Pepperdine Policy Review

Volume 5 Article 9

4-20-2012

Are All R&D Dollars Created Equal?: A Look at the Effect of Federal Investment on Patent Success

Casey J. O'Neil
Pepperdine University, casey.oneil@pepperdine.edu

Are All R&D Dollars Created Equal?

A LOOK AT THE EFFECT OF FEDERAL INVESTMENT ON PATENT SUCCESS

Casey O'Neil

Many have researched how Research and Development (R&D) spending affects innovation, but few have discussed the difference between federal and private investment. Does a dollar invested by the government have a greater or weaker effect than a private dollar on the percentage of successful U.S. patent bids?

This is an important question to answer. Solow's and Romer's growth models can be used to show innovation is responsible for over 40% of an economy's growth, and Arrow's work proves the level of U.S. innovation is still below the social optimum.[1,2,3] Thus, it is important to find the most efficient way to increase innovation. A definitive answer would determine whether the government should be directly funding R&D or if it can rely more on private funding. After all, the purpose of government funding is to step in where the market does not provide an incentive for private funding. If the more than $25 billion spent every year by the federal government is not effective, it can be cut from the budget.

While there are no studies addressing these exact questions, several studies do brush the topic. A 2003 study by Paroma Sanyal concludes that federal and private funds both have a positive influence on the number of patents granted. In particular, the study showed a greater effect by federal investment but still positive effect from private funding.[4] It should be noted, however, that, the work does not address the success of those patents. It is quite possible that federal investment only increased the quantity, not quality, of applications.

In 2008, David Nevy and Nestor Terleckyj found that federal funding has a positive effect on private funding. Specifically, one dollar of federal funding stimulates $0.27 of private

[1] Solow, Robert M, "A Contribution to the Theory of Economic Growth".
[2] Romer, Paul M. "The Origins of Endogenous Growth."
[3] Arrow, Kenneth. "Economic Welfare and the Allocation of Resources for Invention".
[4] Sanyal, Paroma. "Understanding Patents: The Role of R&D Funding Sources and the Patent Office".

investment. [5] While Sanval did not explore this dimension of federal-private investment relationship, Nevy and Terleckyj's study suggests the coefficient on federal funding is higher than that on private funding.

DESCRIPTION OF THE DATA

I constructed a time series dataset using information from the National Science Foundation (NSF), Bureau of Labor Statistics (BLS), United States Patent and Trademark Office (USPTO), and U.S. Census Bureau. The data span 45 years (1963-2007) and specifically provide data on U.S. federal investment, U.S. private investment, U.S. real Gross Domestic Product (GDP), annual U.S. unemployment rates, the quantity of utility patent applications submitted by domestic entities, the number of those applications approved, and the previous stock of approved U.S. utility patents regardless of the application origin. These are all national level data, not state or industry level; thus, policy recommendations that draw from the analysis must also be nationally focused. The reasons for including these variables will be explained in the next section.

One strength of this dataset is that it provides information on precisely what we seek to observe: utility patents granted to U.S. inventors. It also provides information on potential influencers of patent success - thus allowing us to more accurately isolate the coefficient on the variables of interest – federal investment and private investment. The last strength of the dataset is that it provides hard numbers. This will produce a more unbiased analysis than survey data would.

There are also several weaknesses inherent in the dataset. Firstly, while it is possible the effect of funding depends on its structure, the data do not differentiate among different types of federal and private funding. Does non-profit R&D spending produce different results than other private funds? What about federal funds released through government contracts, or through university research? The dataset also does not provide educational data before 1992. I had hoped to include the number of bachelor degrees in the labor force as a regressor, but the Bureau of Labor Statistics only began tracking that information in 1992, thus including what is available would not suffice for a serious analysis. The data also fall short in that they do not address other

[5] Levy, David M., and Nestor E. Terleckyj. "Effects of Government R&D on Private R&D Investment and Productivity: A Macroeconomic Analysis."

patent types or applications of foreign origin. Any recommendations will only be applicable to the United States.

Ideally, I would like to have application level data. With details to that extent it would be possible to estimate the effects using a logit model – thus predicting an individual patent application's probability of success. It would also be helpful to differentiate among different funding scenarios. Identifying the best way to distribute funding would ensure effective R&D spending – regardless of whether it originates from a federal or private source. I would also have liked to have data covering a longer time frame. Because the USPTO did not differentiate between U.S. and foreign applicants prior to 1963, the data simply are not available.

ECONOMETRIC MODEL

I ran a robust ordinary least squares (OLS) multiple regression on the data. The equation for the model is specified below:

$$y = \alpha + \beta x_1 + \ldots + \beta x_n + \varepsilon$$

In this model, y represents the Portion of successful U.S. utility patents,   represents the constant,   through   represent the regressors (Logged Federal RD, Logged Private RD, Logged GDP, Logged Unemployment, Logged Quantity of U.S. Utility Patent Applications, Logged Number of Past Utility Patents, and Lagged Portion of successful U.S. utility patents), and   is the error term.

For those unfamiliar with econometrics, the y variable is the dependent variable - its value is dependent on the x variables. To put it another way, the x variables are the factors theorized to affect y. Because the model is an estimate of best fit, epsilon then represents the difference between an estimated and observed value.

This model was used because the data meet all the OLS assumptions. They are normally distributed, independently and identically distributed, and the expected value of the error term is zero. The dataset did not provide enough information to perform other types of analyses. Panel data would be needed to perform fixed or random effects and a logit model would be appropriate if application-level data were available.

Before performing the regression, I produced Kernel Density (kdensity) plots to observe the skewness of the variables. The dependent variable, Portion of successful U.S. utility patents, is normally distributed and was thus left as-is. Most of the regressors (specifically Federal RD, Private RD, GDP, Unemployment, Quantity of U.S. Utility Patent Applications, and Number of

Past Utility Patents) however, were skewed to the right. I logged them so as to bring their distributions closer to normal, and thus increase the accuracy of the model.

I then performed a cross-sectional time-series feasible generalized least squares (FGLS) regression to test heteroskedasticity (non-uniform variance) and autocorrelation (correlation of observations within a variable). The results did show heteroskedasticity, but the robust option on the OLS regression can control the effects of that. The same analysis tests for autocorrelation, but none was found. Lastly, I tested for stationarity (statistical properties are consistent over time) of the dependent variable, Portion of successful U.S. utility patents. The Dickey-Fuller test showed we cannot reject non-stationarity. To account for this “walk” around the trend, a one-year lag of the dependent variable was added as a regressor.

I then regressed Portion of successful U.S. utility patents on the variables: Logged Federal RD, Logged Private RD, Logged GDP, Logged Unemployment, Logged Quantity of U.S. Utility Patent Applications, Logged Number of Past Utility Patents, and Lagged Portion of successful U.S. utility patents. The results of this are shown in TABLE 1.

A brief description of these variables and reasons for their inclusion in the regression is provided below:

Portion of successful U.S. utility patents: Continuous variable between 0 and 1. It provides the percentage of submitted utility patent applications that were ultimately approved. This is the regression’s dependent variable.

Logged Federal RD: Logged values of federal R&D spending – by source of funds. Federal investment is a theorized positive influence, and is one of the primary variables of interest.

Logged Private RD: Logged values of private R&D spending – by source of funds. It is a theorized positive influence, and is the other primary variable of interest.

Logged GDP: Logged values of the United States’ annual real GDP. It is a theorized positive influence, and thus, is used as a control variable.

Logged Unemployment: Logged values of the annual U.S. unemployment rates. It is a theorized positive influence, and thus, used as another control variable.

Logged Quantity of U.S. Utility Patent Applications: Logged values of the quantity of patent applications submitted each year. A greater number of submissions could be due to an increase in low quality applications and would thus decrease the percentage of successful applications. For this reason, it is included as a control variable.

Logged Number of Past Utility Patents: Logged values of the annual stock of existing patents. It is included in order to control for theorized positive impact a larger stock of knowledge and for the expected negative impact from having opportunities for a valid patent (the comparatively low-hanging fruit has been picked). It is thus included as another control variable.

Lagged Portion of successful U.S. utility patents: As discussed before, the dependent variable is non-stationary, and exhibits a "walk" around a trend. Including this one-year lag as a regressor makes the estimation more accurate.

EMPIRICAL RESULTS

Regressing Portion of successful U.S. utility patents on the variables described above yields the values shown in TABLE 1.

TABLE 1

Regression of *Portion of successful U.S. utility patents* on Federal R&D and Private R&D

RESULTS	OLS (no lag)		OLS (with lag)	
logFederal	-0.073	(0.19)	-0.136	(0.19)
logPrivate	-0.230*	(0.11)	-0.097	(0.13)
logGDP	1.542**	(0.54)	1.099*	(0.60)
logUnemployment	0.024	(0.09)	0.035	(0.09)
logPatApps	-0.054	(0.12)	-0.028	(0.12)
logPatStock	-1.746*	(0.90)	-1.360	(0.92)
Lagged Portion of Successful U.S. utility patents			0.187	(0.16)
Constant	17.628**	(8.46)	14.277	(8.65)
Observations	45		44	
Adj. R-squared	0.749		0.763	
Standard errors in parentheses				
** p<0.05, * p<0.10				

From TABLE 1, we see the coefficients on the variables interest, Logged Federal RD and Logged Private RD, are not statistically significant. This suggests that neither federal nor private investment, at least at a general level, increase the percentage of successful U.S. utility patents.

The results do not contradict Sanyal's claim that both federal and private funds increase the quantity of patent applications. Assuming the quality mix remains constant, an increase in applications would have no effect on the percentage of successful applications and would, in turn, mean that the coefficients on Logged Federal RD and Logged Private RD would be insignificant. Indeed, Logged Quantity of U.S. Utility Patent Applications is not significant, and neither is Logged Federal RD or Logged Private RD.

Since federal spending is not observed to have a greater association than private spending, the results appear to contradict the Levy and Terleckyj's study. If, as Levy and Terleckyj found, federal funding stimulated a fraction of private funding, and as Sanyal states, both increase patent applications, the coefficient on Logged Federal RD would be greater than that on Logged Private RD. This is not the case. It is possible, though, that this is due to the placement of the stimulated investment. If it is distributed equally across a constant quality mix (stimulated investment is not spent any more or less wisely than previous investment), the coefficients in our regression would be as observed.

IMPLICATIONS AND CONCLUSIONS

The regression results suggest that neither policies increasing federal funding nor those encouraging private investment will increase (or decrease) the percentage of successful U.S. utility patent bids. As mentioned before, however, this study looks only at general investment. It does not address targeted funding strategies.

It is logical, then, to look towards policies encouraging targeted investment. Further studies should also be focused in this area. We need more and better data – specifically application-level data, information on different funding scenarios, and data for more years. By obtaining these data and identifying other potential determinants of patent success, one could better recommend policies to increase the efficacy of federal R&D funding.

REFERENCES

Arrow, Kenneth. "Economic Welfare and the Allocation of Resources for Invention". Santa Monica, CA: RAND Corporation, 1959.

Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey. 2011. Raw data.

Levy, David M., and Nestor E. Terleckyj. "Effects of Government R&D on Private R&D Investment and Productivity: A Macroeconomic Analysis". The Economics of Innovation Policy. Volume 2, 19-29. Elgar Reference Collection. International Library of Critical Writings in Economics, vol. 222. Cheltenham, U.K. and Northampton, Mass.: Elgar, 2008. EconLit, EBSCOhost (accessed September 27, 2011).

National Science Foundation, Division of Science Resources Statistics. 2007. National Patterns of R&D Resources: 2006 Data Update. NSF 07-331. Brandon Shackelford and John E. Jankowski, project officers. Arlington, VA.

Romer, Paul M. "The Origins of Endogenous Growth." 1994. Journal of Economic Perspectives, 8(1): 3–22.

Solow Robert M, “A Contribution to the Theory of Economic Growth”. 1956, Quarterly Journal of Economics 70(1): 65–94.

Sanyal, Paroma. "Understanding Patents: The Role of R&D Funding Sources and the Patent Office". Economics of Innovation and New Technology 12, no. 6 (December 2003): 507-529. EconLit, EBSCOhost (accessed September 28, 2011).

Scherer, F. M. New Perspectives on Economic Growth and Technological Innovation. Washington, D.C.: Brookings Institution, 1999. Print.

United States Patent and Trademark Office. Patent History. 2011. Raw data.

U.S. Census Bureau. U.S. Real GDP. 2011. Raw data.