Scherer, Frederic Michael

Working Paper

The size distribution of profits from innovation

ZEW Discussion Papers, No. 96-13

Provided in cooperation with:
Zentrum für Europäische Wirtschaftsforschung (ZEW)

Suggested citation: Scherer, Frederic Michael (1996) : The size distribution of profits from innovation, ZEW Discussion Papers, No. 96-13, http://hdl.handle.net/10419/29444

Discussion Paper

Discussion Paper No. 96-13

The Size Distribution of Profits from Innovation

Frederic M. Scherer

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Industrial Economics and International Management Series

Discussion Paper No. 96-13

The Size Distribution of Profits from Innovation

Frederic M. Scherer

The Size Distribution of Profits from Innovation*

by
F. M. Scherer

Harvard University

May 1996

Abstract
The research reported in this paper seeks to determine how skewed the distribution of profits from technological innovation is -- i.e., whether it conforms most closely to the Paretian, log normal, or some other distribution. The question is important, because high skewness makes it difficult to pursue risk-hedging portfolio strategies. This paper examines data from several sources -- the royalties from U.S. university patent portfolios, the quasi-rents from marketed pharmaceutical entities, and the returns from two large samples of high-technology venture startups. The evidence reveals a distribution closer to log normality than Paretian. Preliminary hypotheses about the underlying behavioral processes are advanced.

JEL Classification : 031

Acknowledgements
The support from a Sloan foundation grant is gratefully acknowledged.

* Paper presented at the International Conference on the Economics and Econometrics of Innovation, Strasbourg, June 3, 1996.

1 Introduction

It is now widely recognized that the size distribution of profit returns from technological innovation is strongly skewed to the right. The most profitable cases contribute a disproportionate fraction of the total profits from innovation. Much less well understood is the exact form the distribution function takes. In an early analysis (Scherer 1965), I found that the reported profits from a small survey of U.S. patents conformed tolerably well to the Pareto-Levy distribution:

(1) $N = kV^{-\alpha}$,

where V is the value of profits from an innovation, N is the number of cases with value V or greater, and k and α are parameters. For that sample, α appeared to be less than 0.5. Since then economists have sought to discern the size distribution's form by analyzing the rate at which patents have been allowed to expire before full term due to non-payment of the periodic maintenance fees imposed by many national patent offices. Early evidence analyzed by Pakes and Schankerman (1984, p. 78) favored the Pareto-Levy distribution, but later work by Schankerman and Pakes (1986) found mixed but stronger support for the log normal distribution.[1]

The difference between distributions is important. When the distribution is Pareto-Levy and α is less than 2.0, the variance is not asymptotically finite, and for $\alpha < 1$, the mean is also not asymptotically finite. What this implies is that as one draws ever larger samples, there is an increasing chance that some unprecedentedly large value (e.g., an extraordinarily large profit) will be included, overwhelming the observations drawn previously and forcing the mean and variance upward -- in the limit (to be sure, never attained) to infinity. With finite variances and means, log normal and similar skew distributions are better behaved statistically. Still the more rightward-skewed the distribution is, whether Pareto-Levy, log normal, or some related form, the more difficult it is to hedge against risk by supporting sizable portfolios of innovation projects. The potential variability of economic outcomes with Pareto-Levy distributions is so great that large portfolio draws from year to year can have consequences for the macroeconomy. In a simulation experiment, Nordhaus (1989, p. 324) discovered that aggregated Pareto-distributed productivity effects from samples approximating in size the number of patents issued annually in the United States mimicked long-term productivity fluctuations actually experienced by the U.S. economy between 1900 and 1985.[2]

---

[1] See also Pakes (1986, p. 777), Schankerman (1991); and Lanjouw (1992), all of whom find distributions less skewed than the Pareto-Levy.

[2] On the aggregation properties of Pareto-Levy distributions, see Mandelbrot (1963).

Crucial to the portfolio properties of large invention samples is the value distribution of observations in the right-hand (most valuable) tail. On this, studies of patent renewals provide only limited insight. Even in nations with relatively high maintenance fees that rise over the patent's life span, only 10 to 20 percent of the issued patents survive to full term (i.e., 18 to 20 years) after paying all fees. Such patents are clearly of relatively high value. However, the distribution of values within the full-term cohort is ascertained in renewal studies only by extrapolation, not by direct measurement. Because, as we shall see, it is difficult under even the best of circumstances to discriminate among size distributions on the basis of right-hand tail characteristics, extrapolation is hazardous. Also, the mapping from patents to innovations is far from simple. Many innovations are covered by numerous patents, some with a crucial imitation-blocking role, some not. Although patents cannot add significant value to a worthless technology, they enhance the rewards appropriated from some valuable innovations, but in other cases are unimportant because there are alternative barriers to competitive imitation. See Scherer (1977) and Levin et al. (1987).

This paper reports the first results from an ongoing attempt to surmount the limitations of prior research. It probes the right-hand tail, analyzing detailed innovation data across the full spectrum of positive payoff outcomes. And it examines not only individual patents, but technological innovations construed more broadly. Specifically, evidence will be presented from three patent samples, in two of which complementary patents are bundled together; two nearly exhaustive samples of new pharmaceutical entities introduced into the U.S. market; and two large samples of high-technology venture capital investments.

The approach pursued here is inductive, seeking not to impose an a priori theory upon the data, but merely to see what general patterns the data reveal as a first step toward sorting out plausible from implausible theories. This methodology is common in the early hypothesis-formulating stages of the physical sciences.[3] The data will for the most part be presented graphically, notably, on double logarithmic coordinates (on which a Pareto-Levy distribution is linear) and log normal probability coordinates (on which the cumulative log normal distribution has a linear plot). Statistical goodness-of-fit tests among alternative distributions and the exploration of behavioral processes leading to the observed distributions are planned for later papers.

2 Patent Data

Figure 1 plots on log-log coordinates the profitability data analyzed in my 1965 article. The profit estimates were drawn from a survey of 600 U.S. patents by Sanders et al. (1958, p. 355), among which 74 useable responses with positive profits were received. (For 149 additional patents, net losses, typically small, were quantified.) The data were

---

[3] See Ijiri and Simon (1977, pp. 4-5 and 109-111), who explain their approach to analogous problems in similar terms; Scherer (1986) (on Kepler, Einstein, Watson, and Crick), and more generally Hanson (1961).

reported as patent counts over seven profit intervals. Consistent with the Pareto distribution assumptions of equation (1), the dotted line plots the data points at lower interval bounds. Since the lowest (under $5,000) interval was bounded at zero, for which no logarithm exists, a lower bound of $1,000 was assumed. The distribution function's implied left-tail shape is sensitive to that assumption. The highest profit interval, "over $1 million," included five patents. That interval is unbounded, and the right-hand tail's configuration could be sensitive to the distribution of values within it. The solid line plots data points at the geometric means of class intervals, assuming conservatively a mean value for the highest-profit interval of $2.24 million, i.e., the geometric mean of $1 million and $5 million. A value of $3,000 was assumed for the lowest-profit interval. Since two of the seven intervals are in principle unbounded, arbitrariness is inescapable.

For the dotted lower-interval-bound plot, a straight line fitted by least squares yields an α value of 0.37, with r^2 of 0.914. For the solid geometric mean plot, $\alpha = 0.45$ and $r^2 = 0.949$. Both α values are within the range where neither means nor variances are asymptotically finite. However, both plots reveal a modest degree of concavity inconsistent with the Pareto-Levy hypothesis. The data are far too fragmentary to support statistical tests of alternatives to linearity in the logarithms.

For a first extension of the patent value analysis, a novel data source was tapped. The Bayh-Dole Act of 1980 changed U.S. law, permitting researchers in universities and other non-profit institutions supported by federal government grants to apply for and receive patents on inventions resulting from their government-funded research. Many universities established technology licensing offices to apply for patents on such inventions and to negotiate licenses with private sector enterprises for their commercial exploitation. By 1993, the royalty revenues received by 117 U.S. universities from their outstanding technology licenses had reached an annual rate of $242 million.[4] Among those 117, the top ten institutions had royalty revenues of $171 million, or 70.6 percent of the total.

One of the top ten on this list was Harvard University, my employer. The Harvard Office of Technology Licensing kindly provided to me a detailed confidential tabulation of the total royalties received between 1977 and May 1995 on 118 technology "bundles" with non-zero royalties whose patents had been applied for by the end of 1990. Among the 118 bundles, 27 included more than one patent, and six included five or more patents. The twelve bundles with the largest cumulative royalties originated nearly 84 percent of total portfolio royalties -- characteristic evidence of high royalty distribution skewness.

Figure 2 plots the royalty income (multiplied by a constant disguise parameter) from individual invention bundles. The plot is clearly not linear, as would be expected with a Pareto-Levy distribution, but shows considerable concavity to the origin. If a straight line

---

[4] Aggregated university data were provided by the Association of University Technology Managers.

(in the logarithms) is forced by least squares to the data, the indicated α is -0.41, with r^2 of 0.865. However, α is evidently larger in the higher-value tail of the distribution. A linear regression on the 50 most valuable invention bundles yields an α of 0.69, with the standard error of the coefficient being 0.013 and $r^2 = 0.984$. A further analysis of all 118 Harvard technology bundles revealed the fitted Pareto-Levy line to be insignificantly influenced by the number of patents in the bundles and (surprisingly) to the age of the patent bundles.

Figure 3 plots the Harvard royalty data on log normal prob-ability coordinates, with the cumulative probability given on the horizontal axis and the cumulative number of invention bundles required to reach that probability, starting from the most valuable bundle (i.e., from the right-hand tail), on the vertical axis. The linear fit one would expect if the distribution were log normal is absent. The first- and second-highest royalty cases have values too similar to yield a straight-line fit (as is evident also in Figure 2). The remainder of the distribution is slightly concave downward. Thus, the Harvard royalty distribution is much less skew than one would expect under a Pareto-Levy law and somewhat less skew than predicted by a log normal law.

Eleven university technology licensing offices, including the top ten royalty recipients of 1993, were asked to provide information on the distribution of their technology license royalties, divided into nine value ranges, and on total royalty income, for each of their fiscal years 1991 through 1994.[5] Six of the eleven, with total royalties of nearly $83 million in 1993 on 466 positive-royalty cases, responded favorably. All six accounted for their licensed technologies as bundles of patents and disclosures rather than single patents, although, as in the Harvard case, most of the bundles contained only one patent. The largest royalty interval, "more than $5 million," was open-ended. Because data on total royalty receipts per year were obtained, however, it was possible to approximate with tolerable accuracy the values of individual bundles in that tail of the distribution. The approximation was carried out by assigning to closed intervals the mean value of Harvard patents in that interval (which in every case was close to the geometric mean of the interval extremes). Interval totals were found by multiplying the number of patents in an interval by the mean values, and the sum of such totals for all closed intervals was subtracted from total royalties to estimate royalties in the right-hand tail (with at most one observation per university per year).

[5] In interpreting data published by the Association of University Technology Managers, one must be careful to eliminate royalties from trademark licensing, e.g., from firms printing university seals on their garments.

The fractions of total sample royalties contributed by the top six technology bundles in each year were as follows:

	1991	1992	1993	1994
Percent of royalties	66.2%	70.9%	75.6%	74.4%
Total number of royalty-paying bundles	350	408	466	486

The most lucrative bundle licensed by any university held the process and product patents on gene splicing methods assigned in 1980, 1984, and 1988 to Stanley Cohen of Stanford University and Herbert Boyer of the University of California (and administered by the Stanford Technology Licensing Office). At the end of fiscal year 1994, 290 non-exclusive licenses to that bundle had been issued. Over the four years covered by our sample, the bundle yielded royalty payments of roughly $75 million. See Winston-Smith (1996). Since licenses to the Cohen-Boyer patents, which had a revolutionary impact on the biotechnology industry's development, carried only modest royalty payments,[6] the social surplus contributed by the inventions was vastly in excess of royalties appropriated by the patent-holding institutions.

Figure 4 arrays the six universities' royalty distributions on double logarithmic coordinates. Over the four years sampled, the distributions are reasonably stable. They are clearly not linear as predicted under the Pareto-Levy law; considerable concavity is evident. However, it would be premature to reject the linearity hypothesis for the most valuable tail. Fitting log-linear regressions to bundles with annual royalties of $50,000 or more, the results are as follows:

	1991	1992	1993	1994
Estimated α	0.665	0.660	0.583	0.649
Standard error	(.038)	(.036)	(.039)	(.029)
r^2	0.963	0.971	0.949	0.974
Observations (N)	14	12	14	15
Bundles included	64	75	65	76

Again, the α values are in the range within which, asymptotically, Pareto-Levy distributions have neither finite means nor variances.

---

6 The original terms called for a $10,000 advance payment plus royalty rates ranging from 0.5 percent (on the sale of end products such as injectable insulin) to 1-3 percent on bulk products and 10 percent on basic genetic vectors and enzymes.

3 The Profitability of Approved New Pharmaceutical Entities

New chemical entities for use as pharmaceuticals in the United States must undergo a rigorous series of clinical tests before being approved by the Food and Drug Administration. On average, 17.5 new chemical entities (NCEs) received FDA approval per year between 1970 and 1986.[7] Only about 23 percent of the new chemical entities entered into human trials emerged with marketing approval from the FDA. Counting both successes and failures, but ignoring the time value of invested funds, the average pre-clinical and clinical research and development cost of new drugs appearing on the market during the late 1970s and early 1980s was nearly $100 million (in 1987 dollars). See DiMasi et al. (1991).

Henry Grabowski and John Vernon (1990, 1994) used detailed data on drug sales to estimate the gross profitability (before deduction of R&D costs) of new chemical entities (other than cancer drugs) approved by the FDA during the 1970s and early 1980s whose development was carried out by industrial companies in the United States. Subtracting estimated production and marketing costs from sales revenues, domestic and foreign, they obtained for each drug what are best described as Marshallian quasi-rents to R&D investment. These quasi-rents were discounted at a real discount rate of 9 percent to the date at which the drugs were first marketed. The drugs were divided into deciles in descending discounted quasi-rent order, leading to the frequency distribution shown in Figure 5 for average quasi-rents of drugs introduced during the 1970s. A similar analysis with nearly identical results was conducted for drugs introduced between 1980 and 1984. Also estimated was the average research and development investment per approved new chemical entity, including the cost of failed experiments, brought forward at compound interest to the time of marketing -- in Figure 5, $81 million (in 1986 dollars) per new drug introduced during the 1970s. Drugs in the top decile -- the so-called blockbusters -- generated 55 percent of total 1970s sample quasi-rents, i.e., 5.6 times the average R&D costs underlying their market entry.[8] Drugs in the second decile yielded double their R&D investments, drugs in the third decile essentially broke even, and drugs in the seven lowest deciles brought in discounted quasi-rents less than their average R&D investments.

A high degree of skewness is evident in Figure 5. To permit a more detailed analysis, Grabowski and Vernon supplied the quasi-rent data for individual NCEs, multiplied to

---

[7] Pharmaceutical Manufacturers Association, Statistical Fact Book (August 1988), Table 2-4. The typical new drug is protected by one product patent and sometimes by a few process patents. In 1976, U.S. pharmaceutical manufacturers obtained at least 868 patents and, allowing for incomplete sample coverage, as many as 1,000. See Scherer (1983, p. 110). Thus, there is far from a one-to-one correspondence between new patent counts and new product counts.

[8] One might expect R&D costs to be higher for the most lucrative drugs, but the available evidence fails to provide support. See DiMasi et al. (1994).

maintain confidentiality by an undisclosed disguise parameter. (Multiplication by a constant does not distort the size distri-bution parameters in which we are interested.) Figure 6 plots on double log coordinates the data for 98 NCEs introduced during the 1970s. (Two observations with negative quasi-rents are omitted.) Figure 7 does the same for 66 NCEs introduced between 1980 and 1984. As with the quite differently constituted patent data, the distribution is much too concave to be Pareto-Levy. However, if one focuses only on the most successful third of all the new products, a log linear regression fits tolerably well:

	1970s NCEs	1980s NCEs
Estimated α	1.14	1.18
Standard error	(.04)	(.075)
r^2	0.964	0.926
Number of NCEs	33	22

Here, for the first time, with samples that cover almost exhaustively the relevant population of domestically developed and approved new chemical entities in their time frames, we find α values in the tail exceeding the unit threshold below which Pareto-Levy first moments are asymptotically infinite. From the evidence analyzed thus far, we begin to suspect a sampling bias on estimated α values. Given concavity of the innovative reward distribution on log-log coordinates, the smaller and less exhaustive the sample is relative to the potential universe, the less likely it is that extreme outliers will be drawn in, and hence the lower (in absolute value) estimated α coefficients for the right-hand tail will be.

Figure 8 plots the NCE data for the 1970s on log normal probability coordinates. As with the Harvard invention bundle sample, the plot deviates visibly from linearity. Its consistent concavity suggests more equality among the observations, and hence less skewness, than one would expect if the data conformed to a log normal process.

4 High-Technology Venture Firm Startups

An institution that contributes enormously to America's prowess in high-technology fields is its venture capital industry. Hundreds of new firms are founded each year to develop and commercialize promising ideas emerging from university laboratories, independent inventors, and industrial corporations that for some reason chose not to pursue the opportunities internally. See Roberts (1991). Typically, initial experiments and bench model development are supported using the technically trained entrepreneur's own funds and seed money raised from acquaintances (who as high-technology "angels" may sustain many such early investments). When this low-cost preliminary activity yields promising results, the fledgling enterprise turns to a high-technology venture fund for financial support, which ranges from a few hundred thousand to several million dollars. The venture capital fund raises money from an array of investors -- in the industry's early history, from wealthy individuals, but more recently, from pension funds and university endowments -- and attempts to pool its risks by investing in dozens of startup enterprises. If an individual venture succeeds in marketing one or more new products with good prospects, it "goes public" -- i.e., it floats an initial public offering (IPO) of its common stock; or (somewhat more frequently) its investors sell out their shares to a well-established company. The venture fund investors then "cash in" their proceeds or reinvest them in the new publicly-traded company shares.

The first modern U.S. high-technology venture capital fund was the American Research and Development Corporation (ARDC), founded in Boston shortly after the close of the Second World War. Figures 9a and 9b, drawn from Willmann (1991), trace ARDC's early portfolio history. Figure 9a shows the number of individual startup companies in which ARDC invested annually (light dotted line) and the total number of companies in its portfolio (solid line). During the 1950s, its portfolio contained from 23 to 30 companies. Its investment target count rose into the mid 40s by the 1960s. Figure 9b traces the net value of ARDC's investment portfolio. During the mid-1950s, a few successes (such as High Voltage Engineering Company and Airborne Instruments Inc.) fueled an appreciable portfolio value increase. In 1966, however, the portfolio value exploded. By decomposing the portfolio into two parts -- Digital Equipment Company (DEC) and more than 40 other companies -- Figure 9b shows that most of the explosion was attributable to ARDC's $70,000 investment (in 1957) in DEC. DEC's great success came with the introduction of the first time-sharing computer, the PDP-6, in 1964, and a powerful but inexpensive minicomputer, the PDP-8, in 1965. An initial public offering of DEC's common stock was floated on the New York Stock Exchange in 1966.

From the history of ARDC, we see considerable skewness in the returns from high-technology investments and the volatility such skewness can impart, despite venture investors' attempts to hedge against risk by forming sizeable portfolios. To determine how well the history of ARDC generalizes, two additional data sources were tapped.

One study of venture capital performance was conducted by the leading source of information on U.S. venture funds, Venture Economics, Inc. (1988, Chapter II). Venture Economics analyzed the success of 383 individual startup company investments made by 13 U.S. venture portfolios between 1969 and 1985 whose investment cycles had been largely completed by 1986.[9] Figure 10 arrays the individual investments by the multiple of terminal value relative to original investment outlays. The 26 individual startups returning ten times or more the funds' initial stakes accounted for 49 percent of total terminal portfolio values. More than a third of the ventures returned less than their

[9] During the early 1980s, venture capital funds began to invest in real estate deals, leveraged buyouts, and other targets as well as high-technology startups. The fraction of the investments attributable to genuinely high-technology startups in the portfolios analyzed here was not disclosed.

original fund investments. Figure 11 plots the distribution function on double log coordinates in two ways, one (solid line) using interval average values per portfolio investment as the horizontal axis variable (and assuming "total losses" to return 10 percent of the initial investment); the other using the lower threshold values for the intervals. Since the average return in the highest-return interval was reported to be 21.6 times original investment values, the threshold approach suppresses important information on the distribution's upper tail. With both methods, the distribution function is concave, not linear as implied by the Pareto-Levy hypothesis. If a linear function is forced upon the full centered observation data set by least squares, the implied α value is 0.60. However, if the regression is limited to the four highest-value (i.e., right-hand tail) interval means, the fitted α is 0.974, with standard error of 0.044 and r^2 of 0.996.

Figure 12 summarizes the results of a similar study by a San Francisco venture capital house, Horsley Keogh Associates (1990). Included in the analysis were 670 distinct investments (totalling \$496 million) in 460 companies made between 1972 and 1983 by 16 venture capital partnerships. The ultimate portfolio value was calculated as of December 1988, at which time the funds had distributed to their partners \$822 million and retained assets of \$278 million. The 34 investments that yielded ten times or more their original value contributed 42 percent of the portfolios' total terminal value. Slightly more than half of the investments entailed some loss. Figure 13 plots the distribution function on double-log coordinates, again using both mean values (solid line) and interval threshold values (dotted line). (Investments in the most lucrative interval returned on average 19.25 times their initial value.) Again we find concavity over the full range of observations, but near linearity in the right-hand tail. For all centered data points, the fitted least-squares line has an α of 0.77. However, for the four highest-value centered observations, $\alpha = 0.998$, its standard error = 0.083, and $r^2 = 0.986$.

5 Ongoing Empirical Research

To supplement the insights achieved through the studies summarized above, two additional empirical research projects are underway.

One analyzes the stock price histories of 131 U.S. high-technology companies, initially backed by venture capital funds, that floated initial public stock offerings between 1983 and 1986. The 131 companies are believed to be an exhaustive sample of such IPOs in the relevant time frame. Monthly changes in their common stock values will be tracked to the end of 1995. It is evident from preliminary work that the distribution of returns is highly skew.[10]

[10] For similar research on the 13 most merger-prone U.S. conglomerate corporations' common stock price performance between 1965 and 1983, see Ravenscraft and Scherer (1987, pp. 38-44). The fitted α for gains in those companies' stock values was 0.58.

A second project probes the tail of the distribution of values associated with the approximately 21,000 patents applied for in 1977 and eventually issued by the German Patent Office.[11] From that population 4,349 patents, including 1,435 patents of domestic German origin, paid all renewal fees and expired after running their full statutory 18-year term in 1995. The German renewal fees, it is worth noting, are among the highest and most progressive in the world. Through a mail and telephone survey, rough preliminary value estimates will be obtained from companies holding the full-term patents of domestic origin. During the summer of 1996, officials of companies holding the 250 patents found to be most valuable within that larger surviving cohort will be interviewed face-to-face to obtain detailed information on profitability, invention characteristics, and the role patent protection played in appropriating the returns from innovation.

6 Behavioral Implications

With the completion of that essentially descriptive research, attention will shift to another question: what behavioral processes give rise to the observed distribution of returns from technological innovation? It is already clear that there are important regularities in the distribution of returns. The distributions are almost uniformly concave to the origin on doubly logarithmic coordinates, and not linear, as postulated by the Pareto-Levy hypothesis. They are highly skewed, however, with long tails approximately linear on log-log coordinates. Values of α fitted to tail observations indicate that it is most difficult to achieve high predictability through portfolio averaging strategies.

If such regularities prevail in the distribution of outcomes, it is reasonable to suppose that there are more or less well-defined behavioral processes generating the outcomes. Thus, subsequent research will focus on formulating and testing what Ijiri and Simon (1977, p. 109) call "extreme hypotheses" -- i.e., assertions that a particular specific functional relation leads to the emergence of a particular distribution of outcomes.

One possible extreme hypothesis is that some Supreme Power regularly strews about the industrial landscape a distribution of raw profit potentials for technological innovation that is highly skew, just as the distribution of petroleum reservoirs within a land mass, and hence the opportunity for profiting from exploratory well drilling, is believed to be log normal. See Adelman (1972, p. 35). The profit opportunities from innovation might be roughly proportional to the size of markets, which, we know from the distribution of sales or value added across conventionally defined industries, is skew-distributed. However, a plot on log-log coordinates of the distribution of value added in 450 U.S. four-digit manufacturing industries for 1987 revealed much more concavity, and hence less skewness, than has been observed for any of the innovation profit distributions

[11] The research is being conducted jointly with Dietmar Harhoff of the Zentrum für Europäische Wirtschaftsforschung, Mannheim.

investigated here. At the very minimum, something must be adding skewness to the innovation return distributions.

An alternative extreme hypothesis is that the distribution of returns from innovation results from some variant of a Gibrat process, under which numerous chance events interact multi-plicatively, reducing the profits actually realized from an innovative potential which may or may not be skewed initially. If P_o is the initial potential and ε_i is the i^{th} stochastic multiplier affecting the amount of value that can be appropriated by an innovator, the ultimate innovator's quasi-rent is:

$$(2) \quad V = P_o \varepsilon_1 \ldots \varepsilon_i \ldots \varepsilon_n;$$

where the typical ε consists of an expected value less than unity plus an error component. The ε's represent inter alia the initial probability of technical success, the time at which the innovator arrives on the market with its new product and hence the strength of first-mover advantages, the strength of the innovator's patent protection, the finesse with which initial marketing efforts are conducted (crucial e.g. in the competition between anti-ulcer drugs Tagamet and Zantac), and the extent to which the market is fragmented by imitators in each subsequent year of commercial sales (which is likely to be correlated with the strength of first-mover advantages). Taking logarithms, we have:

$$(3) \quad \ln V = \ln P_o + \ln \varepsilon_1 + \ldots + \ln \varepsilon_i + \ldots + \ln \varepsilon_n;$$

which, by the central limit theorem, is normally distributed with sufficiently large n. Thus, under the logic of Gibrat's law, a log normal distribution of V might be anticipated.

Our initial data suggest that the distribution of returns from innovation may be less skew than log normal, but with a long log-linear tail. Thus, the data may conform better to a Yule distribution than the log normal distribution. Yule distributions, Simon and Ijiri have shown, result from Gibrat-type processes in which the initial population is not fixed, but into which entry occurs, concentrated in the lower-value tail. That the distribution of returns from innovation conforms more closely to the Yule than to the log normal is therefore an alternative extreme hypothesis. Less than log normal skewness, Ijiri and Simon have suggested (1977, pp. 88-93), may also result from a "splitting" process corresponding to the stochastic processes underlying Bose-Einstein statistics. In the world of technological innovation, "splitting" may occur endogenously as the most profitable innovations attract more competitive imitators than less profitable innovations.[12]

Thus, there are plausible theoretical links between the stochastic search, experimentation, and market penetration processes associated with innovation and the kinds of return

---

[12] See the reinterpretation of Mansfield's (1977) results in Scherer (1983).

distributions observed thus far, and whose parameters and dynamic evolution we hope to pin down more decisively through further research. When that empirical research is completed, simulation models will be constructed and tested to determine what kinds of plausible stochastic behavioral models generate the observed profit return distributions. It is hoped that in this way a deeper under-standing of the economics of technological innovation will follow.

References

Morris A. Adelman, The World Petroleum Market (Baltimore: Johns Hopkins University Press, 1972).

Joseph DiMasi, Ronald W. Hansen, Henry Grabowski, and Louis Lasagna, "Research and Development Costs for New Drugs by Therapeutic Category," working paper, Tufts University, 1994.

Henry Grabowski and John Vernon, "A New Look at the Returns and Risks to Pharmaceutical R&D," Management Science, vol. 36 (July 1990), pp. 804-821.

Henry Grabowski and John Vernon, "Returns on New Drug Introductions in the 1980s," Journal of Health Economics, vol. 13 (1994), pp. 383-406.

N. R. Hanson, Patterns of Discovery (New Haven: Yale University Press, 1958).

Horsley Keogh Associates, Horsley Keogh Venture Study (San Francisco: privately distributed, 1990).

Yuji Ijiri and Herbert A. Simon, Skew Distributions and the Sizes of Business Firms (Amsterdam: North-Holland, 1977).

Jean Olson Lanjouw, "The Private Value of Patent Rights in Post WWII West Germany," working paper, Yale University, July 1992.

Richard C. Levin, Alvin Klevorick, Richard R. Nelson, and Sidney G. Winter, "Appropriating the Returns from Industrial Research and Development," Brookings Papers on Economic Activity, 1987, no. 3, pp. 783-820.

Benoit Mandelbrot, "New Methods in Statistical Economics," Journal of Political Economy, vol. 71 (October 1963), pp. 421-440.

Edwin Mansfield, John Rapoport, Anthony Romeo, Samuel Wagner, and George Beardsley, "Social and Private Rates of Return from Industrial Innovations," Quarterly Journal of Economics, vol. 91 (May 1977), pp. 221-240.

Ariel Pakes and Mark Schankerman, "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources," in Zvi Griliches, ed., R&D, Patents, and Productivity (University of Chicago Press: 1984), pp. 73-88.

Ariel Pakes, "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," Econometrica, vol. 54 (July 1986), pp. 755-784.

David J. Ravenscraft and F. M. Scherer, Mergers, Sell-offs, and Economic Efficiency (Washington: Brookings, 1987).

Edward B. Roberts, Entrepreneurs in High Technology (New York: Oxford University Press, 1991).

Barkev Sanders, Joseph Rossman, and L. James Harris, "The Economic Impact of Patents," Patent, Trademark, and Copyright Journal, vol. 2 (September 1958), pp. 340-363.

Mark Schankerman and Ariel Pakes, "Estimates of the Value of Patent Rights in European Countries during the Post-1950 Period," Economic Journal, vol. 97 (December 1986), pp. 1-25.

Mark Schankerman, "How Valuable Is Patent Protection? Estimates by Technology Field Using Patent Renewal Data," National Bureau of Economic Research Working Paper no. 3780 (July 1991).

F. M. Scherer, "Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions," American Economic Review, vol. 55 (December 1965, pp. 1097-1123.

F. M. Scherer, The Economic Effects of Compulsory Patent Licensing (New York University Graduate School of Business Administration, Monograph Series in Finance and Economics: 1977).

F. M. Scherer, "The Propensity To Patent," International Journal of Industrial Organization, vol. 1 (March 1983), pp. 107-129.

F. M. Scherer, "R&D and Declining Productivity Growth," American Economic Review, vol. 73 (May 1983), pp. 215-218.

F. M. Scherer, "On the Current State of Knowledge in Industrial Organization," in H. W. de Jong and W. G. Shepherd, eds. Mainstreams in Industrial Organization, Book I (Dordrecht: Kluwer, 1986), pp. 5-22.

Venture Economics, Inc., Venture Capital Performance (Boston: 1988).

Heidi Willmann, "Innovation in the Venture Capital Industry: A Study of American Research and Development Corporation," term paper, John F. Kennedy School of Government, Harvard University, May 1991.

Sheryl Winston-Smith, "The Cohen-Boyer Patent: A Case Study," term paper, John F. Kennedy School of Government, Harvard University, January 1996.

Figure 1

Pareto Plot of Original Sanders Data

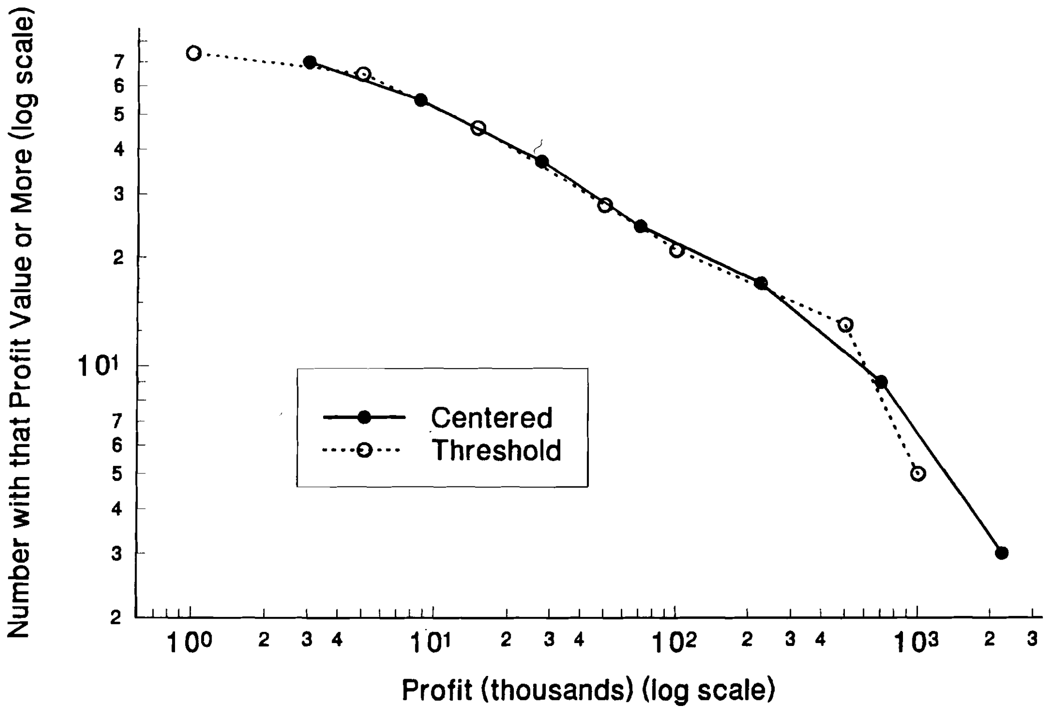

Figure 2

Pareto Plot of Harvard Invention Bundles

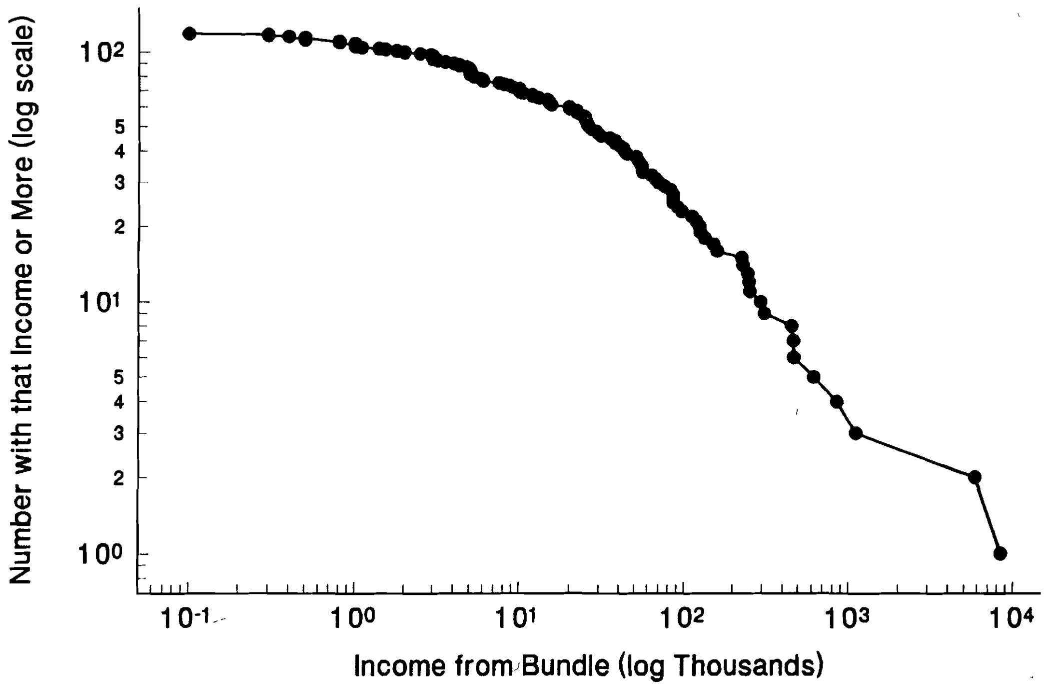

Figure 3

Log Normal Plot of Harvard Invention Bundles

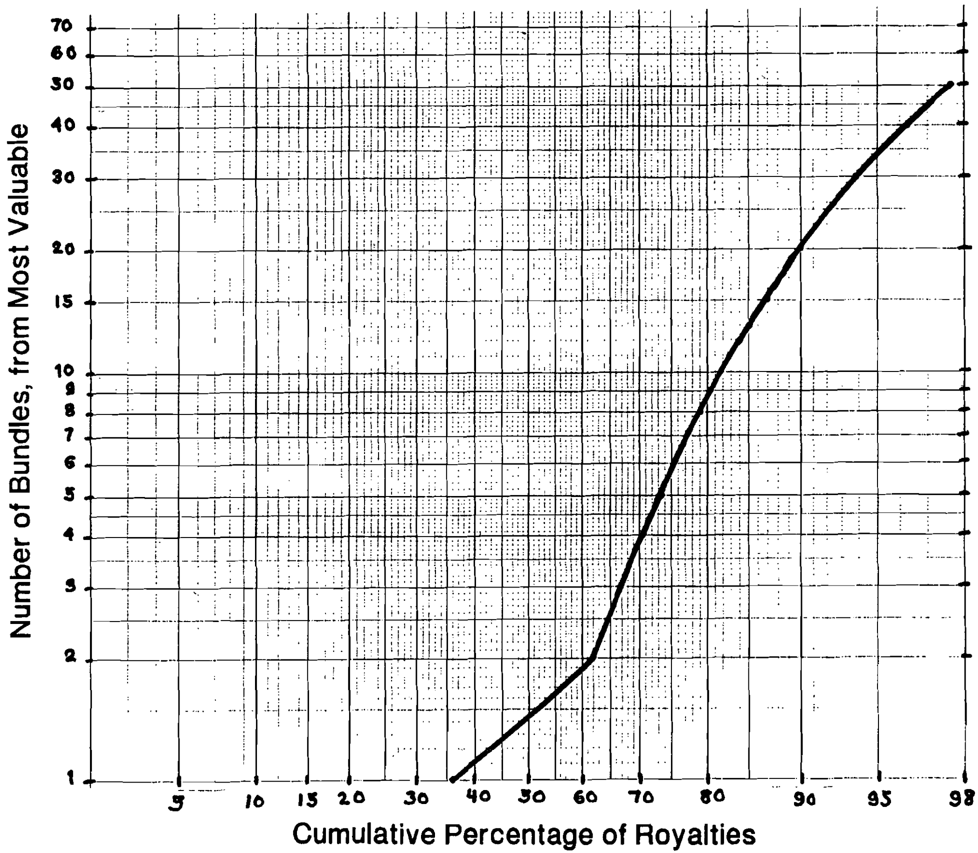

Figure 4

Pareto Plot of Six Universities' Patent Royalties by Year

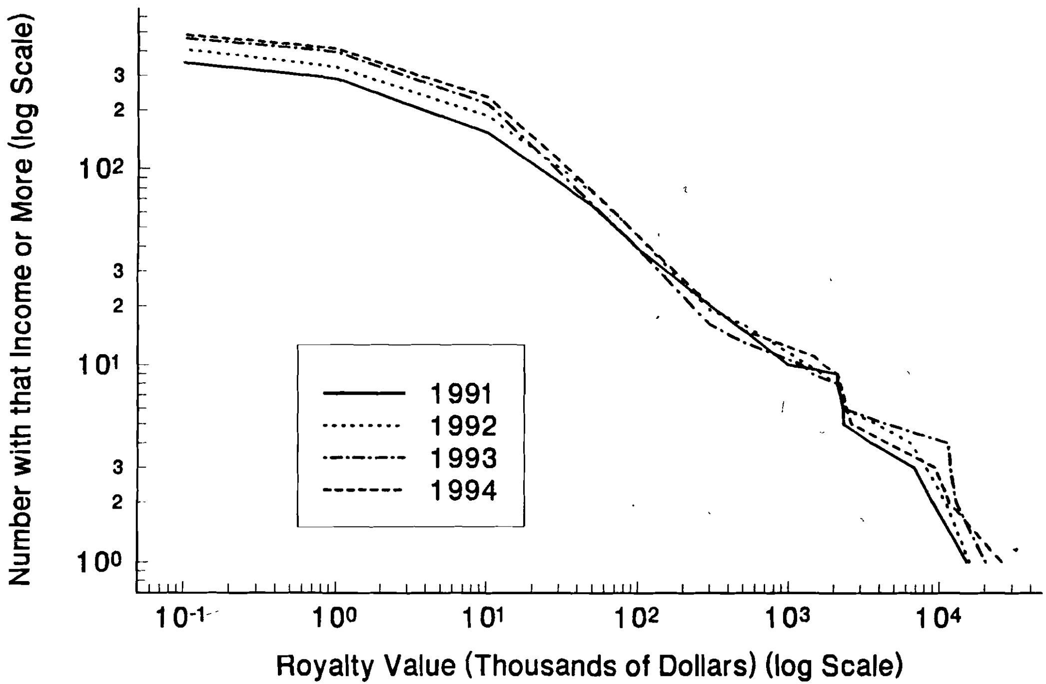

Figure 5
Distribution of 1970s Drug NCE Quasi-Rents

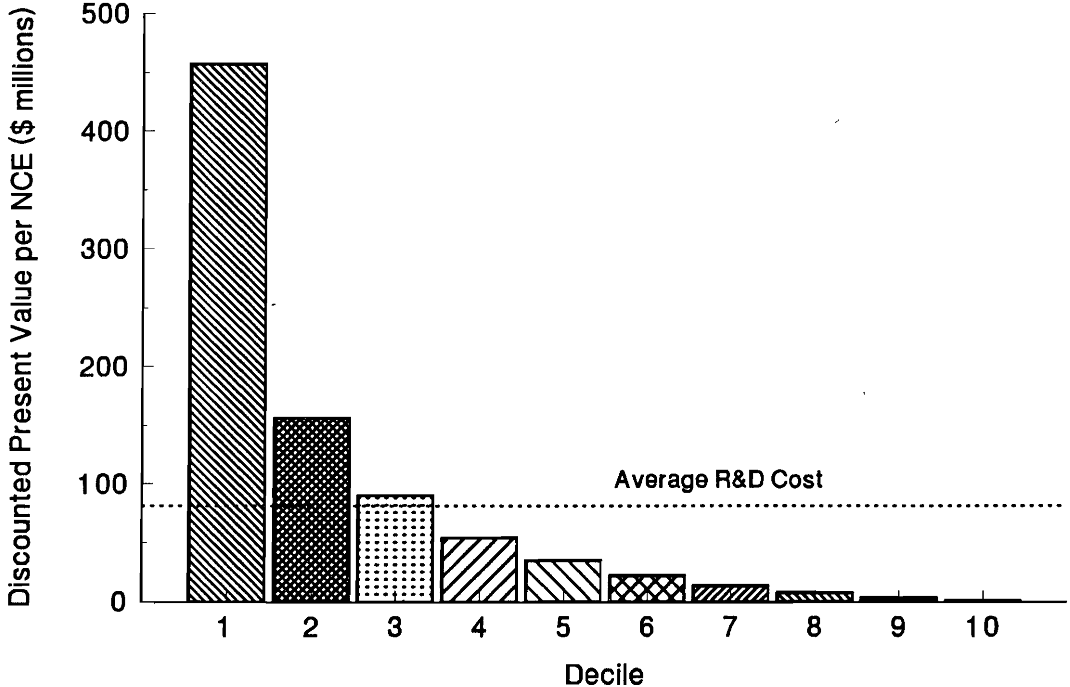

Figure 6

Pareto Plot of 1970s New Drug Entities

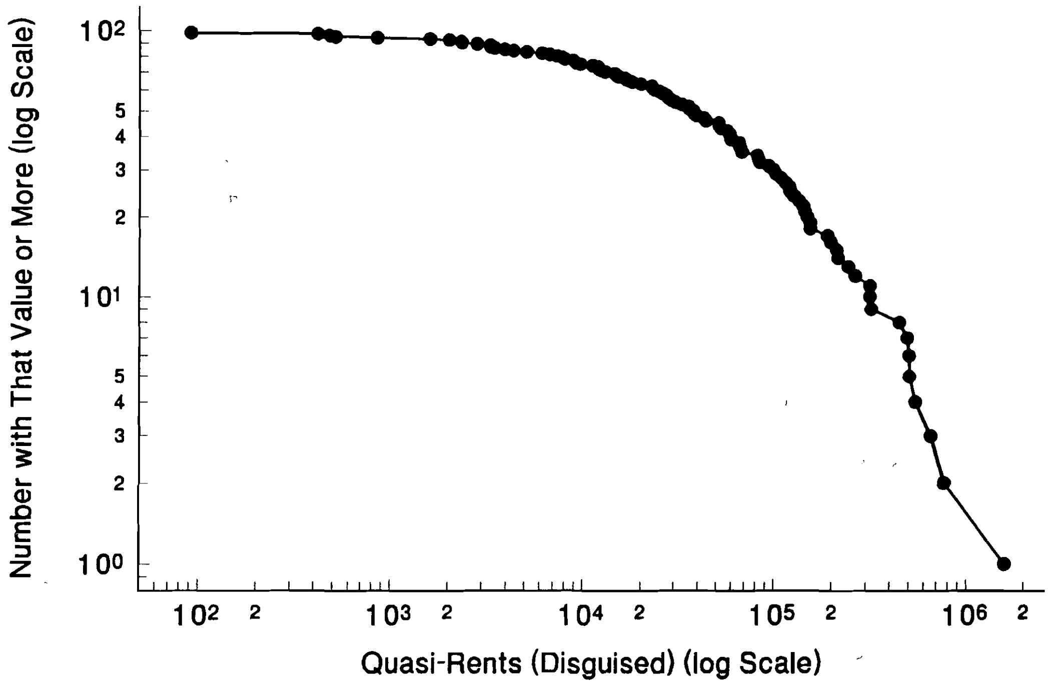

Figure 7
Pareto Plot of 1980s New Drug Entities

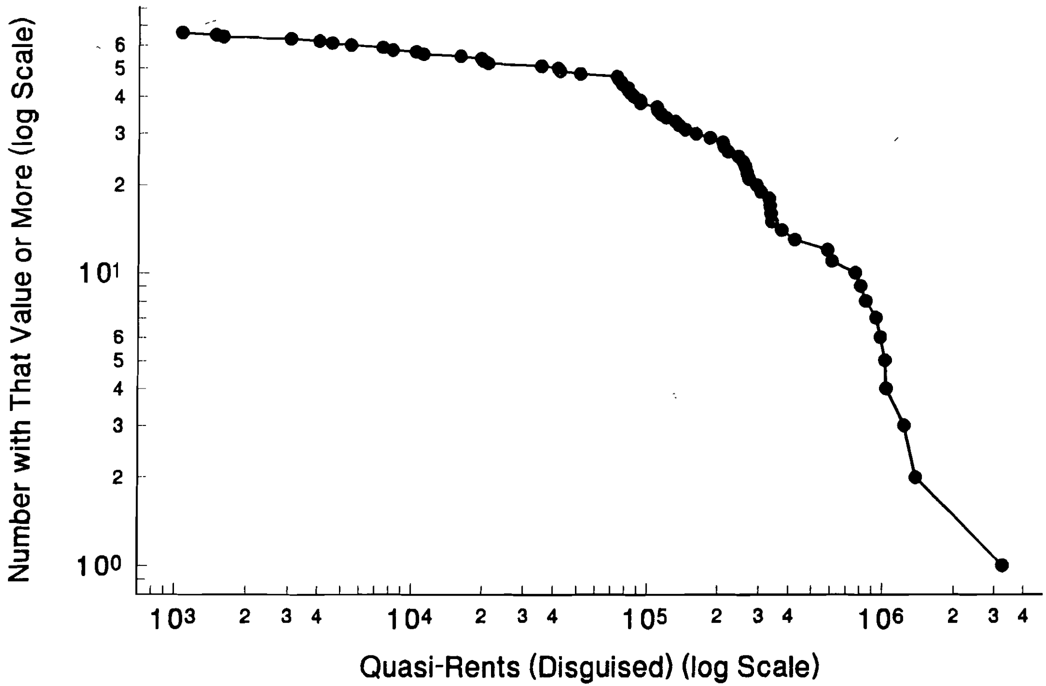

Figure 8

Log Normal Plot of 1970s Drug Quasi-Rents

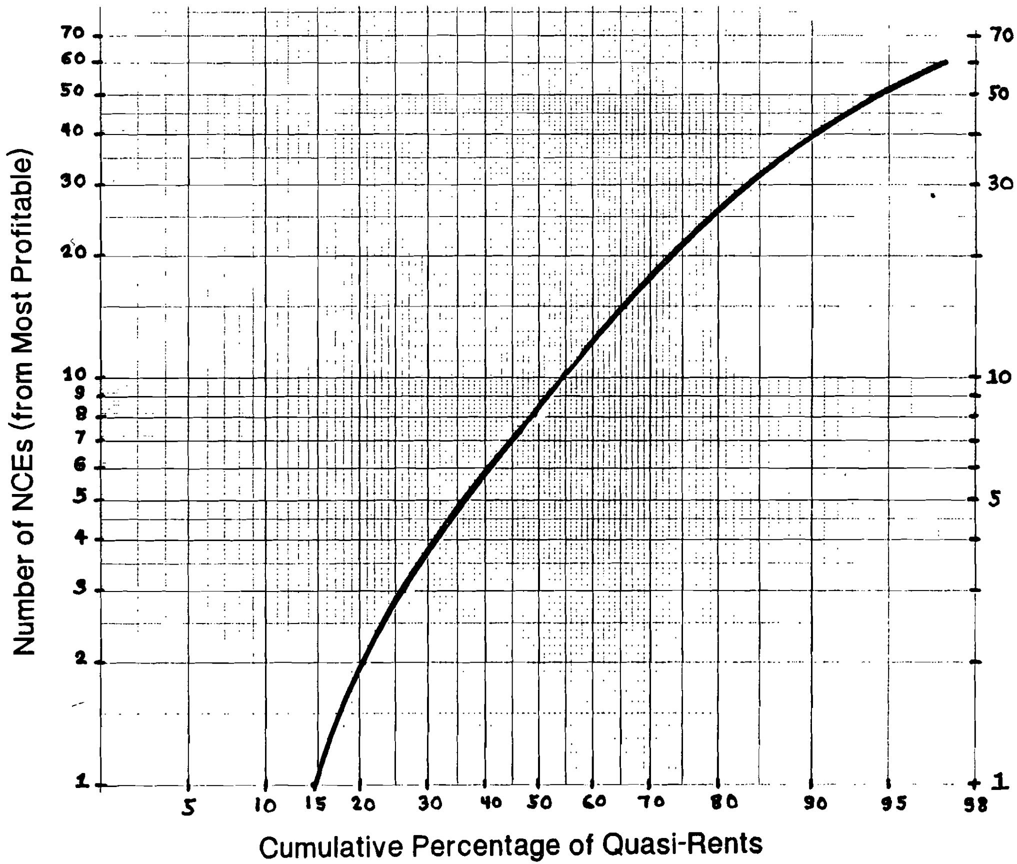

Figure 9a

Company Investments in ARDC's Portfolio

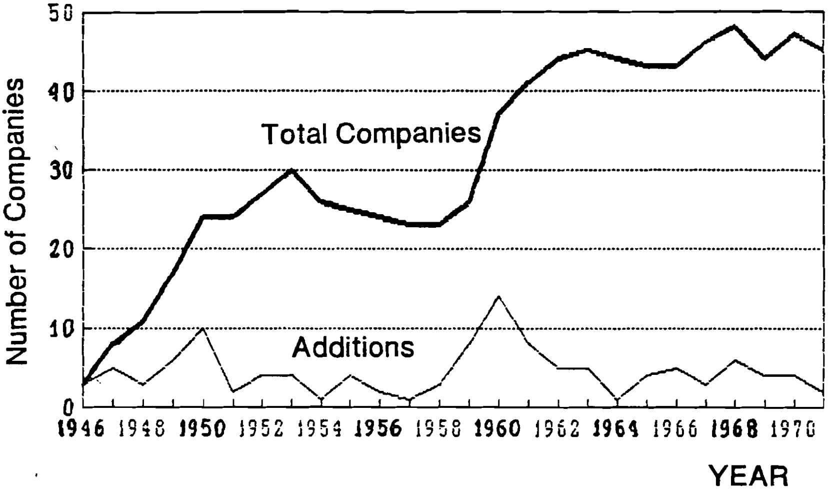

Figure 9b

ARDC's Net Asset Value Per Share

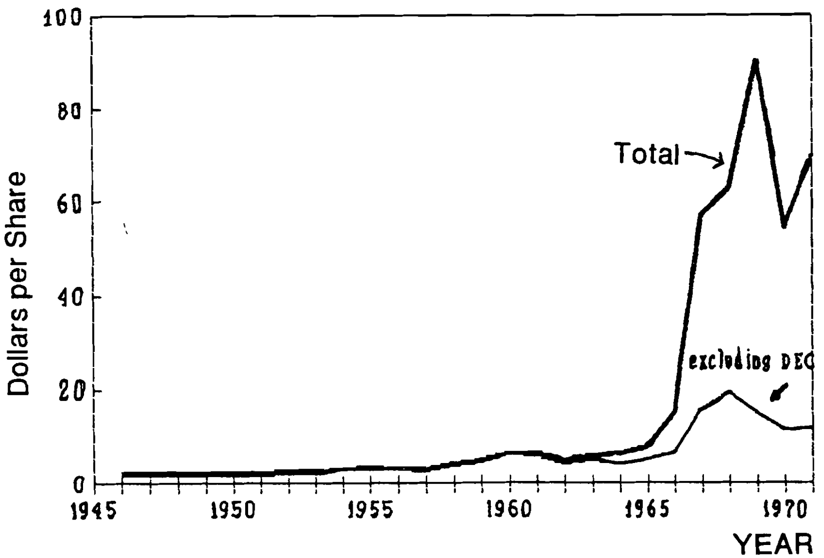

Figure 10

Distribution of Gains on 383 Venture Portfolio Companies

Source: Venture Economics, Inc. (1988)

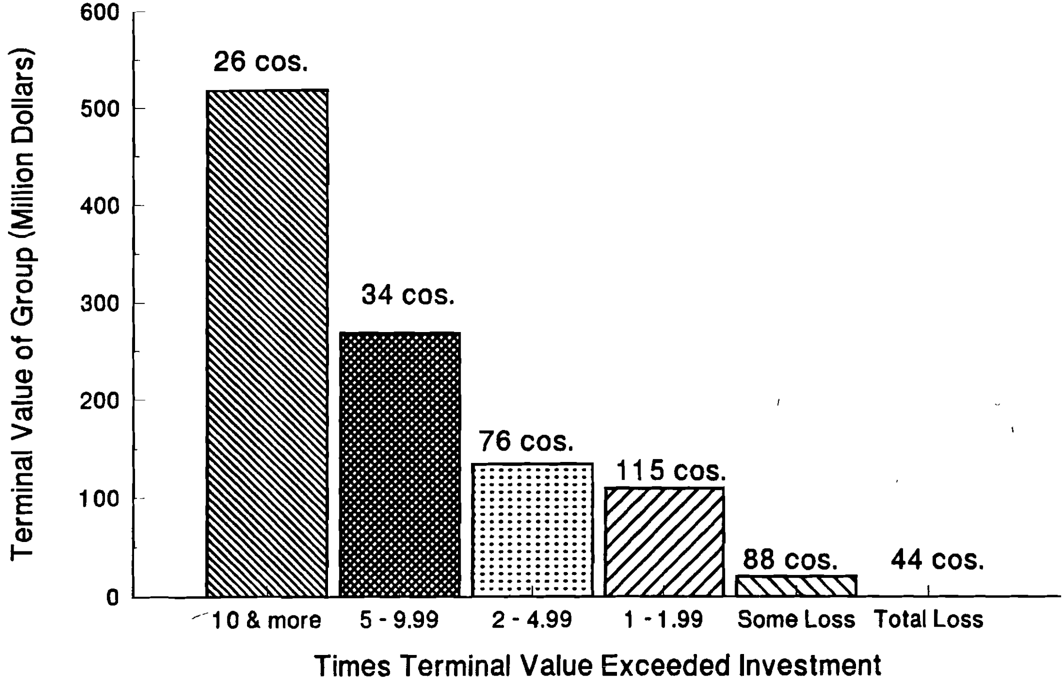

Figure 11
Pareto Plot of Returns from 383 Venture Portfolio Investments

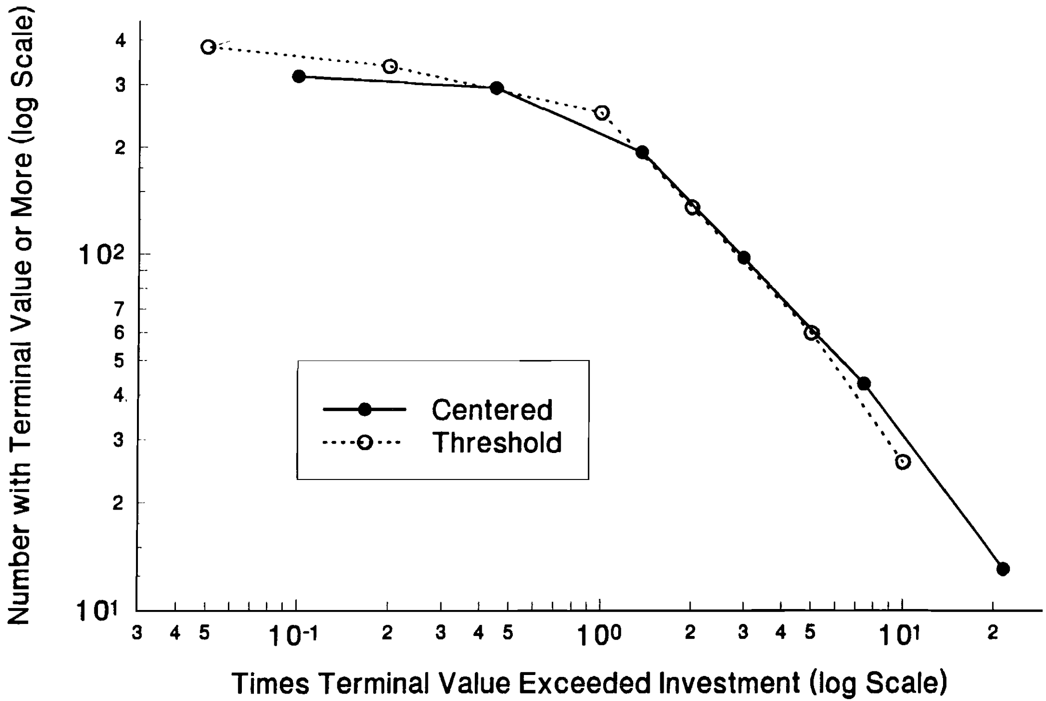

Figure 12

Distribution of Gains on 670 Venture Portfolio Investments

Source: Horsley Keogh (1988)

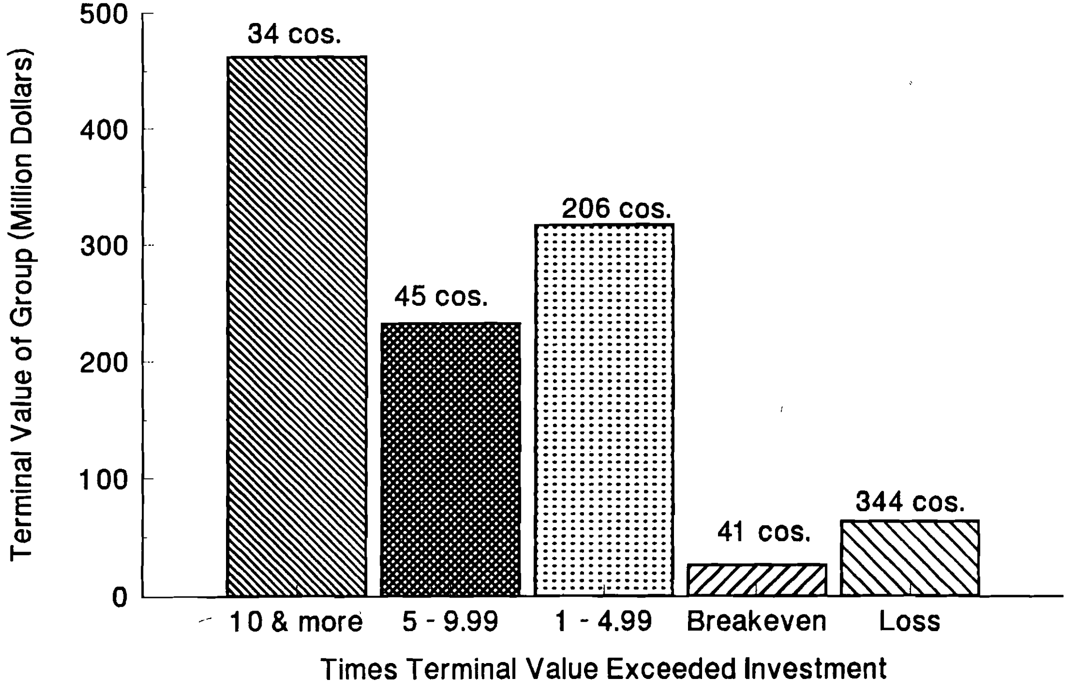

Figure 13
Pareto Plot of Returns from 670 Venture Portfolio Investments