Champernowne, A model of income distribution, *Econ. J.*, 63, (1953), 318-351.

D.G. Champernowne, *The Distribution of Income*, Cambridge University Press, Cambridge, (1973).

R. F. Engle, “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation”, *Econometrica*, 50, (1982), 987-1008.

E. F. Fama, “The Behavior of Stock Prices”, *Journal of Business*, 38, 1, (1965), 34-105.

W. Feller, *An Introduction to Probability Theory and its Applications*, Vol 2., 2nd edn, Wiley, New York, (1971).

P. Gopikrishnan, V. Plerou, LAN Amaral, M. Meyer, HE Stanley, “Scaling of the distribution of fluctuations of financial market indices”, *Phys. Rev. E* 60 (1999) 5305.

D. A. Hsu, Miller, R. B., and Wichern, D. W., "On the Stable Paretian Behavior of Stock-Market Prices", *Journal of the American Statistical Association*, 68, (1974), 34.

H. Kesten, “Random difference equations and renewal theory for products of random matrices”, *Acta Math*. 131 (1973), 207–248.

M. Levy, “Are Rich People Smarter?”, *Journal of Economic Theory*, 110, 1, (2003), 42-64.

M. Levy, “Market Efficiency, the Pareto Wealth Distribution, and the Lévy Distribution of Stock Returns”, in *The Economy as an Evolving Complex System* III, S. Durlauf and L. Blume (Eds.), Oxford University Press, forthcoming.

M. Levy and H. Levy, “Investment Talent and the Pareto Wealth Distribution: Theoretical and experimental Analysis”, *Review of Economics and Statistics*, 85, 3, (2003), 709-725.

M. Levy, and S. Solomon, Power laws are logarithmic Boltzmann laws, *Intern. J. Modern Physics C* , 7, (1996), 65-72.

M. Levy and S. Solomon, New evidence for the power-law distribution of wealth, *Physica A*, 242, (1997), 90-94.

O. Malcai, O. Biham, S. Solomon, “Power-law distributions and Levy-stable intermittent fluctuations in stochastic systems of many autocatalytic elements”, *Phys. Rev. E* 60 (1999) 1299.

B. Mandelbrot, “The Variation of Certain Speculative Prices.”, *Journal of Business*, 36, (1963), 4.

R.N. Mantegna, and Stanley, H. E., “Scaling Behavior in the Dynamics of an Economic Index.” *Nature*, 376, (1995), 46-49.

M. Marsili, S. Maslov and YC Zhang, “Dynamical optimization theory of a diversified portfolio”, *Physica A* 253, (1998), 403.

R. Officer, “The Distribution of Stock Returns”, *Journal of the American Statistical Association*, 67, (1972), 340.

V. Pareto, *Cours d'Economique Politique*, Vol 2, (1897).

V. Pareto, *Manual of Political Economy*, Augustus M. Kelley, New York, (1971), translated from the original 1906 *Manuale d'Ecoonomia Politica*.

M. Pedrosa, and R. Roll, “Systematic Risk in Corporate Bond Credit Spreads”, UCLA Working Paper 7-98 (1998).

J. Persky, Retrospectives: Pareto's law, *J. Econ. Perspect.*, 6, (1992), 181-192.

H. Simon, On a class of skew distributions, *Biometrica*, 52, (1955), 425-440. (Reprinted in Simon, *Models of Man*, 1957).

H. Simon, and CP Bonini “The Size Distribution of Business Frims,” *American Economic Review*, 58(4), (1958), 607-617.

D.J. Slottje, *The Structure of Earnings and the Measurement of Income Inequality in the U.S.*, Elsevier Science Publishers, New York, (1989).

D. Sornette and R. Cont, “Convergent multiplicative processes repelled from zero: power laws and truncated power laws”, *J. Phys. I* France 7, (1997), 431-444.

H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford) (1971).

J. Steindl, *Random Processes and the Growth of Firms - A Study of the Pareto Law*, Charles Griffin & Company, London, (1965).

H. Takayasu, *Fractals in the Physical Sciences*, Wiley, New York, (1990).

H. Takayasu, Sato AH and Takayasu M., "Stable infinite fluctuations in randomly amplified Langevin systems", *Phys. Rev. Lett.* 79, (1997) 966.

J. Teichmoeller, "A Note on the Distribution of Stock Price Changes", *Journal of the American Statistical Association*, (1971), 66.

H. Wold and P. Whittle, A model explaining the Pareto distribution of wealth, *Econometrica*, 25, (1957), 591-95.

G. K. Zipf, *Human Behavior and the Principle of least Effort* (Addison-Wesley Press, Cambridge, MA) (1949).