Process Innovation, Product Innovation and Firm Size[*]

Xin Fang
Department of Economics
University of Illinois at Chicago
601 S. Morgan Street (MC 144)
Chicago, IL 60607
USA
fangxinzj@gmail.com
Telephone: 1-847-668-6629

Abstract

The determinants of a firm's decision to allocate Research and Development funds between process and product innovation are analyzed. Growth in size will initially increase the probability that a firm undertake product innovation. However, after the threshold size is crossed, the fraction of new products in total sales declines with further increase in firm size. Therefore, the overall relationship between share of product innovation and firm size is nonlinear. The empirical results using Mannheim Innovation Panel survey from ZEW show that in most industries, large firms have a larger share of new product sales.

.

JEL Codes: L25, L60
Keywords: Process Innovation, Product Innovation, Firm Size, Cannibalization, Threshold Size.

---

[*] Financial support from University of Illinois at Chicago is gratefully acknowledged. I also acknowledge the data support from Center for European Economic Research (ZEW). As this paper is based on some of my dissertation work, I am grateful to Dr. Richard Peck, Dr. Joshua Linn, Dr. Houston Stokes, Dr. Helen Roberts and Dr. Anthony Pagano at University of Illinois at Chicago for their invaluable suggestions and unwavering encouragement. I also thank Dr. Margaret S. Loudermilk and Dr. Quincy Jin for their helpful comments.

I. Introduction

Innovation is crucial for total factor productivity, industry and economic growth. As a spokeswoman of Motorola stated, “innovation is essential to staying competitive.” The importance of Research and Development (R&D) is underscored by the fact that companies invest about the same in economic downturns as they do in more prosperous times[1]. Firm size, in turn, is a key driver to R&D decisions. As Scherer (1980) states, “size is conducive to vigorous conduct of R&D.” Larger firms have better appropriability and benefit more from the cost reduction or quality increase due to their scale of operation (Schumpeter, 1950). This suggests that larger firms tend to be more innovative.

R&D investments are heterogeneous. An innovative firm may allocate R&D between process innovations that “reduce the cost of producing existing products”, and product innovations that “create new or significantly improved products”[2]. Firm size affects a firm’s allocation of resources between process and product innovation. Several earlier papers examine this issue theoretically and empirically. Most use a linear approach and conclude that the relationship between firm size and a firm’s innovation allocation is monotonic.

For example, Scherer (1991) finds a significant positive monotonic relationship between firm size (measured by sales) and the fraction of R&D effort devoted to process innovation based on the data for large U.S. companies. Fritsch and Meschede (2001) use the data collected by postal questionnaire from manufacturing enterprises in three German regions in 1995 and regress the amount of process R&D expenditures on firm size (measured as number of employees). They show that resources devoted to process R&D rise relatively more with sizes than expenditures on product R&D.

Cohen and Klepper (1996) model a firm’s internal allocation of resources between product and process innovation for price taking firms. Using the same data as Scherer (1991), they derive an indirect empirical evidence of a positive effect of firm size (measured by sales) on the share of process R&D (measured by the proportion of patents that have been classified as representing process innovation) under a constraint of marginal returns to R&D.

Some non-monotonic evidence of the relationship between the firm size and the allocation of innovation funds has also been detected. Pavitt et al (1987) use the data for 4,000 UK manufacturing industry innovations from 1945 to 1983 and calculate process innovation fractions of total R&D by innovating firm-size class. They find that the smallest innovators are least process-oriented, then followed by the largest firms. Midsized firms are most process-oriented.

In the literature discussed above, two critical factors are not taken into account to analyze the relationship between firm size and allocation of R&D resources: cannibalization and threshold size effect. The two factors suggest that a simple monotonic relationship is not sufficient to describe the effect of firm size. In addition, empirical investigations using nonlinear approaches remain largely unexplored. In this paper, I adopt the model of Cohen and Klepper (1996) and suggest a non-monotonic relationship of firm size and share of product innovation by incorporating these two factors. The hypothesis is tested empirically using German Innovation Survey data (1993~2004) from ZEW (Center for European Economic Research). Hurdle Model and Generalized Linear Model (QMLE approach suggested in Papke and Wooldridge, 1996) are used to examine the presence of nonlinearities. Besides fitting a general pattern for all the industries, I also run regressions for each individual industry. The findings confirm the non monotonic hypothesis and suggest large firms are generally more intensive in new product sales, which implies an increasing proportion of product R&D as firm size increases.

This paper is organized as follows: In the next section, the framework of Cohen and Klepper (1996) and the prediction of non-linear effect will be presented. Then the hypothesis is put forward based on the prediction and then empirically tested. Finally, I conclude this paper and discuss the implications.

II. Theory

Cohen and Klepper (1996) model R&D in the following way. Firms are assumed to be price takers and benefit from process innovations mainly through their ex ante output. It is also assumed that firms do not sell or license technologies. Therefore, the expansion of firm size due to process innovation is very limited. In particular, process innovation depends more heavily on the ex ante firm level than product innovation, which implies that larger firms benefit more than smaller ones from the process innovation.

Process innovations lower the production cost and increase the price cost margin, while product innovations raise the price of new or improved products and also the price-cost margin. Firms' positive profits created by innovations are assumed to be temporary. After a period, other firms in the same industry imitate the process or product innovation and profits fall to zero.

Two additional factors not considered in their model are: cannibalization and threshold size. Cannibalization refers to the situation that an incumbent firm creates competitors to its own established products when a new or improved product is invented. Similarly in the innovation literature, high substitutability between newer or improved products and established products will cause "creative destruction", or replacement effect as mentioned in Arrow (1962) and Tirole (1988).

All else equal, cannibalization will hurt large firms more. Assuming the same rate of cannibalization, larger firms will lose more established product sales because of the larger base. Therefore, cannibalization implies that the comparative advantages of larger firms in process innovations are weaker than predicted by Cohen and Klepper (1996). As a result, the proportion of R&D resources devoted to process innovation should not increase as much for larger firms, all else equal. Cannibalization cannot be incorporated in Cohen and Klepper's work due to their perfect competition and price taker setup. However, for a monopolistic competition structure or a model with capacity constraint and ladder shaped demand[3], cannibalization is possible.

Nelson (1967) suggests a threshold size that must be attained for a firm to successfully engage in innovation. He suggests that innovation is subject to significant economies of scale. Large-scale projects are more likely to be carried out by large firms because some programs are very costly. If firms are not large enough, their innovation efforts are nonexistent. Threshold size for innovation also might occur if there are borrowing constraints related to innovations. Aghion and Tirole (1994) and Lerner (2003) argue that financing problems can prevent a small firm from conducting product innovation alone, while cooperation with other firms could be problematic due to the allocation of ownership and control, especially when the capital requirement is substantial relative to the intellectual inputs. Large firms have more internal funds and better access to external funds, so that firm size increases raise the probability of successful innovation.

Threshold sizes for process innovation and product innovation are unlikely to be the same. For example, process innovation is more likely to be subject to a lower size threshold. This is because for an individual firm, the revenue of established products already exceeds or equals production cost. With the success of process innovation, cost will fall and positive profit margin will be enlarged. However, revenue from new products (including those significantly improved products) needs to cover not only the innovation cost, but also the production cost, thus requiring on the margin higher returns to product innovation. Moreover, product innovation projects are usually more costly than process innovation projects, such as the invention of a new drug or HDTV.

When a firm is only capable of undertaking process innovations, there is a corner solution with respect to product innovation. In this case, a firm size increase might make product innovation possible and thus move the firm out from its corner solution. Firm size affects a firm’s profitability. As Nelson (1967) points out, absolute gains from a given percentage of cost reduction or a given percentage improvement in product characteristics are larger for large firms, or can be realized faster. In addition, returns to innovations are closely related to the market size for the products. This is especially the case for product innovation in the presence of cannibalization. Cannibalization hurts process innovations and benefit product innovations more for large firms. This is because large firms have larger existing pool of consumers who will potentially switch from their own established products, resulting in a higher demand for their new products and thus higher returns to product R&D. In addition, with respect to borrowing constraints, being large will facilitate firms to get enough funds for their product innovation projects due to their access to more internal and external funds.

As a result, the manner in which allocation of resources between process R&D and product R&D varies with firm size suggested here is different from the prediction of Cohen and Klepper model (1996). Firm size increase does not necessarily lead to the drop of product R&D share. Initially, for very small firms, the share of product R&D would rise when threshold sizes are crossed, then fall as predicted by the traditional literature. The decline is expected to be less in size than predicted by Cohen and Klepper (1996a) because of the cannibalization effect.

The nonlinear and non-monotonic relationship between firm size and allocation of funds between two types of innovation offers a new explanation to reconcile the conflicts of the fact that large firms are more active new product inventors in contrast to the previous theoretical predictions. I will test the hypothesis empirically.

III. Empirical Tests

This paper argues shares of product innovation will initially rise with firm size due to the threshold sizes firms have for their product innovations. After the threshold sizes are crossed, shares of product innovation will tend to fall. Therefore, we will focus on modeling a nonlinear relationship of a firm’s choice between process and product R&D investments and firm size.

The effects of innovations are reflected in the sales of products. Production innovations can be represented by the sales from new or improved products, while process innovations are applied towards the established products. Therefore, the proportion of new or improved product sales to all sales (new product sales ratio) can be used as a proxy measure for the relative efforts made towards product innovation.

III (i) Data

Firm-level data from the Mannheim Innovation Panel Manufacturing and Mining & Services survey (MIP) that covers the years from 1993 to 2004 is used in the estimation procedure. The surveys were conducted annually by the Center for European Economic Research (ZEW) and focused on all firms located in Germany that have at least five employees and are active in the manufacturing and mining sectors. I use an unbalanced panel data with 11 broad industry categories according to German WZ93 classification for manufacturing and mining industries from year 1992 to year 2003. Participation in the survey is on a voluntary basis and the response rate is between 20.6-23.7% (Janz et al, forthcoming).

For each survey year, process innovations are referred to as any new or significantly improved production processes introduced during the preceding two years. Product innovations are new or markedly improved new products invented in the previous two years. Descriptive statistics are shown in Table I. In the survey, firms report their percentages of sales from new or markedly improved products (new product sales ratio). A firm with a zero new product sales ratio is treated as having no product innovations. For all the 5853 observations, there are 2455 observations of zero new product sales, 41.9% of the full sample. Box Plots of new product sales ratios of each industry in Figure 1 show that Mining has the lowest average new product sales ratio and high standard deviation. Chemicals, Machinery, Electrical Equipment, Medical and other Instruments, and Transportation Equipment can be observed to be active in product innovations.

<TABLE I here>

<Figure 1 here>

We categorize firm size in deciles for each industry and plot the new product sales ratio as shown in Figure 2. Most industries exhibit positive relationships between new product sales ratio and firm size. To control for the confounding effects of other variables, we need to run multivariate regressions to derive the effects of an increase in firm size.

<Figure 2 here>

III(ii) Empirical Model

As discussed previously, a threshold size effect implies that the probability of a firm engaging in product R&D rises with firm size, that is, a rising new product sales ratio initially when firms grow in size. Then new product sales ratio tends to fall with further firm size increase as a result of a firm's inclination to choose process R&D over product R&D. To test the hypothesis, new product sales ratio is estimated as a nonlinear function of firm size controlling for other factors. I expect to obtain a bell shaped curve.

ZEW surveys construct the variable for new product sales ratio. To protect the privacy of participating firms, instead of a value, a range is quoted within which the value of the company lies. Values of proportion of new and improved product sales are represented by discrete variables from 0 to 9.[3] I take the median of the intervals as the constructed dependent variable, denoted as $salesnp_t$:

$$(1) \qquad salesnp_t = \frac{npsales_t}{sales_t}$$

where $npsales_t$ denotes the sales from new and improved products at time t, and $sales_t$ represents the total sales of the firm at time t.

Number of employees is used to represent the size of firms ($size_t$). I use this instead of firm sales to avoid a mechanical relationship between sales and new product sales ratio ($salesnp_t$). Firm's decision on process and product innovation are assumed to be determined by the ex ante firm size. I lagged the firm size (number of employees) for two years to evaluate the size effect upon a firm's decision to innovate. The ex ante firm size is denoted as $size_{t-2}$.

Depreciation rate of established products is included in the regressions，denoted as dos_t. Because $sales_t$ is by functional form related to the dependent variable, I lag the variable for one period as shown in equation (2):

$$(2) \qquad dos_{t-1} = sales_{t-1} - npsales_{t-1} - sales_{t-3}$$

New products are developed during the preceding two years. Accordingly, $sales_{t-1} - npsales_{t-1}$ represent sales of products developed in and before *t-3* (established products) as of time *t-1*. $sales_{t-3}$ stands for sales of these established products at *t-3*.

Depreciation rate of established products is included for several reasons, First, cannibalization incorporated in our model work in the same direction as the obsolescence of established products for a firm brought on by its own new products. Given the product life cycle, a firm without product innovation and its resulting cannibalization may still face declining sales of its existing products if they are at late stage of product life cycle. This is because better substitutes produced by other firms will emerge as time passes (Brockhoff, 1967). Depreciation of established product sales is typically negative, as indicated below in TABLE II. It tends to make the new product sales ratio higher, so dos_{t-1} is negatively correlated with new product sales ratio. Absolute sales are also expected to decline more for larger firms. Therefore, the size effects of interest will probably be upward biased if this variable is not controlled. The results would be unreliable since larger firms end up with higher sales ratio of new products might because their sales of "old" products fall more.

The estimated coefficient of dos_{t-1} is expected to be negative. This variable will capture partial effects of $size_{t-2}$ since firm size is related to the change of established product sales.

However, there is no serious collinearity problem between $size_{t-2}$ and dos_{t-1} according to the variance inflation factors.

I also include fixed effect dummies for 12 years and 11 industries. Fixed effects control for specific factors, such as industry wide technology advances that are assumed exogenous in contrast to individual firm advances. Another factor controlled by industry dummies is market structure or competition measured as concentration ratio.

Firm's productivity in process and product innovation and other shocks to new product sales should be considered in estimating a variable related to productivity. However, due to the data limitation, I cannot use intermediate products or R&D investments towards new products as the proxy for productivity shocks and it would possibly bias the estimation of firm size coefficient if shocks have differential effects for firms' new product sales ratio with different sizes.

Letting *i* index industries, the following model will be estimated:

$$salesnp_{it} = \phi(size_{it-2}, dos_{it-1}) \qquad (3)$$

<TABLE II here>

III (iii) Estimation Methods

Small firms may not be capable of making profits from innovation because their size lies below the threshold size. Firm size increase raises the probability of positive new product sales, that is, extensive margin of firm size on new product sales ratio will be positive. For innovating firms over the threshold size and capable of both types of innovations, larger firms lean towards process innovation, causing a negative intensive margin, that is, new product sales ratio will drop eventually with the increase in firm size. This type of relationship justifies the usage of Hurdle Model, or two-tiered model (Wooldridge, 2002, P. 536-538) that allows estimation of extensive and intensive margins separately.

The estimation proceeds in two steps. In the first step to get the extensive margin γ, as shown in (4), marginal effect of firm size on the probability of product innovations given γ will be calculated, and coefficient of firm size is expected to be positive.

(4) $P(salesnp_{it} = 0 \mid size_{it-2}, dos_{it-1}) = 1 - \phi(\gamma_0 + \gamma_1 size_{it-2} + \gamma_2 dos_{it-1} + \varepsilon_{it})$

Conditioned on firms that have positive new product sales, new product sales ratio and other explanatory variables are assumed to follow a log normal distribution. β 's can be estimated using OLS regression in the second step as shown in (5). Intensive margin is expected to be negative, implying a lower new product sales ratio with firm size increase for product innovators.

(5) $\log(salesnp_{it}) = \beta_0 + \beta_1 size_{it-2} + \beta_2 dos_{it-1} + \varepsilon_{it}$

Fitted values for new product sales ratio can be derived from the following formula given by Wooldridge (2002, P537):

(6)
$P(salesnp_{it} \mid size_{it-2}, dos_{it-1}) = \phi(\gamma_0 + \gamma_1 size_{it-2} + \gamma_2 dos_{it-1}) \exp(\beta_0 + \beta_1 size_{it-2} + \beta_2 dos_{it-1} + \sigma^2 / 2)$

where the variance, σ^2, can be estimated consistently from the OLS estimator of log-normal model.

The overall relationship is predicted to be nonlinear and non monotonic with a bell shape expected. Therefore, other nonlinear specifications, such as OLS with squared term for firm size is also used:

(7) $salesnp_{it} = \beta_0 + \beta_1 size_{it-2} + \beta_2 size_{it-2}^2 + \beta_3 dos_{it-1} + \varepsilon_{it}$

Considering new product sales ratio a fraction between zero and one, I adopt Quasi Maximum Likelihood Estimation of Papke and Wooldridge (1996). The Generalized Linear Model estimated is as follows:

(8) $E(salesnp \mid size_{it-2}, dos_{it-1}) = G(\beta_0 + \beta_1 size_{it-2} + \beta_2 size_{t-2}^2 + \beta_3 dos_{it-1})$

where G(.) is a cumulative distribution function which fits the limit dependent variable situation. Logistic function and standard normal cumulative distribution function forms for G(.) will be both attempted. The derived standard errors are robust to possible variance misspecifications.

Our estimation will be organized in the following order: First, the model is estimated using a pooled sample and standard errors are clustered at industry level. In addition, structural difference tests justify the regressions for each individual industry to release the constraints on the slope coefficients and allow for industry specific pattern, regressions will be run for each industry with time fixed effects.

OLS regressions, followed by the Wooldridge approach of Hurdle Model and OLS and QMLE (Logit and Probit) with squared firm size term will be run for both pooled sample and individual industries.

III (iv) Results

TABLE III shows the estimation results for Hurdle Model of pooled sample. The Probit regression as the first part of the Hurdle Model yields a positive and significant extensive margin, or a positive marginal effect of firm size on firms' chances of success in product innovation. The second part of Hurdle Model shows that the intensive margin is negative and statistically significant in the Robust Regression, which suggests that new product sales ratio for firms with positive new product sales declines as firm size increases. The marginal effect of firm size at the mean firm size (347.2178) is 1.10514e-04 and the fitted new product sales ratio peaks for firms with 50,000 employees. The results from Hurdle model confirm our hypothesis. In addition, the new product sales ratio starts to fall only for extremely large firms, which means, large firms' sales are more intensive in new products most of the time.

<TABLE III here>

OLS and QMLE (Logit and Probit as the Link Functions in Generalized Linear Model) estimation results are shown in TABLE IV. In QMLE (Logit), G(.) is the logistic distribution function, while in QMLE (Probit), G(.) denotes the standard normal cumulative distribution function. A nonlinear relationship between size and new product sales ratio is shown in the results. The negative coefficient of squared term of size indicates negative marginal effect of firm size on new product sales ratio after a certain level of firm size. The coefficient is statistically significant in QLME (Probit) estimation.

<TABLE IV here>

At the sample mean of $size_{t-2}$ (*347.2178*), and dos_{t-1} (*-44.71403*), marginal effect of $size_{t-2}$ in OLS is 1.46e-05. Calculated values in QMLE (Logit) and QMLE (Probit) are 3.56e-06 and 5.58e-06, respectively. The positive marginal effect implies that an increase in firm size results in an increase in the new product sales ratio at mean size. Similar to the Hurdle model, marginal effects only become negative at very large firm size. By calculation, only firms with more than 42613 employees (QLME Logit) or 45714 employees (QLME Probit) experience a decline in new product sales ratio with firm size increase.

Group difference test (Appendix A) shows each industry has a distinct relationship between firm size and new product sales ratio. Therefore, regressions for each individual industry are run in the following part. Results are shown in TABLE V.

<TABLE V here>

In the Hurdle Model regressions, extensive margins are positive for all the industries. Therefore, a firm size increase tends to raise the probability of a firm engaging in product innovation. Coefficients are not significant at 10% for Mining and Food, Tobacco, Textiles. All the size coefficients as intensive margin estimates are negative. Most of them are statistically significant at 5%, except for Chemicals, which is significant at 10% level. Wood, Paper, Furniture, Plastics, Glass, Ceramics and Transport Equipment do not have significant coefficients. Therefore, larger firms with product innovations are associated with lower the new product sales ratio in most industries, as predicted in our hypothesis.

<TABLE VI here>

The marginal effects of firm size on new product sales ratio are shown in TABLE VI. Column (1) shows positive marginal effects calculated from the Probit estimation for the extensive margin at the mean values of firm size (shown in Column (2)) and other variables. The calculated marginal effects of Mining and Food, Tobacco, Textiles are not statistically significant. Column (3) shows the total marginal effects of firm size at mean values of firm size and other variables. Besides Food, Tobacco, Textiles, all other industries' total marginal effects at mean firm sizes are positive, although small.

<TABLE VII here>

The results for QMLE (GLM Logit) and QMLE (GLM Probit) are shown in Table VII. The coefficients for $size$ and $size^2$ are insignificant for Mining, Chemicals, and Electrical Equipment in QMLE (Logit) regressions. Possible collinearity between size and size^2 (.9746 for Mining, .9654 for Chemicals, .9413 for Electrical Equipment) or measurement error in the data could lead to this result. In the QMLE (Probit) regressions, Mining, Chemicals and Machinery have same situations. For Metal industry, in QMLE regressions, *size* coefficients are not significant, while the coefficients for $size^2$ are significant, also suggesting a nonlinear relationship between new product sales ratio and firm size. In contrast, Plastics industry has an insignificant size^2 coefficient with significant size variable, suggesting a linear relationship.

Industries with similar patterns of relationship between new product sales ratio and firm size are labeled with same group number. Group1 has negative signs for size and positive signs for size^2, suggesting U shaped curves. Group 2 has positive *size* coefficients and $size^2$ coefficients, implying bell shaped curves. Group 3 has negative signs for both variables. Therefore, for valid observations (with positive number of employees), only downward sloping curves will be observed.

The sizes where the marginal effects switch signs are calculated in TABLE VIII. In Group 1 (Mining and Machinery), marginal effects of firm size turn from negative to positive when firm size increases. However, the largest firm in Machinery has 22608.14 employees and the turning point exceeds the maximum. In this case, only negative marginal effects could be observed. In Group 3 (Chemical and Electrical Equipment), marginal effects will remain negative. Marginal effects in Group 2 switch from positive to negative after a certain firm size. The only exception is Transport Equipment industry, where the “threshold size” is greater than the maximum firm size in the industry and only negative marginal effects are present.

For all the industries with nonlinear curves for new product sales ratio, the “threshold size” is large (ranging from 87.5% to 99.4% of largest industry firm size). The findings are consistent with our previous evidence in the pooled regressions.

<TABLE VIII here>

IV. Conclusion

This study argues that a firm’s internal allocation of process and product R&D relates to its size in a nonlinear and non-monotonic way. In contrast to the previous literature, this work uses nonlinear techniques to test our hypothesis. Some evidence that supports the hypothesis is found. The results differ from the previous work also in that large firms invest relatively more in product R&D and relatively less in process R&D in most cases.

The existence of a threshold firm size for profitable product innovation leads to the possibility of zero expenditures on product innovation. Literature suggests that larger firms are more likely to invest and succeed in product innovation. For innovating firms, after taking into account cannibalization, the comparative advantage for larger firms in process innovation is weakened in the presence of new products. This leads to a smaller reduction in the share of product R&D than suggested by Cohen and Klepper (1996).

Composition of sales reflects firm's innovation investment choice and success, and is therefore related to firm size. Probability of product invention increases with firm size when firms are small due to the threshold size effect and the share of new products in total sales (new product sales ratio) tends to fall with further firm size increase. As a result, the new product sales ratio may be non monotonic in firm size.

The hypothesized relationship between new product sales ratio and firm size is tested with data from the MIP survey. Nonlinear econometric models including Hurdle and QMLE with polynomial terms are adopted. In the pooled regressions with full sample, the Hurdle model and QMLE estimates support the hypothesis. In the regressions for each industry, individual regressions reveal obvious nonlinear patterns for most of the industries. In addition, except for Chemicals and Electrical Equipments, the sizes at which the marginal effect switches signs are large for each industry.

Fitted values of new product sales ratio curve and calculated firm size where new product sales ratio reaches maximum show that the share of new products in total sales actually increases with firm size until firm becomes very large. Two factors might be considered responsible: first, threshold size effect dominates the scale effect of firm size; second, for firms undertaking product innovations, cannibalization effect reduces the comparative advantage of larger firms in process innovation, which enhances the positive impact of firm size on new product sales ratio. Therefore, the "threshold sizes" are very large and for most industries, larger firms excel in product innovation, both in absolute and relative terms.

As this paper argues, a monotonic and linear relationship is not sufficient to explain the relationship of firm size and firms' innovation choices. The results of this paper show that it might not be desirable to break a large firm into small pieces. Large firms invest more in proportion in new products and large in size might be desirable if in the sense of bringing more new products into existence.

Footnote:

[1] Justin Scheck, Paul Glader. Wall Street Journal. (Eastern edition). New York, N.Y.: Apr 6, 2009. pg. A.1

[2] Tirole, J., 1988. *The Theory of Industrial Organization*. MIT Press.

[3] In 'Process Innovation, Product Innovation and Firm Size', Dissertation, Xin Fang, 2009.5

[4] 0: 0% 1: <5% 2: <10% 3: <15% 4:<20% 5: <30% 6: <50% 7:<75 8:<100% 9:100%

References

Arrow K. J., 1962, 'Economic Welfare and the Allocation of Resources for Invention', in *Universities National Bureau Committee for Economic Research, The Rate and Direction of Inventive Activity* (Princeton University Press, Princeton).

Aghion P. and Tirole, J., 1994, 'On the Management of Innovation', *Quarterly Journal of Economics*, 109, pp. 1185–1207.

Brockhoff K, 1967, 'A Test for the Product Life Cycle', *Econometrica*, 35, pp. 472-484.

Cohen W.M. and Klepper S., 1996, 'Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D', *The Review of Economics and Statistics*,78, pp.232-243.

Fritsch M. and Meschede M., 2001, 'Product Innovation, Process Innovation, and Size', *Review of Industrial Organization*, 19, pp.335-350.

Janz N., Ebling G., Gottschalk S. and Niggemann H., forthcoming, *The Mannheim Innovation Panels (MIP and MIP–S) of the Centre for European Economic Research (ZEW)* (Schmollers Jahrbuch).

Lerner J., Shane H. and Tsai A., 2003, 'Do equity financing cycles matter? Evidence from biotechnology alliances', *Journal of Financial Economics*, 67, pp. 411-446.

Levin R.C., Klevorick A., Nelson N., and Winter S., 1986, *Survey Research on R &D Appropriability and Technological Opportunity, Part I: Appropriability* (Yale University).

Nelson R.R., Peck M.J. and Kalacheck E.D., 1967, *Technology, Economic Growth, and Public Policy* (Brookings, Washington, D.C.).

Papke L.E. and Wooldridge J.M., 1996, 'Econometric Methods for Fractional Response Variables with An Application to 401(K) Plan Participation Rate', *Journal of Applied Econometrics*, 11, pp. 619-632.

Pavitt K., Robson M., and Townsend J., 1987, 'The Size Distribution of Innovating Firms in the UK: 1945-1983.' *The Journal of Industrial Economics*, 35, pp. 297-316.

Scherer F.M., 1980, *Industrial Market Structure and Economic Performance* (Rand McNally, Chicago).

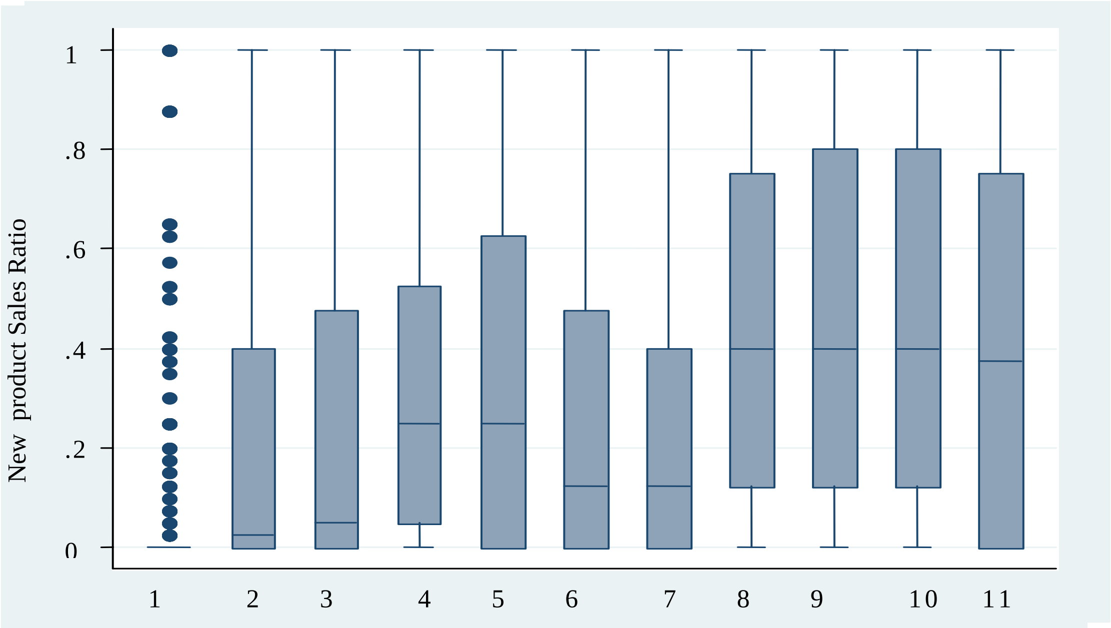
Scherer F.M., 1991, ‘Changing Perspectives on the Firm Size Problem’, in *Innovation and technological change: an international comparison* (Havester Wheatsheaf, New York).

Schumpter J.A., 1950, *Capitalism, Socialism, and Democracy* (Harper and Row, New York).

Tirole J., 1988, *Theory of industrial organization* (The MIT Press).

Wooldridge J.,2002, *Econometric Analysis of Cross Section and Panel Data* (The MIT Press, Cambridge, Massachusetts).

Figure 1 Box Plots by Industry



1. Mining
2. Food, Tobacco, Textiles
3. Wood, Paper, Furniture
4. Chemicals
5. Plastics
6. Glass, Ceramics
7. Metals
8. Machinery
9. Electrical Equipment
10. Medical and Other Instruments
11. Transport Equipment

Figure 2 New Product Sales Ratio with Firm Size in Deciles by Industry

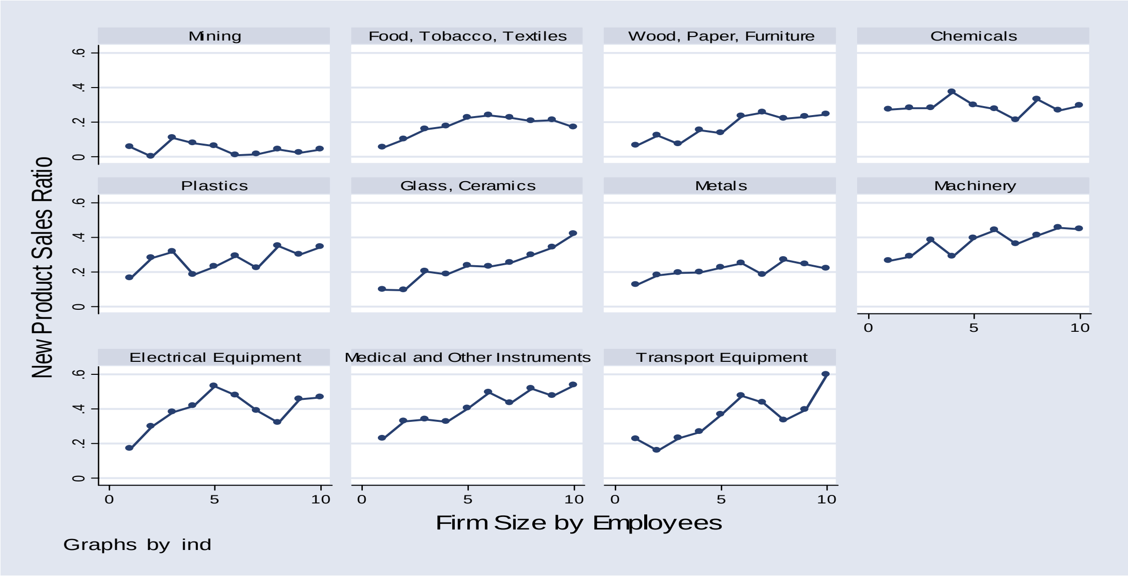


TABLE I
SAMPLE STATISTICS

Industry	Number of Observations	Average New Product Sales Ratio	Average Firm Sales	Average Number of Employees (When R&D is decided)
Mining	921	0.065 (0.194)	441.479 (1898.305)	1905.466 (8512.738)
Food, Tobacco, Textiles	2,995	0.240 (0.329)	75.012 (232.595)	179.125 (388.146)
Wood, Paper, Furniture	2,313	0.260 (0.338)	76.902 (341.776)	206.334 (430.107)
Chemicals	1,782	0.338 (0.315)	445.623 (2995.195)	999.354 (8583.367)
Plastics	1,930	0.321 (0.335)	59.524 (296.524)	178.112 (323.684)
Glass, Ceramics	1,244	0.268 (0.330)	111.422 (688.298)	518.267 (2881.219)
Metals	3,652	0.255 (0.315)	146.046 (1921.049)	294.982 (2477.579)
Machinery	4,073	0.435 (0.346)	134.612 (568.781)	429.864 (1376.103)
Electrical Equipment	2,140	0.445 (0.353)	273.602 (2063.080)	866.955 (6277.828)
Medical and Other Instruments	1,761	0.453 (0.341)	74.198 (449.320)	257.815 (1144.719)
Transport Equipment	1,242	0.402 (0.356)	915.723 (5223.035)	1900.374 (9631.531)
Total	24053	0.333 (0.345)	199.220 (1819.106)	529.070 (4238.647)

TABLE II

SUMMARY STATISTICS

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
salesnp	17749	.3326258	.3454721	0	1
$size_{t-2}$	11313	529.0701	4238.647	.6615753	174604.6
dos_{t-1}	6621	-52.72846	878.7937	-33841.62	8327.23

TABLE III
ESTIMATION RESULTS OF HURDLE MODEL[a]

Variable	**Probit[b] (Hurdle I)**	**OLS of Log-Normal[c] (Hurdle II)**	**OLS of Log-Normal[d] (Hurdle II: Robust)**
size_{t-2}	6.72e-04**	-2.06e-05	-4.12e-05**
	(1.42-e-04)	(1.53e-05)	(6.74e-06)
dos_{t-1}	4.87e-05	-8.51e-05	-1.49e-04**
	(7.05e-05)	(5.29e-05)	(2.30e-05)
constant	-2.65**	-1.09e-00**	-2.39e-00**
	(2.22e-01)	(4.06e-03)	(8.71e-01)
R^2	0.1266	0.2338	-
DF	5752	3297	5752

[a] The quantities in (.) below estimates are the standard errors here and after.
* means the coefficient is significant at 10% level, but not 5% level.
** means the coefficient is significant at 5% level.
[b] Part I of Hurdle Model: Probit (with binary dependent variables) regression
[c] Part II of Hurdle Model: Log-Normal OLS regression
[d] Part II of Hurdle Model: Log-Normal Robust regression

TABLE IV
OLS AND QMLE RESULTS[a]

Variable	**OLS[b]**	**QMLE[c] (Logit)**	**QMLE[d] (Probit)**
size_{t-2}	1.46e-05**	7.73e-05**	4.48e-05**
	(4.30e-06)	(3.22e-05)	(1.39e-05)
size^2_{t-2}	-1.52e-10**	-9.07e-10	-4.90e-10**
	(3.84e-11)	(6.13e-10)	(1.72e-10)
dos_{t-1}	-8.71e-06	-4.05e-05	-2.34e-05
	(9.87e-06)	(4.90e-05)	(2.68e-05)
constant	2.59e-01**	-4.58**	-2.33e**
	(1.04e-03)	(2.01e-02)	(8.06e-03)
R^2	0.2088	-	-
DF	5751	5751	5751

[a] Note: The quantities in (.) below estimates are the OLS standard errors or, for QMLE, the GLM standard errors robust to variance misspecification.
* means the coefficient is significant at 10% level, but not 5% level.
** means the coefficient is significant at 5% level.
[b] OLS regression
[c] QMLE (Logit) regression
[d] QMLE (Probit) regression

TABLE V

HURDLE MODEL REGRESSIONS [a]

Industry	**Probit (Hurdle I)**	**# of obs (Degrees of Freedom)**	**R^2 (Pseudo)**	**Log-Normal (Hurdle II Robust Regression)**	**# of obs (Degrees of Freedom)**	**R^2**	**Ratio of positive new product sales**
Mining	1.86e-04 (1.70e-04)	145 (134)	0.1229	-1.68e-03** (2.17e-04)	16 (6)	0.6213	.124
Food, Tobacco, Textiles	1.91e-04 (2.23e-04)	742 (731)	0.0182	-7.20e-04** (3.37e-04)	304 (293)	0.2240	.410
Wood, Paper, Furniture	9.24e-04** (2.45e-04)	449 (438)	0.0637	-6.20e-05 (1.68e-04)	179 (168)	0.1200	.399
Chemicals	4.59e-04** (1.79e-04)	445 (434)	0.0454	-1.11e-04* (5.67e-05)	298 (287)	0.2343	.655
Plastics	1.30e-03** (3.83e-04)	548 (537)	0.0689	-1.38e-04 (1.90e-04)	311 (300)	0.2241	.568
Glass, Ceramics	1.74e-03** (5.51e-04)	320 (309)	0.1125	-3.74e-04 (2.50e-04)	164 (153)	0.1755	.516
Metals	5.38e-04** (1.36e-04)	891 (880)	0.0671	-2.90e-04** (8.28e-05)	411 (400)	0.2192	.461
Machinery	1.01e-03** (3.05e-04)	1015 (1004)	0.0909	-7.36e-05** (1.88e-05)	739 (728)	0.2585	.728
Electrical Equipment	1.68e-03** (3.89e-04)	525 (514)	0.0743	-8.69e-05** (2.03e-05)	410 (399)	0.3003	.781
Medical and Other Instruments	2.92e-03** (7.52e-04)	506 (495)	0.1277	-2.09e-04** (1.03e-04)	380 (369)	0.2339	.755
Transport Equipment	7.13e-04** (3.00e-04)	267 (256)	0.1232	-3.50e-05 (2.76e-05)	180 (169)	0.2887	.678

[a] The quantities in (.) below estimates are standard errors.
* means the coefficient is significant at 10% level, but not 5% level.
** means the coefficient is significant at 5% level.

TABLE VI

MARGINAL EFFECTS FOR HURDLE MODEL[a]

Industry	**(1)** **Probit Marginal Effect (Hurdle I)**	**(2)** **Mean of size**	**(3)** **Hurdle Marginal Effect** $\partial E[y/x]/\partial x$
Mining	3.14e-05 (4.00e-05)	337.498	8.61e-06
Food, Tobacco, Textiles	7.43e-05 (7.00e-05)	136.486	-7.29e-05
Wood, Paper, Furniture	3.56e-04** (9.00e-05)	190.924	1.67e-04
Chemicals	1.63e-04** (5.00e-05)	337.982	3.61e-05
Plastics	5.02e-04** (1.30e-04)	154.024	2.23e-04
Glass, Ceramics	6.79e-04** (1.80e-04)	209.385	2.89e-04
Metals	2.14e-04** (6.00e-05)	209.815	3.28e-05
Machinery	2.80e-04** (4.00e-05)	390.339	1.08e-04
Electrical Equipment	1.55e-04** (3.00e-05)	757.239	3.30e-05
Medical and Other Instruments	6.25e-04** (7.00e-05)	222.072	3.21e-04
Transport Equipment	1.24e-04** (3.00e-05)	1503.61	6.41e-05

[a] The quantities in (.) below estimates are standard errors.

* means the coefficient is significant at 10% level, but not 5% level.

** means the coefficient is significant at 5% level.

TABLE VII

QMLE (LOGIT&PROBIT) REGRESSIONS[a]

Industry	**Size$_{t-2}$ (Logit)**	**Size$_{t-2}$ (Probit)**	**Size$_{t-2}^2$ (Logit)**	**Size$_{t-2}^2$ (Probit)**	**# of obs (Degrees of Freedom)**
Mining^1	-9.62e-04	-1.81e-04	1.44e-07	2.03e-08	145
	(1.47e-03)	(6.98e-04)	(2.17e-07)	(1.56e-07)	(133)
Food, Tobacco, Textiles^2	3.77e-03**	1.85e-03**	-5.97e-06**	-2.68e-06**	742
	(1.34e-03)	(6.19e-04)	(2.54e-06)	(9.81e-07)	(730)
Wood, Paper, Furniture^2	1.99e-03**	1.14e-03**	-8.37e-07**	-4.43e-07**	449
	(5.46e-04)	(2.73e-04)	(4.00e-07)	(1.46e-07)	(437)
Chemicals^3	-4.38e-05	-2.48e-05	-2.88e-09	-1.20e-09	445
	(1.65e-04)	(9.28e-05)	(1.74e-08)	(9.37e-09)	(434)
Plastics^2	9.07e-04**	5.36e-04**	-4.60e-07	-2.68e-07	548
	(4.55e-04)	(2.66e-04)	(3.03e-07)	(1.74e-07)	(537)
Glass, Ceramics^2	1.08e-03**	6.79e-04**	-5.04e-07**	-2.92e-07**	320
	(3.62e-04)	(2.07e-04)	(1.03e-07)	(5.35e-08)	(308)
Metals^2	2.75e-04	1.85e-04	-1.59e-07**	-9.06e-08**	891
	(2.01e-04)	(1.16e-04)	(5.75e-08)	(3.17e-08)	(879)
Machinery^1	-3.71e-05	-1.83e-05	3.46e-10	1.59e-11	1015
	(6.39e-05)	(3.71e-05)	(4.23e-09)	(2.72e-09)	(1003)
Electrical Equipment^3	-7.05e-05	-3.92e-05*	-3.70e-10	-2.04e-10**	525
	(3.97e-05)	(2.10e-05)	(1.90e-10)	(1.00e-10)	(513)
Medical and Other Instruments^2	7.11e-04**	4.38e-04**	-4.65e-08**	-2.85e-08**	506
	(2.02e-04)	(1.12e-04)	(1.08e-08)	(6.43e-09)	(494)
Transport Equipment^2	2.18e-04**	1.31e-04**	-3.76e-09**	-2.27e-09**	267
	(6.90e-05)	(3.87e-05)	(1.31e-09)	(7.42e-10)	(255)

[a] The quantities in (.) below estimates are standard errors. The subscripts are group numbers for industries.

* means the coefficient is significant at 10% level, but not 5% level

** means the coefficient is Significant at 5% level.

TABLE VIII

FIRM SIZE FOR MARGINAL EFFECTS SIGN CHANGE

Industry	**Sign Change Firm Size (Logit)**	**Sign Change Firm Size (Probit)**	**Maximum Firm Size**	**Sign Change Size/Maximum Industry Size**
Mining[1]	3.34e+03	4.46e+03	71749.13	95.4%
Food, Tobacco, Textiles[2]	3.16e+02	3.45e+02	5039.781	87.5%
Wood, Paper, Furniture[2]	1.19e+03	1.29e+03	6063.933	97.9%
Chemicals[3]	-7.60e+03	-1.03e+04	174604.6	-
Plastics[2]	9.86e+02	1.00e+03	3167.575	97.4%
Glass, Ceramics[2]	1.07e+03	1.16e+03	49259.08	94.8%
Metals[2]	8.65e+02	1.02e+03	102359.4	95.5%
Machinery[1]	5.36e+04	5.75e+05	22608.14	-
Electrical Equipment[3]	-9.53e+04	-9.61e+04	128319.4	-
Medical and Other Instruments[2]	7.65e+03	7.68e+03	18429.68	99.4%
Transport Equipment[2]	2.90e+04	2.89e+04	157435	-

APPENDIX A

We use Wald Test, and Likelihood Ratio Test to test the significance of difference.

The Likelihood Ratio Test Statistics is constructed as mentioned in Green (2002):

(9) $LR = -2(ln\text{-}LR - ln\text{-}LU)$

ln-LR stands for the restricted likelihood ratio, which in our model is the likelihood for the pooled regression. *ln-LU* is the unconstrained likelihood ratio, which is calculated as the sum of all the likelihood ratios for individual industry regression. LR is compared with chi-squared table to check the significance of constraints.

GROUP DIFFERENCE TESTS[a]

	$Size_{T-2}$ ($Size_{T-2}$ & $Size_{T-2}^2$)	Total
Probit (Hurdle I)	chi2(1) = 22.39 Prob > chi2 = 0.0000	chi2(11) = 442.58 Prob > chi2 = 0.0000
Log-Normal (Hurdle II) Robust Regression	F(1, 3376) = 11.37 Prob > F = 0.0008	F(10, 3376) = 8.90 Prob > F = 0.0000
GLM (Logit)	chi2(2) = 8.12 Prob > chi2 = 0.0173	chi2(12) = 453.49 Prob > chi2 = 0.0000
GLM (Logit)	chi2(2) = 14.63 Prob > chi2 = 0.0007	chi2(12) = 471.02 Prob > chi2 = 0.0000

[a] Group I: Mining, Food, Tobacco, Textiles, Wood, Paper, Furniture, Plastics, Glass, Ceramics, Metals

Group II: Chemicals, Machinery, Electrical Equipment, Medical and Other Instruments, Transport Equipment