J Evol Econ (2000) 10: 175–200

Journal of
Evolutionary
Economics


# Uncertainty and the size distribution of rewards from innovation*

**F. M. Scherer[1], Dietmar Harhoff[2], Jörg Kukies[3]**

[1] John F. Kennedy School of Government, Harvard University,
Cambridge, MA 02138, USA (e-mail: mike_scherer@harvard.edu)
[2] Universität München, D-80539 München, Germany
[3] Graduate School of Business, University of Chicago, Chicago, USA

**Abstract.** Previous research has shown that the distribution of profit outcomes from technological innovations is highly skew. This paper builds upon those detailed findings to ask: what stochastic processes can plausibly be inferred to have generated the observed distributions? After reviewing the evidence, this paper reports on several stochastic model simulations, including a pure Gibrat random walk with monthly changes approximating those observed for high-technology startup company stocks and a more richly specified model blending internal and external market uncertainties. The most highly specified simulations suggest that the set of profit potentials tapped by innovators is itself skew-distributed and that the number of entrants into innovation races is more likely to be independent of market size than stochastically dependent upon it.

**Key words:** Innovation – Risk – Uncertainty – Skew distributions – Gibrat's Law

**JEL-classification:** O31, C15

## 1 Introduction

For two years the authors have been compiling and analyzing several new data sets on the profit returns to inventions and high-technology startup

---

Paper presented at the 1998 World Conference of the International Joseph A. Schumpeter Society, Vienna, June 13–16, 1998.

* Financial support from the Sloan Foundation and the Center for European Economic Research, Mannheim, at earlier stages of this project is gratefully acknowledged.

*Correspondence to*: F.M. Scherer

enterprises. The data, combined with evidence available earlier, exhibit striking regularities. The size distribution of rewards from technological innovation is highly skew, with an extended tail on the high-value side. A log normal distribution provides in most instances the best fit. Having pinned down with considerable confidence the structure of rewards ensuing from innovative activity, we advance now to the stage of formulating what Ijiri and Simon (1977, p. 109) called "extreme hypotheses," attempting to determine what behavioral processes, undoubtedly stochastic, led to the observed distributions. In this paper we summarize the empirical evidence and then explore, mainly through simulation exercises, plausible alternative stochastic processes to see which ones conform most closely to the size distributions actually encountered in the world of innovation.

## 2 The empirical evidence

Our research has focused on eight sets of data, seven of which are new to the economics literature (see Scherer, 1998; Harhoff et al., 1999). The most ambitious effort involved collecting useable survey and interview evidence on 772 German and 222 U.S.-origin inventions, on all of which German patent applications were filed in 1977, leading to issued German patents considered sufficiently valuable by their holders to warrant paying annual renewal fees totalling DM 16,075 until their expiration at full term in 1995. These will be called the German and U.S. patent data sets. Two other data sets pertained also to invention patents, one tallying the royalties received between 1977 and 1995 on the 118 invention patent "bundles" licensed by the Harvard University Office of Technology Licensing, the other royalties received in the years 1991 through 1994 by six research-oriented U.S. universities on 350 to 466 licensed patent bundles. These will be called the "Harvard" and "Six University" data sets respectively. Two other data sets tallied the asset value appreciation (or loss) experienced on a total of 1,053 investments in startup companies by U.S. venture capital firms between 1969 and 1988. One sample, covering 383 investments, was compiled by Venture Economics Inc.; the other, covering 670 investments, by Horsely Keogh Associates. They will be referred to by the names of the compiling firms. Still another analysis of startup company experiences was carried out by the authors. It followed the common stock values of 131 high-technology companies previously nurtured by venture capital firms from the time they made initial public stock offerings (IPOs) between 1983 and 1986 through (at least, for 52 surviving entities) 1995. This is called the IPO survey. Finally, we used the data compiled and analyzed previously by Henry Grabowski and John Vernon (1990, 1994) on the discounted present value of profits (or more exactly, quasi-rents) realized on two sets of new pharmaceutical chemical entities marketed in the U.S. market – 98 introduced in the 1970s and 66 between 1980 and 1984. These will be called the "Grabowski-Vernon drug" data sets.

Table 1 characterizes in two main ways the reward outcome distributions observed for these eight samples. The second numerical column estimates the fraction of total observed sample profits, royalties, or market

**Table 1.** Paremeters of the observed innovation reward size distributions

| Data set | Number of observations | Percent of value in top 10 percent | Slope estimates $\alpha$ | |
|---|---|---|---|---|
| | | | All obs. | Right tail[a] |
| German patents | 772 | 88% | .42 | $.87_{(81)}$ |
| U.S. patents | 222 | 81–85% | .32 | $.43_{(63)}$ |
| Harvard patents | 118 | 84 | .41 | $.71_{(30)}$ |
| Six university patents | | | | |
| 1991 Royalties | 350 | 93 | .50 | $.64_{(153)}$ |
| 1992 Royalties | 408 | 92 | .51 | $.65_{(186)}$ |
| 1993 Royalties | 466 | 91.5 | .51 | $.61_{(214)}$ |
| 1994 Royalties | 411 | 92 | .52 | $.65_{(233)}$ |
| Venture economics startups | 383 | 62 | .60 | $.97_{(136)}$ |
| Horsley-Keogh startups | 670 | 59 | .77 | $.90_{(285)}$ |
| IPOs – 1995 value | 110 | 62 | .44 | $.82_{(13)}$ |
| Grabowski-Vernon | | | | |
| 1970s drugs | 98 | 55 | .43 | $1.22_{(23)}$ |
| 1980s drugs | 66 | 48 | .41 | $1.36_{(17)}$ |

[a] The number of right-hand tail observations covered is indicated in subscripted parentheses. For the IPOs, the sample includes only 53 firms surviving at the end of 1995.

value captured by entities ranked among the top ten percent of reward earners. For samples on which the distributional data were available only by class intervals, the Top Ten Percent share was estimated by nonlinear interpolation, assuming a log normal distribution of rewards.[1]

The third and fourth columns in Table 1 summarize the distribution parameters using the Pareto-Levy distribution, which has the particularly simple form:

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

where V is the value of the profits or other rewards from an innovation, N is the number of cases with value V or greater, and k and $\alpha$ are parameters. For data sets on which individual observations were available, the log-linear fit implied by the Pareto-Levy distribution did not emerge; the distribution was characteristically concave to the origin, as illustrated (despite one extreme billion-Mark value) for the German patent data in Fig. 1. However, in many instances, observations in the right-hand (most valuable) tail exhibited a fit close to log linear, and in any case, information on $\alpha$ slope values for the entire set of observations and those in the right-hand tail conveys a good sense of both the distribution's skewness and the degree of

[1] For the most valuable U.S. patents, a range of estimates is presented, assuming alternatively that patents in that open-ended interval averaged $200 million and $250 million. For values in closed distribution intervals, geometric means of the interval bounds were used to estimate interval average values.

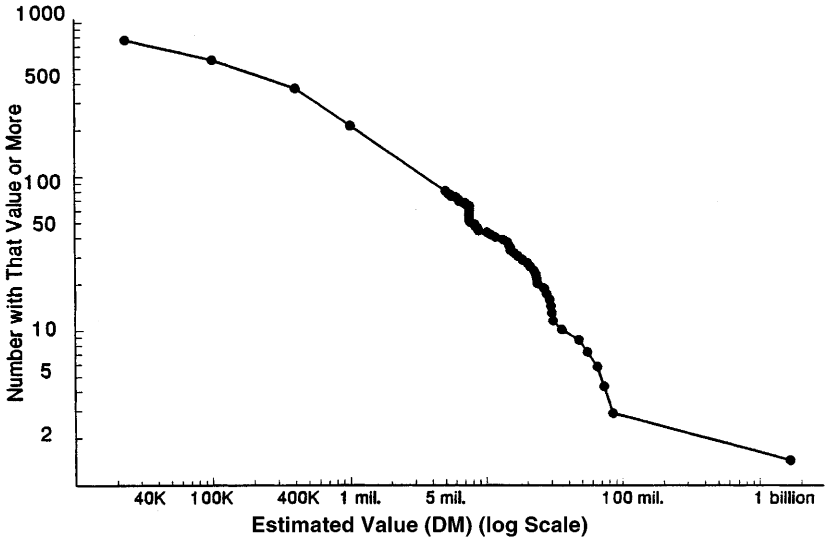


**Fig. 1.** Plot of German renewed patent values on Pareto coordinates

concavity. Therefore, we report in Table 1 values of $\alpha$ estimated by ordinary least squares regression for all available observations and for a subset conforming as closely as possible to the most valuable 25 percent of those observations. For data available only by class intervals, the right-hand tail $\alpha$ values were estimated by reaching out to the include the class boundary including (and over-reaching) the 25 percent point.[2]

Two generalizations stand out. First, in all cases, a relatively small number of top entities accounted for the lion's share of total invention or innovation value. The highest concentration of value is unambiguously for the patents, which tend to cover the narrowest range of innovative subject matter.[3] (Many new products and processes are protected by multiple patents.) The fraction of total portfolio value attributable to the top ten percent of business entities is quite similar for the two sets of venture-fund-

[2] For the German patents, we include only right-hand tail observations on which detailed interview data were available.

[3] A simulation analysis revealed that when there are complementarities or substitution effects among patents, a distribution that was purely Paretian ignoring those effects became concave when the effects were introduced. Three sets of 200 observations were generated, assuming a Pareto distribution with $\alpha = 0.5$. The observations were then randomized, and observation values were assumed to be affected by the values of their nearest neighbors, mediated by some random number. For example, the value of observation $V_i$ became $V_i^* = V_i + (NORM_i\, V_{i+1}) + (NORM_i\, V_{i-1})$, where $NORM_i$ was a normally distributed random variable with mean zero and standard deviations of 0.2 and 0.33. Distinct concavities appeared in the double-log distribution of observations so transformed, e.g., as the largest original values appreciably augmented or reduced the value of smaller neighbors but were little affected by those neighbors' values. A simulation permitting only rectangularly distributed positive complementary effects induced only slight concavity.

backed startup companies and for the IPO companies (whose value gains occur at a later life cycle stage, since venture funds typically liquidate their positions shortly after the companies they have backed launch initial public offerings). The smallest share of total portfolio value, and hence the least skewness, is found for the new drug samples, perhaps because many drugs are protected by multiple patents but possibly also because of biases in the way the estimates were made.[4]

Second, substantial skewness is shown both by large fractions of total sample value attributable to the top 10 percent "winners" and by relatively flat Pareto function slope estimates (i.e., low absolute values of the slope parameter $\alpha$). The simple correlation between top 10 percent value shares and right-hand tail $\alpha$ estimates for the twelve samples is $-0.83$. Only for the two new pharmaceutical product samples, with the lowest top ten percent value shares, are the right-hand tail $\alpha$ estimates greater than unity – the threshold above which Pareto distributions have asymptotically finite means (see Mandelbrot, 1963; Scherer, 1998). For right-hand tail slope values below unity, reducing risk by forming ever-larger portfolios is problematic, since, in the pure Pareto case, means do not converge, following the weak law of large numbers, toward stable values with rising sample size. The near ubiquity of slope values below unity in absolute value reveals a high degree of skewness indeed.

## 3 Further insights from the IPO sample[5]

Because the stock market values of firms included in our IPO sample could be tracked monthly over a period of from nine to 13 years following the initial issuance of shares between January 1983 and December 1995, particularly rich insight into the evolution of high-technology enterprises' early economic success could be achieved. The 131 IPOs were believed to be an exhaustive sample of IPOs launched during 1983–86 that had been backed previously by venture capital funds and that operated in pre-specified high-technology industries, excluding health care provision. Monthly stock prices were obtained, initially from the Center for Research on Securities Prices files and, to fill numerous gaps, by manual search. No record of actual market transactions was found for 21 of the IPOs, and so the final sample included only 110 companies. Of those 110, 52 survived to the end of 1995, 23 disappeared by merger (only five of which outperformed the NASDAQ index prior to their acquisition), and 35 were delisted, often accompanied by bankruptcy. Our basic analytic approach was to set aside \$1000 for investment in each IPO, parking the funds in the NASDAQ index

[4] The estimates were based upon sales data, to which a non-varying quasi-rent margin fraction was applied before returns were discounted back to the date of product introduction. If the best-selling drugs (the "blockbusters") carried higher profit margins, which seems plausible, the degree of distribution skewness will be underestimated.

[5] Joerg Kukies was primarily responsible for the data collection on this sub-project. He was assisted by Jesus Viejo Gonzales, Christopher Choi, and Anne Sohns. Valuable leads to data sources were provided by Josh Lerner.

from January 1, 1983, until the time of IPO. Stock splits and dividends were tallied, and cash dividends (a fairly rare event) were reinvested (with no allowance for tax) at the time of payment. When companies exited by merger or delisting, the cash proceeds were parked in the NASDAQ index until December 1995.[6] The total value of the 110 investments as of December 31, 1995, was as follows:

| | |
|---|---|
| 52 surviving companies | \$ 417,002 |
| 23 acquired companies | 96,400 |
| 35 delisted companies | 21,178 |
| Total terminal value | \$ 534,580 |

Had the same initial \$110,000 been invested in the NASDAQ index in January 1983, investors would have had \$501,908. Thus, our sample of high-technology companies fared only slightly (and statistically insignificantly) better than the NASDAQ index generally (compare Brav and Gompers, 1997).

Figure 2 illustrates how company investment values evolved between June 1986 and December 1995. It is confined to ten companies, including the five most successful full-term survivors and five others selected randomly. The share values for the random choices cluster so closely in the \$0–2000 range that they are for the most part indistinguishable. Had one invested \$1000 in Adobe Systems, Concord Computing, or Amgen, on the other hand, one would have had shares valued at \$77,565, \$74,130, and \$55,980 respectively by the end of 1995.

Figure 3 shows how cross-sectional value distributions on doubly logarithmic coordinates evolved between 1987 and 1995. To the 52 full-term survivors from our main sample, one additional high-technology IPO backed by a venture fund but missed by our initial search – Microsoft – has been added for the analysis that follows.[7] As time advances, the distribution becomes increasingly skew, as shown by slope ($\alpha$) coefficients for the entire cross-sectional distribution falling in absolute value from 0.96 at the end of 1987 to 0.47 by the end of 1995. For the top 13 companies (i.e., the top 25 percent) at any given point, the slope values declined from 1.25 in 1987 to 0.95 in 1991 and 0.72 in 1995.

The value trajectories shown in Fig. 2 follow quite "noisy" paths between 1987 and 1995, among other things with frequent rank order changes. That what occurred approximated a random walk is shown by estimating

[6] The value of delisted companies is almost surely over-estimated, since it was based upon the last reported end-of-month value, and since later quotations probably fell, it is doubtful that investors could have liquidated and reinvested their stakes at the reported price.

The top ten percent company share tallied in Table 1 includes the reinvested proceeds from one company that merged on particularly attractive terms along with 12 full-term company values.

[7] Microsoft "went public" in March 1986. A parked 1983 investment of \$1000 accumulated to \$129,958 by December 1995. Adding it was important to our test for path dependence.

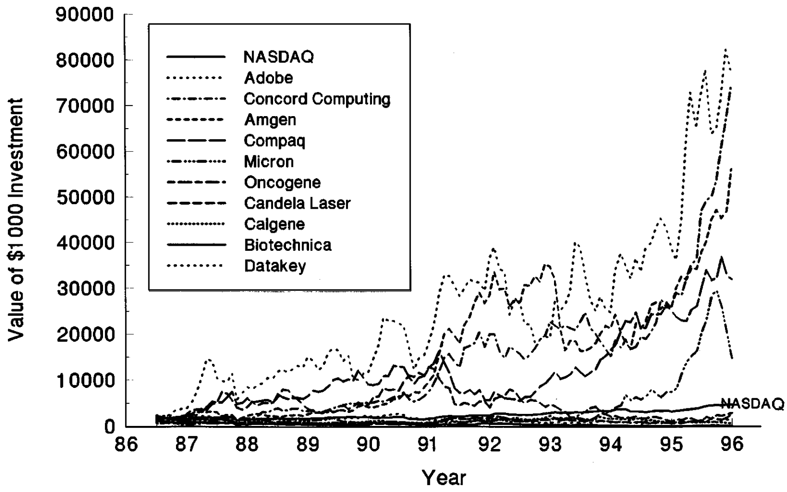


**Fig. 2.** Evolution of the value of $1000 investments in ten IPOS

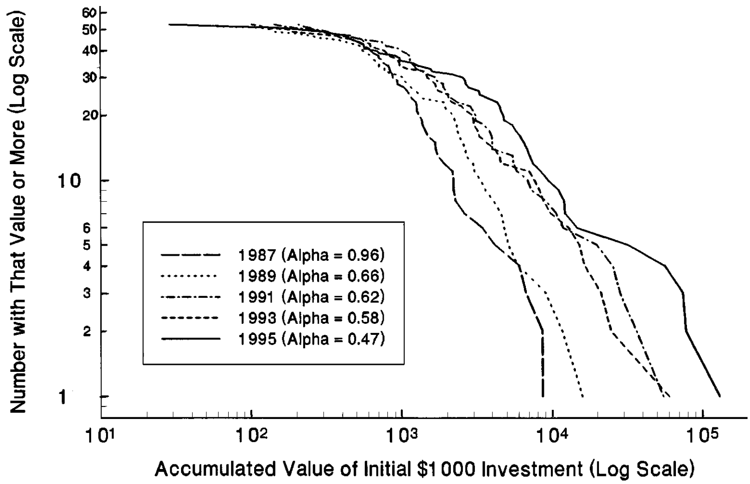


**Fig. 3.** Pareto plot of 53 IPO values, 1987–1995

for each of the 53 full-term companies a simple monthly market model regression of the form:

$$\delta P_t = a + \beta(\delta NASDAQ_t) + e_t \ , \qquad (2)$$

where $\delta P_t$ is the month-to-month percentage change in a company's stock value, $\delta NASDAQ_t$ is the corresponding change in the NASDAQ index, $e_t$ is a random error, and $\beta$ measures the stock's market risk. The mean value of $\beta$ was 1.16, showing that the new high-technology stocks on average exhibited more systematic risk than the overall NASDAQ market portfolio. The Durbin-Watson coefficients for the 53 price change regressions averaged 2.13, with a standard deviation of 0.22. Only nine of the 53 regressions had Durbin-Watson coefficients outside the 1.64 to 2.36 range, beyond which one might infer at the 95 percent confidence level negative autocorrelation (eight cases) or positive autocorrelation (one case).[8] Thus, the visual impression of value changes that exhibit much "white noise" is confirmed.

Since value evolutions governed by path dependence can also generate highly skew distributions (see DeVany and Walls, 1996), further tests for possible path dependence were conducted.

An intercept term $a$ in equation (1) significantly different from zero in equation (1) implies systematic upward or downward price drift. Only four of the 53 market model regressions had t-ratios on $a$ (all positive) of 2.00 or more. Thus, there was little evidence of significant upward drift, although four of the most successful companies (in descending t-ratio order, Microsoft, Concord Computing, Adobe, and Compaq) were exceptions.

As an additional test for path dependence, price levels (not percentage changes) during months 2 through 37 were regressed on an integer time variable. Extrapolated predictions from that regression were then added to regressions for the months remaining through December 1995 of price levels on the value of the NASDAQ index. In 34 out of 53 cases, addition of the extrapolated trend series variable led to $R^2$ increments significant in F-ratio tests at the 5 percent level. Among 19 cases with positive first-three-year price trends, the trend variable had significant positive explanatory effects for two computer-oriented companies (Microsoft and Compaq) and four biotech companies (Amgen, Chiron, Molecular Biosystems, and AL Pharmaceuticals) and significant negative effects (indicating a reversal of the early trend) in 13 cases. Among nine cases with negative first-three-year trends, a significant continuation of the downward trend was found in six cases and a significant reversal in three cases. Altogether, only six of the 53 regressions exhibit significant continuation of a positive initial trend consistent with the type of path dependence that would increase skewness over time. Since random walks produce what appear in hindsight to be cycles or trends, some incidence of correlated growth between early and later periods is almost inevitable. Thus, we are led to conclude that the incidence of path dependence was at best modest. It almost surely existed in the case of Microsoft because of the cumulative software lock-in effects resulting from the Microsoft operating system's early selection for IBM's personal computers.

## 4 Technical uncertainty vs. product market uncertainty

Initial public stock offerings are seldom made until one or more products have been shown by research and development to be technically feasible (although for pharmaceutical products, clinical testing often continues with funds raised through IPOs). The skew distributions that emerge following an IPO suggest that there must be substantial risks associated with consumers' (and regulators') response to new products.

[8] Negative autocorrelation implies non-random alternation between high and low changes; positive autocorrelation non-random sequences of positive or negative price changes.

Insight into this point can be sharpened by re-examining the findings from pioneering research by the late Edwin Mansfield and his students (1977, pp. 22–32). He and his colleagues secured information on the outcomes of individual R&D projects in 16 chemical, pharmaceutical, electronics, and petroleum corporations. They quantified three different success probabilities: (1) the probability that a project's technical goals will be met; (2) the probability that, given technical success, the resulting product or process will be commercialized; and (3) the probability that, given commercialization, the project yields a return on investment at least as high as the "hurdle rate" applied by the firm's decision-makers on investment projects. For all 16 companies combined, the average conditional probabilities were:

| | |
|---|---|
| Probability of technical success | 0.57 |
| Commercialization, given technical success | 0.65 |
| Financial success, given commercialization | 0.74 |

The firms' overall average success rate is found by multiplying the three conditional probabilities, i.e., $0.57 \times 0.65 \times 0.74 = 0.27$.

The question arises, could a three-stage stochastic process with these average success probabilities lead without more to the kind of skew reward distributions revealed by our data? As one source of insight, a Monte Carlo simulation model was tested. Each of 156 "firms" was assumed to make three successive stochastic draws from rectangular distributions, one with mean probability 0.57 and range 0.14 to 1.00, one with mean probability 0.65 and range 0.30 to 1.00, and one with mean probability 0.74 and range 0.48 to 1.00. The product of the three probability draws was multiplied by a uniform profit potential of 1000. The resulting distribution of profits was highly concave on doubly logarithmic coordinates, with a slope coefficient (corresponding to the Pareto $\alpha$) of $-1.21$ over all observations and $-4.33$ over the most profitable 39 observations. Top 10 percent outcomes accounted for 23.8 percent of total profits. Thus, the distribution was much less skew than any we have observed for cross sections of firms or individual inventions.[9] This could be because the simulation here abstracts from possible skewness in the size distribution of projects sampled by Mansfield et al. and perhaps also because the projects, mostly from large, well-established corporations, were less risky on average than those covered by our various surveys.

## 5 Gibrat processes

When the market value of a representative company grows or declines randomly from month to month (or year to year) by a stochastic multiplier, some version of a Gibrat process may be operating (see Gibrat, 1931; Ijiri

[9] Even less skewness emerged when the sampling distributions around Mansfield's three means were symmetrically normal (and bounded above at unity) rather than rectangular, e.g., with a Paretian slope of 6.65 for the most profitable 25 percent of sample observations.

and Simon, 1977). With the simplest Gibrat process, an initial value $V_0$ is altered sequentially over time by a string of random and independent multipliers $\epsilon_i, i = 1, T$, so that in the final period the value $V_T$ is:

$$V_T = V_0 \ \epsilon_1 \ldots \epsilon_i \ldots \epsilon_T \ . \tag{3}$$

Taking logarithms, one obtains:

$$\ln V_T = \ln V_0 + \ln \epsilon_1 + \cdots + \ln \ \epsilon_i + \cdots + \ln \ \epsilon_T \ . \tag{4}$$

Under the central limit theorem, the sum of a string of random variables is asymptotically normally distributed. Thus, following what is called Gibrat's Law, ln $V_T$ should be normally distributed, given sufficiently large T.

The weight of evidence from our innovation reward samples is that a log normal distribution fits the data better than plausible alternatives. The time series evidence from our IPO study suggests that a random walk of the Gibrat type operates. The two pieces of evidence together point toward some kind of Gibrat process as a source of differing innovation payoffs, and in particular, as the source of the observed skew distribution of market values resulting from equal-sized investments in individual high-technology IPOs.

We must nevertheless inquire whether there are sufficiently many independent random shocks $\epsilon_i$ to satisfy the "large numbers" requirement for asymptotic log normality. From the estimation of equation (2) above for 53 companies, the mean standard deviation of month-to-month stock value growth (in percentage terms) was found to be 18.61 percent, with a range of from 10.1 to 30.8 percentage points. This reveals a value evolution process that was on average quite noisy.

To see whether a strict Gibrat process with monthly random shocks of that magnitude would yield the observed IPO value distributions, a Monte Carlo experiment was conducted. An initial investment of $1000 was made in 100 hypothetical IPOs. Each investment thereupon grew on average by 1.03 percent per month, plus or minus a normally distributed error term with mean zero and standard deviation of 18.6 percent, for a total of 120 months.[10] The shares of the ten value leaders were computed and Pareto $\alpha$ slope values were estimated for all 100 companies and the 25 value leaders for each two-year interval, with results as follows:

| | Top 10 Value Share | Pareto $\alpha$ $N = 100$ | Pareto $\alpha$ Top 25 |
|---|---|---|---|
| After 24 months | 31.3% | 0.93 | 2.29 |
| After 48 months | 48.7% | 0.63 | 1.23 |
| After 72 months | 57.6% | 0.51 | 1.07 |
| After 96 months | 62.0% | 0.47 | 1.09 |
| After 120 months | 67.0% | 0.41 | 1.00 |

Plainly, considerable skewness emerges as the Gibrat process unfolds. At the end of ten years, the firms comprising the top ten percent of the value

[10] I.e., the monthly multiplier was $1.0103 + \epsilon$ , where $\epsilon$ was normally distributed with mean zero and standard deviation 0.186.

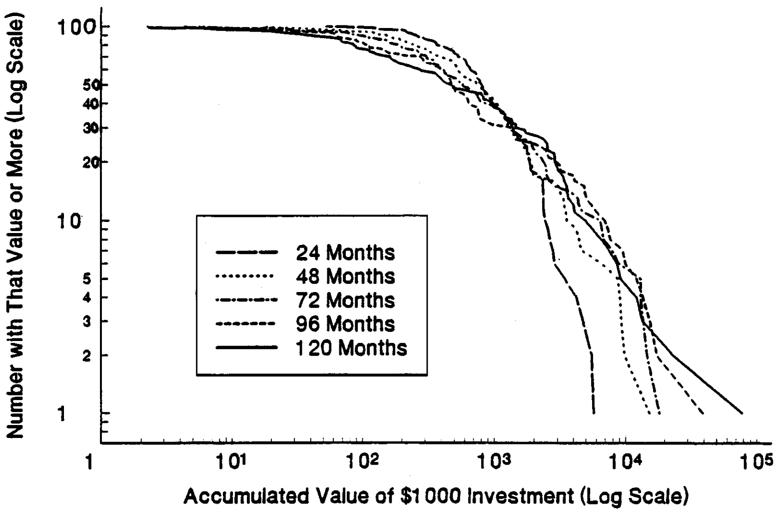


**Fig. 4.** Plot of Gibrat experiment outcomes on Pareto coordinates

distribution account for 67 percent of total sample value, compared to 62 percent for our IPO sample in 1995 when the universe is considered to be all 110 firms or 61 percent (for five firms) when only the 52 full-term survivors are included. The evolution of full-sample $\alpha$ estimates for the Monte Carlo experiment tracks closely the experience of the actual IPO companies. Slopes for the most successful 25 percent differ somewhat more, e.g., 0.72 for the actually observed full-term IPOs in 1995, compared to 1.00 after the experiment has run its course. Considerable similarity (but not identity, especially in the high-value tail) of distribution evolution patterns is suggested by comparing Fig. 3, which tracks 53 IPOs at five time benchmarks, and Fig. 4, which summarizes the Gibrat experiment's cross-sectional distribution.

If the distribution of company values resulting from the experiment were in fact log normal, the skewness parameter $\sqrt{\beta_1}$ would be insignificantly different from zero and the kurtosis parameter $\beta_2$ would be insignificantly different from 3.0 when logarithms are taken of the resulting value observations.[11] For the Gibrat experiment and 53 actual IPOs (including Microsoft), the measured parameters at roughly comparable time points were as follows:

| | Gibrat experiment | | 53 actual IPOs | |
|---|---|---|---|---|
| | Skewness | Kurtosis | Skewness | Kurtosis |
| 6 years, 1991 | −0.27 | 3.33 | 0.27 | 2.85 |
| 10 years, 1995 | −0.22 | 2.91 | −0.03 | 3.09 |

[11] Where $\mu_k$ is the k[th] distribution moment, the coefficient of skewness is defined to be $\sqrt{\beta_1} = \mu_3/\mu_2^{3/2}$. The coefficient of kurtosis is defined as $\beta_2 = \mu_4/\mu_2^2$.

**Table 2.** Evolution of individual firm values in the Gibrat experiment

| Firm rank at month 48 | Firm value at month | | | | Firm rank at month 120 |
|---|---|---|---|---|---|
| | 48 | 72 | 96 | 120 | |
| 1 | 15,150 | 18,064 | 15,270 | 22,934 | 2 |
| 2 | 9,748 | 13,489 | 7,555 | 1,715 | 28 |
| 3 | 9,181 | 14,454 | 38,848 | 76,848 | 1 |
| 4 | 8,987 | 12,915 | 13,475 | 5,528 | 9 |
| 5 | 8,610 | 12,931 | 9,850 | 8,437 | 6 |
| 6 | 6,455 | 6,502 | 5,912 | 3,517 | 14 |
| 7 | 4,513 | 1,453 | 758 | 1,433 | 30 |
| 8 | 4,369 | 2,221 | 6,869 | 3,256 | 17 |
| 9 | 4,150 | 7,846 | 3,123 | 2,801 | 22 |
| 10 | 3,493 | 2,464 | 4,803 | 1,034 | 38 |
| 11 | 3,456 | 6,312 | 1,361 | 933 | 40 |
| 12 | 3,342 | 1,289 | 2,564 | 4,022 | 11 |
| 13 | 3,292 | 2,440 | 1,429 | 12,016 | 4 |
| 14 | 3,101 | 1,565 | 571 | 459 | 50 |
| 15 | 3,022 | 8,758 | 2,404 | 2,833 | 20 |
| 16 | 2,702 | 843 | 1,381 | 1,090 | 37 |
| 17 | 1,951 | 4,228 | 4,743 | 2,098 | 27 |
| 18 | 1,932 | 4,661 | 4,964 | 3,501 | 15 |
| 19 | 1,899 | 132 | 64 | 66 | 86 |
| 20 | 1,833 | 1,323 | 743 | 917 | 41 |
| 21 | 1,830 | 6,754 | 9,400 | 4,847 | 10 |
| 22 | 1,821 | 388 | 962 | 3,583 | 13 |
| 23 | 1,663 | 1,879 | 7,245 | 9,055 | 5 |
| 24 | 1,655 | 1,189 | 2,899 | 2,920 | 19 |
| 25 | 1,655 | 975 | 748 | 241 | 61 |

For both the experiment and the actual IPO data, the measured skewness and kurtosis coefficients lie well within the 95 percent confidence bounds consistent with normality of the underlying distribution (see D'Agostino and Stephens, 1986, pp. 376–385). Thus, we conclude that, given the observed month-to-month changes in high-technology IPO market values, an approximately log normal cross sectional distribution of company common stock values could and plausibly did emerge after six to ten years of random growth in the marketplace.

Table 2 focuses on the 25 most valuable Gibrat experiment companies after 48 months have transpired, ranked in descending market value order, and shows how the market values of those companies evolved in the experiment's remaining years. Three of the top four companies at the 120-month mark had already emerged among the top 13 at the 48-month mark. Path dependence is suggested, when in fact what happened is that some companies experienced an early run of good fortune, after which there was (by the assumption of independent random growth draws in any given month) an equal probability that they would continue to grow as rapidly as other firms. The Pearsonian correlations among full-sample company values at various experimental time points were as follows:

| | 24 months | 48 months | 72 months | 96 months |
|---|---|---|---|---|
| 48 months | .500 | 1.000 | .908 | .654 |
| 72 months | .432 | .908 | 1.000 | .735 |
| 96 months | .276 | .654 | .735 | 1.000 |
| 120 months | .197 | .537 | .587 | .852 |

All inter-correlations exceed 95 percent statistical significance thresholds; and the correlations between distributions separated by two years increase systematically after the first 24-month sorting-out has occurred. This facet of the experiment reveals that one must be wary of inferring path dependence causation after-the-fact when an early growth spurt is followed by continued rank leadership (see Feller, 1957, pp. 77–85; Scherer and Ross, 1990, pp. 141–142).

## 6 A more structured model

It would be premature, however, to conclude that the skew distributions observed in our various data sets arise solely from Gibrat processes of the simple form modelled thus far. The closest consonance has been found to the stock market performance of companies during the ten years after they have floated initial public offerings. Because of disclosure lags, the connection between the underlying events determining company profitability and common stock prices may be erratic. Many of the IPOs included within our sample sold numerous products, not always closely related, during the later years covered by our analysis, but most of our data cover single well-defined inventions or innovations. And at bottom, the Gibrat process is a "black box" under which random events of unspecified origin affect company value. We ask therefore, is it possible to identify stochastic models that yield invention and innovation reward size distributions like those we have observed, but that have a richer underlying structure reflecting the kinds of technical and market events known to occur when new products and processes are developed and commercialized?

To address this question, a much richer stochastic model was developed and tested in 20 initial variants covering some 2,400 innovation histories involving approximately 10,000 simulated company participants.[12] Extensions were then tested in 20 further runs. The model is essentially Gibratian in flavor, with a series of random draws affecting innovating and imitating firms' profit realizations multiplicatively. However, the stochastic events have been structured to conform as closely as possible to what has actually been recorded in numerous histories of patented inventions and innovations. By way of introduction, a profit potential of varying size comes into existence, firms compete through research and development to tap the profit potential, some succeed more rapidly than others and enter the market in a stochastic order that determines market shares, price competition may or

[12] All simulations were performed using FORTRAN programs modified for each set of varying assumptions.

may not appear to erode profit realizations, and new technologies emerge to render the subject technology obsolete.

### 6.1 The model parameters

*Profit potential.* The logical starting point for any realistic model is an annual profit potential that can be tapped by firms through innovation. Two quite different assumptions were modelled. In eight first-stage cases, the $i^{th}$ innovation's potential was determined by random draws from a rectangular distribution over the range 0–1. Each sampled variate was multiplied by \$1000 to determine the annual profit (or more exactly, quasi-rent) potential. However, draws of less than 0.1 were multiplied by 10 so that only about one case in 100 had a profit potential of less than \$100. Thus, most of the values were distributed rectangularly over the range \$100–1000. In 12 cases, random draws with replacement were made from a skew distribution replicating the distribution of {value added less payroll outlays} reported for four-digit manufacturing industries in the 1987 U.S. *Census of Manufactures*. Fifty "NEC" (not elsewhere classified) industries and 109 other industries below a specified size threshold were deleted, leaving 298. After rescaling, the range of annual skew quasi-rent potentials was from \$40 to \$2590, with mean of 210. The distribution was nearly log normal, with skewness coefficient of 1.00 and kurtosis coefficient of 3.56. However, it was less skew than any of the innovation reward distributions from our empirical samples; the top 10 percent of observations by number accounted for only 45 percent of total sample potential profits, and for the top 25 percent, the Pareto slope coefficient was 1.41.

*Number of competing firms.* Two approaches, exogenous and endogenous, were used to determine the number of firms seeking to exploit the profit potential. With exogenous entry, the number of firms was independent of quasi-rent potential, being set as 5.0 plus a normally distributed random variable with mean zero and standard deviation of 2.0. Values below 1.0 were truncated at 1.0 and values above 12.0 were truncated at 12.0. With endogenous entry, the number of firms N varied stochastically with the size of the innovation's quasi-rent potential. In rectangular potential cases, $N = (1.33 + .00667 \text{ POTENTIAL})\Phi$, where $\Phi = (1 + \sigma/4)$, $\sigma$ being a normally distributed variate with mean 0 and standard deviation 1. Figure 5 illustrates the distribution of firm numbers as a function of the potential from simulation run 13. The number of firms per thousand dollars of quasi-rent potential increases less than proportionately with the potential, i.e., with an elasticity of roughly 0.75, reflecting the tendency (not modelled explicitly) for R&D costs per firm to rise when more firms are contending for a prize (see e.g. Reinganum, 1989). In the skew potential case, $N = (1.75 + .00625 \text{ POTENTIAL})\Phi$, where $\Phi = (1 + \sigma/4)$.

*Date and order of entry.* The dates T at which firms begin to tap the profit potential, and hence the order of entry, were determined, consistent with

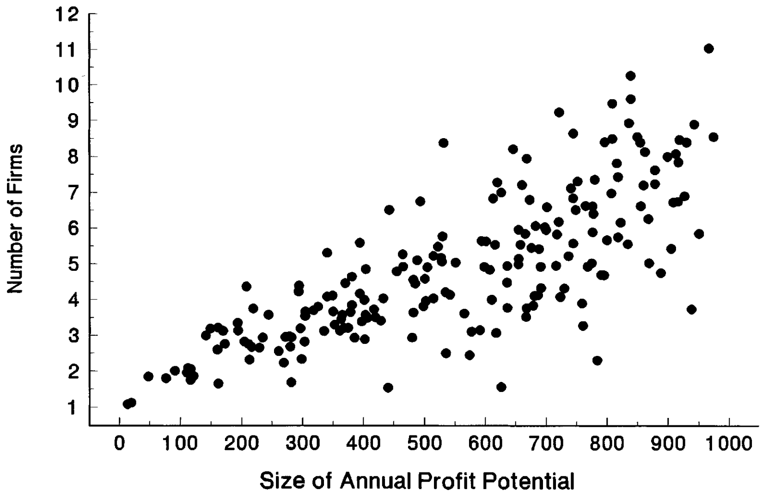


**Fig. 5.** Illustration of endogenously determined firm numbers

many models of research and development rivalry (see e.g. Reinganum, 1989, pp. 855–856), as the expected value of a hazard function $T = 1/h$, where h is 0.667 times a rectangularly distributed random variate over the interval 0–1. Before addition of one-half year for production startup, rounding, and truncation, T lay in a 90 percent confidence interval from 1.58 years to 30 years. Fractional values were rounded downward to the nearest integer.

*First mover advantages and market shares.* Firms were ranked according to their time of innovation (before rounding), and the earliest entrants enjoyed first mover advantages whose strength depended upon a random draw from a rectangular distribution. For the first 20 runs, strong first mover advantages applied with probability 0.1; middling advantages with probability 0.5; and weak advantages with probability 0.4. In the strongest first mover case, the market (and hence profit potential) share of a firm ranked $j + 1$ was 36.8 percent of firm j's share; with intermediate first mover advantages, 60.6 percent; and with the weakest advantages, 77.9 percent. Market shares were normalized to sum to unity in each time period. No changes in entry ranks were effected during any given 20-year run for the first 20 runs, although all firms' shares fell as additional firms entered.

*Profit potential penetration rate.* Most models of R&D rivalry have assumed that the full profit potential is appropriated as soon as at least one firm completes its R & D (for an exception, see Scherer, 1967). This is highly unrealistic; it takes time to penetrate a new product's potential market. Thus, in all runs, the rate of quasi-rent potential penetration for market participants following innovation by the first firm was determined by the function $(1 - e^{-pt})$, where p is a random variable drawn from a rectangular distribution bounded between 0.3 and 0.75 and t is a running time variable initialized at the first entrant's R&D completion date. Thus, if p = 0.5, the fraction of quasi-rents appropriated is 39.3 percent in the first year of sales, 63.2 percent in the second year, 91.8 percent in the fifth year, etc.

*Market structure and price competition.* Price competition among market participants can deplete the market's profit potential. For each year in which firms participated in an innovation market, a Herfindahl-Hirschman index was calculated. For HHI values greater than 0.333, joint profit maximization was assumed, with no depletion of the quasi-rent potential. For lower HHI values, firms were assumed to engage in Cournot rivalry, with the quasi-rent potential being dissipated more, the lower the HHI index. Thus, with a numbers-equivalent structural index (1/HHI = 0.250) implying four equal-size firms, 43.8 percent of the maximum profit potential is appropriated; with a numbers-equivalent of 12, 15.6 percent of the potential is appropriated.

*Obsolescence.* Paralleling the firm entry approach, the impact of competition from superior outside goods (Schumpeter's "creative destruction") was modelled in two ways. In the first ten runs, obsolescence was stochastically exogenous. From year six on, random draws were made from a 0–1 rectangular distribution, and in the first year with a value less than or equal to 0.1, obsolescence commenced. From that time on, the quasi-rent pool was eroded by $e^{-0.15\tau}$, where $\tau$ is the number of years elapsing from the onset of outside goods competition. On average, erosion started in year 11, but with considerable dispersion.[13] In the next ten runs and later variants, obsolescence was endogenous. Both the date at which outside competition commenced (but not earlier than year 3) and the rate at which it eroded the subject innovation's profitability were assumed to rise stochastically with the size of the potential (undissipated) profit pool. Figure 6a,b illustrates for Monte Carlo run 13 the dependence of starting years and erosion rates (vertical axis) upon the size of the quasi-rent potential (horizontal axis).

*Discount rate.* In all runs, individual firm quasi-rents were discounted to present value as of year zero at a 10 percent real interest rate.

### *6.2 Results*

Table 3 summarizes the results of the first 20 Monte Carlo experiments embodying various combinations of the assumptions described above. In each run, 120 innovation histories were simulated. Some innovations, especially those with only one or two participating firms, dropped out because no firm completed its R&D project in time to have a discounted

[13] No explicit assumption is made concerning the source of outside innovations. They could come from outside firms or from one or more of the modelled firms.

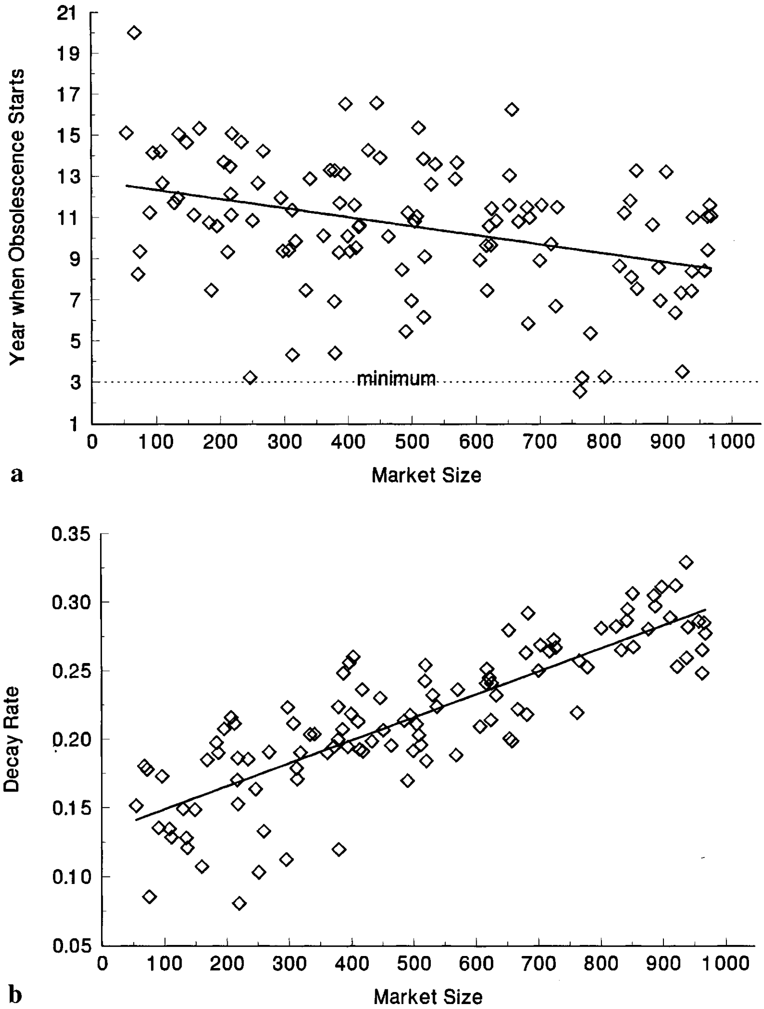


**Fig. 6. a** Endogenous determination of obsolescence starting year. **b** Endogenous determination of obsolescence decay rate

quasi-rent present value of 5 or more. (The maximum DPV for a first mover was 11,144; the mean 783.)

In Table 4, the discounted present value of quasi-rents for the first firm into the market is taken as dependent variable in an ordinary least squares regression with key experimental parameters as explanatory variables. The sample comprises 940 observations from eight runs, each representing the first run embodying a particular constellation of assumptions. The most powerful explanatory variable by a substantial margin, with a partial $r^2$ of 0.500, is the size of the quasi-rent potential firms competed to tap. The second-highest t-ratio (and partial $r^2$) is for the Herfindahl-Hirschman index (measured at midpoint year 10), in part because higher HHI values imply higher market shares for leading firms and also because values below

**Table 3.** Summary of the first 20 Monte Carlo experiments

| Run | Potential | Entry | Obsol. | Parameters for first firm to innovate | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | No. | Top 10% | $\alpha_{ALL}$ | $\alpha_{25\%}$ | Skew | Kurtosis |
| 1 | Rect. | Exog. | Exog. | 120 | 27.3% | 1.02 | 3.08 | −0.15 | 2.44 |
| 2 | Rect. | Exog. | Exog. | 118 | 27.1% | 0.90 | 3.19 | −0.40 | 2.50 |
| 3 | Rect. | Endog. | Exog. | 119 | 24.9% | 1.56 | 2.77 | 0.29 | 2.50 |
| 4 | Rect. | Endog. | Exog. | 118 | 23.5% | 1.39 | 3.41 | −0.35 | 4.02 |
| 5 | Skew | Exog. | Exog. | 118 | 54.5% | 0.80 | 1.10 | 0.61 | 3.38 |
| 6 | Skew | Exog. | Exog. | 120 | 44.9% | 0.87 | 1.31 | 0.29 | 3.57 |
| 7 | Skew | Exog. | Exog. | 118 | 50.5% | 0.85 | 1.12 | 0.41 | 3.69 |
| 8 | Skew | Endog. | Exog. | 114 | 36.0% | 1.11 | 1.53 | 0.38 | 4.24 |
| 9 | Skew | Endog. | Exog. | 115 | 40.2% | 1.11 | 1.49 | 0.66 | 3.94 |
| 10 | Skew | Endog. | Exog. | 111 | 28.8% | 1.27 | 2.67 | 0.04 | 3.61 |
| 11 | Rect. | Exog. | Endog. | 118 | 27.0% | 0.88 | 2.82 | −0.95 | 4.91 |
| 12 | Rect. | Exog. | Endog. | 118 | 24.2% | 0.90 | 4.23 | −0.48 | 2.32 |
| 13 | Rect. | Endog. | Endog. | 117 | 23.8% | 1.61 | 2.87 | −0.06 | 3.56 |
| 14 | Rect. | Endog. | Endog. | 118 | 22.8% | 1.63 | 3.12 | 0.01 | 2.95 |
| 15 | Skew | Exog. | Endog. | 118 | 42.1% | 0.85 | 1.61 | 0.35 | 2.44 |
| 16 | Skew | Exog. | Endog. | 118 | 31.8% | 0.97 | 2.18 | −0.14 | 2.62 |
| 17 | Skew | Exog. | Endog. | 118 | 44.8% | 0.88 | 1.37 | 0.38 | 2.84 |
| 18 | Skew | Endog. | Endog. | 116 | 28.9% | 1.45 | 2.19 | 0.49 | 3.21 |
| 19 | Skew | Endog. | Endog. | 115 | 25.7% | 1.47 | 2.50 | 0.06 | 3.20 |
| 20 | Skew | Endog. | Endog. | 115 | 30.0% | 1.24 | 2.26 | 0.10 | 3.14 |

Codes: Rect. = rectangular profit potential distribution, Endog. = endogenous, Exog. = exogenous.

0.333 precipitated profit-eroding price competition.[14] HHI values at year 10 fell below the 0.333 threshold in 42 percent of the cases. Thirteen percent of the first movers remained pure monopolists at that milestone. The third-highest t-ratio emerges for the year at which obsolescence began to set in. A variable measuring the erosion rate due to obsolescence is also significant. Together, the two variables imply an important role for Schumpeterian creative destruction. Because in most cases some firm's hazard function outcome let the firm complete its innovation and enter the market at year two or three, variations in initial starting dates had only a modest impact on quasi-rent values.[15] A year's delay implies a sacrifice amounting to 12.5

[14] The strength of first mover advantages was highly correlated with HHI indices and had little incremental explanatory power once HHI indices were included in the regression. A regression relating the HHI values to the number of firms and the strength of first mover advantages was as follows:

$$\text{HHI} = \underset{(81.42)}{0.979} - \underset{(33.59)}{0.067}\ \text{NFIRMS} - \underset{(24.99)}{0.120}\ \text{FIRSTMOVE},\ R^2 = 0.716;$$

where NFIRMS is the number of firms participating in an innovation rivalry and the FIRSTMOVE index is scaled at 1 for the largest market share differential between first-moving and lower-ranked firms and 3 for the smallest. For pure monopoly cases (NFIRMS = 1), FIRSTMOVE was scaled at zero.

[15] The mean first-completion date, excluding values above 20, was at year 2.67; the median at year 2.30; the mode at year 2.01.

**Table 4.** Regression of firm 1's discounted present value on model parameters

| Variable | Coefficient | t-ratio | Mean |
|---|---|---|---|
| Intercept | −465.48 | 3.72 | – |
| Profit potential | 1.79 | 30.51 | 391 |
| Starting date | −98.07 | 5.44 | 2.67 |
| Penetration rate | 431.33 | 3.29 | 0.53 |
| Obsolescence date | 32.48 | 7.17 | 11.12 |
| Obsolescence rate | −1591.94 | 4.20 | 0.17 |
| HHI index, year 10 | 1514.21 | 17.72 | 0.45 |
| Endogenous entry | −166.62 | 4.83 | 0.50 |
| Skew distribution | −4.98 | 0.14 | 0.50 |

N = 940; mean dependent variable value = 783; standard deviation = 808; $R^2$ = 0.607

percent of average first mover discounted quasi-rents for the full sample. With endogenous entry, quasi-rents are reduced on average by 21 percent relative to the full-sample mean value, all else equal. Skew profit potentials reduce first-mover discounted present values by about one quarter relative to full-sample means, but with considerable variance.

We return to Table 3 to inquire which experimental conditions led to quasi-rent size distributions most closely consonant with those we have observed empirically. For each run, innovations were ranked in descending order of the quasi-rents realized by the first mover. Examining first the shares of total discounted quasi-rents in any given run appropriated by the most profitable ten percent of the first movers, we find results approaching those observed empirically only in runs 5–7, 15, and 17 – all with skew profit potentials and exogenous entry. In only two cases (runs 5 and 7) did the leading firms secure 50 percent or more of total sample quasi-rents, i.e., roughly the same share as the drug innovators, and five to eleven percentage points less than the high-technology venture investments and IPOs (compare Table 1). Those five cases, and especially the skew potential, exogenous entry cases, are the only ones with Pareto slope coefficients for firms in the right-hand quarter of the size distribution anywhere near those observed empirically – i.e, in the range of 1.10 to 1.37. None of the simulations yielded right-hand tail slope coefficients as low as those observed for the venture startups and IPOs. Both endogenous entry and endogenous obsolescence (Schumpeterian creative destruction) reduce leading firm quasi-rent shares and steepen distribution function tail slopes.

When logarithms are taken of discounted first-mover quasi-rents, the skewness coefficients for their size distributions scatter fairly tightly about the zero value consistent with a log normal distribution.[16] The average skewness coefficient over 20 runs is +0.08. The distributions exhibit a slight tendency toward higher kurtosis coefficients than the 3.0 value that would prevail if the size distributions were strictly log normal (last column of Table 3).

[16] A 90 percent confidence interval around the zero expectation extends to plus or minus 0.36 with 118 observations.

None of the 20 simulations generated a concentration of quasi-rents as high as was observed empirically for individual patents or related patent bundles. To explore whether some important stochastic element was overlooked, 20 additional runs embodying three changes were computed.

One possible limitation may have been the assumption that first movers' relative market positions persisted throughout the experiment, with no displacement from later movers offering marginally superior products or sustaining more effective marketing campaigns. To pursue this lead, runs 21 through 28 accepted the basic parameter assumptions of runs 13–14 and 18–20, but added a variant of Gibrat-like random growth. For each year after at least one firm completed its R&D project, the individual firm market shares determined on the basis of first-mover principles were subjected to a random annual multiplier $(1 + \mu)$, where $\mu$ was normally distributed with mean zero and standard deviation 0.1. The altered market shares were then normalized to sum to unity. Random growth increments or shortfalls experienced in any given year were carried forward into later years, so the Gibrat effects were cumulative. The results are presented in the first panel of Table 5. If anything, less skewness emerges than in otherwise similar runs 13–14 and 18–20. The top ten percent shares are lower on average, especially for the runs tapping a skew quasi-rent potential, and the right-hand tail distribution slopes are steeper on average. The highest first mover share and the least steep slope are found in each category among the simulation runs without random market share changes. Evidently, Gibrat processes do not necessarily increase skewness when high concentration materializes at the outset as a result of first mover advantages and/or random sampling of quasi-rent potentials. For the most dominant firms in such cases, the probability of *further* market share gains through particularly lucky random draws is small.

In some of the Gibrat runs, the market shares of first movers were eventually exceeded by those of the second or third movers. Despite discounting, later movers' total discounted quasi-rents exceeded those of first movers in from one to eight cases per run. Because of such rank changes and also because firms that are not the first to complete their R&D projects can make significant innovative contributions, quasi-rents were merged and arrayed in seven cases for both the first and second movers, and in two illustrative cases, for the first three entrants. The results are summarized in the second and third panels of Table 5. The total quasi-rent share of first- and second-moving firms ranked among the combined top ten percent increased in every case relative to the top ten percent shares for first movers only. For upper quartile slopes, on the other hand, the differences between first-mover-only and combined first and second mover distributions were erratic.

In runs 1–20, the probability that first movers would achieve the largest proportional disparity between their market shares and those of later movers was 0.1. Six additional runs, numbers 29–34, increased the probability of the most extreme disparity to 0.3, with assignment to the intermediate and weakest first mover cases each accorded probabilities of 0.35. These stronger first mover assumptions were simulated for analogues

**Table 5.** Summary of further simulation run results

| Run | Potential | Entry | Obsol. | Parameters for first firm to innovate | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | No. | Top 10% | $\alpha_{ALL}$ | $\alpha_{25\%}$ | Skew | Kurtosis |
| With Gibrat growth process added | | | | | | | | | |
| 21G | Rect. | Endog. | Endog. | 119 | 22.2% | 1.37 | 3.28 | −13.51 | 152 |
| 22G | Rect. | Endog. | Endog. | 119 | 21.5% | 1.07 | 4.04 | −2.58 | 18.9 |
| 23G | Rect. | Endog. | Endog. | 120 | 25.1% | 1.08 | 3.35 | −0.71 | 4.35 |
| 24G | Rect. | Endog. | Endog. | 120 | 22.4% | 1.28 | 3.85 | −0.40 | 3.22 |
| 25G | Skew | Exog. | Endog. | 120 | 35.5% | 0.95 | 2.00 | 0.16 | 2.64 |
| 26G | Skew | Exog. | Endog. | 120 | 40.1% | 0.89 | 1.60 | 0.21 | 2.83 |
| 27G | Skew | Exog. | Endog. | 120 | 38.9% | 0.95 | 1.54 | 0.14 | 2.92 |
| 28G | Skew | Exog. | Endog. | 119 | 32.4% | 0.95 | 2.11 | −0.08 | 2.56 |
| First and second entrants | | | | | | | | | |
| 5–2 | Skew | Exog. | Exog. | 229 | 56.1% | 0.78 | 1.15 | 0.49 | 3.64 |
| 13–2 | Rect. | Endog. | Endog. | 226 | 25.4% | 1.22 | 2.83 | −1.16 | 9.61 |
| 14–2 | Rect. | Endog. | Endog. | 230 | 25.5% | 1.10 | 2.71 | −1.68 | 11.92 |
| 17–2 | Skew | Exog. | Endog. | 228 | 45.6% | 0.90 | 1.41 | 0.46 | 2.89 |
| 25–2 | Skew | Exog. | Endog. | 233 | 39.1% | 0.97 | 1.96 | 0.37 | 2.79 |
| 26–2 | Skew | Exog. | Endog. | 231 | 42.1% | 0.88 | 1.61 | 0.16 | 3.22 |
| 27–2 | Skew | Exog. | Endog. | 230 | 40.0% | 0.91 | 1.62 | 0.13 | 3.02 |
| First three entrants | | | | | | | | | |
| 14–3 | Rect. | Endog. | Endog. | 317 | 30.0% | 1.00 | 2.65 | −1.02 | 6.62 |
| 17–3 | Skew | Exog. | Endog. | 320 | 48.6% | 0.81 | 1.38 | 0.28 | 3.20 |
| Stronger first mover advantages | | | | | | | | | |
| 29 | Rect. | Exog. | Endog. | 118 | 25.4% | 0.96 | 3.44 | −0.55 | 3.08 |
| 30 | Rect. | Exog. | Endog. | 119 | 24.9% | 1.04 | 3.53 | −0.41 | 2.66 |
| 31 | Skew | Exog. | Endog. | 118 | 39.1% | 0.84 | 1.77 | 0.03 | 2.75 |
| 32 | Skew | Exog. | Endog. | 119 | 37.0 | 0.97 | 1.68 | 0.16 | 3.06 |
| 33G | Skew | Exog. | Endog. | 119 | 32.5 | 0.98 | 2.23 | 0.00 | 2.51 |
| 34G | Skew | Exog. | Endog. | 120 | 37.6 | 0.79 | 1.86 | 0.01 | 2.40 |
| PIMS market share determination | | | | | | | | | |
| 35 | Rect. | Endog. | Endog. | 116 | 21.0 | 1.66 | 4.00 | −0.27 | 2.88 |
| 36 | Rect. | Endog. | Endog. | 119 | 21.2 | 1.38 | 3.82 | −0.90 | 5.60 |
| 37 | Rect. | Endog. | Endog. | 118 | 19.7 | 1.29 | 4.87 | −1.19 | 6.12 |
| 38 | Skew | Endog. | Endog. | 113 | 27.9 | 1.25 | 2.10 | −0.12 | 3.69 |
| 39 | Skew | Endog. | Endog. | 115 | 26.6 | 1.30 | 2.98 | 0.04 | 2.65 |
| 40 | Skew | Endog. | Endog. | 116 | 29.8 | 1.25 | 2.48 | 0.22 | 3.13 |
| 35–3 | Rect. | Endog. | Endog. | 313 | 28.6 | 0.87 | 3.00 | −0.70 | 4.07 |
| 38–2 | Skew | Endog. | Endog. | 210 | 32.3 | 0.83 | 2.25 | −0.48 | 3.42 |

of runs 11–12, 15–17, and 25–28. The results are presented in the fourth panel of Table 5. Surprisingly, there was very little difference in top ten percent quasi-rent shares or upper quartile slopes between the averages for runs 29–34, with stronger first mover advantages, and their earlier counterparts.

Finally, a referee suggested that an implausibly high number of innovation participants resulted from the endogenous innovation race participation equation and that the distribution of market shares might be more realistically determined by using an empirical formula derived from PIMS (Profit Impact of Market Strategy) studies:[17]

$$MS_j = \left(\frac{1}{n}\right)\sum_{i=j}^{n}\left(\frac{1}{i}\right); \quad j = 1, \ldots, n \ , \tag{5}$$

where $MS_j$ is the market share of the jth firm among a total of n firms. To explore these possibilities, the coefficients on POTENTIAL in the endogenous firm number determining equations were reduced by 30 percent for the rectangular case and 20 percent for the skew case, leading to decreases in both the maximum and average number of firms.[18] Market shares were determined precisely according to equation (5) with probability 0.5. To allow differences in the strength of first mover advantages, market shares generated by equation (5) were raised to the exponent 0.8 with probability 0.25 (reducing the asymmetry of normalized market shares) and to the exponent 1.25 with probability 0.25 (increasing the asymmetry of shares). Other assumptions were as in simulation runs 13–14 and 18–20. The last panel of Table 5 reports the first mover quasi-rent results of six simulation runs embodying these changes, three for rectangular profit potential cases and three for skew potential cases. Data for two representative runs including the first, second, and third entrants are also presented. The runs for first movers alone exhibit less skewness than their original counterparts. The three-firm simulations (including those not reported in Table 5) had full-sample $\alpha$ values as low as those obtained in earlier simulations, but right-hand tail $\alpha$ values were higher and top ten percent share values were lower. Thus, the revised assumptions did not provide a superior approximation to observed real-world innovation reward size distributions.

Figure 7a–d graphs on Pareto coordinates four representative size distributions from the first 20 simulation runs – three with distribution parameter outcomes most closely approximating those observed empirically and one (from run 13) with one of the lowest top ten percent first mover quasi-rent shares. All are concave to the origin, as were all distributions constructed from actual data. None exhibits an outlying value some distance removed from the main cluster of observations, as emerged with the detailed German patent value data (Fig. 1) and also the Harvard University patent royalty and drug profit distributions reported in Scherer (1998). However, such outliers appeared in earlier trial simulations (not explicitly reported here) and cannot be ruled out for alternative random draw constellations. Endogenous entry run 13 stands out for its consistently steep right-hand tail slope.

[17] See e.g. Buzzell and Gale (1987), who do not, however, report the market share formula identified by the referee.

[18] As a result, the average number of firms was reduced to 3.52 in the rectangular case and 2.42 in the skew case, compared to 4.47 and 2.63 in comparable previous simulations.

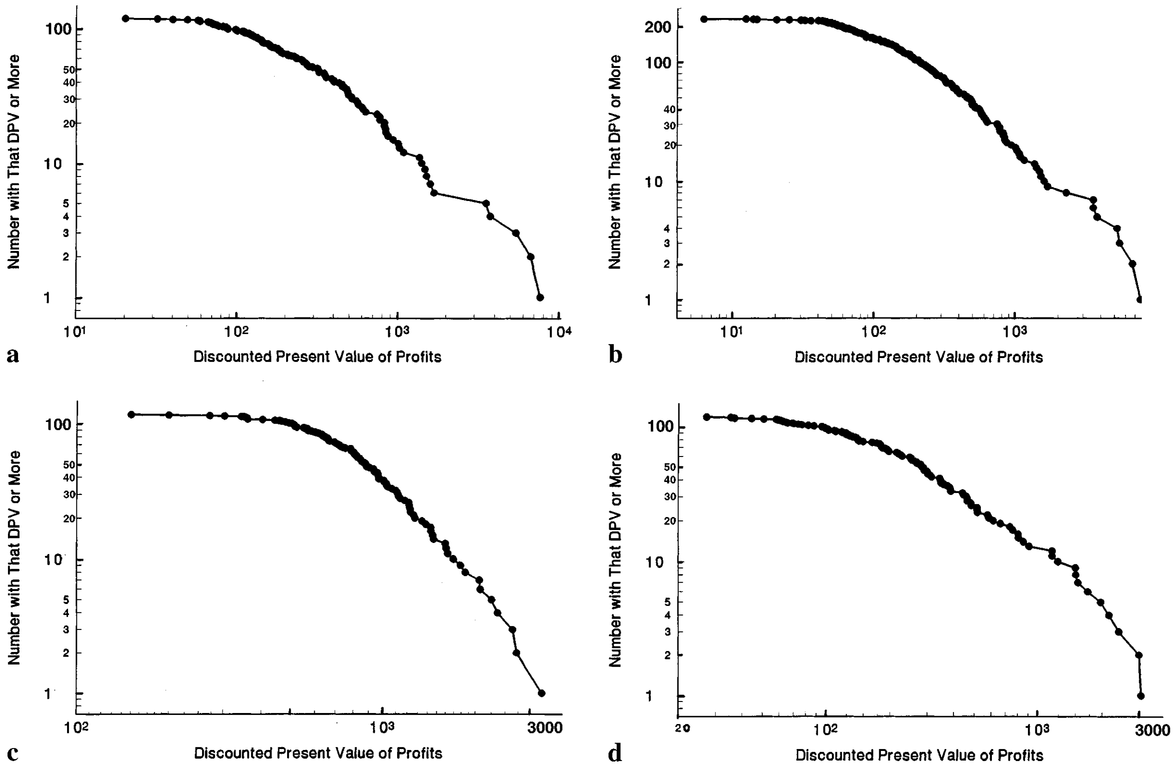


**Fig. 7.** **a** Plot of run 5: Exogenous entry, skew potential, first firm profits only. **b** Plot of run 5: exogenous entry, skew potential, first and second firm profits. **c** Plot of run 13: endogenous entry, rectangular potential, first firm profits only. **d** Plot of run 17: exogenous entry, skew potential, first firm profits only

### *6.3 Evaluation*

To the extent that the simulation runs track with reasonable fidelity the kinds of uncertainties encountered in real-world innovation histories, it would appear that ex ante innovation profit potentials are more likely to be skew-distributed than rectangularly distributed, and entry into innovation "markets" appears to be approximated better by the assumption of exogeneity than endogeneity.

Skewness of the innovation profit potentials seems eminently plausible. The distributions of manufactured product sales at the four- or five-digit level of disaggregation are known to be skew, and although most innovation markets are probably narrower subsets of five-digit product categories, skewness appears to persist at the quite narrow seven-digit level of disaggregation.

Because inventive and innovative efforts are known to be responsive to the pull of demand (see Schmookler, 1966; Scherer, 1982) and because would-be innovators plainly search aggressively for profit-earning opportunities, our finding that an endogenous entry assumption yields less realistic size distributions than an exogenous entry assumption merits deeper critical scrutiny. If innovative profit potentials are skew-distributed and contests to appropriate them are of a winner-take-all character, skew outcome distributions could result even if innovation races attracted sufficiently many entrants to leave zero ex ante profit expectations, as a large class of theoretical models postulates (see Reinganum, 1989). However, empirical studies by Mansfield et al. (1981, p. 627), Gort and Klepper (1982, p. 629), and Utterback (1994, Chapter 2) reveal that in the great majority of cases, first movers do not appropriate all the profits from an innovation race. It is frequently possible to invent around first movers' patents, and most innovations attract a stream of rival firms, some early, some late, offering variants and imitations. Two phenomena appear to explain the apparent failure of entry to rise systematically (even if stochastically) with the size of an innovation profit potential. For one, it may be so difficult to predict in advance the existence and size of an innovation profit potential that matching of entry attempts to profit prospects operates quite imperfectly – so much so that little or no correlation exists between the profit potential and the number of firms, holding R&D costs constant (see also on this problem Richardson, 1960). Second, as Henderson and Cockburn conclude from a detailed analysis of rivalry in new drug development (1994a, b), firms' capabilities to undertake innovations are often heterogeneous and specialized. An imperfect match between capabilities and opportunities undermines the correlation between profit potentials and entry.

## 7 Conclusion

From our research and research by others, e.g., on invention patent renewal rates (see e.g. Lanjouw et al., 1996), it is clear that strong regularities pervade the size distribution of rewards realized on technological innovations.

The distribution is highly skew, so that a relative handful of successful innovators reap the lion's share of profits from innovative efforts. For the richest data, a log normal distribution provides the best overall fit. In this paper we have attempted to determine what stochastic processes might lie behind the observed size distributions. Some kind of iterated proportional stochastic growth process of the sort first proposed by Robert Gibrat seems the most plausible candidate. The simplest Gibrat processes yield over plausible time intervals log normal distributions with parameters approximating those we have observed in the real world. However, the stochastic events experienced in the evolution of a technological innovation almost surely have a richer stochastic structure than does the simplest Gibrat process. In this paper we have attempted to characterize that structure and determine through simulation experiments whether realistic size distributions emerge. Some of the experiments, and especially those which tap skew profit potentials and assume stochastically exogenous entry into innovation rivalries, yield size distributions approximating those observed in the real world. The simulated distributions, however, do not probe the extremes of skewness found for patented inventions. To approximate even more closely what occurs in the uncertain world of innovation rivalry, further stochastic elements may have to be identified.

## References

Brav A, Gompers PA (1997) Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies. Journal of Finance 52: 1791–1821

Buzzell RD, Gale BT (1987) The PIMS principles: Linking strategy to performance. Free Press, New York

D'Agostino RB, Stephens MA (eds) (1986) Goodness-of-fit techniques. Marcel Dekker, New York

DeVany A, Walls WD (1996) Bose-Einstein dynamics and adaptive contracting in the motion picture industry. Economic Journal 106: 1493–1514

Feller W (1957) An introduction to probability theory and its applications, 2nd edn. Wiley, New York

Gibrat R (1931) Les inegalities economiques. Recueil Sirey, Paris

Gort M, Klepper, S (1982) Time paths in the diffusion of product innovations. Economic Journal 92: 630–653

Grabowski H, Vernon J (1990) A new look at the returns and risks to pharmaceutical R&D. Management Science 36: 804–821

Grabowski H, Vernon J (1994) Returns on new drug introductions in the 1980s. Journal of Health Economics 13: 383–406

Harhoff D, Scherer FM, Vopel K (1998) Exploring the tail of patent value distributions. Working paper, Center for European Economic Research, Mannheim

Harhoff D, Narin F, Scherer FM, Vopel K (1999) Citation frequency and the value of patented inventions. Review of Economics and Statistics 81: 511–515

Henderson R, Cockburn I (1994a) Racing to invest: The dynamics of competition in ethical drug discovery. Journal of Economics and Management Strategy 3: 481–519

Henderson R, Cockburn I (1994b) Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal 15: 63–84

Ijiri Y, Simon H (1977) Skew distributions and the sizes of business firms. North-Holland, Amsterdam

Lanjouw JO, Pakes A, Putnam J (1996) How to count patents and value intellectual property: Uses of patent renewal and application data. Working paper, Yale University, New Haven

Mandelbrot B (1963) New methods in statistical economics. Journal of Political Economy 71: 421–440

Mansfield E, Wagner S et al. (1977) The production and application of new industrial technology. Norton, New York

Mansfield E, Schwartz M, Wagner S (1981) Imitation costs and patents: An empirical study. Economic Journal 91: 907–918

Reinganum JF (1989) The timing of innovation: Research, development, and diffusion. In: Schmalensee RL, Willig RD (eds) Handbook of industrial organization, vol I, pp. 849–908. North-Holland, Amsterdam

Richardson GB (1960) Information and investment. Oxford University Press, Oxford

Scherer FM (1967) Research and development resource allocation under rivalry. Quarterly Journal of Economics 81: 359–394

Scherer FM (1982) Demand-pull and technological innovation: Schmookler revisited. Journal of Industrial Economics 30: 225–238

Scherer FM (1998) The size distribution of profits from innovation. Annales d'Economie et de Statistique no. 49/50: 495–516

Scherer FM, Ross DR (1990) Industrial market structure and economic performance, 3rd edn. Houghton-Mifflin, Boston

Schmookler J (1966) Invention and economic growth. Harvard University Press, Cambridge, MA

Utterback J (1994) Mastering the dynamics of technological innovation. Harvard Business School Press, Boston