What Explains R & D Efficiency Differences across U.S. States

Mohammad Ashraf*

ABSTRACT. In this study we use data envelopment analysis to calculate relative efficiency indices of states with regard to the use of research and development (R&D) funds. Then we use regression analysis to explore the contributing factors to the observed efficiency differences. Our estimates indicate that number of firms, and college and university enrollment ratios have significant impact on how efficiently a state uses its R&D funds. Regional factors also play an important role in an efficient use of R&D funds (O31; O5).

I. Introduction

The role of R&D in increasing economic growth has been well documented (Aghion and Howitt 1998; Griliches 1988; Ha and Howitt 2007; Kumar and Russell 2002; Leahy and Neary 1997). Number of patents is often used as a measure of R&D output (Griliches 1990; Gurmu and Perez-Sebastian 2008; Ha and Howitt 2007). However, as we show below, output of R&D as measured by number of patents in the case of total R&D, and number of papers published in the case of academic R&D, differs from state to state. In this paper we find out which states use R&D funds more efficiently, and explore factors that may affect the efficient use of R&D funds. We use state-level data for the U.S. to perform our analysis.

This question is important for at least two reasons. One, R&D grant-making agencies, public and private, would like to know where their funds are used more efficiently. And secondly, states that are relatively inefficient would like to find out which factors affect the efficiency in a positive manner and replicate the success stories.

Although the literature on the topic of economic growth is rich and deep, to our knowledge, few studies have looked at what factors affect the efficiency of R&D spending. This study contributes to the literature on economic growth by filling this gap.

---

*Professor of Economics, School of Business, The University of North Carolina at Pembroke. Email: ashraf@uncp.edu. I want to thank the anonymous referees for their invaluable comments. All errors belong to the author.

In this study we use a two-step procedure. First, we use data envelopment analysis (DEA) to measure the relative efficiency of states with regard to the use of R&D funds. The natural logs of relative efficiency measures of states using DEA are then used as dependent variables in our regression analysis. The DEA approach was pioneered and formalized by Charnes, Cooper, and Rhodes (1978). The use of the DEA technique allows us to measure the relative efficiency without having to impose any functional form restrictions.

The rest of the paper is organized as follows. Section II provides a review of relevant literature, Section III provides details of the model used, Section IV provides details of data and data sources, Section V presents and discusses the results, and Section VI concludes the paper.

II. Literature Review

One of the most important questions in economics, and indeed in many other disciplines (i.e., public policy, sociology, political science, history, social work, etc.), is to figure out the sources of economic growth. The knowledge about the sources of economic growth becomes even more important when economies go through recessions. There is a rich body of literature in the area of economic growth. A comprehensive review of the literature is beyond the scope of this study. We refer the curious reader to Islam (2003) for an excellent review of literature on economic growth. Here we merely outline, to our knowledge, some of the most influential studies on the topic.

Starting with the seminal paper by Robert Solow (Solow 1956), which considers technological progress as the principal source of economic growth researchers have looked at the topic of economic growth from a number of different angles. For instance, papers by Robert Lucas (Lucas 1988), and Paul Romer (Romer 1986, 1990), consider broader measures of human and physical capitals as engines of economic growth.

R&D had been identified as one of the most important factors affecting economic growth (Acemoglu and Zilibotti 2001; Aghion and Howitt 1998; Barro 1990; Basu and Weil 1998; Bernard and Jones 1996; Feser, Renski, and Goldstein 2008; Frantzen 2002, 2004; Färe, Grosskopf, Norris, and Zhang 1994; Griliches 1988; Kumar and Russell 2002; Romer 1986, 1990; Scherer 1984; Zachariadis 2003, 2004). Studies point out that R&D expenditures have significant positive impact on output growth. According to the Coe and Helpman (1995) estimates, the return on R&D expenditure in terms of productivity gains for the G-7 countries is as high as 123% (as reported in Zachariadis, 2004, p.424).

Marios Zachariadis in his (2004) study uses data from 1971 to 1996 for "ten advanced OECD countries." He reports that "…increasing aggregate R&D intensity [measured as the ratio of R&D expenditures to output] by one percentage point increases output growth in aggregate economy by 0.38 percentage points." (Zachariadis, 2004, p.424).

However, to our knowledge, few studies have explored the factors that may affect the efficiency of R&D spending. This study attempts to fill this gap. In order to measure the efficiency of states in using R&D funds we construct a relative efficiency index using the DEA approach. Although a number of studies in the field of operations research have used the DEA approach (Chambers and Fare 2008; Olesen and Petersen 2002; Ray 2004), the use of DEA in economics has been rather slim (Barnum and Gleason 2008; Fanchon 2003; Jerzmanowski 2007; Kumar and Russell, 2002). An excellent review of literature on the use of DEA is presented in Førsund and Sarafoglou (2002).

In the next section we provide details of our model.

III. Model

We use a two-step procedure in this study. In the first step we calculate relative efficiency indices of states in utilizing R&D funds. We use the DEA technique to calculate these indices, which are later used in regression analysis. We divide this section into two subsections—Subsection A and Subsection B, which provide details of the DEA model and the regression model, respectively.

A. THE DEA MODEL

DEA is a non-parametric technique which constructs a piece-wise hull around data points. This allows one to measure the relative efficiency of entities. We follow the DEA literature convention and refer to the entity whose relative efficiency is being measured a decision making unit (DMU). The details of the R&D input and output measures are provided in the Data and Data Sources section where we elaborate on the data and data sources used to calculate the efficiency measures.

We use an output-oriented measure to calculate efficiency of different economic units. In an output-oriented measure the question asked is "by how much output(s) can be changed without changing the input(s)?" On the other hand, in an input-oriented measure the question asked is "by how much input(s) can be changed to achieve the same level of output(s)?" The choice between input- and output-oriented DEA depends upon the variable over which a DMU has more control. However the relative efficiency ranking of DMUs is not affected by the choice of an output-oriented measure versus an input-oriented measure. The output-oriented approach is appropriate in the present context because a DMU often has less control over the allotment of R&D funds. This is especially true when grants are the main sources of R&D. Figure 1 explains the input- versus output-oriented techniques graphically.

Assume that there are three DMUs, *a*, *b*, and *c*, and that all DMUs are using one input, *x*, to produce one output, *y*.

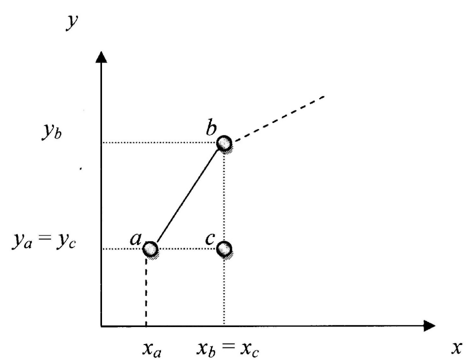

Figure 1. Input- and Output-Oriented Technical Efficiency

From an output-oriented point of view, DMUs *b* and *c* are using the same level of input, but DMU *c* is producing output y_c, with $y_c < y_b$. This makes DMU *c* a relatively inefficient DMU.

From the point of view of an input-oriented measure DMU *a* is a relatively efficient DMU. This is because both *a* and *c* DMUs are producing the same level of output, however DMU *c* is using input level x_c, with $x_c > x_a$.

In formulating the DEA model we make the following assumptions which are fairly unrestrictive. Here we present the basic setup of the model and refer the reader to Charnes et al. (1978); Ray (2004); and

Coelli et al. (2005) for details of the DEA technique.

To begin we assume that there are *J* DMUs producing *M* outputs using *N* inputs. The c^{th} DMU, for $j = 1, \dots, c, \dots, J$, uses a vector $x^c = (x_{1c}, \dots, x_{nc})$ of inputs to produce an output vector $y^c = (y_{1c}, \dots, y_{mc})$. With this setup in mind, we further assume that:

1. Each observed input output bundle is feasible. That is, an observed bundle (x^j, y^j) is a feasible bundle.
2. The production possibility set is convex. That is, if two bundles, (x^a, y^a) and (x^b, y^b), used by DMU *a* and *b*, respectively, are feasible bundles, then any bundle $(\bar{x}, \bar{y})$, which is a weighted average of these bundles is also feasible. We define $\bar{x} = \lambda x^a + (1 - \lambda) x^b$, and $\bar{y} = \lambda y^a + (1 - \lambda) y^b$, for $0 \leq \lambda \leq 1$.
3. Inputs and outputs are freely disposable. That is, if (x^o, y^o) is a feasible bundle then, for any $x \geq x^o$, (x, y^o) is also a feasible bundle. And if (x^o, y^o) is a feasible bundle then, for any $y \leq y^o$, (x^o, y) is also a feasible bundle.

Following Charnes et al. (1978) we define a DEA problem for the c^{th} DMU as:

$$\max hc = \frac{\sum_{m=1}^{M} u_m y_{mc}}{\sum_{n=1}^{N} v_n x_{nc}} \tag{1}$$

Subject to:

$$\frac{\sum_{m=1}^{M} u_m y_{mc}}{\sum_{n=1}^{N} v_n x_{nc}} \leq 1; \quad \forall j$$

Where *v* and *u* are the input and output weights, respectively and:

$v_n \geq 0$; $u_m \geq 0$; $n = 1, \dots, N$; $m = 1, \dots, M$

The data on inputs and outputs for all DMUs are used and the efficiency of the c^{th} DMU is rated relative to all the DMUs included in the dataset. Note that the problem presented in (1) is a fractional program, which is rather hard to solve. However, (1) can be converted into an ordinary linear programming problem using a solution suggested by Charnes et al. (1978) Ray (2004), and Coelli et al. (2005). In the interest of brevity we omit the detailed derivation of the solution technique and refer the interested reader to Charnes et al. (1978), Ray (2004), and Coelli et al. (2005).

We solve this linear programming problem for each DMU and obtain the efficiency scores which are used in the regression analysis as the dependent variable. In this study we use data for 51 DMUs (50 states of the US plus the District of Columbia). We use one input and one output for each of the DMUs. That is, $J = 51$, and $M = N = 1$. (More on the data used in the Data and Data Sources section.)

B. THE REGRESSION MODEL

In the second step we run regression using the relative efficiency index calculated in the first step as the dependent variable. Our population regression model takes the following form.

$$lnE = \beta_0 + \boldsymbol{\beta}' \boldsymbol{X} + \varepsilon \tag{6}$$

Where the dependent variable *lnE* is the natural log of the relative efficiency measure calculated in the first step, β_0 is the intercept, $\boldsymbol{\beta}'$ is a row vector of coefficients to be estimated, $\boldsymbol{X}$ is a column vector of the independent variables, and ε is the error term with constant variance, σ^2, and zero mean.

The variables included in the $\boldsymbol{X}$ vector are the number of colleges and universities in a state, the number of students enrolled, and number of total firms in a state. These variables were converted into per capita terms using state populations. We also included regional dummies to control for the regional effects on the efficiency measures. States included in each region are presented in Table 1.

TABLE 1–U.S. Regions

West	Midwest	South	Northeast
Alaska	Illinois	Alabama	Connecticut
Arizona	Indiana	Arkansas	Maine
California	Iowa	Delaware	Massachusetts
Colorado	Kansas	Florida	New Hampshire
Hawaii	Michigan	Georgia	New York
Idaho	Minnesota	Kentucky	New Jersey
Montana	Missouri	Louisiana	Pennsylvania
Nevada	Nebraska	Maryland	Rhode Island
New Mexico	North Dakota	Mississippi	Vermont
Oregon	Ohio	North Carolina	
Utah	South Dakota	Oklahoma	
Washington	Wisconsin	South Carolina	
Wyoming		Tennessee	
		Texas	
		Virginia	
		Washington, DC	
		West Virginia	

Source: US Census Bureau (www.census.gov)

IV. Data and Data Sources

This section is divided into three subsections. Subsection A provides details of the data used in the DEA, Subsection B provides details of data used in the regression analysis, and Subsection C gives details of data sources.

A. DATA USED IN THE DEA

In order to calculate the relative efficiency index using the DEA technique we use two datasets. In the first dataset we use total real R&D expenditure by a state as a percentage of real gross state product (hereafter total R&D ratio) at time t as the input variable, and the total number of patents granted to the residents of a state at time t+4 as the output measure. Both variables are used in per capita terms.

Our choice of using R&D expenditure as input measure and number of patents as output measure is guided by the economic growth literature. For instance, Zachariadis (2003) considers R&D expenditure to be an input in the "production of patents." (p.570). Madsen (2008) uses number of patents to measure "innovative activity" in OECD countries. More recently, Ashraf and Mohabbat (2010, p.40) argue that number of patents "represent relatively more closely the 'productivity' of R&D" expenditure. See also Griliches (1990) on this issue.

Following Zachariadis (2004), among others, we use real R&D expenditure as a percentage of real gross state product (GSP) as opposed to R&D expenditure in absolute terms. This is because a state spending a bigger *portion* of its GSP on R&D is in effect using higher amount of input than a state which spends a smaller portion of its GSP on R&D, even if the former state, as compared with the latter state, is spending a smaller absolute amount on R&D.

The choice of using a four-year lag in the case of total R&D ratio is guided by the evidence in economic growth literature. Studies show that the lag between R&D expenditure and its impact on output is about five years (Aghion and Howitt 1998; Griliches 1988). Studies also points out that achieving a patent is a step between R&D expenditure and its impact on output (Griliches 1990), and that the lag between achieving a patent and output is about one year or less (Ashraf and Mohabbat 2010). These facts led us to use a four-year lag.

To further solidify our analysis, before we used total R&D ratio and patents to calculate our relative technical efficiency estimates using the DEA technique, we used regression analysis to determine the time it takes for total R&D ratio to have the most impact on patents. Akaike's Information Criterion (AIC) and Schwartz Baysian Information Criteria (SBC) statistics pointed to a four-year lag as an optimal lag length for this dataset. See Geweke and Meese (1981) on use of AIC and SBC to determine lag length. To conserve space these regression results are not provided here. However, these results are available from the authors.

We use data for 2002 for total R&D ratio and 2006 for patents. That is, $t = 2002$, and $t+4 = 2006$. There are two reasons for choosing data for 2002 and 2006. First, we wanted to use the most recent data in our analysis. Second, we wanted to avoid the recession which started in the fourth quarter of 2007 (www.nber.org). The 2007 recession by most accounts is an outlier. For instance, according to the Bureau of Labor Statistics (www.bls.gov) unemployment report, unemployment rate reached 10.2 % in October 2009. This high unemployment rate had not been witnessed since December 1982 when the unemployment rate reached 10.8 %. We did not want our results to be influenced by an outlier data point.

In the second dataset we use the total real academic R&D expenditure, as a percentage of real gross state product (hereafter academic R&D ratio), as the input variable and the total number of peer reviewed papers published per one million dollars of total academic R&D as the output variable. Here again, before we used the DEA approach to calculate relative efficiency, we used regression analysis to determine the time it takes for academic R&D ratio to have the most impact on peer reviewed article output. The AIC and SBC statistics pointed to a three-year lag to be optimal in this case. In this dataset the latest year for which data are available is 2005. As a result, for academic R&D ratio our *t* is 2002 and *t*+3 is 2005 for the number of peer reviewed papers published.

B. DATA USED IN THE REGRESSION ANALYSIS

Using these two efficiency measures, with each one serving as our dependent variable, we run two different sets of regressions. In the first relative efficiency measure total R&D ratio and total number of patents are used as input and output variables, respectively. In regression analysis we use the natural log of this measure and represent it by lnE_{tot}.

In the second relative efficiency measure we use academic R&D ratio and the total number of peer reviewed papers published as input and output variables, respectively. In regression analysis we use the natural log of this measure as dependent variable and represent it by lnE_{aca}. Values of both relative efficiency measures range between 0 and 1, with 0 indicating the least efficient DMU and 1 the most efficient DMU, in relative terms.

As mentioned in the Model section, the right-hand-side variables include the number of colleges and universities in a state, the number of students enrolled, and number of total firms in a state, in per capital terms. To control for regional effects we divided states into four regions: West, Midwest, South, and Northeast. In dividing states into four regions we follow US Census Bureau (www.census.gov). See Table 1 for states included in each region. We provide more details on the structure of independent variables in Section V where we present and discuss the regression results.

C. DATA SOURCES

The data source for R&D expenditure is the National Science Foundation database (TABLE 11. Research and development expenditures, by state, performing sector, and source of funds: 2002, Millions of Current Dollars). The URL is www.nsf.gov.

In order to convert real R&D data as a percentage of state income, we used real gross state product data for state income. The data source is the US Bureau of Economic Analysis. The URL is www.bea.gov.

The data source for the number of degree granting institutions and the student enrollment is the US National Center for Education database (Table 271: Degree Granting Institutions, Number and Enrollment by State: 2000). The URL is http://nces.ed.gov.

The source for the patent data is the US Patent and Trademark Office, Technology Assessment and Forecast database. The URL is www.uspto.gov.

In order to convert these data into per capita terms we used the state population data. The data source is the US Census Bureau (Table 1: Annual Estimates of the Population for the United States, Regions, States, and Puerto Rico: April 1, 2000 to July 1, 2007 (NST-EST2007-01) Source: Population Division, U.S. Census Bureau, Release Date: December 27, 2007). The URL is www.census.gov.

In the next section we present and discuss the results of our estimations.

V. Regression Results

We first present and discuss the results where the dependent variable is lnE_{tot}—the natural log of the relative efficiency measure, where the relative efficiency measure is calculated using total R&D ratio and total number of patents as input and output variables, respectively. The results are presented in Table 2.

The structure of Table 2 is as follows. The first column lists the variable names. Parameter estimates, standard errors, t-stats, p-values, and standardized beta estimates are listed in columns two, three, four, five, and six, respectively. The last column, column seven, provides the variance inflation factor (VIF)–a measure of multicollinearity of the independent variables. A value greater than 5 indicates high degree of multicollinearity and demands attention (Studenmund 2006).

TABLE 2–Total R&D Expenditure and Total Number of Patents

Dependent Variable: $1nE_{tot}$ (natural log of the relative efficiency measure, where the relative efficiency measure is calculated using total R&D ratio and total number of patents as input and output variables, respectively)

Variable	β Estimates	HC-Std. Error	t-value	p-value	Standardized β Estimate	VIF
Intercept	5.931[a]	2.19	2.71	0.009	-	-
ln(enroll/ institutions)	0.599[b]	0.256	2.34	0.024	0.287[b]	1.193
ln(total firms_pc)	2.226[a]	0.565	3.94	0.000	0.517[a]	1.426
Northeast	0.633[a]	0.224	2.83	0.007	0.337[a]	1.493
Midwest	0.45[b]	0.174	2.59	0.013	0.259[b]	1.355
West	0.517[b]	0.219	2.36	0.023	0.315[b]	1.592

$\bar{R}^2$=0.51; F-stat = 11.24 (p-value F-stat= 0.00); HC-Std Error = Heteroscedasticity Consistent Standard Errors; Significance Levels: a = 1%, b= 5%, c=10%; Null hypothesis of homoscedastic errors: χ^2-stat = 17.16 (p-value χ^2-stat = 0.248).

In this table we present the standardized coefficient estimates, along with the "regular" estimates. The benefit of standardizing the estimates is that in a multiple regression one can judge the relative importance of each variable. A significant coefficient estimate indicates that a one standard deviation change in the independent variable will lead to a $\hat{\beta}$ standard deviations change in the dependent variable, where $\hat{\beta}$ is the coefficient estimate value. While discussing the results we will refer to the standardized estimates.

In this model our dependent variable is lnE_{tot}. Recall that we calculate this relative efficiency measure with total R&D ratio and total number of patents as input and output variables, respectively. The independent variable *ln*(*enroll/institutions*), hereafter, enrollment ratio, is the natural log of the ratio of the total number of students enrolled and the total number of degree granting institutions in a state (in per capita terms). The second independent variable *ln*(*total firms_pc*) is the natural log of the total number of firms in a state (in per capita terms). Northeast, Midwest, and West are the regional dummies, with the South being the control variable. States included in each region are listed in Table 1 above. With regard to the independent variables, we experimented with a number of variants of the model. The present model was selected on the basis of coefficient significance, F-stat, $\bar{R}^2$, and VIF.

The results present a rather interesting picture. First of all, note that lal the variables included in the model are significant at least at the 5% level. The significance of the coefficient estimates is measured using heteroscedasticity consistent standard errors. See Stock and Watson (2003) on the use of heteroscedasticity consistent standard errors. Note also that the null hypothesis of homoscedastic errors is not rejected. A non-rejection of the null hypothesis is an indication of errors being homoscedastic. The χ^2-stat is 17.16 with p-value=0.248. As Stock and Watson (2003) note, it is preferable to use heteroscedasticity consistent standard errors even if the null hypothesis of homoscedastic standard errors is not rejected.

The overall fit of the model, as represented by the F-stat, is significant at the 1% level. The value of $\bar{R}^2$ = 0.51 indicates that about 51% of the variation in the dependent variable is explained by the independent variables included in the model.

Starting with the enrollment ratio, we find that it has a significant and positive impact on how efficiently a state uses its R&D funds. The standardized coefficient estimate implies that a one standard deviation increase in the enrollment ratio will lead to 0.287 standard deviations increase in the efficiency of a state in utilizing the R&D funds.

Recall that enrollment ratio measures the number of students enrolled per university or college. A relatively higher enrollment ratio implies relatively larger universities and colleges with regard to student population. Usually at larger universities and colleges research by faculty members is one of the main requirements. Larger universities also have basic physical capital equipment in terms of laboratories and human capital in terms of faculty members and graduate students in place. As a result R&D funds can be put to use relatively quickly and more efficiently in larger universities. Just as certain countries may be better suited for a given technology (Los and Timmer 2005), larger institutions of higher education may be better suited for utilzing R&D grants and performing research than smaller ones.

Next recall that number of patents awarded to the residents of a state is the output measure in the calculation of relative efficience index, lnE_{tot}–the dependent variable in this model. Some of the research conducted in universities and colleges may result in acquiring patents. This is especially true in the case of physical sciences, i.e., physics, chemistry, and biology. In the present context this means that states with larger universities seem to use R&D funds more efficiently.

Another possible channel of relatively efficient use of R&D funds is the cooperation between universities and industry. Because of the availability of larger knowledge pool in the form of faculty and graduate students, firm managers are also likely to be gravitated towards larger universities. Some of this cooperation may result in acquiring patents. (We included the enrollment ratio and number of firms interaction term in the model. However, the VIF value pointed to collinearity. As a result, we dropped the interaction term from the final model.)

The next variable is the natural log of total number of firms (in per capita terms). It also has a positive significant impact on the ability of a state to utilize the R&D funds efficiently. The standardized coefficient estimate implies that a one standard deviation increase in the natural log of the total number of firms per capita will lead to 0.517 standard deviations increase in the dependent variable, lnE_{tot}, which measures the efficiency of a state in utilizing R&D funds.

One explanation of this result may be that the larger the industrial base in a state, the better the infrastructure in place to take advantage of R&D funds, and the more efficient the use of R&D funds. Another explanation may be that a larger number of firms will lead to inter firm competition. This may lead to an efficient use of R&D funds. Yet another explanation may be that a larger industrial base will attract larger, deeper, and more competent human capital pool. Competition among workers may also lead to an efficient use of given R&D funds. Furthermore, an interaction among workers may lead to further depending and broadening the knowledge pool resulting in even more efficient use of R&D funds. Indeed this is one of the channels of economic growth postulated by endogenous growth theory (Aghion and Howitt 1998; Lucas 1988; Romer 1986, 1990).

The coefficient estimates associated with the regional dummies are also of great interest. Recall that the South is in the control region. The coefficient estimates imply that compared to the South, all three regions, West, Northeast and Midwest, use R&D dollars more efficiently. This result may be due to the relatively broad industrial base in the West, Northeast, and Midwest.

In Table 3 we present the regression results where the dependent variable is lnE_{aca}—the natural log of relative efficiency measure which is calculated using academic R&D ratio and the total number of peer reviewed paperes published as input and output variables, respectively. The structure of Table 3 is the same as that of Table 2.

TABLE 3–Total Academic R&D Expenditure and Total Number of Papers Published

Dependent Variable: lnE_{aca} (natural log of the relative efficiency measure, where efficiency measure is calculated using academic R&D ratio and total number of peer reviewed papers published as input and output variables, respectively)

Variable	β Estimate	HC-Std. Error	t-value	p-value	Standardized β Estimate	VIF
Intercept	3.851[a]	0.087	44.26	0.000	0	0
ln(enroll/ institutions)	0.26[a]	0.0697	3.73	0.001	0.382[a]	1.068
Northeast	0.146[b]	0.064	2.28	0.027	0.238[b]	1.254
Midwest	0.105	0.064	1.64	0.108	0.185	1.281
West	-0.156[c]	0.087	-1.79	0.08	-0.29[c]	1.302

$\bar{R}^2$=0.22; F-stat = 4.46 (p-value F-stat= 0.004); HC-Std Error = Heteroscedasticity Consistent Standard Errors; Significance Levels: a =1%, b=5%, c=10%; Null hypothesis of homoscedastic errors: χ^2-stat = 6.26 (p-value χ^2-stat = 0.619).

Here again we find some interesting results. The F-stat and the $\bar{R}^2$ point to a rather good fit of the model. The value of $\bar{R}^2 = 0.22$ implies that about 22% of the variation in the dependent variable is explained by the independent variables. As in Table 2, coefficient estimate significance levels are measured using heteroscedasticity consistent standard errors. The null hypothesis of homoscedastic errors is again not rejected. The X^2-stat is 6.26 with a p-value of 0.619. This result points to errors being homoscedastic.

The independent variables included in the model are the enrollment ratio and the regional dummies: Northeast, Midwest, and West. The South serves as the control variable. As in Table 2, the model presented here was selected on the basis of coefficient significance, F-Stat, $\bar{R}^2$, and VIF values.

The enrollment ratio is again significant at the one percent level. The standardized coefficient estimate implies that a one standard deviation increase in the natural log of the enrollment ratio will lead to a 0.382 standard deviations increase in lnE_{aca}—the natural log of the efficiency measure of a state in utilizing academic R&D funds.

As mentioned above, larger universities are relatively more research oriented. Publications in leading peer reviewed journals and the ability to secure grants are required for promotion and tenure. To take economics as an example, a cursory review of job advertisements in publications such as *Job Openings for Economists* will show that larger universities expect candidates to have high quality research and the ability to secure grants as one of their top priorities. It stands to reason then, that faculty members at these universities will be more involved in research. As a result, more peer reviewed articles will be published.

Interestingly, regional dummy estimates indicate that as compared with the South, only the Northeast uses its academic R&D funds more efficiently. The coefficient estimate is significant at the 5% level. We do not find any statistically significant difference between the South and the Midwest. And as compared with the South, the West uses its academic R&D dollars less efficiently. The coefficient for the West is negative and significant at the 10% level. The implication of this result is that, when it comes to producing peer reviewed publications, the Southern universities do not lag behind their Midwestern and Western counterparts.

VI. Conclusion

This paper contributes to the literature by providing insights into the factors that may affect the efficient use of R&D funds by states. We use state level data for the U.S. and make use of the DEA technique to calculate relative efficiency measures. Natural logs of these efficiency measures are then used as dependent variables in the regression analysis. Our results indicate that enrollment ratio and number of firms along with the regional factors are the major contributors to the efficient use of R&D funds. From the point of view of R&D funds granting agencies, larger institutions of higher learning are better suited for their funds. This result does not bode well for smaller institutions.

Future research may also utilize the time-series dimension of data. In this respect a panel estimation approach may be more illuminating in deciphering factors that make a state more efficient user of R&D funds.

References

Acemoglu, D., and F. Zilibotti. 2001. "Productivity Differences." *The Quarterly Journal of Economics,* 116(2): 563-606.

Aghion, P., and P. Howitt. 1998. *Endogenous Growth Theory*. Cambridge, MA: The MIT Press.

Ashraf, M, and K. A. Mohabbat. 2010. "Output Convergence and the Role of Research and Development." *Annals of Economics and Finance*, 11(1): 35-71.

Barnum, D. T., and J. M. Gleason. 2008. "Bias and Precision in the DEA Two-stage Method. *Applied Economics,* 40(16-18): 2305-2311.

Barro, R. J. 1990. Government Spending in a Simple Model of Endogeneous Growth. *The Journal of Political Economy,* 98(5, Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems): S103-S125.

Basu, S., and D. N. Weil. 1998. Appropriate Technology and Growth. *The Quarterly Journal of Economics,* 113(4): 1025-1054.

Bernard, A. B., and C. I. Jones. 1996. Comparing Apples to Oranges: Productivity Convergence and Measurement across Industries and Countries. *The American Economic Review,* 86(5): 1216-1238.

Chambers, R. G., and R. Fare. 2008. A 'Calculus' for Data Envelopment Analysis. *Journal of Productivity Analysis,* 30(3): 169-175.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research,* (2): 429-444.

Coe, D. T., and E. Helpman. 1995. International R&D Spillovers. *European Economic Review,* 39: 859–887.

Coelli, T. J., D. S. P. Rao, , C. J. O'Donnell, and G. E. Battese. 2005. *An Introduction to Efficiency and Productivity Analysis* (2nd ed.). New York: Springer.

Fanchon, P. 2003. Variable Selection for Dynamic Measures of Efficiency in the Computer Industry. *International Advances in Economic Research,* 9(3): 175-186.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. 1994. Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review,* 84(1): 66-83.

Feser, E., H. Renski, and H. Goldstein. 2008. Clusters and Economic Development Outcomes: An Analysis of the Link between Clustering and Industry Growth. *Economic Development Quarterly,* 22(4): 324-344.

Førsund, F. R., and N. Sarafoglou. 2002. On the Origins of Data Envelopment Analysis. *Journal of Productivity Analysis,* 17(1): 23-40.

Frantzen, D. 2002. Intersectoral and International R&D Knowledge Spillovers and Total Factor Productivity. *Scottish Journal of Political Economy,* 49(3): 280-303.

Frantzen, D. 2004. Technological Diffusion and Productivity Convergence: A study for Manufacturing in the OECD. *Southern Economic Journal,* 71(2): 352-376.

Geweke, J., and R. Meese. 1981. Estimating Regression Models of Finite but Unknown Order. *International Economic Review,* 22(1): 55-70.

Griliches, Z. 1988. Productivity Puzzles and R & D: Another Nonexplanation. *The Journal of Economic Perspectives,* 2(4): 9-21.

Griliches, Z. 1990. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature,* 28(4): 1661-1707.

Gurmu, S., and F. Perez-Sebastian. 2008. Patents, R&D and Lag Effects: Evidence from Flexible Methods for Count Panel data on Manufacturing Firms. *Empirical Economics,* 35: 507-526.

Ha, J., and P. Howitt. 2007. Accounting for Trends in Productivity and R&D: A Schumpeterian Critique of Semi-endogenous Growth Theory. *Journal of Money, Credit and Banking,* 39(4): 733-774.

Islam, N. 2003. What have We Learnt from the Convergence Debate? *Journal of Economic Surveys,* 17(3): 309-362.

Jerzmanowski, M. 2007. Total Factor Productivity Differences: Appropriate Technology vs. Efficiency. *European Economic Review,* 51(8): 2080-2110.

Kumar, S., and R. R. Russell. 2002. Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence. *American Economic Review,* 92(3): 527-548.

Leahy, D., and J. P. Neary. 1997. Public Policy towards R&D in Oligopolistic Industries. *The American Economic Review,* 87(4): 642-662.

Los, B., and M. P. Timmer. 2005. The 'Appropriate Technology' Explanation of Productivity Growth Differentials: An empirical approach. *Journal of Development Economics,* 77(2): 517-531.

Lucas, R. E. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics,* 22(1): 3-42.

Madsen, J. 2008. Semi-endogenous versus Schumpeterian Growth Models: Testing the Knowledge Production Function Using International Data. *Journal of Economic Growth,* 13(1): 1-26.

Olesen, O. B., and N. C. Petersen. 2002. The Use of Data Envelopment Analysis with Probabilistic Assurance Regions for Measuring Hospital Efficiency. *Journal of Productivity Analysis,* 17(1): 83-109.

Ray, S. C. 2004. *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*. New York: Cambridge University Press.

Romer, P. M. 1986. Increasing Returns and Long-run Growth. *Journal of Political Economy,* 94(5): 1002-1037.

Romer, P. M. 1990. Endogenous Technological Change. *The Journal of Political Economy, 98*(5, Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems): S71-S102.

Scherer, F. M. 1984. *Innovation and Growth: Schumpeterian Perspective*. Cambridge, MA: The MIT Press.

Solow, R. M. 1956. A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics,* 70(1): 65-94.

Stock, J. H., and M. W. Watson. 2003. *Introduction to Econometrics* Addison-Wesley.

Studenmund, A. H. 2006. *Using Econometrics: A Practical Guide* (5th ed.) Addison-Wesley.

Zachariadis, M. 2003. R&D, Innovation, and Technological Progress: A Test of the Cchumpeterian Framework without Scale Effects. *The Canadian Journal of Economics /Revue Canadienne d'Economique,* 36(3): 566-586.

Zachariadis, M. 2004. R&D-Induced Growth in the OECD? *Review of Development Economics,* 8(3): 423-439.