International Journal of Business and Economics, 2009, Vol. 8, No. 1, 7-19

Relation of Firm Size to R&D Productivity

Jinyoung Kim
Department of Economics, Korea University, Korea

Sangjoon John Lee*
College of Business, Alfred University, U.S.A.

Gerald Marschke
Department of Economics, University at Albany—SUNY, U.S.A.
NBER and IZA

Abstract

Many studies have shown that small firms generate more patents per R&D dollar than large firms. Does this mean that small firms are more efficient innovators than large firms? In this paper we exploit a unique data set to reexamine the firm size-innovation relationship. Because firm-reported R&D expenditures may be a biased measure of R&D activities due to under-reporting by small firms, we use the number of inventors in the firm's employ as a measure of R&D inputs. We focus on the pharmaceutical and semiconductor industries, two industries that are prolific generators of homogenous innovations. As has been found elsewhere in the literature, we find that patents per R&D dollar decline with firm size for both industries. This contrasts with the relationship between patents per inventor and firm size. The average number of patents per inventor increases with size in the semiconductor industry. In the pharmaceutical industry, we find no relationship between the number of patents produced per inventor and firm size.

Key words: patents; innovation; labor productivity; research; firm size

JEL classification: O30; O32; O34; J21; J24

1. Introduction

Many empirical studies of the relationship between innovative output and firm size have found that small firms generate more innovations per dollar of R&D. Researchers have puzzled whether this is evidence that small firms are more efficient innovators than large firms. A credible alternative explanation is that small

Received October 29, 2008, revised August 4, 2009, accepted August 6, 2009.
*Correspondence to: College of Business, Alfred University, 1 Saxon Drive, Alfred, NY 14534, U.S.A. E-mail: sangjoon@alfred.edu. Kim was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2007-332-B00062). Lee is grateful to the National Bureau of Economic Research for the NBER-Sloan Science and Engineering Workforce Project Fellowship.

firms under-report their R&D expenditures. Understanding the innovation enterprise, a proven and powerful engine for improving living standards in developed economies, is essential for crafting effective government policies. If, as these empirical studies suggest, small firms are more efficient innovators, then perhaps the enormous government R&D subsidies that presently go mainly to large firms are misallocated. If instead large firms or firms that control a large share of their market are more productive, then it calls into question those antitrust policies that aggressively attempt to limit the size of dominant firms in markets. Such a finding would suggest that aggressive antitrust enforcement comes at the expense of more innovation.

In this paper we exploit a unique data set to examine the firm size-innovation relationship. Because firm-reported R&D may be a biased measure of R&D activities, we use the number of inventors in the firm's employ—which we identify from the inventor field on patent applications—as a measure of R&D inputs. While inventors as a measure of R&D inputs exclude capital and other non-labor inputs, they have the potential advantage of being less subject to the size-related under-reporting problem. We focus on the pharmaceutical and semiconductor industries, two industries that are prolific generators of innovations and patents. As has been found elsewhere in the literature, we find that patents per R&D dollar decline with firm size for both industries. This contrasts with the relationship between patents per inventor and firm size. The average number of patents per inventor increases with size in the semiconductor industry. In the pharmaceutical industry, we find no relationship between the patent-inventor relationship and size.

The paper is organized as follows. The next section summarizes the literature that measures the variation in innovation across firms of different sizes and discusses the challenges of measuring firm R&D inputs and outputs. Section 3 describes our data and spells out the empirical estimation method. Section 4 discusses the empirical results. Section 5 offers some concluding remarks.

2. Literature Review

Since Schumpeter's 1942 seminal work, economists have been interested in the relationship between a firm's R&D enterprise and its size. Schumpeter (1976, Chapter VIII) argued that firms have an advantage in R&D in the markets in which they have market share because monopoly power enables them to capture the returns to innovation. Schumpeter and others (e.g., Cohen and Klepper, 1996; Panzar and Willig, 1981) have suggested that size may be an important determinant of R&D because of either economies (due to fixed costs to mounting an R&D effort) or diseconomies of scale (e.g., because of inflexibility imposed by bureaucratization of the R&D enterprise) and because size may enable better access to external sources of financing. Large firms with many product lines may be better able to exploit unexpected innovations.

A large body of empirical research has examined the relationship between the productivity of R&D in firms and firm size. Such research has examined how patent (or citation-weighted patent) yields from R&D activities (usually measured by R&D expenditures) vary with firm size. While the empirical literature fails to generate a consensus view, a number of studies report that the patent yield from R&D expenditures falls with firm size (see, for example, Acs and Audretsch, 1991; Bound et al., 1984; and Hausman et al., 1984). Some scholars, however, have found an advantage to size (e.g., Cohen and Klepper, 1996), others have found that both large and small firms have higher R&D productivity (Tsai and Wang, 2005), and still others have found that there is no relationship between size and R&D productivity; see Symeonides (1996) for a survey.

Another line of research has investigated the relationship between the investment in the R&D enterprise and firm size. The reasoning is that if large firms enjoy an R&D productivity advantage over small firms, say, then we should observe that R&D expenditures in large firms are proportionally larger than in small firms. While there are exceptions, the literature has found no overall effect of size on R&D intensity (e.g., Acs and Audretsch, 1987, 1990; Bound et al., 1984; Cohen et al., 1987; and Scherer, 1984). That is, it appears that overall R&D intensity increases with firm size more or less proportionately. Patel and Pavitt (1992) found variation in this relation across manufacturing industries: in 13 of the 16 industries they analyzed (including the pharmaceutical industry) they found that R&D intensity rose proportional to firm size, while in the three remaining industries (chemicals, mining, and motor vehicles) they found R&D expenditures rose more than proportional to firm size.

Many scholars have discussed the limitations of the various measures of R&D inputs and outputs, including patents for output and R&D expenditures and scientific employment for inputs (e.g., see Cohen and Levin, 1989). R&D expenditures if measured correctly are a relatively good measure of inputs in the innovation production function since it includes both capital and labor inputs. Moreover, R&D expenditures capture both the quantity and quality of R&D inputs. Because small firms often conduct their R&D informally, lacking R&D departments and separate R&D staff, however, they likely under-report their R&D expenditures (Kleinknecht, 1987; Kleinknecht and Verspagen, 1989; and Schmookler, 1972). This under-reporting could account for some or all of the apparent size disadvantage in innovative productivity.

This paper considers the number of the firm's research scientists—as identified in the US patent records—as an alternative measure of the firm's R&D inputs. While the number of research scientists excludes capital and other labor and non-labor inputs, it has the potential of being less sensitive to the size-related under-reporting problem. Also, R&D expenditures include purchase prices of durable capital equipment rather than its depreciation. These facts give the number of research scientists the advantage of representing an input flow (see Cohen and Levin, 1989, for a discussion of the strengths and weaknesses of R&D expenditures and employment as measures of R&D activity).

3. Empirical Analysis and Data Creation

We use panel data on firms in the pharmaceutical and semiconductor industries to test whether firm size is a determinant of R&D productivity. We use the patent production function approach, which was first introduced by Pakes and Griliches (1980) to determine firm's R&D productivity. In our basic specification we test whether our measure of R&D productivity is sensitive to firm size, holding constant other firm characteristics, such as the firm age, the capital-labor ratio, the number of product lines, and a citation-based measure of patent quality.

Our data are taken from five sources: (1) Patent Bibliographic data (Patents BIB) released by the US Patent and Trademark Office (USPTO) that contain bibliographic information for US utility patents issued from 1969 to 2002, (2) Compact D/SEC database since 1989 which contains firm information taken primarily from 10-K reports filed with the Securities and Exchange Commission, (3) Standard & Poor's Annual Guide to Stocks—Directory of Obsolete Securities which includes a history of firm name changes, (4) NBER Patent-Citations data collected by Hall et al. (2001) which contain all citations made by patents granted in 1975–1999, and (5) Thomas Register data which report the firm's founding year.

As the first step for creating the sample we use for our analysis, we choose all firms whose primary SIC code is 2834 (pharmaceutical preparation) or 3674 (semiconductor and related devices) in the Compact D/SEC data. We select these two industries for our study because firms in these industries are active in patenting and produce homogenous products relative to other industries. By limiting our analysis to two relatively homogenous industries, we avoid problems due to the incomparability in utility and marketability of innovations, and in patent propensities across industries. Because patents are typically assigned to the firm (the *assignee*) that employs the inventors, we identify the inventors' employers in the Patents BIB data by patent assignees. Note that we select only the years 1989 through 1997 for our study because the Compact D/SEC data before 1989 are unavailable to us and we found that starting with application year 1998 the patent time series tailed off due to the review lag at the USPTO.

Because parent firms sometimes patent under their own names and at other times under the names of their subsidiaries, merging the Patents BIB data with firm-level data in the Compact D/SEC data is not straightforward. Mergers and acquisitions at both the parent firm and subsidiary levels, common in these two industries during the 1990s, and name changes further complicate linking the patent to firm-level data. (The USPTO does not maintain a unique identifier for each patenting assignee at the parent firm level nor does it track assignee name changes.) Thus, to use the firm-level information available in the Compact D/SEC data, the names of parent firms and their subsidiaries and the ownership of firms must be tracked over the entire period of the study, which is accomplished based on the subsidiary information in the Compact D/SEC data.

Since the Compact D/SEC data do not report old names of the firms that change their names (in many cases, after mergers), we use the S&P data to track the history of name changes of each assignee and link firm-level information in the Compact D/SEC data before and after a name change.

After merging the Patents BIB data with the firm-level information in the Compact D/SEC data, we then link the patent inventors to the firms in the Compact D/SEC data by the final firm name to produce a data set on inventors and patents that includes firm-level data (e.g., R&D expenditures, sales, and employment level) on the patents' assignees. To obtain firm-level patent and inventor data, we aggregate up from the patent assignee-level USPTO data.

Information on all citations is from the NBER Patent-Citations data collected by Hall et al. (2001). In these data each citing patent that was granted between 1975 and 1999 is matched to all patents cited by the patent. According to Hall et al. (2001), 50% of all citations are made to patents at least 10 years older than the citing patent and 5% of citations refer to patents that are at least 50 years older than the citing patent. Thus, we do not observe the total number of citations for the typical patent in our data. We construct the projected number of all citing patents for each patent in our data as follows. Based on the average number of citing patents per cited patent in each ensuing year since the application year for pharmaceutical and semiconductor patents in the USPTO data, we calculate the percentage of all citing patents that cites the target patent in the first N years after the cited patent's application year ($N = 1,\ldots,30$). For example, the number of citing patents in the first 5 years after application is shown to comprise 38.1% of all citing patents in the pharmaceutical industry. For a pharmaceutical patent for which our data are censored 5 years after its application year, we estimate the total projected number of citing patents by multiplying the actual number of citing patents for the patent in the 5 year period by 100/38.1. We choose patents in subcategories of 14, 19, and 31–33 in the USPTO classification for the pharmaceutical industry and subcategories of 21, 22, 24, 41, and 46 for the semiconductor industry. Definitions and summary statistics of variables used in our analysis are reported in Table 1.

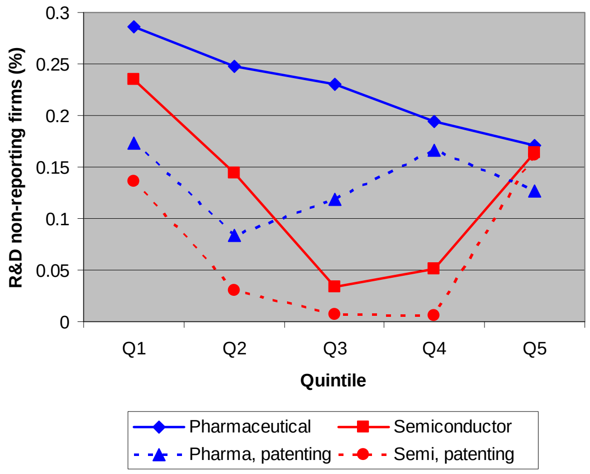
Figure 1 shows the fractions of firm-years in the pharmaceutical and the semiconductor industries that report zero R&D expenditures for each quintile of firms based on employment size. The two solid lines present the fractions of firm-years with zero R&D expenditures in all firm-years in the two industries, while the two dotted lines show the fractions of zero R&D firm-years among those with a positive number of patent grants. Consistent with our expectation that there would be proportionally fewer zero R&D firms among those who patent than among all firms, a dotted line lies above a solid line for both industries.

One interesting finding from the figure is that the largest firms in the semiconductor industry are less likely to report R&D than other firms except very small firms, which is contrary to what is usually reported in the literature. On the other hand, the pharmaceutical industry show a monotonically declining fraction of zero R&D firms among all firms as the firm size gets larger.

Table 1. Variable Definition and Sample Statistics

	Definition	Mean (Standard Deviation) Pharmaceutical	Semiconductor
PAT	Number of patents granted to a firm by application year	33.54 (64.21)	40.33 (126.2)
R&D	Real R&D expenditures in 1996 dollars	2,098 (3,853)	693.3 (2,250)
INV	Number of inventors named on the firm's patents in year *t*	61.35 (114.7)	41.57 (113.4)
PAT/R&D	PAT divided by R&D	0.1655 (1.833)	0.1178 (0.7497)
PAT/INV	PAT divided by INV	0.6251 (0.3665)	0.7865 (0.3355)
SALES	Real sales volume in 1996 dollars	20,277 (37,098)	6,521 (27,237)
K/L	Capital-labor ratio, or deflated plant and equipment over the number of employees	0.6369 (0.7492)	1.256 (7.548)
FAGE	Years elapsed since the founding year of the firm	58.66 (47.59)	19.46 (20.62)
NSIC	Number of secondary SICs assigned to the firm	2.370 (1.751)	1.554 (0.9671)
CITE	Average projected number of citations per patent	4.600 (7.117)	7.707 (9.422)

Figure 1. Fraction of Firms by Employment Size Reporting No R&D



Notes: There are 1,620 firm-years and 760 firm-years with patents in the pharmaceutical industry and 1,111 firm-years and 693 firm-years with patents in the semiconductor industry.

In the semiconductor industry, the fraction of zero R&D firms among patenting firms shows a similar pattern by firm size as in the fraction of zero R&D firms among all firms, but the zero R&D share among the largest firms is higher than any other size groups. Note that nearly 20% of the firms who patent in the largest-size quintile do not report R&D. Unlike the zero R&D share among all firms, there is no particular pattern in the relationship between the zero R&D fraction among patenting firms and firm size in the pharmaceutical industry.

Table 2. Number of Firms Reporting R&D Expenditures by Employment Size

Employment Size	Number of Firms	Pharmaceutical	Semiconductor
<10	179 (0.07)	145	22
10–99	822 (0.32)	581	246
100–999	979 (0.39)	460	525
1000+	559 (0.22)	308	252
Total	2,539 (1.00)	1,494	1,045

Notes: Proportions are in parentheses.

4. Empirical Findings

4.1 Patents and R&D Expenditures

Tables 3 and 4 report the results of our estimation of the determinants of patenting productivity in the pharmaceutical and semiconductor industries, respectively. Marginal effects are shown in Table 5. These regressions relate the firm's granted patents applied for in year *t* (PAT) and the firm's R&D expenditures (R&D) in that year. Our use of contemporaneous R&D, as opposed to lagged R&D, follows the extensive literature estimating patent production functions (e.g., Hall et al., 1986). Evidence suggests that R&D activities and innovations occur somewhat simultaneously. Moreover, if a firm attempts to patent an innovation, it files the application while the innovation is being developed or very shortly afterwards (Hall et al., 1986).

Models 1a through 5a report the results of regressions using R&D expenditures as a measure of innovative input, and Models 1b through 5b report the results using the number of inventors (INV) as the input measure. Models 1a through 3a and 1b through 3b report the results of linear specifications, where the response variable is the log of the patent-R&D and patent-inventor ratios, respectively. In these models, the patent yield measure is regressed on the log of sales. In Models 2a and 3a and 2b and 3b, we include random firm effects. In Models 4a and 5a and 4b and 5b, the response variable is PAT and the principle predictor variable of interest is either log R&D or log INV. In these specifications we use a negative binomial-based maximum likelihood model (see, for example, Cameron and Trivedi, 1986). We favor the negative binomial because the number of patents granted to a firm in a particular year is a nonnegative count variable. The coefficient estimates in these models allow us to test for scale economies with respect to the size of the R&D enterprise.

Models 1a through 3a in Tables 3 and 4 directly test Schumpeter's proposition that large firms are more efficient innovators. Model 1a is a simple least squares regression of the log of the patent-R&D ratio on the log of sales. Model 2a re-estimates this model assuming firm-specific random effects. Model 3a includes (in addition to sales), the firm's capital-labor ratio (K/L), years since the firm was founded (FAGE), and the number of business lines in the firm (NSIC), measured by the number of secondary SICs identified with the firm. We include the log of the capital-labor ratio (K/L) as a regressor because, given R&D expenditures, a highly capitalized firm may have a stronger incentive to patent than less capitalized firms.

Table 3. Patents, R&D Expenditures, and Inventors: Pharmaceutical Industry

	log(PAT/R&D)			PAT		log(PAT/INV)			PAT	
Response	**(1a)**	**(2a)**	**(3a)**	**(4a)**	**(5a)**	**(1b)**	**(2b)**	**(3b)**	**(4b)**	**(5b)**
	OLS	**OLS, RE**	**OLS, RE**	**NB**	**NB, RE**	**OLS**	**OLS, RE**	**OLS, RE**	**NB**	**NB, RE**
log(SALES)	-0.1485***	-0.1887***	-0.2589***			-0.0035	-0.0132	0.0206		
	-10.41	-8.34	-4.56			-0.62	-1.34	0.91		
log(R&D)				0.6456***	0.4608***					
				40.66	6.98					
log(INV)									0.9741***	1.0535***
									91.59	30.38
log(K/L)			0.2253		0.0208			-0.1330**		-0.0702
			1.51		0.18			-2.22		-1.18
log(FAGE)			0.1676		0.1040			-0.0467		-0.0842*
			0.95		0.80			-0.56		-1.66
log(NSIC)			0.0651		0.0220			-0.0762		-0.1225***
			0.35		0.20			-1.13		-2.82
CITE			0.0340		0.0116			0.0177**		0.0406***
			1.38		0.63			2.21		4.30
$CITE^2$			-0.0007		-0.0009			-0.0008**		-0.0020***
			-0.71		-0.73			-2.39		-3.07
Adj. R^2	0.1789	0.1805	0.2494			0.0008	0.0008	0.0164		
Log Lik.				-1912	-669				-1472	-550
Obs.	494	494	161	560	167	494	494	161	560	167

Notes: Two rows for each regressor report the estimated coefficients and inverse coefficients of variation. Constant terms are not reported. *, **, and *** denote significance at the 10%, 5% and 1% levels.

For example, a patent infringement lawsuit that leads to production stoppage will be more destructive for a firm that has made a large capital investment in a state-of-the-art physical plant. Such vulnerability may encourage the firm to develop a diverse portfolio of patents that it can use as a bargaining chip to ward off infringement suits (Cohen et al., 2000; Parr and Sullivan, 1996). We include NSIC as a regressor to isolate and control for economies of scope.

Our empirical exercise is designed to measure the effect of the size of the firm or R&D enterprise on real innovative output, which patents proxy. The economic values of the innovations underlying patents, however, vary considerably from patent to patent.

Moreover, the economic value of many patents is close to zero. Evidence exists that a patent's citation by a subsequent patent—each patent documents the "prior art" upon which the new innovation builds—is an indicator not only of the importance of the underlying innovation but its economic value as well (see Trajtenberg, 1990). We created the variable CITE, the estimated average number of citations per granted patent by the firm in year t, to capture the quality of the firm's patents. Model 3a includes CITE as an additional regressor. Thus Model 3a generates an estimate of the effect of size on patent counts holding constant patent quality.

Table 4. Patents, R&D Expenditures, and Inventors: Semiconductor Industry

	log(PAT/R&D)			**PAT**		**log(PAT/INV)**			**PAT**	
Response	**(1a)**	**(2a)**	**(3a)**	**(4a)**	**(5a)**	**(1b)**	**(2b)**	**(3b)**	**(4b)**	**(5b)**
	OLS	**OLS, RE**	**OLS, RE**	**NB**	**NB, RE**	**OLS**	**OLS, RE**	**OLS, RE**	**NB**	**NB, RE**
log(SALES)	-0.1409***	-0.1375***	-0.1368***			0.0626	0.0561	0.0686		
	-5.56	-3.89	-2.30			6.12	3.67	2.74		
log(R&D)				0.8834***	0.3782***					
				35.70	7.74					
log(INV)									1.0914	1.0871
									93.43	48.86
log(K/L)			0.0766		0.0420			0.0378		-0.0135
			1.24		1.04			1.34		-1.40
log(FAGE)			0.2314		0.3384***			-0.0253		-0.0045
			1.65		3.01			-0.44		-0.09
log(NSIC)			-0.0756		-0.1121			-0.0124		-0.0260
			-0.62		-1.35			-0.22		-0.95
CITE			-0.0125		-0.0675***			-0.0064		-0.0199
			-1.25		-8.77			-1.40		-5.84
$CITE^2$			-0.0001		0.0007***			-0.0001		0.0002
			-0.60		3.42			-0.52		2.26
Adj. R^2	0.0528	0.0546	0.0254			0.0636	0.0653	0.1099		
Log Lik.				-1912	-1355				-1421	-1085
Obs.	538	538	380	542	381	538	538	380	542	381

Note: Two rows for each regressor report the estimated coefficients and inverse coefficients of variation. Constant terms are not reported. *, **, and *** denote significance at the 10%, 5% and 1% levels.

Consistent with earlier studies, the results in Tables 3 and 4 indicate that small firms (classified by sales) publish more patents per R&D dollar, contrary to Schumpeter's hypothesis. R&D expenditures is not used as a size measure in these regressions since the denominator of the dependent variable is R&D and having R&D as a regressor may result in a spurious negative relationship between the response variable and the regressor by construct. Models using the number of employees as the measure of size yield results that are qualitatively similar to the models that use the variable SALES. Our findings in Models 1a through 3a in both tables are robust to the specifications used. For the pharmaceutical industry the elasticity estimates of PAT/R&D with respect to SALES is negative and significant and ranges from −0.15 to −0.26. In the semiconductor regressions, the elasticities are approximately −0.14.

In the negative binomial regressions (Models 4a and 5a), we find evidence of decreasing returns to scale. Note that in negative binomial regressions the coefficients on the log scale have an elasticity interpretation. Model 5a, which includes covariates other than R&D, controls for patent quality, allows for firm-specific random effects (see Hausman et al., 1984), and produces elasticity estimates that are significantly less than one. In the negative binomial model, the coefficient estimates of the log-transformed regressors have an elasticity interpretation These estimates are somewhat larger than estimates reported in other studies. Bound et al. (1984) report elasticity estimates for firms from a broader sample of industries that range from 0.32 to 0.38 (see their Table 2.8, p. 41).

Table 5. Marginal Effects in Elasticity

Response	Pharmaceutical				Semiconductor			
	(3a)	(5a)	(3b)	(5b)	(3a)	(5a)	(3b)	(5b)
	OLS, RE	NB, RE	OLS, RE	NB, RE	OLS, RE	NB, RE	OLS, RE	NB, RE
	log(PAT/R&D)	PAT	log(PAT/INV)	PAT	log(PAT/R&D)	PAT	log(PAT/INV)	PAT
SALES	-0.2589		0.0206		-0.1368		0.0686	
R&D		0.4608				0.3782		
INV				1.0535				1.0871
K/L	0.2253	0.0208	-0.1330	-0.0702	0.0766	0.0420	0.0378	-0.0135
FAGE	0.1676	0.1040	-0.0467	-0.0842	0.2314	0.3384	-0.0253	-0.0045
NSIC	0.0651	0.0220	-0.0762	-0.1225	-0.0756	-0.1121	-0.0124	-0.0260
CITE	0.1938	0.0654	0.1011	0.2284	-0.1002	-0.5414	-0.0514	-0.1601
$CITE^2$	-0.0464	-0.0538	-0.0488	-0.1265	-0.0177	0.0936	-0.0071	0.0323

4.2 Patents and Inventors

Models 1b through 5b of Tables 3 and 4 report the estimated relationship between the number of patents and the number of inventors in the pharmaceutical and the semiconductor industries, respectively. Models 1b through 5b replicate the specifications 1a through 5a, substituting INV for R&D. In both Models 1b and 2b for the pharmaceutical industry the estimate of the coefficient for log sales is negative but not significant. When the firm covariates and measures of patent quality are added, the coefficient estimate becomes positive, but it remains small relative to its standard error. In the negative binomial regressions (Models 4b and 5b) the coefficient estimate is close to and statistically indistinguishable from one, thus suggesting constant returns to scale.

In the semiconductor regressions, the coefficient estimates for log sales in Models 1b through 3b are significant and positive, the opposite of the findings we observe in the R&D regressions. In addition, Models 4b and 5b suggest increasing returns to hiring inventors—the coefficient estimates are statistically significantly greater than one.

Note that for most of the non-size-related covariates, the coefficient estimates are not robust to the specification used. The coefficient estimate for log K/L is only significant in Model 3b of Table 3. Here, contrary to expectations, the coefficient estimate is negative. The capital-labor ratio may be picking up differences across firms in technology or product mix that affect the firm's ability to extract patentable output from R&D activities. The coefficient estimate for log NSIC is negative and significant in Model 5b of Table 3, suggesting diseconomies of scope. Where the estimated coefficient for CITE is significant in the pharmaceutical industry regressions, the results indicate an inverted U-shaped association between CITE and PAT that peaks at CITE close to 11. Because about 90% of the observations lie below this CITE level, the number of patents mostly increases with CITE, suggesting patent quantity and quality are complements in this industry. Interestingly, the semiconductor regressions show a U-shaped relationship that peaks at CITE close to 49, below which 99% of the observations lie. This result is consistent with quality-quantity substitutability.

5. Discussion and Conclusion

Consistent with a number of papers in the firm size-R&D productivity literature, we find that the patent yield of an R&D dollar falls with firm size in the pharmaceutical and semiconductor industries. We investigate whether this relationship is robust by using the number of inventors employed at the firm in place of R&D expenditures. Using this measure of R&D activity, we find no relationship between R&D productivity and firm size in the pharmaceutical industry. In the semiconductor industry, we find that R&D productivity increases with firm size, consistent with Schumpeter's hypothesis and contrary to the results obtained using R&D expenditures. These findings are robust across specifications that control for, among other things, patent quality.

Our findings using the number of inventors as a measure of innovative inputs may be consistent with a true size effect, such as the ability of large firms to specialize inputs. Kim et al. (2004b) provide some evidence that large firms specialize their scientific labor more than small firms. On the other hand, the number of inventors employed by a firm does not capture R&D capital (e.g., laboratory equipment), materials, and support personnel. These non-inventor inputs may vary systematically with firm size so that our finding of a positive relationship between size and R&D productivity may be due to richer non-inventor inputs in large firms. The threshold effect that inhibits undertaking R&D activity (Gonzalez and Jaumandreu, 1998) and systematic co-variation between size and inventor quality may also account for our results.

References

Acs, Z. J. and D. B. Audretsch, (1987), "Innovation, Market Structure, and Firm Size," *The Review of Economics and Statistics*, 69(4), 567-574.

Acs, Z. J. and D. B. Audretsch, (1990), *Innovation and Small Firms*, Cambridge, MA: MIT Press.

Acs, Z. J. and D. B. Audretsch, (1991), "R&D, Firm Size and Innovative Activity," in *Innovation and Technological Change: An International Comparison*, Z. J. Acs and D. B. Audretsch eds., Ann Arbor, MI: University of Michigan Press.

Bound, J., C. Cummings, Z. Griliches, B. H. Hall, and A. Jaffe, (1984), "Who Does R & D and Who Patents?" in *R&D, Patents, and Productivity*, Z. Griliches ed., Chicago, IL: University of Chicago Press.

Cameron, A. and P. Trivedi, (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests," *Journal of Applied Econometrics*, 1, 29-53.

Cohen, W. M. and S. Klepper, (1996), "A Reprise of Size and R&D," *The Economic Journal*, 106(437), 925-951.

Cohen, W. M. and R. C. Levin, (1989), "Empirical Studies of Innovation and Market Structure," in *Handbook of Industrial Organization*, R. Schmalensee and R. D. Willig eds., New York, NY: North-Holland.

Cohen, W. M., R. C. Levin, and D. C. Mowery, (1987), "Firm Size and R&D Intensity: A Re-Examination," *Journal of Industrial Economics*, 35, 543-565.

Cohen, W. M., R. R. Nelson, and J. P. Walsh, (2000), "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," *NBER Working Paper*, No. 7552.

Gonzalez, X. and J. Jaumandreu, (1998), "Threshold Effects in Product R&D Decisions: Theoretical Framework and Empirical Analysis," *Studies on the Spanish Economy*, No. 78.

Hall, B., Z. Griliches, and J. Hausman, (1986), "Patents and R&D: Is There a Lag?" *International Economic Review*, 27(2), 265-284.

Hausman, J. A., B. H. Hall, and Z. Griliches, (1984), "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica*, 52(4), 909-938.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg, (2001), "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," *NBER Working Paper*, No. 8498.

Kim, J., S. Lee, and G. Marschke, (2004a), "R&D Scientists in Pharmaceutical and Semiconductor Industries: Data Construction and Description," *Working Paper*.

Kim, J., S. Lee, and G. Marschke, (2004b) "Research Scientist Productivity and Firm Size: Evidence from Panel Data on Inventors," *Discussion Paper*, University at Albany, Department of Economics, No. 04-06.

Kleinknecht, A., (1987), "Measuring R & D in Small Firms: How Much Are We Missing?" *The Journal of Industrial Economics*, 36(2), 253-256.

Kleinknecht, A. and B. Verspagen, (1989), "R&D and Market Structure: The Impact of Measurement and Aggregation Problems," *Small Business Economics*, 1(4), 297-301.

Lanjouw, J. O. and M. A. Schankerman, (2004), "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators," *Economic Journal*, 114(495), 441-465.

Panzar, J. C. and R. D. Willig, (1981), "Economies of Scope (in Sustainability Analysis)," *The American Economic Review*, 71(2), 268-272.

Pakes, A. and Z. Griliches, (1980), "Patents and R&D at the Firm Level: A First Report," *Economics Letters*, 5(4), 377-381.

Parr, R. L. and P. H. Sullivan, (1996), *Technology Licensing: Corporate Strategies for Maximizing Value*, New York, NY: John Wiley & Sons, Inc.

Patel, P. and K. Pavitt, (1992), "The Innovative Performance of the World's Largest Firms: Some New Evidence," *Economics of Innovation and New Technology*, 2, 91-102.

Scherer, F. M., (1984), *Innovation and Growth: Schumpeterian Perspectives*, Cambridge, MA: MIT Press

Symeonidis, G., (1996), "Innovation, Firm Size and Market Structure: Schumpeterian Hypotheses and Some New Themes," *OECD Economic Studies Working Papers*, No. 27.

Trajtenberg, M., (1990) "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *The Rand Journal of Economics*, 21(1), 172-187.

Tsai, K. and J.-C. Wang, (2005), "Does R&D Performance Decline with Firm Size? A Re-Examination in Terms of Elasticity," *Research Policy*, 34(6), 966-976.