A Cross-sectional Analysis of Research and Development Expenditures.

Beau Buchmann, author

pp. 19-27

Oshkosh Scholar, Volume II, April 2007

University of Wisconsin Oshkosh
Office of Grants and Faculty Development
800 Algoma Blvd.
Oshkosh, WI 54901
(920) 424-3215
www.uwosh.edu/grants

Conclusion

Our regression results agreed with our hypothesis, which states that an increase in spending on public education at the K–12 level will reduce poverty rates. However, there are some weaknesses with the data, and they need addressing in order to create a more robust model.

One of the main weaknesses is that the delay in the lag variable is only for one year. This may be a reason why the lag variable does not appear to be extremely statistically significant. If the delay between the lag variable data and current year data could have a difference of at least 10 years, the results might show an increased level of significance. A delay of only one year for the lag variable does not allow for the benefits or effects from the increase in education expenditures to take their full effects.

Moreover, spending on education is an investment in human capital. Thus, the returns on this investment will not be seen immediately or even after a few years. The returns from this investment will not be evident until after these children move on into adulthood and into the workforce.

However, the problem with the education expenditure data is that the data does not go back more than a few years. Thus, the possibility of testing a 10-, 15-, or even 20-year lag is not feasible as of yet. I hope that in the following years, data will be available, and a minimum 10-year delay in the lag variable will be possible. Nevertheless, we should not disregard the results of this model and view this paper as insignificant. My hope is that this paper sparks some discussion on public education expenditure reform. In addition, I hope that society begins to view the value of education and understand that education is an investment in all of our futures. Therefore, when budget cuts occur and affects the education sector, we should be worried about how this will affect our children and us.

References

Bureau of Labor Statistics. (2000). County-level data for unemployment rates. Retrieved April 12, 2006, from ftp://ftp.bls.gov/pub/special.requests/la/laucnty00.txt

Jha, R., Biswal, B., Biswal, U.D. (2001). An empirical analysis of the impact of public expenditures on education and health on poverty in Indian states. *Queen's University Economic Department Working Paper* (Working Paper No. 998).

Jung, H.S., & Thorbecke, E. (August 2001). The impact of public education expenditures on human capital, growth, and poverty in Tanzania and Zambia: A general equilibrium approach. *International Monetary Fund Working Paper* (Working Paper 01/106).

Centers for Disease Control. (2000). Interactive atlas of reproductive health: Geodataset selection. Retrieved April 6, 2006, from http://apps.nccd.cdc.gov/gisdrh/(ieowtm45ele0dh45mt2qii45)/Default.aspx

National Center for Education Statistics. (2000). Build a table with common core of data. Retrieved April 12, 2006, from http://nces.ed.gov/ccd/bat/

United States Census: Small Area Income and Poverty Estimates. (2000). County-level estimates on median family income and poverty rates. Retrieved April 15, 2006, from http://www.census.gov/hhes/www/saipe/county.html

A Cross-sectional Analysis of Research and Development Expenditures

Beau Buchmann, author
Dr. Marianne Johnson, Economics, faculty adviser

Abstract

Long-term economic growth is sustained by research and development activities conducted by private firms, and therefore, understanding how firms make R&D choices is important. In this study, Compustat data is used to econometrically analyze firm choices about research and development (R&D). Examined are both the total amount of money firms spend on R&D activities and the portion of sales revenue firms spend on R&D. Our interpretations are as follows. The model predicts that larger firms, determined by sales revenue and number of people employed, spend more money on R&D activities. In addition, firms whose operations include more business segments commit fewer revenues to innovation, while firms that operate in more geographic areas typically spend more on R&D. Finally, if a firm owns more intangible assets, such as franchise rights and patents or copyrights, it will likely spend more on R&D.

Introduction

It is widely recognized that innovation is one of the most important factors that drive long-term economic growth (Schumpeter, 1994). One common way to measure the amount of innovation in an economy is by looking at the amount of money private firms report as research and development (R&D) expenditures. This study used cross-sectional data from 2005 to examine the characteristics of firms that conduct R&D. To identify the key characteristics of firms, we developed two least-squared regression models that examine different factors that could influence R&D choices.

The first model studied inter-firm variation of R&D expenditure in attempt to better identify the characteristics that indicate which types of firms fund the most innovation. The second model worked to explain the variation in normalized research and development expenditures. We calculated normalized research and development expenditures by dividing the amount of money each firm spent on R&D in 2005 by the sales revenue the firm earned in the same year. Thus, normalized R&D expenditure was thought of as the fraction of revenue a firm spent on R&D. This approach had the advantage of controlling for firm size, and allowed us to more carefully examine what types of firms conduct R&D.

First, however, we provide a brief survey of some recent research on R&D. Second, we describe the data analyzed in the study. Third, we develop two different models to explain R&D decisions. Finally, in the conclusion, we examine some of the most interesting results of our analysis and make policy recommendations.

Literature Review

Schumpeter suggested that entrepreneurship and innovation are the keys to economic growth over the long run (1994). However, while there is extensive literature on innovation going back in economic history to the 1920s, we focused on three recent articles that are related to the particular econometric approach used in this paper.

One particular policy problem is that identified by Shah (2006). He explained firms cannot capture all of the benefits from R&D activities. Specifically, private corporations cannot prevent competitors from utilizing the fruits of the innovative discoveries. Even though sharing new innovations would likely produce greater economic gains, firms in competitive situations should intuitively keep their advances private in order to maximize profits. Thus, Shah deduced private firms are driven to perform innovative activities in pursuit of competitive advantages, but that firms are deterred by their inability to completely exclude competitors from utilizing the new innovations.

Much like Shah (2006), Moris (2005) stressed the importance of policies, both direct and indirect, that can be implemented to stimulate private R&D investment. Direct policies included grants and government contracts. Indirect policies included a wide variety of different types of tax credits and allowances. In his discussion of tax credit claims, Moris showed that corporations in the manufacturing, information, and professional, scientific and technical-services industries accounted for the most tax credit claims. In all three of these industries, most of the tax credit claims were in sectors adopting or producing advanced technologies, such as computers and computer-related technologies.

Wu (2005) wrote about policies which, at the state-level, induced marginal increases in R&D investment. Evidently the most significant policies Wu studied were state investment tax credits and the substantial state funding of higher education. While Wu discussed the merit of tax credits, he proposed the federal government would provide a more effective service to the private sector by creating a national standard and eliminating the incentives for firms to continually move to different states to take advantage of new and better R&D policies. Schumpeter (1994) made a similar argument.

Given the findings of these recent studies, we decided to include many similar explanatory variables in our model. For example, based on Shah's findings, we included a measure of intangible assets in our econometric model. We also included dummy variables for different industries in our model because of Moris's (2006) study. However, while using several similar variables, our use of a very large cross-sectional Compustat data set allowed us to examine these purported relationships in a new and different context.

Data Collection and Description

This study examined cross-sectional data collected from Standard and Poor's Compustat data set. The Compustat data set contains a vast amount of economic and financial information on nearly 10,000 companies. Ultimately, eight quantitative variables were chosen for this study: total assets, book value, number of employees, intangibles, sales, normalized R&D expenditure (created from R&D and sales data), number of business segments, and geographic region. The definitions of these variables are given in Table 1.

The variables in Table 1 were either used directly in a final regression equation or combined with other variables from Table 1 to generate compound variables (rates, proportions, etc.). In summary, Table 2 displays the mean value of each variable used, along with the standard deviation. Note, from Table 2, some of the data have significant variation. We considered regressions both with and without the outliers. Elimination of outliers did not fundamentally change the regression results and, therefore, we chose to leave all the values in the model, following the methodology of Wooldridge (2004). Excluding the number of business segments and the number of geographic regions, the standard deviation for each variable was quite large compared to its respective mean value. This fact is not surprising because there are thousands of small firms and many fewer extremely large firms.

R&D expenditures were far from evenly distributed across industries. Therefore, we employed a dummy variable for each industry, as defined by the North American Industry Classification System (NAICS).

Characteristic of Firms that Engage in R&D

The first model we examined is designed to identify the characteristics of firms that spend the most on R&D in absolute terms (Model 1). The natural logarithm of each company's R&D expenditure, measured in millions of dollars (LNXRD), is used as the dependent variable. The log-log model has a number of advantages, including ease of interpretation, because all variables can be interpreted as percentage changes (Wooldridge, 2004).

The first key independent variable of this model was the natural log of each firm's sales revenue, in millions of dollars (LNSALE). Sales figures are used as a proxy for the size of the firm. In a study that used time-series data, sales, profits, or retained earnings could be used to portray the relationship between profitability in one year and R&D expenditures in the next. Since the data used in this study is cross-sectional, sales is used as a metric of each firm's size.

As shown in Table 3, the dollar value of a company's sales in 2005 was statistically significantly related to the amount of money it spent on R&D ($p < 0.01$). Therefore, we can confidently state that larger companies tend to spend more money on R&D than smaller companies. A 1% increase in sales will be associated with a 0.56% increase in R&D expenditures. This suggests that polices developed to grow the size of companies may have a positive impact on the amount of R&D private firms conduct. As an example of a policy contrary to encouraging R&D, businesses in some European Union countries are taxed on a highly progressive scale to encourage small business ownership. Based on this study, such policies would be considered detrimental to R&D expenditures.

Another key explanatory variable was the natural logarithm of the value of all intangible assets owned by a company, in millions of dollars (LNINTAN). This variable represented the unamortized value of assets such as patents, copyrights, licenses, franchises and franchise rights, and shelter from import quotas. As evident from Table 3, the amount of intangible assets owned by a company was positively and statistically significantly related to the amount of money a company spends on R&D ($p < 0.01$). Therefore, policies should aim to increase ownership of intangible assets, make such assets more accessible to private organizations, or at least not discourage firms from owning intangible assets. For example, if a firm's patents are protected by statute, the firm would likely invest more money in R&D, leading to the development of new business opportunities, all else equal. However, perhaps more daunting obstacles are the excessive disclosure requirements and lax enforcement faced by firms abroad (Schumpeter, 1994).

The number of business segments in which an organization operates (SEGNUM), as identified by Standard and Poors, is negatively and statistically significantly related to the amount of money a firm spends on R&D ($p < 0.01$), holding other factors constant. Thus the estimation results of Model I suggest that the firms investing the most in R&D are large organizations that specialize in only a few business or industry segments.

In the Model 1 regression, the number of geographic regions (GEONUM) in which a company operates is positively correlated with the amount of money the company will spend on R&D, as determined by the sign on the estimated coefficient. However, this relationship is not statistically significant. The number of geographic regions is measured as a categorical variable, ranging from 1 to 5. Thus we do not log this variable.

The final key explanatory variable examined in Model I is the number of workers each company employs (EMP), measured in thousands of employees. Despite the positive correlation between the number of employees and the amount of money spent on R&D (as determined by the estimated coefficient), the relationship is not statistically significant. Other explanatory variables were included in the regression and Table 3 to control for firm size, market, and sector, but we do not discuss them because the results are not central to understanding R&D choices.

Normalized R&D Expenditures and Firm Characteristics

So far we have discussed characteristics of the firms that are associated with the most innovation activity. This section of the paper focuses on which firms spend the largest portion of their sales revenue on R&D. In this model, we utilized each firm's normalized R&D expenditures as the dependent variable. Normalized R&D expenditures were calculated as total R&D for a firm, divided by the firm's total sales.

Before conducting additional regression analysis, Tables 4 and 5 are examined to determine whether any broad relationship can be identified between the size of a firm and the amount it spends on R&D. Note that the tables display summary statistics for all of the firms in our dataset that reported a positive value of R&D expenditures. The data in the tables are ranked by sales revenue, ascending, and the amount of R&D expenditures reported. From the median values in Table 4, the smallest firms, with respect to sales revenue, account for the highest values of normalized R&D spending. From Table 5, we understand that this phenomenon is not due solely to sales revenue. If this were the case, when the firms are ranked by R&D expenditures, the firms with the largest median normalized R&D expenditure would account for the largest R&D expenditure.

The log of the percent of sales spent on R&D, or normalized R&D expenditure (LNXRDSALE), was our dependent variable in Model II. The first explanatory variable of Model II was the logarithm of the ratio of sales divided by the number of people a firm employs (LNSALEMP). This ratio describes how much sales revenue each employee generates. This is a widely used measure of firm productivity. We found that this productivity measure is negatively correlated with normalized R&D

expenditures, as determined by the estimated coefficient reported in Table 6. This relationship is statistically significant ($p < 0.01$). Perhaps an explanation is that although innovation leads to impressive efficiency gains, the process by which to achieve innovation is highly inefficient due to high upfront investment costs and the indeterminacy of future benefits.

A second explanatory variable was the ratio of assets-to-sales (LNATAL), which is an important measure in financial studies. According to financial theory, more efficient firms should have smaller assets-to-sales ratios because this denotes economically efficient production. This result was confirmed in our study. The estimated coefficient of assets-to-sales is positive and statistically significant (see Table 6).

A firm's book value measures the accounting value of its equity. We use the log of this ratio (LNCSTK) as an explanatory variable. Book value was negatively related to a firm's normalized innovation expenditures. In Table 6, we find that a 1% increase in book value was associated with a 0.09% decrease in R&D expenditures. By revisiting Tables 4 and 5, this relationship is easily identified. Obviously, small firms will have lower book values than large firms. Since we expect normalized R&D expenditure to decrease as firms' sales revenues increase, this concept should be intuitive.

Last, just as in Model I, Model II introduces industry sector dummy variables, along with regional and other firm characteristic controls. The industry sector dummy variables are informative, indicating that firms in high-tech industries are much more likely to engage in R&D than firms in other industries, holding other factors constant. This result is statistically significant. Again, as with Model I, these variables are important for econometric reasons, but are not especially interesting to discuss in relation to R&D decisions.

Conclusions

From the industry dummy variables, the most obvious conclusion we draw from this study is the differing propensity of firms to invest in R&D across different industries. In particular, firms within the manufacturing industry, the information industry, and the professional, scientific, and technical services industry are more likely to engage in R&D. This is similar to the finding of Moris (2005). These industries spend more on innovation than companies in other industries, controlling for size, number of employees, and other firm characteristics. Moreover, companies from these three industries invest a greater portion of their sales revenue in R&D, holding other factors constant.

Another key finding is that firms owning more intangible assets are more likely to conduct R&D activities, or those which conduct more R&D own more intangible assets. In addition, a firm's book-value-per-share is found to be inversely related to a firm's level of R&D intensity. One way to understand this finding is that firms conducting innovative research incur significant risk from an investment with undetermined cash flows. Furthermore, the cost of R&D will be expensed much earlier than any return a firm generates from its innovative activities.

When combining the results from the first and second models estimated, we identify a new and intriguing relationship—what we call a growth or buy-out effect.

When a company is small, in terms of sales revenue, a large portion the sales revenue is reinvested in growing the business. Thus the firm's R&D intensity should be greater early in a company's life. Likewise, innovative firms should *appear* less efficient than less innovative firms due to delayed returns to investment of capital.

Over time, however, we speculate innovative firms will likely become more profitable or be purchased by large multi-sector conglomerates. As sales revenues and available capital grow, R&D expenditures may not increase as fast as sales at this point in a firm's lifespan. Moreover, larger firms have the advantage of scale, and as the innovative products mature, the firm will become more profitable. This hypothesis may explain the inverse relationship between normalized R&D and the measures of efficiency in the second model. This is an interesting result not previously examined in the literature, and more research should look into this topic. However, an important caution is that this is a cross-sectional study, with no method to examine the influence of time. Further research using time-series data could outline a growth pattern for firms that would help evaluate our theory of a growth or buy-out effect.

In terms of policies to promote growth, the protection of intangible assets should continue to be important, both in the U.S. and abroad, as our model finds that intangible assets are a key explanatory variable in determining R&D expenditures. Trade partners, such as China, should be encouraged to adopt legislation that facilitates the protection of intangible assets and intellectual property. With augmented trade liberalization and protection of intangible assets abroad, businesses could grow to serve more geographic segments. Therefore, through access to a larger market, businesses could both take advantage of increasing returns to scale and comparative differences in production. Ultimately, increased international trade would likely lead to further specialization, which the first model suggests is positively correlated with R&D expenditure.

References

Moris, F. (2005). *InfoBrief: The U.S. research and experiementation tax credit in the 1990s* (Science Resources Statistics). Washington, D.C.: National Science Foundation.

Schumpeter, J. (1994). *History of economic analysis*. New York: Oxford University Press.

Shah, A. The economics of research and development: How research and development capital affects production and markets and is affected by tax incentives. (World Bank Policy Research Working Paper No.1325, 2006).

Wooldridge, J. (2005). *Introductory Econometrics*. New Jersey: Southwestern.

Wu, Y. (2005). The effects of state R&D credits in simulating private R&D expenditure: A cross-state empirical analysis. *Journal of Policy Analysis and Management, 24*, 785-802.

Appendix

Table 1
Description of the Variables

Variable (Abbreviation, Units)	Description
Total Assets (AT, $ Million)	Current assets plus net property, plant, and equipment plus other non-current assets.
Book Value (CSTK, $ Million)	The value of assets, less intangible assets and debt.
Employees (EMP, Thousands)	The number of employees each firm has reported to shareholders.
Intangibles (INTAN, $ Million)	The value of intangible assets such as patents, blueprints/ design, copyrights, licenses, etc.
Sales (SALE, $ Million)	The amount of money a company spent to develop new products or services.
Research and Development Expenditures (XRD, $ Million)	The amount of money a company spent to develop new products or services.
Business Segments (SEGNUM, Number)	The number of industry segments available to a company, as determined by Standard and Poor's, from one to ten.
Geographic Segments (GEONUM, Number)	The number of geographic segments available to a company, as determined by Standard and Poor's, from one to ten.

Table 2
Descriptive Statistics *

Variable	Mean	Standard Deviation
AT	8199.54	655.44
CSTK	147.34	902.12
EMP	7.88	36.76
INTAN	603.68	3820.71
SALE	2595.79	1228.79
XRD	93.10	523.58
XRD/SALE	7.48	152.84
SEGNUM	2.15	1.63
GEONUM	3.12	1.23
SALE/EMP	552.20	252.31
AT/SALE	22.61	754.31

*Abbreviated names for variables are provided in Table 1.

Table 3
Regression Results of Model 1

Dependent Variable: LNXRD
Number of Observations: 527

Variable	Estimated Coeff.	Std. Error	t-statistic	p-value
C	-1.239813	0.228875	-5.416994	0.0000
GEONUM	0.052810	0.046963	1.124499	0.2613
SEGNUM	-0.129866	0.036266	-3.580915	0.0004
EMP	0.007468	0.002152	3.470406	0.0006
LNSALE	0.565411	0.039360	14.36497	0.0000
LNINTAN	0.212332	0.029200	7.271570	0.0000
*MAN3133	0.281263	0.113378	2.480755	0.0134
*INFO51	0.364274	0.164721	2.211452	0.0274
*FININ52	-4.121875	1.917874	-2.149190	0.0321
*PROSER54	0.547133	0.253383	2.159314	0.0313
*HEALTH62	-0.552263	0.403685	-1.368054	0.1719
αAR(1)	0.383457	0.042785	8.962315	0.0000
R-squared	0.699291			
Adjusted R-squared	0.692868			
Log likelihood	-820.5838	F-statistic		108.8744
Durbin-Watson stat	2.063738	Prob (F-statistic)		0.000000

Table 4
Average R&D by Sales, Ranked by Total Sales (Ascending by Quartile)

	Arith. Mean	Std. Dev.	Maximum	Minimum	Median
1Q	37.51519	342.61501	7098.00000	0.00012	**0.53740**
2Q	0.27266	0.69158	11.96657	0.00012	**0.11409**
3Q	0.10657	0.13216	1.26322	0.00051	**0.06829**
4Q	0.05203	0.07416	0.63103	0.00015	**0.02138**
1H	18.89392	242.98048	7098.00000	0.00012	0.19062
2H	0.07941	0.11059	1.26322	0.00015	0.03730
Total	9.49020	172.10316	172.10316	172.10316	172.10316

Table 5
Average R&D by Sales, Ranked by Total R&D (Ascending, by Quartile)

	Arith. Mean	Std. Dev.	Maximum	Minimum	Median
1Q	1.93200	16.51518	292.00000	0.00012	0.04869
2Q	17.05397	285.75051	7098.00000	0.00049	0.08746
3Q	11.28776	121.00463	2830.25000	0.00039	0.13387
4Q	7.68434	147.57031	3755.00000	0.00077	0.09691
1H	9.49298	202.53450	7098.00000	0.00012	0.07036
2H	9.48021	134.89365	3755.00000	0.00039	0.11632
Total	9.49020	172.10316	7098.00000	0.00012	0.09024

Table 6
Regression Results of Model II

Normalized RD Expenditures With Sector Dummy Variables
Dependent Variable: LNXRDSALE
Number of Observations: 1730

Variable	Estimated Coeff.	Std. Error	t-Statistic	p-value
C	-2.508717	0.260182	-9.642146	0.0000
LNSALEMP	-0.287021	0.041496	-6.916904	0.0000
LNATSAL	1.111356	0.037712	29.46947	0.0000
LNCSTK	-0.089290	0.010016	-8.915061	0.0000
AG11	-1.269797	0.811460	-1.564830	0.1178
INFO51	1.448761	0.148237	9.773255	0.0000
MAN3133	1.068970	0.124906	8.558167	0.0000
PROSER54	1.006416	0.185294	5.431444	0.0000
R-squared	0.616069			
Adjusted R-squared	0.614508			
Log likelihood	-3017.668	F-statistic		394.7401
Durbin-Watson stat	1.812978	Prob (F-statistic)		0.000000