MALL BUSINESS RESEARCH IN A WORLD OF SKEWED RETURNS

Major Toby Edison and Kevin Carman*

ABSTRACT. This paper highlights in a review of small business evaluations the results of the universal presentation of small business support program effects (such as the Federal program, Small Business Innovation Research (SBIR)) as averages. The paper focuses on observations of skewed returns in related research fields: technology policy and stock market returns. The paper highlights a body of small business program evaluations which help lay the foundation for the possibility that outcomes from small business support programs are skewed in a similar nature to stock market returns and technology investments.

The paper concludes with some recommendations on empirical and non-empirical strategies for researchers to deal with the skewed nature of outcome observations to better craft policy recommendations reducing the impact of extreme data points or "high flyers".

** Major Toby Edison, is a Professor of Program Management for Defense Acquisition University (DAU), West Region, Los Angeles, CA. Toby received his doctorate at the RAND Graduate School in Public Policy Analysis. His thesis was an evaluation of the DOD SBIR (Small Business Innovation Research) program. He has a MBA from UC Irvine. He has held several key acquisition positions within the US Air Force in program management while stationed at Hanscom AFB, MA. and at Los Angeles AFB, CA. Among the programs he has been actively involved in were the Space Radar and Joint STARS programs. He is currently on active duty with the US Air Force deployed overseas. He is married with two children.*

Major Toby Edison, PhD Professor of Program Management, Defense Acquisition University 222 N. Sepulveda Blvd, Suite 1220. El Segundo, CA 90245-5659toby.edison@dau.mil.Ph: 310-606-5924, Fax: 310-606-5925.

Kevin Carman, is Chair, Executive Development and a professor of Program Management at the Defense Acquisition University (DAU), West Region, San Diego, CA. He has a doctorate from Alliant International University in Strategic Management. Dr. Carman has more than 30 years' experience in procurement, contracting, and acquisition. He is a retired Captain from the Navy supply corps. He held several key Program Management positions in the Navy and was a senior manager with Raytheon working on several major defense programs. He is married with two grown children.

Dr. Kevin Carman, DBA ,Chair, Executive Development, Professor, Program Management, Defense Acquisition University, 33000 Nixie Way, Bldg. 50, San Diego, CA. 92147, kevin.carman@dau.mil, Ph: 619-524-0831 Fax: 619-524-4794

"(D)ata analysis, like calculations, can profit from repeated starts and fresh approaches...there is not just one analysis for a substantial problem" (Tukey)

INTRODUCTION

This paper observes that outcomes from small business support programs are likely to have a small fraction of observations with significant positive outcomes which can impact program evaluations and conclusions. Current evaluations of small business support programs generally estimate the efficacy of the program using an estimation of an average outcome or an average treatment effect. Because there are so few small business programs that have been designed for evaluation, researchers will often report a simple average outcome based on a survey response of treated firms, or in very rare cases, report the effectiveness of the program as an average treatment effect based on a quasi-experimental or experimental design. Unfortunately, presenting the small business programs' impact as an average, might be understating the magnitude of the impact of the program. The actual distribution of the outcome of the program is probably not normally distributed, and probably skewed in the positive direction. In light of the possibility that small business programs produce skewed outcomes, this paper recommends that the small business evaluations include discussions on the distribution of outcomes to better inform policy makers, administrators and participations on the nature of the outcomes.

This paper will review small business evaluations which will highlight the universal presentation of small business support program (since as the Federal program, Small Business Innovation Research (SBIR)) effects as averages and the result this has on interpreting these effects. The paper outlines the observations of skewed returns in related research fields: technology policy and stock market returns. The paper highlights a body of small business program evaluations which identifies the possibility that outcomes from small business support programs are skewed in a similar nature or pattern as stock market returns and technology investments. The paper recommends some empirical and non-empirical strategies for researchers to deal with the skewed nature of outcome observations to better craft policy recommendations.

IF A SMALL BUSINESS PROGRAM EVALUATION REPORTS AN OUTCOME IT'S AN AVERAGE

This section of the paper reviews the literature on small business support program evaluations concluding that generally if an effect from program participation is estimated, that normally is presented as an average effect. The findings from two literature reviews on small business support programs are presented. The first a 2008 working paper by Gu, Karoly and Zissimopoulos that summarized the then current literature on small business support program evaluations mostly in the United States. The second review, a dissertation literature review chapter from 2010 by Edison, summarizes the then current literature on the Small Business Innovation Research program. Both document that most program effects are presented as an average treatment effect or as an average of survey responses.

Small businesses are believed to play a vital role in the economy, generating new jobs and fueling innovation. Because of this, many local, state and federal small business support programs have been created to finance, incubate, and educate entrepreneurs in the hopes that new firms will create new jobs and increase regional prosperity. Despite the strong support, and billions of dollars in annual funding for set-aside contracts, subsidized financing, education and business zones there is relatively little rigorous research conducted on small businesses. Moreover, even when rigorous research is conducted it often reveals very little information about how and why programs are effective. (Gu, Karoly, & Zissimopoulos, 2008).

Gu, Karoly and Zissimpolous (2008) review 22 small business support program evaluations which exclusively report program effects as averages of the outcomes observed. The researchers documented that of the 22 studies, 20 of the studies employ econometric analysis or means comparison methods, with two studies reporting descriptive summaries of surveys of the participants. The vast majority of the evaluations were published without comparing the average effect with a treatment groups. This review documented only two studies that employed experimental methods. Benus (1994) estimated a treatment effect from a randomized control trial that offered business training on to randomly selected groups of from a population of individual's interest in self-employment; and Bellotti (2006) documents the experimental design Project GATE (Growing America Through Entrepreneurship) which is used to estimate a treatment effect. Furthermore, the Gu (2008) literature review documents only two econometric quasi-experiments which used matched control groups. Lerner (2006) and Sanders (2002); reinforcing the finding that very few small business support program evaluations, use rigorous estimation models.

Edison (2010) reviewed the literature documenting the methods used in the 39 published reports evaluating the Small Business Innovation Research (SBIR) program from 1996 to 2010. The SBIR is a federal R&D program that mandates large Federal agencies set aside 2.5% of their external R&D budget to be awarded to small businesses through the SBIR program. The program has been studied extensively by the GAO (1999, 2005, 2006), RAND, National Academies of Sciences and various academic researchers. Of the 39 published evaluations of the SBIR program, only six used quasi-experimental econometric methods to estimate an average treatment effect, the remainder of the reports presented the effects of the SBIR program with a simple mean, typically from a self-reported survey. A handful of the reports used other qualitative methods to describe the effects of the program, typically case studies on a handful of participating firms or of SBIR program administrators in the various federal agencies. Regardless of the level or rigor employed in the evaluations, they unanimously find a positive average (treatment) effect for firms that participate in the program. The state of SBIR program evaluations can be summarized by the following 2005 GAO report conclusion, …"an issue that continues to remain somewhat unresolved after almost two decades of program implementation is how to assess the performance of the SBIR program." (GAO, 2005, p.1)

Despite the preponderance of average effects published in the SBIR literature documented by Edison (2010), the literature review documents a subset of analysis that document a phenomenon of extreme outliers in outcomes in Held (2006).[†] The most obvious documentation of extreme outliers is in the analysis of cohorts of first time DOD contractors who won their first DOD contract through the SBIR program. For these cohorts of firms (256 in 1995, and 220 in 1999), non-SBIR defense contract awards for each firm were observed three years following their initial SBIR award. In each cohort more that 95% of the entire cohort's non-SBIR DOD contracts were won by just three firms (figure 1). Edison observes that the extreme outlier phenomenon is consistent with the Scherer and Harloff (2000) generalization that returns to innovation investment appear to be concentrated in a top few observations. Scherer and Harloff (2000) observe the consistent pattern of 50-95% of returns from innovation being concentrated in the top 10% of study population observations, they observe that this pattern is consistent in patents, Venture Capital investments, IPOs, startup firms and pharmaceutical launches.

† The author/researcher of this article was also a member of the Held (2006) research team.

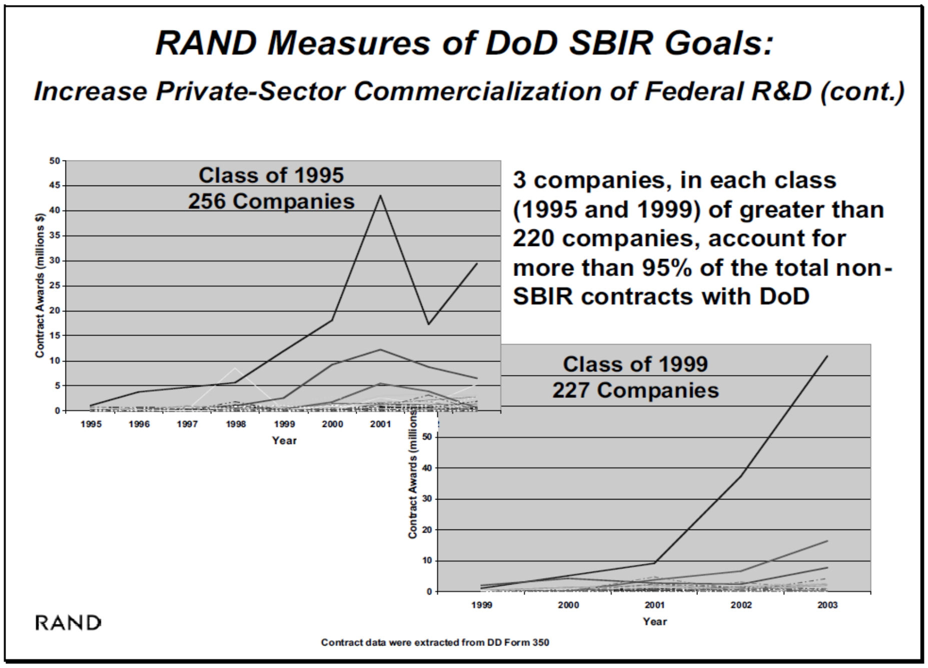

FIGURE 1: GRAPHICAL ANALYSIS OF SBIR 'HIGH FLYERS' (HELD, 2006)

This section summarized the current paradigm of evaluating small business support programs using generally accepted econometric analysis and mean comparison methods. Both methods have the strength in that they are easily understood by informed laymen and advance researchers. Unfortunately both methods rely on the assumption of normality or ‘approximate normality’ to estimate average effects. The Edison (2010) observation that the rare analysis by Held (2006) which observed effects as containing a few high performing outliers (rather than an average), raised doubts about the ‘approximate normality’ of the outcome variables, and thus the validity of using averages to estimate program effects. The next section further discusses Scherer and Harloff's (2000) seminal observations of skewed returns to innovation investments and presents related observations on skewed returns from the financial analysis research community and from the network traffic analysis community.

EXCEPTIONAL OBSERVATIONS LITERATURE REVIEW

A paradigm of research agenda focusing on skewed returns is growing in the diverse fields of innovation policy, financial analysis and network traffic analysis. This paradigm shift has been levered by simultaneous improvements in data collection and data processing which have enabled researchers to collect and analyze an entire census of population observations rather than samples of the population. The researchers analyzing the 'skewed-distributed outcomes' (Scherer, Harloff, 2002) or 'heavy tailed' distributions (Mandelbrot, 2003) (Resnick, 1997)

Peck and Scherer's (1962) research on skewed returns presents evidence from 8 data sets (figure 1) that summarize the skewed-distributed outcome from a diverse range of innovation investments (patents, pharmaceutical introductions, initial public offerings (IPO's) of firms, firm startups). In each of these populations a small fraction of observations (~10%) contain a majority of the returns of the entire population. The authors document implications this generalization has for policy makers: to ensure success a large variety of projects must be supported rather than a small number of 'national champions'; and to judge measure success policy makers must be able to support a program that might show a 90% failure rate. Scherer and Harloff's observations appear to have been incorporated into some mainstream research and development policy evaluations. (Roessner, 2002)

Scherer and Harloff's (2000) concluded that 'a small minority of innovations yields the lion's share of all innovations.' Total economic value also has implications for researchers, which this paper will elaborate:

- If researchers are using sampling methodology, there is a high probability that they will not sample some of the most significant observations, which will dramatically change the results of the evaluation.

- Standard econometric analysis, and mean comparison methods, which rely on normality assumptions, might not be valid.
- Lag effects and spillover effect might be extremely large and potentially macro-economically disruptive.
- Standard sampling methods might miss the few most important observations in the population.
- Matching algorithms might drop key observations.
- Analysis of the returns from the top few performers might be sufficient to perform a cost benefit analysis. (Roessner, 2002)

Table 1
Proportion of innovation samples' total value realized by the most valuable 10% of innovations

Data set	Number of observations	Percent of value in top 10%
German patents	772	84
US patents	222	81–85
Harvard patents	118	84
Six university patents		
1991 royalties	350	93
1992 royalties	408	92
1993 royalties	466	91.5
1994 royalties	411	92
Venture Economics startups	383	62
Horseley–Keogh startups	670	59
Initial public stock offerings (IPOs) — 1995 stock value	110	62
Grabowski–Vernon		
1970s drugs	98	55
1980s drugs	66	48

TABLE 1: RETURNS TO INNOVATION PORTFOLIOS (SCHERER AND HARLOFF, 2000).

LIMITATIONS OF SCHERER AND HARLOFFS' SKEWED-OUTCOME OBSERVATIONS

In a real world sense, Scherer and Harloffs' (2000) observations that returns to innovation investments have a long skewed outcome tail, is fundamental for policy makers and analysts who need to understand the nature of the innovation process. Unfortunately this real-world observation has a significant limitation when analysts are attempting to use standard quantitative tools to describe what they are seeing the data. If what analysts are seeing can be characterized as having a finite 'thin tail' that approximates a log-linear distribution, then researchers can use standard data analytic tools by transforming the outcome observations with a log function. The log-linear functions have the analytical strength of converging to stable means and variances with large numbers of observations. However, if what analysts are seeing is an infinite and or 'fat tail' distribution, then this distribution is better described with a Pareto-Levy function, which has the analytically difficult challenges of having unstable means and variances, especially when larger numbers of observations are analyzed.

Another difficulty both log-normal and Pareto-Levy distributions have is that they are undefined for observations with zero or negative values. In many situations when analyzing returns to innovation, this is not an issue, because some most innovations have some value greater than zero. For example, in the case of Initial Public Offerings (IPO), the stock price of a firm represents an approximation of the value of assets of a company, which will probably have some non-negative worth. However, in the case of patents, some patents might not have any estimated commercial value. This limitation is even more worrisome when judging the impacts of small business support programs, because some interventions might possibly yield no increase in future sales, employment or earnings, and other interventions could actually have a negative impact on business outcomes.

A careful reading of Scherer and Harloff (2000) reveals a worrisome assumption. They postulate that the returns to innovation 'most likely' adhere to a log normal law (a long thin tail) but open the possibility that their research 'may not have captured the most extreme private values'. The authors leave the door open to the possibilities of skewed-outcomes of a Pareto-Levy distribution. For most research applications with a small number of observations, narrowly defined outcome measures, and short time horizons the log-normal approximation will provide researchers and policy makers with useful insights. Unfortunately, the models employing log-normal assumptions might not sufficiently represent the impacts of disruptive innovations, which have the potential to have significant macro-economic impacts. Disruptive innovations are actually, in many cases, what innovation and small business support programs are actually seeking.

In aggregate, the literature review on small business support evaluation in the previous section of this paper confirms that the prevailing paradigm is to assume normality and present an average effect. Given this current paradigm it would be difficult, for researchers to publish program evaluations which abandoned this paradigm, to make it past a peer reviewer or bureaucratic sponsors with a report that presented their results as a distribution of outcomes which was non-normal and skewed. To motivate a paradigm shift in small business research, a brief introduction to the current paradigms in financial research, and network traffic analysis modeling is presented.

HOW OTHER DISCIPLINES EMBRACE NON-NORMALITY

Research in finance and telecommunications network modeling has long embraced the heavy-tailed, skewed distribution phenomenon. The stylized observations or non-normality stemming from heavy-tailed distributions, have caused significant re-examination of the analytical foundations of these disciplines.

The seminal paper on non-normality in network data analysis from these disciplines might be useful for small business support policy analysis is Leland, Taqqu, Willinger, & Wilson (1994) which describes the research team's analysis of Ethernet LAN traffic as 'self-similar' or fractal. This observation counters the then prevailing paradigm of analyzing Ethernet traffic with 'Poisson-like' models. The authors observe that regardless of how small or large the time interval Ethernet traffic reveal a pattern of 'burstiness', which contain an extremely small percentages of extra-ordinarily large packages of information. The researchers then introduce several different advanced statistical methods to represent the behavior of the data. The roots of these analytical methods are found in the body of work by Mandelbrot, which will be discussed shortly. What is more important to small business small business support program analysts is the three implications the authors recommend for engineering implications on how to design systems.

The first implication the 'self-similar' nature of Ethernet data has for system design is that the nature of the sources, destinations and nature of the individual data packets needs to be analyzed. This observation can certainly be appropriated by small business policy researchers by focusing inside the 'black-box' of select firms in the population to discover the nature of their founding story, the nature of the market they are selling to, and their internal structure. The second observation in Leland (1994) is that researchers need to account for and describe the behavior of the variation in network traffic as longer and longer time intervals are observed. Because each population under analysis is necessarily a finite data-set, a mean and a variance can always be calculated, but when the time intervals are increased, the mean and variance might not converge to stable estimates. This observation can also be useful to small business policy researchers, who should describe the effect of increasing sample size or time intervals have on the mean and variance of the outcome of interest. Finally, Leland (1994) details the implications that self-similar traffic pattern have for engineering predictive data traffic models. With better knowledge of the behavior of the traffic data, the engineers can build synthetic models of proposed Ethernet networks which will experience better traffic flow, less congestion, and be optimally sized to reduce excess capacity. The observation that better system design is possible with a better understanding of the nature of the network traffic is directly applicable to small business policy researchers to better describe the nature of the observed outcomes resulting from program designs so that the delivery of the small business support can be more appropriately tailored to provide the right amount of support for small businesses.

This seminal paper by Leland (1994) has many connections to Mandelbrot's (2003) research on non-normal behavior of stock market returns, portfolio risk and optimal portfolio design. The Handbook of Heavy Tailed Distributions in Finance edited by Rachev (2003) contains a good summary of the research and theory that are the foundation of non-Gaussian (non-normal) financial analysis. Rachev (2003) offers these stylized facts regarding the distribution of financial data:

- Non-Gaussian, heavy-tailed and skewed distributions
- Volatility clustering
- Temporal dependence of the tail behavior
- Short and long-range dependence

These theories regarding the behavior of financial markets are directly applicable to small business policy researchers because small business support programs inherently involved financing business ventures with the hopes of those ventures increasing in value and improving the regional or national economy.

This section has some paradigm shifting observations regarding evaluations of small business support programs: specifically that skewed-outcome observations can invalidate normality assumptions, and therefore call into question the recent advancements in small business policy research which have made significant advances to estimate average effects. More research is needed to determine whether the log-normal assumption holds and extreme positive outcomes in small business performance can be characterized as a thin tail or a fat tail. It is the intention of this paper to motivate future researcher to detail the distribution of the outcomes (especially those skewed outcomes) in their evaluations, even if average effects are also presented.

A DEEPER EXAMPLE OF EXCEPTIONAL OUTCOME DATA ON SMALL BUSINESS OUTCOMES

Edison (2010) details the distributions of pre- and post-treatment outcomes for firms who applied for DOD SBIR contracts in 2003. The distributions for pre-treatment non-SBIR DOD contracts clearly adhere to the skewed-outcome distribution described by Scherer and Harloff (2000) (figure3). Likewise, the post-treatment observations (table 2) illustrate another example of skewed, long-tail outcomes.

Given the previous presentation on the skewed nature of outcomes from innovation investments, and financial markets, it seems plausible that the patterns presented from this evaluation of the DOD SBIR program would also be evident in other evaluations of Small Business programs. The only way to verify that this phenomenon is not just an anomaly is for researchers to include an analysis of the structure of the datasets in their publications. Additionally reviewers of research on small businesses should also request their authors to include graphical and statistical summaries of the skewed nature of the outcome variables. Through these subtle actions a better understanding of the nature of the impacts from small business support programs can be characterized and help improve support programs.

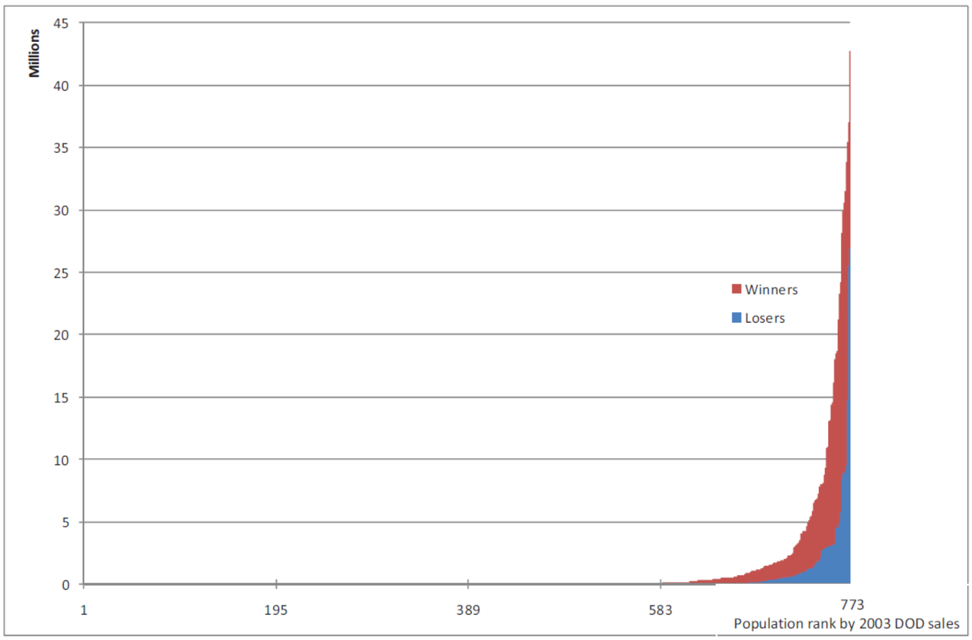

FIGURE 2: PRE-TREATMENT DOD SALES (SCHERER & HARLOFF, 2000)

	loser	winner
Observations	773	687
Mean	220,779	649,822
Stdev	2,121,602	4,468,850
Max	35,700,000	61,600,000
Min	(7,620,000)	(23,800,000)
% Negative	12%	18%
% Positive	15%	29%
% zero	73%	53%
% extreme neg (<2stdev)	0.39%	0.29%
% extreme pos (<2stdev)	1.29%	2.04%
% extreme	1.68%	2.33%

TABLE 2: POST-TREATMENT INCREASE IN DOD SALES (2004-2003) (SCHERER & HARLOFF, 2000)

RECOMMENDATIONS FOR RESEARCHERS

The following recommendations are provided for researcher, especially those researching or using data that describe the effects of small business support programs such as the Federal program, Small Business Innovation Research (SBIR). The recommendations are meant to assist in the more meaningful uses and treatment of these data to interpret the effectiveness of these program on their activities:

1. Present the distribution of the outcome data, with both normality (or log-normality) assumptions and non-normality assumption
2. Present average outcome (e.g. from survey) or average treatment effect from (experiment or quasi-experiment)
3. Do case studies on 'average firm'; well below average firm; and exceptional firm.
4. Unleash the creative forces of policy researchers to invent a way to perform program evaluations in a world of skewed returns.

Peck and Scherer (1962) have noted in numerous studies that returns on innovation are highly skewed. This paper surveys literature on small business evaluation programs, finding that most estimate an average treatment effect. This paper also reviews a subset of the small business literature in the Department of Defense small business programs, concluding that the returns to these programs appear to be skewed as well. Researchers and policy makers should consider the possibility that returns to small business support programs might be skewed, therefore in evaluating the programs a small average treatment effect might be hiding a large effect concentrated in a few high return successes.

Policy makers considering small business interventions should be aware of phenomenon of skewed returns in building their support interventions, and in the rare instance the policy makers choose to evaluate the intervention, they should consider augmenting average treatment effect with more focused case studies on the 'high flyers' or extreme data points.

REFERENCES

Edison, T. R. (2010). *Estimation of the Department of Defense Small Business Innovation Research Program Treatment Effect.* Santa Monica: RAND.

Gansler, J. S. (2004). *Capitalizing on Science, Technology, and Innovation: An Assessment of the Small Business Innovation Research Program: Project Methodology.* Washington D.C.: The National Academies Pre

GAO. (1999). *Evaluation of SBIR Can Be Strengthened. .* Washington D.C.: Government Accountability Office.

GAO. (2005). *Observations on the Small Business Innovation Research Program: Statement for the Record of Anu K. Mittal, Director Natural Resources and Environment Team.* Washington D.C: Government Accountability Office.

GAO. (2006). *Small Business Innovaiton Research: Agencies Need to Strengthen Efforts to Improve the Completeness, Consistency, and Accuracy of Awards Data.* Washington D.C: Government Accountability Office.

Gu, Q., Karoly, L. A., & Zissimopoulos, J. (2008). *Small Business Assistance Programs in the United States*. Santa Monica: RAND.

Held, B. E. (2006). *Evaluation and Recommendations for Improvement of the Department of Defense Small Business Innovation Research Program.* RAND: Santa Monica.

Leland, W. E., Taqqu, M. S., Willinger, W., & Wilson, D. V. (1994). On the Self-similar Nature of Ethernet Traffic (Extended Version). *IEEWACM TRANSACTIONS ON NETWORKING,* , 1-15.

Peck, M. & Scherer, F.M. (1962). *The Weapons Acquisition Process: An Economic Analysis.* Division of Research Graduate School of Business Administration, Harvard University.

Racher, S.T. (2003). *Handbook of Heavy Tailed Distributions in Finance.* Elsevier/North Holland.

Resnick, S. (1997). Heavy Tail Modelling and Teletraffic Data. *The Annals of Statistics* , 1805-1869.

Roessner, D. (2002). Outcome measurement in the United States: State of the art. *American Association for the Advancement of Science.* Boston.

Scherer, F. M. (1965). Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. *American Economic Review* , 1097-125.

Scherer, F. M., and Harloff, T.S. (2000). Technology Policy for a World of Skew-Distributed Outcomes. *Research Policy, 29* (4-5), 559-56.

Scherer, F. M. (1964). The Theory of Contractual Incentives for Cost Reduction. *The Quarterly Journal of Economics* , *78* (2), 257-280.

Tukey, J. W. Data analysis and behavioral science or learning to bear the quantitative man's burden by shunning. In L. W. Jones, *The Collected Works of John W. Tukey.* Monterey: Wadsworth.