Retrospectives
Pareto’s Law

Joseph Persky

This feature addresses the history of economic words and ideas. At a minimum, the hope is to deepen the workaday dialogue of economists. At best, such discussion may cast new light on ongoing questions. If you have comments or criticisms about this column or suggestions for future topics or authors, please write to Joseph Persky, c/o Journal of Economic Perspectives, Department of Economics (M/C 144), The University of Illinois at Chicago, Box 4348, Chicago, Illinois 60680.

Introduction

The nature, causes and consequences of economic inequality have preoccupied economists at least since Adam Smith. Smith himself thought the distribution of income was sensitive to the institutions of a society, but largely independent of the progress or decline of its economy—all incomes moved up or down more or less proportionally (Smith, 1776 [1937], p. 63, 141). From a somewhat different perspective, John Stuart Mill feared that progress would be likely to increase the share of the middle classes, but might well leave the poorest portion of society no better off (Mill, 1849 [1929], p. 699). Of course, the classical economists carried on their discussions in blissful ignorance of the facts. All of this changed dramatically toward the end of the nineteenth century.

The successes of statistics in studying a wide range of phenomena in biology and psychology laid the groundwork for an inductive approach to
economics in general and the income distribution in particular. The long-standing debates over the distribution of income provided a natural starting point for serious statistical analysis. Those debates had sharpened with the rise of significant socialist and trade union movements in western Europe and the United States. Passions flared on all sides. An inductive statistical approach promised scientific answers to the central questions of the day. And newly enacted income and property tax laws in several countries provided rich data sources.

Most 19th century economists had paid lip service to the notion of verifying their theories with empirical research. A few had even attempted to collect relevant data. But it was Vilfredo Pareto with his researches on the distribution of income who first showed what a seriously inductive economics might look like. Pareto considered data for England, a number of Italian cities, several German states, Paris and Peru. Plotting the cumulative distributions of income for these countries on double logarithmic paper Pareto claimed that in each case the result was a straight line with about the same slope. Thus, he argued

1Gini (1936, p. 73) gives credit to Otto Ammon in 1895 for making a connection between the shape of the distribution of individual incomes and the distribution of individual abilities studied by Francis Galton. Moore (1908) argued for an inductive, statistical complement to theoretical economics.
that these distributions could be well summarized by his now famous curve:

\[ \log N = A - \alpha (\log x), \]

where \( N \) is the number of households with incomes greater than \( x \), \( A \) is a parameter and \( \alpha \) is (the absolute value of) the slope (Pareto, 1896 [1965], pp. 1–7; 1897 [1964]). Pareto observed that the values of \( \alpha \) clustered around 1.5.\(^2\) Thus in Figure 1, which is based on Pareto's original illustration, the slope of the curve for England is parallel to that for "Italian Towns." While Pareto used no quantitative measures of goodness of fit, visual inspection suggested that these linear equations worked quite well.\(^3\) The constancy of \( \alpha \) and the quality of the fit encouraged Pareto to assert a law of income distribution (1896 [1964], p. 7).

The frequency distribution corresponding to Pareto's law is so skewed as to be actually one-sided. Pareto called this distribution the "social pyramid" and drew it as in Figure 2. The frequency distribution begins at a minimum income (\( h \) in Figure 2), declining monotonically throughout.\(^4\) From the start, Pareto emphasized the heavily asymmetric character of his distribution and hence its fundamental difference from a normal curve.

\(^2\)There is a correspondence between Pareto curves and Lorenz curves as pointed out quite early by Dalton (1920b). The curve for \( \alpha = 1.5 \) implies that the top 10 percent of all income recipients receive 46 percent of all income and the top 20 percent of recipients receive 58 percent of all income. An \( \alpha \) of 1.5 corresponds to a Gini coefficient of 0.5.

\(^3\)Based on the data Pareto used, I calculate an \( R^2 \) of .997 and .998 for England 1879–80 and the Italian Towns respectively.

\(^4\)There is a necessary mathematical relation between \( h \) and the constant \( A \) in Pareto's curve: \( h = (A/P)^{1/\alpha} \), where \( P \) is the total population.
I can find no evidence on exactly how Pareto stumbled across this formulation. However Wesley Mitchell et al. (1921), in discussing Pareto’s law, suggested that double logarithmic paper was commonly used by engineers. Pareto trained as an engineer and for several years practiced that profession. Perhaps this background influenced his choice of graph paper.

Pareto’s Theory of Income Distribution

Pareto was overjoyed with his new law, and set out to explain it. According to Pareto, the possible sources of income inequality included chance, social institutions and human nature. Hypothesizing a simple binomial process of chance accretions to individual incomes, he argued that the resulting frequency distribution would hardly square with his new law. Chance would produce neither the right skew nor the right curvature (Pareto, 1896–1897 [1964], pp. 316–17).

Using a mass of empirical evidence, Pareto asserted that the observed similarity across time and space in income distributions ruled out social institutions as important contributors. Citing data from such diverse societies as ancient Peru, the Cherokee Indians and Prussia, Pareto claimed that they all had quite similar distributions of wealth and income. Given the major institutional differences between these societies, Pareto argued that his law would remain even after profound changes in economic organization. Hence he concluded that the distribution of income depended primarily on human nature (Pareto, 1896–97 [1964], p. 363).

Pareto remained somewhat vague as to exactly what in human nature determined the law. At places he argued that it corresponded to an underlying distribution of human abilities, truncated by the necessities of survival. Those with abilities considerably lower than the mean must either die or be supported by others. Hence, the observed distribution had a truncated lower tail (Pareto, 1906 [1971], p. 284).

In a quite different vein, Pareto put forth a highly original argument that focussed not on the distribution of abilities but rather on the will of the most successful to command the resources of society. Different environments and economic institutions could powerfully affect the level of material prosperity reached in a given society. And these differences would undoubtedly imply that different societies would reward different characteristics. Pareto suggested that those who won in the struggle for society’s rewards would have a tendency to accept a certain share (as measured by his distribution) rather than fight without end.

Human nature here determined the appetite of the victors in the social struggle, whatever their abilities, to appropriate the product of society, whatever its character (Pareto, 1896–97 [1964], pp. 371–72). Presumably, this appetite was hardwired into human beings. Social engineering of one sort of
another could never alter this fundamental fact. “This curve gives an equilib-
rium position, and if one diverts society from this position, automatic forces
develop which lead it back there” p. 360). This argument, while highly specula-
tive and a bit ad hoc, at least made clear the forces that prevented a legislatively
redistribution of income. Whatever the institutional structure, those best
adapted to it would eventually demand a well-defined share of the total income.

Pareto did acknowledge some drift in the data toward lower values of $\alpha$
which he interpreted as a reduction in inequality.\footnote{Pareto's definition of income inequality has led to a continuing debate and a good deal of confusion about the relation of $\alpha$ and inequality. A printer's error in the Cours substituted an "increase in inequality" for a "reduction in inequality" in a crucial passage, although a footnote made clear the original meaning. (See discussion by Pareto, 1906 [1971], p. 290.) Pareto correctly concluded that in terms of his (intended) definition of inequality lower values of $\alpha$ were associated with a reduction of inequality. Dalton (1920b) and Gini (1936) demonstrated that in terms of Lorenz curves and Gini coefficients, lower values of $\alpha$ are associated with increases of inequality. John Chipman (1974) came back and showed that for additively separable social welfare functions, lower $\alpha$ values, holding $A$ constant, imply an increase in welfare which presumably supported Pareto's original claim that reducing $\alpha$ is a good thing (Cowell, 1977, p. 97). If that's all clear, consider the fate of M. Allais, who according to Kirman (1987), starts from Pareto's original misprint, and compounds it with the assertion that Pareto thought a reduction of $\alpha$ raises inequality.} He proved a theorem: given
that the income distribution must follow a Pareto curve, then an increase in the
mean income holding $\alpha$ constant implies an increase in the minimum income
($h$). Alternatively, holding minimum income constant, an increase in the mean
implies a reduction in $\alpha$ and hence in inequality.\footnote{This theorem is based on Pareto's definition of inequality and thus has been caught up in the controversies discussed in the previous note. Dalton (1920b, p. 359) points out that given the more common definition of the Gini coefficient, the result would be opposite; that is, a growth in mean income, holding minimum income constant will increase inequality. More generally, Dalton finds the notion that growth in average incomes would have a predictable effect on inequality naive, given the range of institutional forms.} Either way, the interests of
the lower classes were intrinsically linked to the overall level of production.
Seeing little reason to expect large changes in $\alpha$, Pareto encouraged society to
focus on raising mean incomes, rather than pursuing short-sighted efforts to
change the income distribution. That distribution rested on fundamental struc-
tural realities of human nature.\footnote{The term “structural” here is used advisedly. Like Walras and Marx, Pareto discounted the multiplicity of social institutions and forms asserting that all of these ultimately reflected deeper necessities of human organization. Modern structuralism traces its roots to the linguistics of Ferdinand Saussure, who, according to no less an authority than Jean Piaget (1970, p. 77) drew his inspiration in part from Walras and Pareto.}

**Pareto's Law and Pareto Optimality**

Pareto's interpretation of his empirical law anticipated his broader sociol-
ogy. Since the law implied a preset distribution for society, the role of elites or
aristocracies in stimulating economic growth became essential. Pareto envisioned history as an alternating cycle of two elites: the foxes and the lions. A soft but guileful elite (the foxes) might present itself as sympathetic to the masses but would lay claim to just as large a share as a more forceful elite (the lions). The greed and weakness of the foxes inevitably produced slow growth and social chaos. Society then welcomed the lions back to power. Ultimately, Pareto's theory of elites provided an intellectual veneer for fascist positions; in fact, Vincenzo Fani called Pareto the Karl Marx of Fascism. It is a sorry tale of the early 20th century how the liberal, pacifist Pareto ended as window dressing for Mussolini's fascists. But more relevant to economists of today are the connections between Pareto's Law and Pareto optimality.

Welfare theorists have criticized Pareto for not emphasizing that a distinct optimum is associated with every possible distribution of income. Pareto's expression for the economic optimum, maximum ophemality, seemed to suggest an optimum optimorum, not one of an infinite number of efficient allocations. This criticism is not completely fair because Pareto, in his discussion of the collectivist state, acknowledged explicitly that to achieve socialist objectives the question of distribution must be solved before the question of resource allocation (Pareto, 1906 [1971], p. 267). However, he was generally unwilling to devote serious attention to this matter.

In part this reluctance reflected his assertion, shared with many first generation neoclassicists, that rigorous interpersonal comparisons of utility were impossible. Given Pareto's sensitivity to the egalitarian demands of socialism, he must have been particularly wary of the radical implications of combining diminishing marginal utility with interpersonal comparisons.

Still, I think there is more to Pareto's reluctance to consider the range of potentially efficient allocations. In general, Pareto's discussions of efficiency have been considered separately from his income distribution law. However, since Pareto was confident that the distribution of most societies looked quite similar, there was little need to discuss alternative efficient points. Socialist leaders would ultimately generate about the same distribution as the capitalists, although they would probably reduce the average income. As a result, the optimum achieved through a freely competitive market was just about the best that could be hoped for. I suspect that many economists today hold a similar view.

---

8Quoted in Bucolo (1980, pp. 285–86). Or as Adrian Lyttleton put it, "If Pareto's theory had not existed, fascism would have had to invent it." Quoted in Bellamy (1987, p. 33).

9Pareto did become disillusioned with liberalism in his later years and longed for a sort of moral rearmament. He welcomed Mussolini's rise to power but died soon after in 1923 (Bellamy, 1987, pp. 30–33). Most scholars now assert that Pareto would not have approved of Mussolini's economic policies. Bucolo (1980) provides a treatment of these matters sympathetic to Pareto.

10Pareto used his own term, "ophemality," to refer to the satisfaction people get from economic goods. He insisted on saving the word "utility" for a broader notion of satisfaction derived not only from goods but from social and political institutions as well. "Utility" in Pareto also becomes confused with the modern concept of externalities. See Pareto (1916 [1963], pp. 1459–72).
The Attack on Pareto

Not surprisingly, Pareto’s sweeping assertion of his law drew considerable attention, as well as criticism that continues even today. Pareto participated only in the first round of debate and then seems to have lost interest. He engaged in a minor altercation with Edgeworth and another with the anarchist Sorel (Busino introduction to Pareto, 1896 [1964]). One of the first comments on Pareto’s Law was that of Lorenz (1905), who used his criticism of Pareto—“logarithmic curves are more or less treacherous” (p. 217)—as a springboard to developing his own famous curve.

Pareto had presented his law as derived from simple empiricism. His supporters suggested that his procedure was similar to those of Boyle, Gay-Lussac, and Avogadro (Moore, 1908, p. 28). Thus, it is not surprising that much of the criticism of Pareto focussed on the empirical base of his law. In 1912 A.C. Pigou mounted a serious attack along these lines. According to Pigou, Pareto’s $\alpha$ varied considerably in both the data Pareto had used and in subsequent data generated by Bowley. Hence there was no single distribution that could be called a law. Corrado Gini (1936) offered a similar critique.

Early econometricians also had a go at Pareto’s empirical argument. Warren Persons (1909) applied least-squares to Prussian data similar to Pareto’s original. Although the Pareto curve fit well in its logarithmic cumulative form, Persons argued that such a curve hid a number of sins. From the estimated curve he calculated the expected frequencies for nine income classes. Comparing these to actual frequencies he found considerable errors ranging from −37 percent to +35 percent. He observed that “an error in a logarithm gives a much larger error in the natural number” (p. 425), and concluded that the accuracy of Pareto’s law was “apparent and not real” (p. 427). The question of how well the law fits the data became a perennial one. Others took up Person’s argument, including the staff of National Bureau (Mitchell et al., 1921) in their first monograph, where Frederick Macaulay (p. 124) “put aside ‘Pareto’s Law’ as having at the present time little more than historical interest.”

Pareto’s critics also challenged his claim that social institutions had negligible impact on income distributions. For example, Pigou (1912) argued that the legal structure of inheritance laws had demonstrable effects on inequality. He went on to assert that a number of other institutional factors might also influence income distributions—most importantly access to education and training.12

11Boyle’s law is that the volume of a body of gas is inversely proportional to the pressure, given a constant temperature. Gay-Lussac’s law is that the volumes of two or more gasses that combine to make a new gas are in the proportion of small whole numbers to each other and to the volume of the resulting gas. Avogadro’s number states that the number of molecules contained in one mole of a substance is $6.02252 \times 10^{23}$, where a mole is the quantity of a substance where the weight in grams is numerically equal to its molecular weight.

12Hugh Dalton (of the Pigou-Dalton condition) picked up the attack, cataloging many ways political and social reforms could change the distribution of incomes. Dalton (1920a, p. 352) listed “the
Finally Pigou and Dalton attempted to make the inductive character of Pareto's law a liability, rather than a virtue. Since Pareto had generalized his law from a narrow range of real historical institutions, its application must properly be limited to that range alone. “...[E]ven if the statistical basis of the “law” were much securer than it is, the law would but rarely enable us to assert that any contemplated change must leave the form of income distribution unaltered,” Pigou (1920, p. 655) wrote. He went on, “As things are, in view of the weakness of its statistical basis, it can never enable us to do this.”

The Counterattack

Further attacks were mounted on Pareto’s Law in the 1930s, most notably that of Shirras (1935) in the Economic Journal. However, as D. H. MacGregor (1936) suggested the following year, “Economics has not so many inductive laws that it can afford to lose any.” MacGregor (p. 80) viewed Pareto’s Law as “a stage in investigation, like Boyle’s Law or Darwin’s Law.” All of these “although amended ... remain authoritative as first approximations not to be lightly gone back on.”

The last half of the 1930s saw a renaissance of Paretian thinking, as well as considerable interest in Pareto’s sociology (Bucolo, 1980, p. 288). Pareto’s disillusionment with liberalism played well in the midst of the Great Depression. A need for strong political leaders was voiced in many quarters. The rise of fascism in Europe might also have enhanced interest in Pareto’s work.

Perhaps the chief resource of the Paretian partisans was the fact that despite all the nitpicking, those double logarithmic curves still looked good. In a thoughtful piece, Norris Johnson (1937) reviewed the attacks on Pareto’s Law. Johnson avoided completely the most egregious Paretian claims and ideology. He gave up any effort (always a bit tentative even in Pareto’s work) to explain the lower tail of the income distribution. While acknowledging that fits were not always perfect and that the $\alpha$ parameter might vary some, he concluded with a strong positive endorsement (p. 26): “It cannot be denied that Pareto developed a fundamental yardstick. He found a useful simple description of the scheme of income distribution in the upper brackets and a most interesting limitation on the range of measured inequality in the distribution.”

Johnson’s approach was careful and cautious, but others remained eager to champion a more sweeping interpretation of Pareto’s Law. Chief among these were Harold T. Davis and Carl Snyder. Davis, one of the pioneers of econometrics, helped to found Econometrica, for which he served as Associate Editor for

judicious increase of civil rights, the improvement of education, the extension of the field of employment of women, the modification of the laws of inherited wealth, various forms of taxation and the regulation of the currency, in such a way as to secure a steady and gradual increase in the value of money.”

13Also see Dalton (1920a, p. 127).
26 years. He also worked on the staff of the Cowles Commission in its early days in Colorado Springs. Davis (1941, p. 571) strongly asserted that statistical techniques provided a way "to test the thesis that there is an economic basis to history, and to measure with some accuracy the actual significance of this factor in the march of historical events."

Davis embraced Pareto's Law as one of his prime examples of economic determinism. While recognizing that the Law was originally put forth in a "rigid and uncompromising form," he felt confident (p. 395) that "no one, however, has yet exhibited a stable social order, ancient or modern, which has not followed the Pareto pattern at least approximately." Starting with Pareto's data, working through those of the National Bureau and citing the studies of Johnson, Davis argued for the general validity of the law.

Davis also attempted a theoretical rationalization of Pareto's Law. He offered a "law of the distribution of special abilities" which asserted that the probability of an additional unit of ability was independent of the level of ability. Davis concluded that for high levels of ability the distribution would approximate the Pareto (1941, p. 427). He cited data on billiard playing, mathematics writing and chemical abstracts in support of his thesis. Davis also cited the work of C. H. Boissevain (1939) on the distribution of abilities that could be represented as the product of two or more factors each of which followed the binomial distribution. This is still a common explanation for the skew in abilities and the distribution of income, although Boissevain seldom receives credit for the notion.

And then Davis went whole hog. If Pareto's curve was a determinate social phenomenon, substantial digressions from the established pattern must lead to social disorganization. Davis speculated that both the French and Russian revolutions were "aggravated, if not actually caused, by an undue concentration of wealth and income." On the other hand, the slow recovery of the American South after the Civil War might have been the consequence of the levelling of the income distribution created by the abolition of slavery. Similarly, the lack of will shown by the French in their embarrassing loss to Nazi Germany was in part caused by the demoralization of the upper classes caused by too even a distribution of wealth (pp. 435–36).14

Davis was joined in his speculations by Carl Snyder, who presented his ideas to the 1937 Cowles Commission Research Conference and expanded on them in his 1940 manifesto, *Capitalism the Creator*. Brushing aside criticisms of Pareto's Law, Snyder endorsed the notion that Pareto's Curve represented a "general expression for practically every kind of highly specialized ability," an idea he apparently arrived at independently. According to Snyder, "Pareto's Curve is destined to take its place as one of the great generalizations of human knowledge," since it reflected the economic manifestations of our inability to create "by legislation or any fanciful law . . . a Newton, a Shakespeare or an

---

14I first became aware of Davis' writings from the charming notes in Cowell (1977). The spirit of my discussion above (although not the specific examples) draws heavily on Cowell.
Edison,” (1937, p. 61). Snyder then claimed that any effort to redistribute away from this natural distribution of ability resulted in “the ruthless destruction of civilization by these barbarian hordes.” Leveling incomes would gain little for the masses in the short run and destroy the talents that drove the system. Snyder failed to explain exactly why those talents were so frail, or why capital accumulation was such a fragile business. Nevertheless, society had to shun the “seductive altruism” that disguised the “fierce urge to destroy, to level up, to abase the talented, the capable and the rich: one of the strongest drives within those vast lower levels of our neolithic population” (1940, p. 418). Snyder out-Paretoed Pareto.

Several More Rounds

There’s more to the bibliography of Pareto’s law, much more. This literature has traveled in many directions. Zipf (1949) enunciated a law of “least effort” and interpreted the Pareto distribution of income as an equilibrium between the greed of superiors and the need to maintain loyalty among inferiors.15 Champernowne (1953) and Wold and Whittle (1957) designed Markov chains in stochastic processes that converged on Pareto distributions. Stiglitz (1969) provided a rationalization of Champernowne’s model in the context of neoclassical growth theory with random returns to capital. Mandlebrot (1960) explored how the Pareto-Levy family of distributions generalized the stability property of the normal curve. Samuelson (1972) and Chipman (1974) returned to the question of how changes in \( \alpha \) affect welfare.16

In one of the most Paretian of the modern contributions to this literature, Herbert Simon (1957) developed a simple model of the hierarchy of control for large corporations. Simon noted that most hierarchies defined a span of control for each executive within quite narrow limits and that a “rule of proportionality” held between the salaries of an executive and that executive’s immediate subordinates. Simon suggested a ratio of 1.25 to 2 adding that this rule’s “correctness as a norm is accepted more or less as a truism” (1957, p. 33). Although Simon did not expand on this model, the notion of a fixed norm for the ratio of salaries in a hierarchy recalls Pareto’s suggestion of a deeper social equilibrium lying behind his income curve.

Despite the analytical sophistication in the work of Simon and others, there remains in the modern discussion of Pareto’s law a tension quite similar to that found early in the century. Samuelson (1972, p. 408) concluded, “An alleged constancy of this coefficient [\( \alpha \)]—I write alleged because careful research shows that Pareto and his disciples overdid their claims for its empirical

15Zipf disputed the value of \( \alpha \), claiming it should be 2.0, not 1.5. Zipf gave relatively little credit to Pareto, whose sociology of equilibrium he used extensively. See Zipf (1949, Chapter 11).
16See Cowell (1977), Kirman (1987), and Steindl (1987) for discussions of these and additional references.
invariance—tells us nothing about the ability of society to alter its inequality of wealth or income . . ." On the other hand, Chipman (1974, p. 227) in a footnote to an otherwise highly analytic piece, took time to lambast an anonymous referee. According to Chipman, the referee assumed "that there is some self-appointed academic elite ("we") who have sufficient political power to be able to impose at will any income distribution that conforms to their preconceived ethical notions, even though such an income distribution may never have been observed in the past; and that these ethical notions are so self-evidently superior to those held by others that the empirical evidence as to the possibility of imposing such an income distribution can be completely ignored and even disdained."

For all the excesses of the Paretian camp followers, there remains the significant insight that the history of all hitherto existing society is a history of social hierarchies. There is the feel of structure behind income distributions. Almost all income distributions are continuous, unimodal and highly skewed. We have no examples of uniform distributions or egalitarian distributions or strikingly trimodal distributions. Something is going on here.

Those who are uneasy with this empirical observation, and who also expect that a solipsism lies embedded in the denial of interpersonal comparisons of utility, have searched for limitations on Pareto's claim. They continue to hope that institutions can be designed to promote both equity and efficiency. On occasion, that hope may lead to wishful thinking. Nevertheless, I suspect that an important measure of any society is the extent to which it lives up to its moral insights despite the imperatives of structure.

A postscript. In his fine discussion of the problems of the middle class in the 1980s, Frank Levy (1988) concludes that it was the slowdown in the national growth rate and not any drastic redistribution of income that created the anxieties of the "vanishing middle." Pareto would have approved.

— Special thanks again to Carl Shapiro, Joseph Stiglitz and Timothy Taylor whose highly useful and stimulating comments ran almost as long as this essay.

References


Snyder, Carl, Capitalism the Creator. New York: Macmillan, 1940.