WP 2009-32
October 2009

Working Paper

Department of Applied Economics and Management
Cornell University, Ithaca, New York 14853-7801 USA

Understanding the Entrepreneur: An Index of Entrepreneurial Success

Harold O. Fried and Loren W. Tauer

UNDERSTANDING THE ENTREPRENEUR: AN INDEX OF ENTREPRENEURIAL SUCCESS

Harold O. Fried, Union College, USA
Loren W. Tauer, Cornell University, USA

ABSTRACT

A measure of entrepreneur success is important to identify current and future successful ventures, to further our understanding of the entrepreneurial process and to guide public policies to improve the success rate of start-ups. In this paper we propose an index of entrepreneur success that accommodates multiple inputs and outputs, that is predicated on inputs and that mitigates the impact of outliers. We relate the index to characteristics of the entrepreneur and the venture: age, experience, gender, race, competitive advantage, education, and birthplace. The data are from the Kauffman Firm Survey. The index is calculated for 2,863 firms in 2006.

INTRODUCTION

The entrepreneur is critical to the vitality of a market economy. The better we understand the determinants of successful entrepreneur ventures, the higher will be our standard of living and the brighter our future. In this paper, we propose an index of entrepreneurial success.

The classical view of resource allocation in a market economy is a dynamic process in which resources adjust in response to the appearance of profitable opportunities. As consumer wants, resources and technology change, opportunities emerge for firms to produce new products and services. The motivation of firms is economic profit. It is in their self-interest to engage in the production of new products and thereby earn more than in their current activity. Over time, competition erodes these extraordinary profits. Consumers benefit as the economy's scarce resources are "automatically" reallocated to areas in which they generate the most benefit. The private self-interest serves the social interest. This is Adam Smith's "invisible hand" at work. However, the process is not necessarily nearly as automatic, smooth, or predictable as it is often presumed (Christensen 1997, 2003 and 2004). The adjustment of resource allocation in response to the appearance of profitable opportunities does not just happen. It requires an entrepreneur to act, and not all entrepreneurs are equally successful.

The movement to a new equilibrium is an iterative process: new ventures emerge, old ventures die, new ventures succeed and new ventures fail. It is market performance that is the ultimate judge of successful new ideas, not the passion and the commitment of the entrepreneur. When new ventures fail, resources were misallocated and society is worse off. This is an *ex post* result. To the extent that we can improve our understanding of the entrepreneur and the drivers of successful ventures, the adjustment of resources over time will be a more socially efficient process.

It is important to be able to measure the extent of success of individual entrepreneur ventures. This is not straightforward. In particular, success might be multidimensional, and success is dependent upon the inputs available. We develop an index of entrepreneur success at the level of the individual firm that accommodates multivariate measures of success and, at the same time, controls for resources at the disposal or use of the entrepreneur. Based upon our index, we investigate some determinants of successful ventures.

We begin with a discussion of our method in the context of existing work and the particular application to startup firms. Next, we describe the data. We utilize the Ewing Marion Kauffman Foundation Firm Survey for 2005 and 2006, consisting of 2,864 firms that started in 2004. The third section presents the results. We end with a conclusion.

THE RESEARCH METHOD

A large amount of research has been completed on determining what leads to successful entrepreneurship, and the characteristics of the successful entrepreneur (Blanchflower, 1998; Evans and Leighton, 1998). Given firm data on the characteristics of entrepreneur ventures, the standard choice from the economist's toolkit is to run probit or logit regressions (Harada, 2003). The dependent variable is a binary measure of entrepreneur success or no success, and the independent variables are potential explanations for success. In some cases a continuous measure of success is constructed and linear regression is employed. The magnitudes of the estimated coefficients and their degrees of statistical significance provide information on the relative and absolute importance of the explanatory variables on entrepreneur success.

Our objective is to generate an index of entrepreneur success. We adopt a production function approach. The individual entrepreneur transforms inputs into multiple measures of entrepreneur success, subject to the influence of variables outside of his/her control. This requires an alternative specification than a simple single measure of success such as profit or even survival. We use a frontier method that can accommodate more than one measure of success, and where extreme observations are assumed to contain particularly interesting information rather than represent error in measurement.[1] It is the extreme observations that might reveal the determinants of successful entrepreneurship.

These observations represent "best-practice" for the entrepreneur. The goal is to construct an index of entrepreneur success that is based upon best-practice and is multivariate.

Figure 1 illustrates the distinction between average practice and best-practice. A measure of entrepreneur success is on the vertical axis; a possible determinant of entrepreneur success, say start-up experience, is measured on the horizontal axis. The ordinary least squares line though the "middle" of the data estimates average practice. Deviations from the line are assumed to be white noise. The stair case frontier captures best-practice under the assumptions that only observed production points are feasible and input is freely disposable[2]. This concept is typically defined as the free disposable hull frontier (FDH) (Deprins, D., Simar, L. and H. Tulkens, 1984). The graphical illustration is limited to two dimensions with one measure of success and one measure of resources. In practice, the FDH is calculated as an integer programming problem and can accommodate multiple measures of success and multiple measures of resources.

In order to construct an index of entrepreneur success, it is necessary to quantify the performance of each venture relative to best practice. Points on the FDH frontier illustrated in Figure 1 represent best-practice in the sense that there are no other firms in the data set represented by the various points which are able to produce more or the same level of entrepreneur success with fewer or the same inputs. These firms receive an index value equal to 100. Firms below the FDH frontier are inefficient in achieving entrepreneur success; these firms could produce more entrepreneur success given their inputs, were they to produce according to the best-practice frontier. The measure of inefficiency (below best-practice performance) is actual relative to optimal performance. For firm A, this is (E_AA/E_AA^*) 100, where E_A represents an experience amount and A* represents the firm that produces the highest success from the quantity of experience E_A. Suppose this value equals 60. The interpretation is that Firm A is achieving only 60 percent of the entrepreneur success that it could have achieved were it to perform at best practice at level A*, given its level of experience, E_A. The performance of all firms in the data set can be quantified to produce an index that ranges between 0 and 100, where 100 is complete success. We construct this index.

There are some important characteristics of the index to highlight:

(1) The index is specific to an input – output specification for entrepreneur success that is multidimensional; it can accommodate multiple inputs and multiple outputs. In this context, a 60% index value for a firm means that best practice corresponds to a 67%[3] expansion of *each* output.

(2) As is apparent from Figure 1, inefficient firms are often expanded to a flat portion of the best-practice frontier FDH. The implication is that Firm A, for example,

could achieve a 67% expansion of each output *and* contract the experience input horizontally to the producing firm to the left. This is slack in the integer programming problem. Although it is not apparent from the two dimensional illustration, there can be similar "slack" on the output side when there is more than one output. The index value is an upper bound measure of entrepreneur success.

(3) The index is conditional on inputs. A venture can be best-practice with low values for the success measures if it achieved this success with few inputs to work with. A venture can be below best-practice with high values for the success measures if it had a large amount of inputs to work with. Entrepreneur success cannot be determined based upon output measures alone.

(4) There is no single generally accepted input – output specification for entrepreneur success. Our results are sensitive to the defined outputs and inputs. If important measures of success are not included as an output, firms do not receive "credit" for performing well in this dimension and their success is measured too high. If important resources are not included as inputs, firms that utilize these omitted resources relatively intensively appear more successful than they in fact are.

There are numerous frontier techniques that identify best-practice. The stochastic frontier approach modifies the error term of a standard regression with the addition of a second error that results in a fitted line that attaches more importance to the outlying observations (Kumbhakar and Lovell, 2000). This has the advantage of permitting statistical inference, but has the disadvantages of strong parametric assumptions and generally is limited to a single left hand side variable. Data envelopment analysis (DEA) is similar to the free disposal hull (FDH) except that it assumes that production can occur along linear line segments that span frontier observations. We choose not to make this assumption, although we do compute the DEA value for comparison purposes.[4]

Daraio and Simar (2005) introduced the order-m approach that embeds the free disposal hull (FDH) into a probabilistic framework that mitigates the influence of outlying observations while maintaining the advantages of a flexible functional form and multiple variables on the left hand side. We use this approach to construct an index of entrepreneur success. This is useful and essential because some entrepreneurs are extraordinarily successful and that success can mask firms that are simply successful.

The order-m concept is based upon a probabilistic specification of the production process and is specified here as the process of successful entrepreneurship. The entrepreneurship process is described by the joint probability measure (X,Y), where X are inputs and Y are measures of success. This joint probability completely characterizes the probabilistic entrepreneurship process. Under an output orientation, this joint probability can be written as:

$$F_{Y|X}(y|x) = \text{Prob}(Y \leq y \mid X \leq x) \qquad [1]$$

The expected order-m frontier for a fixed integer value of m≥1 is the expected value of the maximum of m random variables $Y^1,\ldots,Y^m$ drawn from the conditional distribution function of Y, given that X≤ x. Essentially, a firm's efficiency is computed in reference to a random sample of m other firms drawn with replacement who use the same or fewer inputs than the firm being evaluated. This can be done by Monte-Carlo methods, or more efficiently by numerical integration.

The estimator by integration is given by:

$$\theta_m(x,y) = E[\max(Y^1,\ldots,Y^m)|X\leq x] = \int_0^{\infty}(1-[F_{Y\mid X}(y\mid x)]^m)dy \qquad [2]$$

or

$$\theta_m(x,y) = \theta(x,y) + \int_0^{\theta(x,y)}(1-[F_{Y\mid X}(y\mid x)]^m)dy, \qquad [3]$$

where $\theta_m(x,y)$ is the order-m efficiency estimate for each firm, which is computed from the FDH output efficiency estimate plus the defined integral. These can be computed using nonparametric integration methods as shown by Daraio and Simar (2005).

The production process for entrepreneur success includes both conventional inputs that are under the control of the entrepreneur (x variables) and exogenous variables (z variables). For example, total cost influences entrepreneur success; it is under the control of the entrepreneur; and it is clearly classified as an input. In contrast, experience of the entrepreneur might also influence success, but it is not under the control of the entrepreneur, at least in the short run. It is exogenous. We can accommodate a z variable in the index, and we can identify the relationship between a z variable and performance.

To estimate efficiency conditional upon an exogenous characteristic of the startup (z), the equations are modified so that the output y is not only conditional upon the inputs x, but also conditional upon z. Equation [3] is then modified as:

$$\theta_m(x,y|z) = \theta(x,y|z) + \int_0^{\theta(x,y|z)}(1-[F_{Y\mid X,z}(y\mid x,z)]^m)dy. \qquad [4]$$

A nonparametric estimate requires a kernel estimator for z with a bandwidth. A triangle distribution is used as the kernel and a bandwidth (h) is utilized for the z variable. To identify the relationship between z and performance, we construct the ratio of unconditional efficiency to conditional efficiency and regress this ratio against the exogenous variable using a nonparametric approach.

The advantageous and strength of the order-m technique is that it incorporates the stochastic nature of entrepreneurship success and still allows a non-parametric relationship between determinants and success. The stochastic allowance is important because some entrepreneurs are successful because they are lucky. These occurrences can be differentiated because of the probabilistic specification of the order-m technique. The probabilistic nature of the order-m technique is a function of the size of m relative to the size of the sample. In the extreme, where m equals the sample size, the order-m approach is non-stochastic and approaches FDH. As m becomes smaller, the technique is more stochastic: the results are based upon the expected value of small subsamples which are more likely to exclude outlying observations.

THE MODEL AND THE DATA

The results are only interesting and useful to the extent that the input – output specification is plausible. The general principle is to identify a core model that includes quantifiable inputs and outputs that are under the control of the entrepreneur. There are clearly other variables that determine entrepreneur success, but are more or less exogenous to the entrepreneur. For example, various personality traits could certainly influence the ability to transform inputs into outputs, but personality is exogenous. We omit exogenous variables from the calculation of the core index and investigate their influence in terms of how they impact the index. If the exogenous variable is dichotomous, we can compare group means. If the exogenous variable is continuous, we can calculate the conditional order-m frontier and fit a nonparametric regression to the ratio of unconditional to conditional efficiencies. We have data on both dichotomous and continuous exogenous variables.

The Core Model

The inputs are total cost and owner hours. Total cost measures the resources used by the enterprise. The variable owner hours measures the commitment and effort of the entrepreneur(s) to the enterprise. We have data on up to ten owners. The variable total owner hours is the sum over all owners. The greater is the total cost and owner hours, the more outputs the enterprise must produce in order to be on the frontier and receive an index value equal to 1.00. This is the sense in which the measure of entrepreneur success is conditional on the resources available to the enterprise. Other firms in the sample with the same or less total cost and owner hours and more of each output will render this firm below best practice and an index value less than 1.00. It is possible for firms with few inputs to receive an index value equal to 1.00 even though there are likely to be many firms in the sample with better output numbers, it is just that these other firms utilize more inputs.

The outputs are total revenue and revenue growth. Total revenue is a measure of current success. Revenue growth is a measure of successful growth. To the extent that a startup firm is already large, it is doing well; to the extent that a startup firm is growing, it may be larger in the future. Both revenue and revenue growth are important. High revenue and high revenue growth will render a firm efficient only if there are no other firms in the data set with the same or more revenue and revenue growth and the same or fewer of both inputs. Firms with high revenue and revenue growth and firms with low revenue and revenue growth can be efficient, depending upon their input use relative to other firms in the sample.

The core model includes profit implicitly since revenue is included as an output and total cost is included as an input. Notice that to be efficient, a firm must produce more total revenue and revenue growth with less total cost and owner hours than other firms in the sample. Since profit is total revenue minus total cost, efficient firms produce more profit.

Figure 2 illustrates the considerations that underlie the calculation of the index value for startup firm A. The firm being evaluated is at the origin; inputs are on the horizontal axes; outputs are on the vertical axes. Firms in the northeast quadrant have more inputs and more outputs. These firms are irrelevant to determining the index value of Firm A; they are simply larger. Firms in the southwest quadrant have fewer inputs and fewer outputs than Firm A. These firms are also irrelevant to the evaluation of Firm A; they are simply smaller. Firms in the southeast quadrant have more inputs and fewer outputs. Firm A dominates these firms, rendering them inefficient, but they do not enter into the evaluation of Firm A. Firms in the northwest quadrant have fewer inputs and more outputs than Firm A. These firms dominate Firm A and are relevant to the evaluation of Firm A.

Under standard FDH which is deterministic, Firm A would be efficient only if the northwest quadrant were empty. However, under order-m, Firm A could be efficient even if there were some firms in the northwest quadrant since it is possible that all of the m samples might miss these firms. This illustrates the probabilistic nature of the order – m technique. The fewer the number of firms in the northwest quadrant, the more likely they are to be missed. The smaller is the value of m, the more likely that firms in the northwest quadrant are to be missed.

Figure 2 also illustrates the calculation of the index value. The procedure is to calculate the index value of Firm A relative to all firms in the northwest quadrant and choose the smallest index value. Firm A is evaluated relative to the frontier observation in the northwest quadrant. To the extent that there is an outlier in the northwest quadrant, then the order – m technique might miss that outlier and Firm A is rendered more efficient than would otherwise be the case.

Exogenous Variables

Exogenous variables influence the transformation of inputs into outputs and are generally outside the control of the entrepreneur. We calculate the core index without considering the exogenous variables and then examine the influence of exogenous variables on the core index. For dichotomous exogenous variables, we compare group means. For continuous exogenous variables, we fit a nonparametric regression to the ratio of unconditional to conditional efficiencies. We also estimate a truncated regression.

We examine the following dichotomous exogenous variables: gender, U.S. citizen, native born, primary owner ethnicity (white, black, Asian, Hispanic), produce a product. We examine two continuous variables – primary owner age and primary owner work experience.

Data

The data are from the Ewing Marion Kauffman Foundation Firm Survey (KFS). The KFS is a random sample of 4,929 firms taken from the Dunn and Bradstreet data base of 250,000 new businesses that were started in 2004. Interviews were conducted using telephone and a web survey. There are 2,034 high tech firms. There are 2,166 variables. Follow up surveys were conducted in 2005, 2006 and 2007 (Robb and DesRoches 2008). We use data from 2005 and 2006 for 2,864 firms. The data can be found at: http://www.kauffman.org/research-and-policy/kauffman-firm-survey.aspx. We use the non-public data set which includes more detailed information on firm location and characteristics.

The core index is calculated for 2,863 firms. Since order-m and FDH methodologies cannot accommodate negative values, revenue growth is transformed by adding the largest negative value to all of the observations. This does impact the values of the index; it does not impact the relative rankings.

RESULTS

The Order-m500 Entrepreneur Success Index (ESI)

Order-m500 efficiency scores for the core model is an index of entrepreneur success. It ranges from 0 to 1.00 and ranks 2863 firms. The index is multivariate on the output and input sides and essentially evaluates the ability of firms to transform total cost and owner hours into revenue and revenue growth. Profit is implicitly taken into account since dominating firms have more revenue and less total cost, thereby higher profit. The index has a straightforward and meaningful interpretation – an index value equal to 0.8, for example, means that the entrepreneurial venture is achieving 80% of the best-practice revenue and revenue growth, given its total cost and owner hours. It is important to recognize that the underlying methodology evaluates firms based upon output given inputs, which is a very different approach from typical indices based upon output alone. Our approach is grounded in production theory.

Table 1 summarizes the index. The FDH based index is provided for comparison purposes. The mean value of the index is 0.60 with a standard deviation equal to 0.32. There is large variation in the performance of entrepreneurs: on average, an entrepreneur is producing 60% of best-practice revenue and revenue growth, given inputs, although in light of the large standard deviation, many entrepreneurs are doing much better and many are doing much worse. There exists the potential to reap large social gains from improving the performance of entrepreneurial ventures.

Figures 3 and 4 contain the histograms for FDH and order-m500 indices.[5] The large variance in performance is clearly apparent with a mass at zero and spikes around the mean. The two histograms illustrate the significance of employing the order-m methodology. FDH is based upon pure dominance over the full sample without any mitigation for the influence of outlying observations. Mean efficiency is lower than for order-m as one would expect. The higher order-m scores are the result of mitigating the influence of outlying observations since dominance is based upon successive draws of 500 from the sample with replacement and calculating the mean score. This is an appealing feature of our index since otherwise a small number of highly successful ventures would depress the index values for many other ventures.

The Order-m ESI Conditional on Age and Experience

The order-m index is based upon the core model. As a first step to understanding the determinants of the index or the drivers of entrepreneurial success, we calculate order-m efficiency scores conditional on continuous exogenous variables, one at a time.

We begin with the age of the primary owner. We expect an inverted “U” shaped relationship as older entrepreneurs benefit from experience at first, but over time, the benefits from experience become dominated by the physical and mental deterioration associated with age. This is a classic tradeoff. Our methodology permits us to focus on the effects of age on entrepreneur success. We calculate the efficiency score conditional on age with a bandwidth of 1.98 and a logistic kernal function for smoothing. We form the ratio of the unconditional to the conditional efficiency scores (the age efficiency ratio, AER) and fit a nonparametric regression of AER against age. This is illustrated in Figure 5.

Where the unconditional and conditional efficiency scores are equal, the AER equals one. This corresponds to an entrepreneur who is performing equally relative to her age group and the total sample. Handicapping for age does not improve performance. This entrepreneur is managing her age excellently – she can hold her own relative to the entire sample without making any special provision for age. Where the conditional score is higher than the unconditional score, the AER is less than one. Handicapping for age improves the evaluation of performance. This entrepreneur is not managing her age particularly well; the performance of other entrepreneurs outside her age bracket dominates her by more than the performance of entrepreneurs inside her age bracket. Because the order-m approach permits efficiency scores greater than one where the firm under evaluation is not contained in the random sample, there are some values of AER greater than one. We do find evidence of an inverted “U” shaped relationship between age and entrepreneur success. Based upon the non-parametric regression, the optimal age is 48.

We repeat the exercise for work experience. The bandwidth is 1.98 and the kernel function is logistic. The ratio of the unconditional score to the score conditional on years of experience, the experience efficiency ratio (EER), is illustrated in Figure 6. The optimal work experience is 5 years.

There is a large dispersion of the ratios around the fitted non-parametric regression lines, which is not shown due to the confidentially of the data. Although the optimal age is 48 and the optimal work experience is 5 years, many entrepreneurs manage their age and work experience better and worse than the fitted relationship.

To further our understanding of the drivers of entrepreneur success, Table 2 contains the means of the exogenous categorical variables. For all categorical variable pairs, the means are similar and the standard deviations are high. In no cases are the means statistically different.

Truncated Regression

The FDH ESI scores are bounded between zero and one by construction, with many of these scores bunched at zero; many of these startups are simply not successful as measured by revenue and revenue growth, given their inputs. Likewise, order-m500 ESI scores are bounded, but only at zero. These scores can be greater than one since it is possible for the comparison set for a best-practice firm to exclude itself as a comparison point. The order-m500 ESI is censored at the low end. Therefore, truncation regression rather than ordinary least squares was used to investigate the determinants of the ESI in a multivariate context.[6]

The truncated regression was estimated using weighted maximum likelihood. Because this is a non-linear estimator, the estimated coefficients are converted to marginal impacts at the mean values of the independent variables as reported in Table 3. The estimated equation explains very little of the total variation of the index, which simply means that success of a business startup depends upon characteristics that may not be easily quantifiable.

The variables with the most significant statistical impact on the success score are gender and the comparative advantage of the business as assessed by the survey respondent. Gender has a significant impact on the index; businesses solely or primarily owned by men have an estimated success greater than businesses owned by women, increasing the index by 0.049. The role of gender on entrepreneurial success has been discussed elsewhere (Fairlie and Robb, 2007; Robb 2002). Our estimates confirm these findings, which serves to validate our index as a good measure of entrepreneur success.

An affirmative response by the firm respondent to the question of whether the business had a comparative advantage increased the measured success index by 0.031. Causation may be an issue with this variable; clearly a business with a comparative advantage should be more successful, but a business that is successful may also conclude that it has a comparative advantage because it is successful.

The age of the primary owner matters. The impact is quadratic such that success increases with age, reaches a peak, and then decreases. This relationship was earlier discussed with the non-parametric estimation of the ratio of unconditional to conditional on age scores, where the optimal age is 48 years. Although age of the owner is important, it appears that years of experience, also modeled as quadratic, is not important.

The positive coefficient on education suggests that the existence of a college degree or more education contributes to entrepreneurial success. Those degrees increase the success index by 0.024.

The percentage of the business funded by the sole or first owner has a slight negative and significant impact on the success score. Although an owner with a larger percentage of her funds in the businesses might be more motivated to be successful, suggesting a positive coefficient, it is also possible that successful ventures might be more likely to secure outside funding, confirming the negative coefficient. Further analysis beyond this paper is warranted to ferret out this relationship. Since the impact is extremely small and only weakly statistically significant at the 10 percent level, the estimate may simply be spurious.

The other variables are not statistically significant. The existence of more than one owner has an estimated positive but not statistically significant impact on the success index. Possibly, additional owners contribute opinions and expertise. U.S. born has a negative but not statistically significant coefficient. Entrepreneurs born outside the U.S. tend to be more successful.

Race and the business producing a product (as compared to a service) do not appear to have a significant impact on success.

CONCLUSION

A measure of entrepreneur success is important to identify current and future successful ventures, to further our understanding of the entrepreneurial process and to guide public policies to improve the success rate of start-ups. In this paper we propose an index that accommodates multiple inputs and outputs, that is predicated on inputs and that mitigates the impact of outliers.

Statistical exercises are no better than the data. This paper calculates the index for a large sample of firms that started business in 2004. Although tech firms are over-sampled, the data includes all industries ranging from small service firms to ambitious tech-based ventures. It is not clear that all of these firms strive to be large in an entrepreneurial sense. The data captures start-ups, not necessarily entrepreneurial start-ups, depending upon your definition of entrepreneurial. The heterogeneity in the data hampers our efforts to explain the index of entrepreneur success. In further work, we will calculate an ESI for specific industries. High and medium tech firms are of particular interest.

Most importantly, we demonstrate a new and novel methodology for quantifying the success of entrepreneurial ventures. There is more to come.

CONTACT: Hal Fried (friedh@union) and Loren Tauer (lwt1@cornell.edu)

We are grateful to the Ewing Marion Kauffman Foundation for providing financial support and access to the proprietary version of the Kauffman Firm Survey. Alicia Robb guided us in understanding a very large and intimidating data set. Tim Mulcahy was our mentor for working in the secure space of the NORC Enclave. We thank Leopold Simar for providing the Matlab programs to calculate conditional and unconditional Order-m efficiency scores. We take full responsibility for all results.

REFERENCES

Blanchflower, D. G. (1998). What Makes an Entrepreneur. *Journal of Labor Economics* 16(16), 26-60.

Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston: Harvard Business School Press.

Christensen, C. M. &, Rayner, M. E. (2003). *The Innovator's Solution: Creating and Sustaining Successful Growth.* Boston: Harvard Business School Press.

Christensen, C.M. et al. (2004). *Seeing What's Next: Using the Theories of Innovation to Predict Industrial Change.* Boston: Harvard Business School Press.

Deprins, D., Simar, L. & Tulkens, H. (1984). Measuring Labor-Efficiency in Post Offices. In M. Marchand, P. Pestieau, & Tulkens, H. (Eds.), *The Performance of Public Enterprises: Concepts and Measurement*. Amsterdam: North-Holland.

Evans, D. S. & Leighton, L. (1989). Some Empirical Aspects of Entrepreneurship. *The American Economic Review* 79(3), 519-535.

Fairlie, R. & Robb, A.M. (2007). Why are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital. *Journal of Labor Economics* 25(2), 289-323.

Fried, H. O., Lovell, C.A.K. & Schmidt, S. (Eds.) (2008). *The Measurement of Productive Efficiency and Productivity Change*, New York: Oxford University Press.

Harada, N. (2003). Who Succeeds as an Entrepreneur? An Analysis of the Post-entry Performance of New Firms in Japan. *Japan and the World Economy* 15(2), 211-222.

Kumbhakar, S. & Knox , C.A. K. (2000). *Stochastic Frontier Analysis*, Cambridge: Cambridge University Press.

Robb, A. (2002). Entrepreneurial Performance by Women and Minorities: The Case of New Firms. *Journal of Developmental Entrepreneurship* 7(4), 383-397.

Robb, A. & DesRoches, D. (2008). *Kauffman Firm Survey: Baseline/First/Second Follow-Up*. Unpublished manuscript.

Figure 1 - An Illustration of Measuring Best Practice of the Entrepreneur

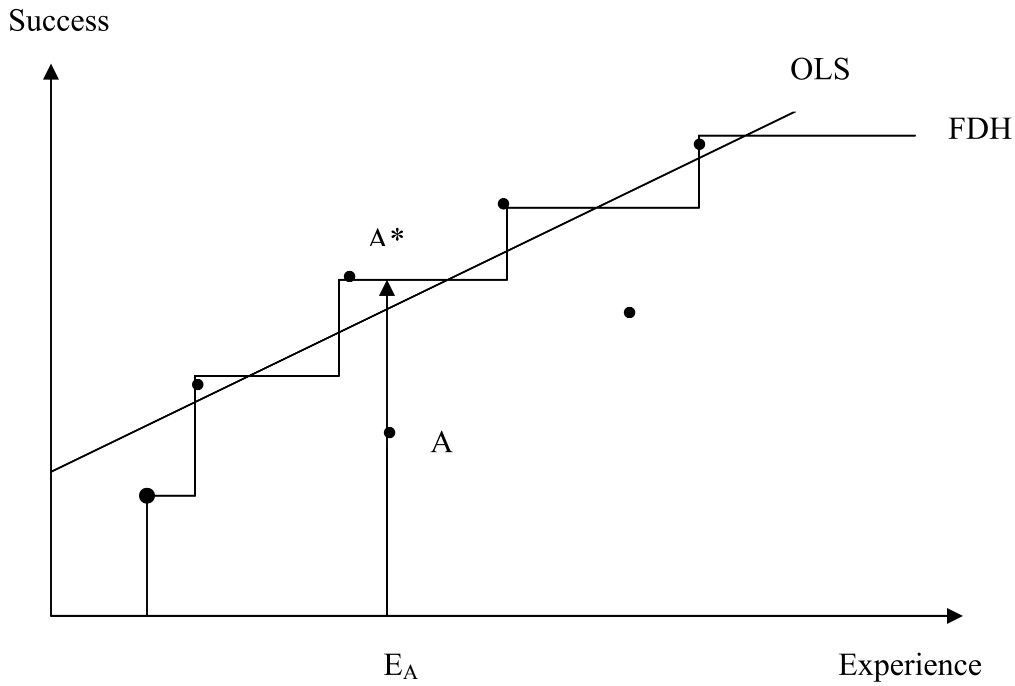

Figure 2 - The Evaluation of Firm A

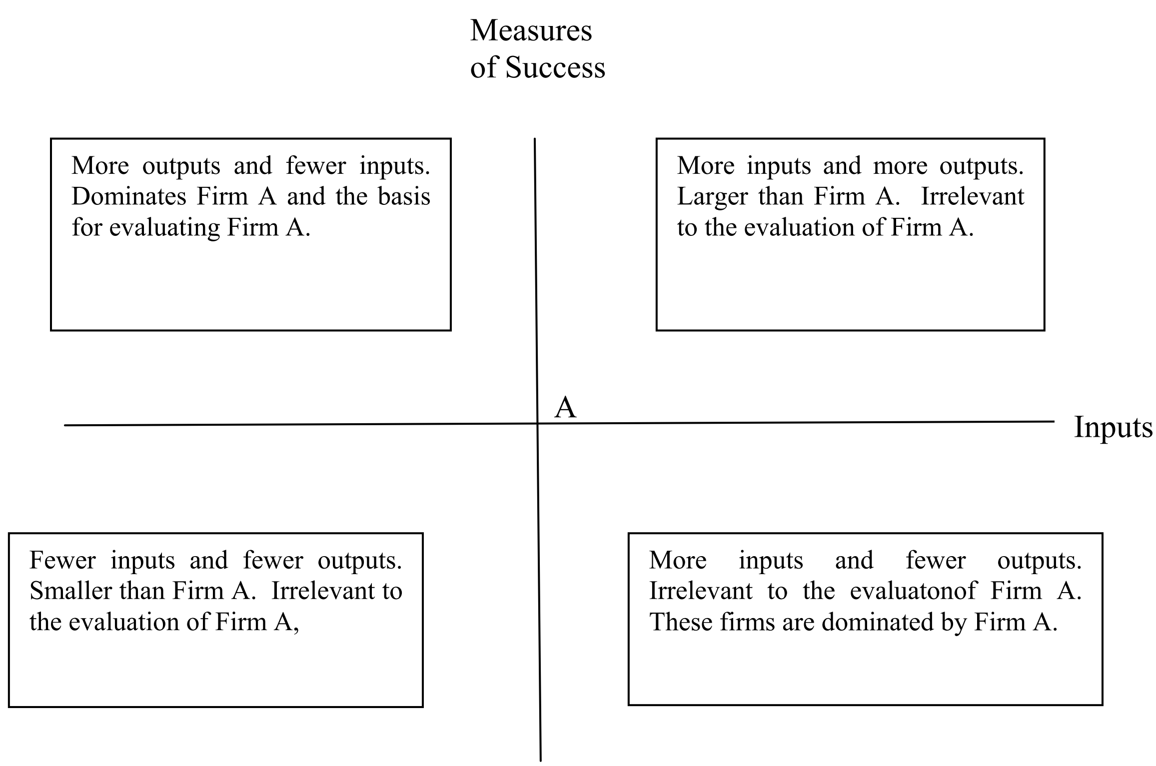

Figure 3 - Histogram of the FDH ESI

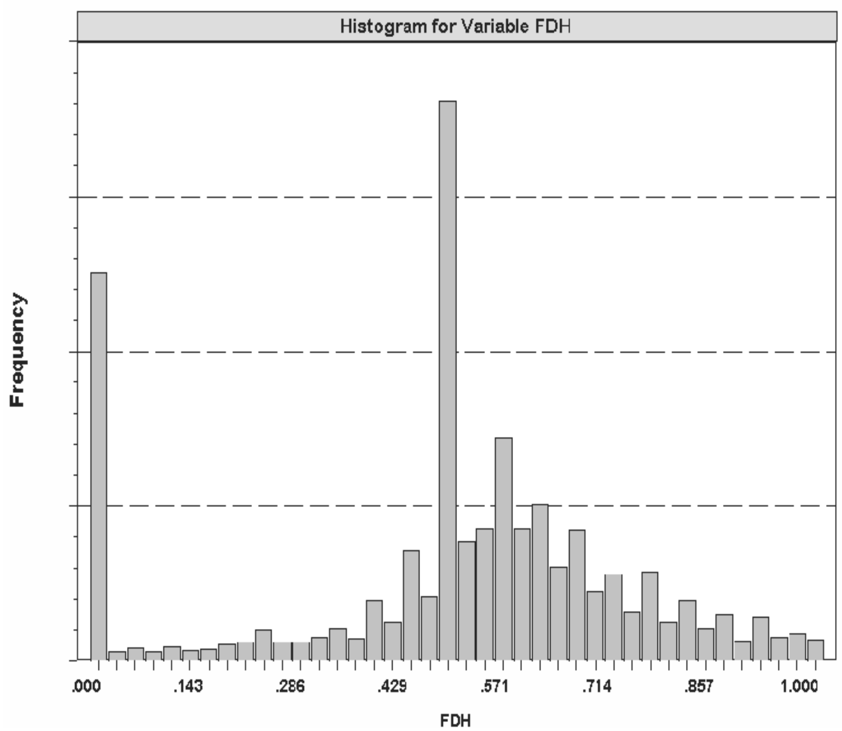

Figure 4 - Histogram of the Order-m ESI

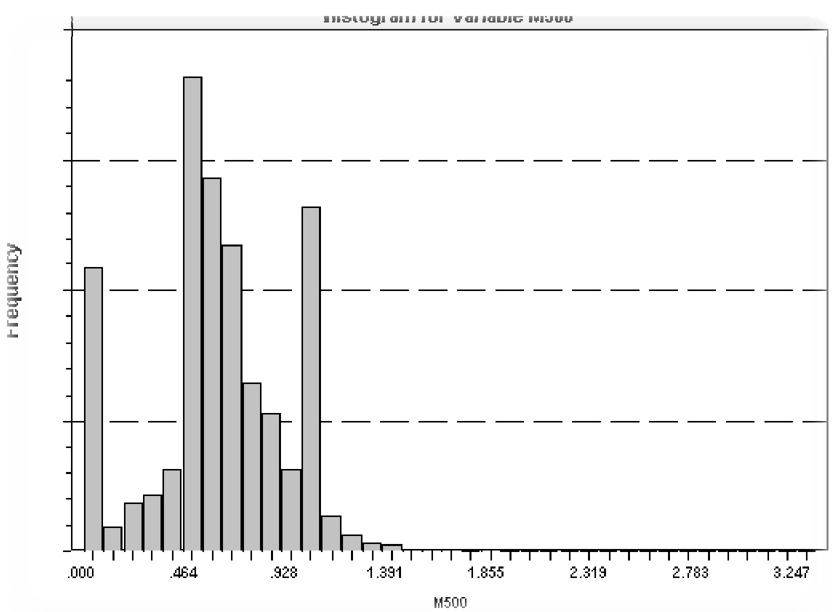

Figure 5 - The AER against Age (M=500 Results)

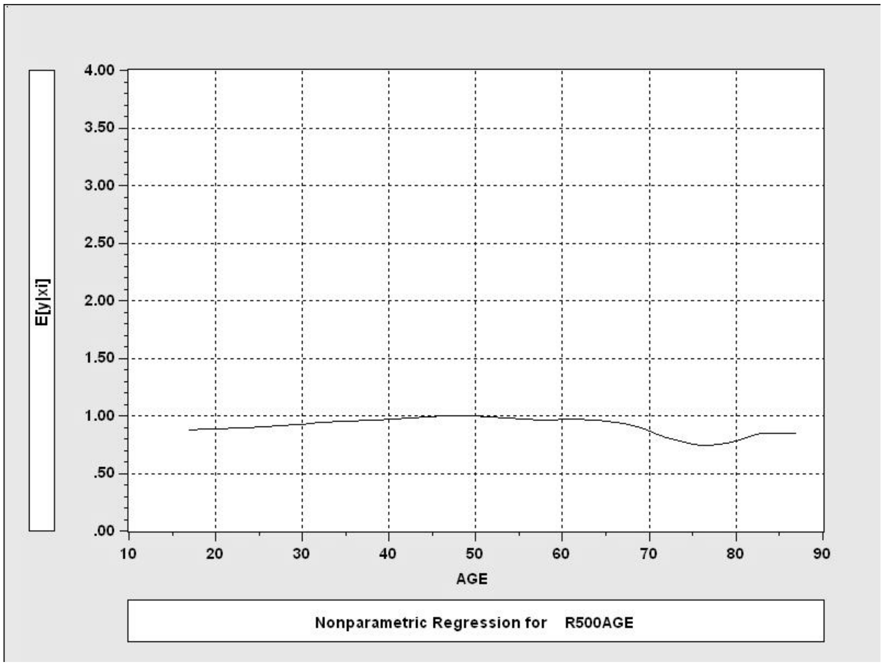

Figure 6 - The EER against Experience (M=500 Results)

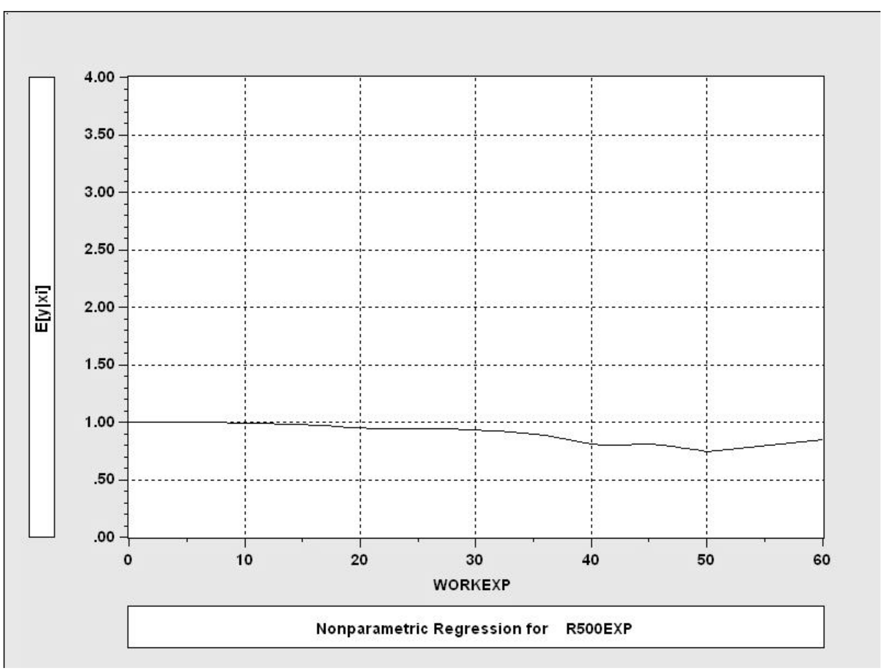

Table 1: Order-m ESI and FDH ESI[7]

	FDH	Order-m500
Mean	0.57	0.60
Standard Deviation	0.29	0.32

Table 2 - Comparing the Means of the Order-m ESI for Exogenous Categorical Variables

	FDH ESI		Order-m500 ESI		Number of Observations
	Mean	Standard Deviation	Mean	Standard Deviation	
Gender					
Men	0.56	0.29	0.60	0.33	2123
Women	0.54	0.30	0.55	0.31	693
Race					
White	0.56	0.29	0.59	0.32	2438
Other	0.54	0.32	0.56	0.34	377
Birthplace					
U.S.	0.55	0.29	0.58	0.32	2529
None U.S.	0.57	0.31	0.60	0.32	287
Education					
CollegeDegree+	0.57	0.29	0.61	0.33	1553
College minus	0.54	0.30	0.57	0.32	1310
Number of Owners					
One Owner	0.55	0.30	0.58	0.33	1670
More Owners	0.57	0.30	0.60	0.32	1193
Comparative Advantage					
Comp Adv	0.57	0.2	0.61	0.32	1921
No Com Adv	0.53	0.31	0.55	0.32	931
Produce a Product					
Yes	0.56	0.29	0.60	0.33	1402
No	0.55	0.30	0.58	0.32	1454

Table 3 - Truncated Weighted Regression of Order-m500 ESI against Causation Factors: Marginal Impacts

	Definition	Mean of Variable	Marginal Coefficient Estimate	Marginal Coefficient/Standard Error
Intercept			0.433***	5.38
Ownership	Percentage	78	-0.0002*	-1.78
Primary Owner Age	Years	45	0.008**	2.29
Age Squared			-0.0001**	-2.45
Work Experience	Years	12	-0.0002	-0.13
Experience Squared			-0.00001	-0.11
Other Owners	1=other owners	0.42	0.017	1.60
Race	1=White	0.85	0.007	0.48
Gender	1=Men	0.71	0.049***	4.103
Education	1=College and Plus	0.51	0.024**	2.34
Produce Product	1=Yes	0.49	-0.001	-0.091
Comparative Advantage	1=Yes	0.67	0.031***	2.79
U.S. Born	1=Yes	0.90	-0.021	-1.15
N after truncation	2456			

Note on significance: *** 1% level; ** 5% level; * 10% level.

[1] For a review of frontier methods, see Fried *et al* (2008).

[2] There is also the assumption of no error in measurement. This assumption is relaxed in the order-m methodology employed in this project and described below.

[3] Note that producing 60 percent of efficient outputs is equivalent to expanding each output by 67 percent.

[4] In practice, the distinction between FDH and DEA tends to be small when the number of observations is large since the staircase FDH frontier tends to have "small" steps. We have a large number of observations (2,864).

[5] Note that the two histograms have different scales on the horizontal axes.

[6] There is some evidence that the standard errors of regressions with efficiency scores as the dependent variable are biased, necessitating a bootstrap to remove the bias. We will explore this further in future work.

[7] For comparison purposes, we also computed DEA efficiency scores. The mean is 0.54 and the standard deviation is 0.27. The similarity with the FDH results is due to the large sample size; the FDH staircase frontier has short steps since the data is dense.

OTHER A.E.M. WORKING PAPERS

WP No	Title	Fee (if applicable)	Author(s)
2009-31	Equity Within and Between Nations		Kanbur, R. and A. Spence
2009-30	Middlemen, Non-Profits, and Poverty		Chau, N., Goto. H. and R. Kanbur
2009-29	Do retail coffee prices increase faster than they fall? Asymmetric price transmission in France, Germany and the United States		Gomez, M. and J. Koerner
2009-28	A Market Experiment on Trade Promotion Budget and Allocation		Gomez, M., Rao., V. and H. Yuan
2009-27	Transport, Electricity and Communications Priorities in Guinea Bissau		Kyle, S.
2009-26	The Macroeconomic Context for Trade in Guinea Bissau		Kyle, S.
2009-25	Cashew Production in Guinea Bissau		Kyle, S.
2009-24	When No Law is Better Than a Good Law		Bhattacharya, U. and H. Daouk
2009-23	Is Unlevered Firm Volatility Asymmetric?		Daouk, H. and D. Ng
2009-22	Conditional Skewness of Aggregate Market Returns		Charoenrook, A. and H. Daouk
2009-21	A Study of Market-Wide Short-Selling Restrictions		Charoenrook, A. and H. Daouk
2009-20	The Impact of Age on the Ability to Perform under Pressure: Golfers on the PGA Tour		Fried, H. and L. Tauer
2009-18	Beyond the Tipping Point: A Multidisciplinary Perspective on Urbanization and Development		Beal, J., Guha-Khasnobis, B. and R. Kanbur
2009-17	Economic Impacts of Soybean Rust on the US Soybean Sector		Gomez, M., Nunez, H. and H. Onal
2009-16	Detecting Technological Heterogeneity in New York Dairy Farms		del Corral, J., Alvarez, A. and L. Tauer
2009-15	Altering Milk Components Produced by the Dairy Cow: Estimation by a Multiple Output Distance Function		Cho, J., Nakane, M. and L. Tauer

Paper copies are being replaced by electronic Portable Document Files (PDFs). To request PDFs of AEM publications, write to (be sure to include your e-mail address): Publications, Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853-7801. If a fee is indicated, please include a check or money order made payable to Cornell University for the amount of your purchase. Visit our Web site *(http://aem.cornell.edu/research/wp.htm)* for a more complete list of recent bulletins.