

Production of Output and Ideas: Efficiency and Growth Patterns in the US

Kyriakos Drivas^a, Claire Economidou^{b,*}, Efthymios G. Tsionas^c

^aDepartment of International and European Economic Studies, Athens University of Economics and Business, Athens 104 34, Greece

^bDepartment of Economics, University of Piraeus, Piraeus 184 35, Greece

^cDepartment of Economics, Athens University of Economics and Business, Athens 104 34, Greece

Abstract

This paper examines efficiency and growth patterns in the production of output and ideas in the US. We employ frontier techniques and jointly estimate the production of output and the production of ideas. We find states to be particularly efficient in the use of their inputs in the production of output process, whereas there is more waste in the use of innovation resources to produce new knowledge. Our results do not lend support to the common perception that richer (more innovative) states are more efficient than less richer (less innovative) states for every dollar spent. Furthermore, an efficient in producing knowledge state is not necessarily efficient in producing goods and services. Finally, a consistent finding is that input accumulation is an important source of growth both for output and for the knowledge.

Keywords: innovation, R&D, patents, efficiency, system of stochastic frontiers

JEL: C11, C24, O32, O51

1. Introduction

Investing in R&D activities to update technological capabilities is an important driver of economic growth and a means for firms, regions, and countries, in general, to face domestic and international competition. Many governments around the world have supported extensive R&D promotion activity of one form or another, ranging from government-sponsored R&D consortia and national R&D laboratories to

^{*}We thank Dimitris Karamanis, Joanna Pagoni and Zoi Georgiopolou for competent research assistance. We gratefully acknowledge helpful comments received at seminars/workshops hosted at the University of Piraeus and the Athens University of Economics and Business. Finally, Kyriakos Drivas gratefully acknowledges financial support from the *National Strategic Reference Framework* No: SH1_4083. The usual disclaimer applies.

^{*}Corresponding author.

Email addresses: dribask@aub.gr (Kyriakos Drivas), economidou@unipi.gr (Claire Economidou), tsionas@aub.gr (Efthymios G. Tsionas)

various fiscal incentives (e.g., subsidies, tax cuts).¹ Innovation activities further affect the marginal productivity of physical capital and labor and therefore shape the production of real output. It becomes even more important for countries' growth when a country is on or approaches the world technology frontier, as there is less room for imitating (copying) already existing technologies.

Nonetheless, additional spending on R&D may be of little, if any, help in enhancing technological progress, if innovation resources which are limited, are used inefficiently. Consequently, the efficient use of resources is of utmost importance in the design of effective policies. In addition, the production of innovation (technological knowledge/ideas) also shapes the production of output (goods and services). For instance, more-(less-) efficient production of knowledge could be associated with more-(less-) efficient production of output as research-oriented firms allocate their resources into the research sector and production of final goods sector (Romer, 1990).

This paper accounts for efficiency in the production of technological knowledge (ideas) as well as in the production of output (goods and services) and apply stochastic frontier analysis; a flexible modeling approach that avoids strong assumptions (e.g. efficient use of resources, etc) of the literature. This modeling strategy adds structure to the unexplained residual and, under reasonable assumptions, disentangles the residual into inefficiency and measurement error. To capture the "jointness in production" of output and ideas - as decisions about knowledge and output productions are simultaneous and, therefore, statistical errors are unlikely to be orthogonal - we estimate jointly the two production frontiers of output and ideas.

Our modeling approach further allows for the decomposition of output and knowledge change into three components, namely technical change, efficiency change, and input change. We aim to assess their individual role in the growth process and discuss important questions in the literature, among others, whether there is any waste in the production of output and ideas, or whether efficient in production of output regions are also efficient in producing ideas.

Consequently, the contribution of this paper is twofold: first, the joint estimation of a production and knowledge function within the stochastic frontier framework, and second, the disentanglement of the drivers of output and knowledge change.

Our paper relates and further adds to large strands of economic literature. For example, the empirical

¹R&D subsidies have been an important tool to support technology policy in the OECD countries. On average, roughly one third of funds for R&D are provided by the government (<http://www.nsf.gov/statistics/seind14/>). More specifically, public support for university research grant programs has increased dramatically in the US since 2009, as a response to the financial crisis. For example, the American Recovery and Reinvestment Act of 2009 (ARRA) has allocated 3 billion dollars to the National Science Foundation (NSF), representing an increase of 50% over the NSF's annual budget to 6 billion. Similarly, the ARRA has allocated 10 billion to the National Institute of Health, representing an increase of more than 30% of the NIH's annual budget to 30 Billion (Whalley and Hicks, 2014). The Israeli experience is also of interest because of its boomed high-tech sector and considerable government-financed R&D activity (Trajtenberg, 2002).

testing of growth models has typically examined output (or productivity) growth ignoring any waste in the use of resources.² A growing body of empirical literature has conducted efficiency analyses along lines similar to that proposed in this paper, but used different modeling approaches. For instance, Färe et al. (1994) and Henderson and Zelenyuk (2007) use data envelopment analysis, while Koop et al. (1999, 2000) and Limam and Miller (2004) use stochastic frontier analysis to examine country-specific inefficiency in a number of developed and developing countries.³ Furthermore, the standard approach in the economics of innovation literature, so far, has been the use of a knowledge (innovation) production function, where the innovative output, the counts of patents, is produced analogously to the production of real output, employing existed knowledge and human capital allowing no waste in their use (Griliches, 1979, 1992; Jaffe, 1989; Feldman, 1994). A still small, body of empirical studies carries out innovation efficiency analyses. For example, the studies of Rousseau and Rousseau (1997, 1998) are among the few attempts, which consider efficiency in innovation process. These studies evaluate innovation performance of a number of, mainly OECD, countries and rank them according to their efficiency scores. A number of subsequent studies, perform similar analysis controlling also for socio-economic factors, institutions, and regulations (Wang, 2007; Wang and Huang, 2007; Cullmann et al., 2012; Franco et al., 2016).⁴ The literature, thus far, has studied separately the production of output and ideas and the possible waste in the use of their inputs. In this paper, we consider the interaction of the two processes and estimate a system of two stochastic frontier productions.

We apply our modeling approach for the case of the US. The latter, has been one of the leading and most prolific nations in innovation research, production of new ideas, which are patented, and also among the most developed countries in terms of output production. Extant cross-country research on either production or knowledge efficiency has considered the US as a homogeneous entity and has not focused on the sub-national level (Koop et al., 1999; Henderson and Zelenyuk, 2007; Wang and Huang, 2007; Cullmann et al., 2012), there is no study hitherto that performs regional (state-level) analysis for the US. While cross-country efficiency analyses have their relevance, the study of efficiency performance at regional level offers valuable insights. Recently an emerging line of research performs efficiency analysis at regional level with the vast majority to focus on innovation efficiency in Germany (Broekel and Schlump, 2009; Brenner and Broekel, 2011; Fritsch and Slavtchevc, 2011; Broekel, 2012; Broekel et al., 2013) and China (Altvater-

²See, for instance, Jones (1995), Coe and Helpman (1995), Aghion and Howitt (1998), Griffith et al. (2004a), Zachariadis (2003), Bottazzi and Peri (2003), Bottazzi and Peri (2007), and Mancusi (2008) among others.

³Various studies also investigate the role of efficiency in explaining growth differentials for a panel of manufacturing industries in the OECD countries. See, for instance, Koop (2001) and Kneller and Stevens (2006).

⁴See Cruz-Cázares et al. (2013) for an updated review on innovation efficiency studies.

Mackensen et al., 2005; Chen et al., 2009; Bai, 2013; Franco and Leoncini, 2013).⁵ The study of Thomas et al. (2011) has been the only attempt so far to assess innovation efficiency for the US states for the period 2004-2008. The authors, however, measure innovation efficiency based on the ratio of R&D outputs (e.g. patents granted or scientific publications) to R&D inputs (e.g. R&D expenditure) concluding that only 14 out of 51 states show modest improvements in innovation efficiency. As states belong in the same country and, therefore, share common institutions, among other things, an interesting issue that arises is whether small differences in states' policies could have different growth implications. Our study consists the first attempt that performs regional analysis and jointly estimates the production of output and knowledge for the states in the US.

Although this paper examines efficiency and growth patterns across states in the US, from the outset of this study, we would like to mention that we do not investigate the sources of states' in(efficiency). The role of institutions and regulations, the cooperation between universities and industry, state-policies related to innovation (taxation, subsidies, noncompetes, etc) are important in shaping efficiency and growth patterns, but a thorough investigation of their impact is beyond the scope of this study.

We, therefore, apply our modeling approach to fifty US states over the period 1993-2006, aiming at answering two key questions: (i) Are the US states efficient in the use of their production inputs? (ii) What drives output and knowledge change in the US?

We find that the states of the US are particularly efficient in the use of their inputs in the production of output process and less efficient in the use of their innovation resources to produce new knowledge; this finding does not come as a surprise as the production of new knowledge is subtler and riskier than the production of goods and services. While the dispersion of output efficiency is smaller than that of knowledge efficiency, the levels of states' efficiency in both types of productions stays pretty much stable during our sample period with a small inclination to decline. Furthermore, an efficient in producing knowledge state is not necessarily efficient in producing goods and services. Finally, a consistent finding is that input accumulation is an important source of growth both for output - a finding we share with related literature (Koop et al., 1999, 2000) - and for the knowledge.

Overall, our efficiency estimates do not lend support to the common perception that wealthier, in terms of state-level gross domestic product, (or more innovative, in terms of R&D spending) states are more efficient than less wealthier (less innovative) states. In fact, it is often the case that being wealthier (or innovative) still allows some room for efficiency improvement. A policy implication that can be derived is

⁵For an extensive exposition of related literature, see Appendix, section A1.

that government efforts (e.g. state subsidization of R&D) should not only be directed toward states with high R&D, but also to states with modest R&D, but which are capable to produce more output or ideas for every dollar spent.

The remaining of the paper proceeds as follows. Section 2 introduces the methodology of the paper. Section 2 presents the data. Section 3 discusses the results. Finally, Section 4 summarizes the findings and concludes.

2. Methodology

In this section, we introduce a model of production, which accounts for inefficient use of inputs. The outcome of the production can be either output (goods and services) or new innovations (ideas/technological knowledge). In doing so, we employ stochastic frontier analysis and estimate the production of output and production of ideas simultaneously, as a system, to account for possible interactions between them, meanwhile, controlling for input endogeneity. At a second stage, we decompose output (knowledge) change into three components: input change, technical change and efficiency change to discuss important questions in the growth literature.

2.1. Stochastic Frontier Model of Production

We model the performance of states' production (of output or new knowledge) by means of a stochastic frontier model.⁶ A frontier production function defines the maximum outcome attainable, given the current production technology and available inputs.

If all states produce output (or knowledge), denoted by Y , on the boundary of a common production set (see point A of Figure 1, below) that consists of an input vector with two arguments, capital stock, $Capital$, and labor, $Labor$, then this process can be described as:

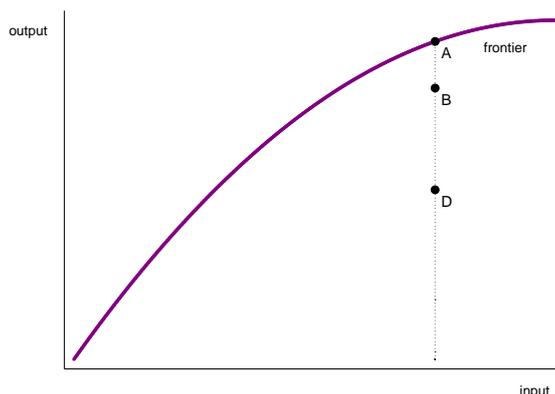
$$Y_{it}^* = f(Capital_{it}, Labor_{it}, t; \beta) \exp\{v_{it}\} \quad (1)$$

where, Y_{it}^* is the frontier (maximum) level of output (or knowledge) in state i at time t , f and parameter vector β characterize the production technology, t is a time trend variable that captures neutral technical change, and v_{it} is an i.i.d. error term distributed as $N(0, \sigma_v^2)$, which reflects the stochastic character of the frontier.

⁶Stochastic frontier analysis (SFA) was introduced by Aigner et al. (1977), Battese and Corra (1977), and Meeusen and van den Broeck (1977).

Some states, however, may lack the ability to employ existing resources efficiently and, therefore, produce less than the frontier output (knowledge). As Figure 1 below shows, for a given technology and set of inputs, there are states that produce at points (B) and (D); in other words their actual output (knowledge) is less than the maximum (frontier) attainable output (knowledge).

Figure 1: A Stochastic Frontier Model



If the difference between maximum attainable (frontier) and actual (observable) output (knowledge) is represented by an exponential factor, $\exp\{-u_{it}\}$, then the actual output (knowledge), Y_{it} , produced in each state i at time t can be expressed as a function of the stochastic frontier output (knowledge), Y_{it}^* or equivalently:

$$Y_{it} = f(\text{Capital}_{it}, \text{Labor}_{it}, t; \beta) \exp\{v_{it}\} \exp\{-u_{it}\}, \quad (2)$$

where, $u_{it} \geq 0$ is assumed to be i.i.d., with a half-normal distribution truncated at zero $|N(0, \sigma_u^2)|$ and independent from the noise term, v_{it} .⁷ Technical efficiency, $\exp\{-u_{it}\}$, is measured as the ratio of actual over frontier output (knowledge) that is, $\exp\{-u_{it}\} = \frac{Y_{it}}{Y_{it}^*}$ such that $0 \leq \exp\{-u_{it}\} \leq 1$ and $\exp\{-u_{it}\} = 1$ implies full efficiency.⁸

Two aspects of equation (2) are important. First, the frontier, as it is defined, represents a set of maximum outputs (knowledge) for a range of input vectors. Therefore, at any moment in time, it is defined by

⁷We decompose the composite residual in equation (2), $\exp\{\varepsilon_{it}\} = \exp\{v_{it}\} \exp\{-u_{it}\}$, and identify its components, $\exp\{v_{it}\}$ and $\exp\{-u_{it}\}$, by re-parameterizing λ in the maximum likelihood procedure, where $\lambda (= \sigma_u / \sigma_v)$, the ratio of the standard deviation of efficiency over the standard deviation of the noise term, and $\sigma (= (\sigma_u^2 + \sigma_v^2)^{1/2})$ is the composite standard deviation. The frontier is identified by the λ for which the log-likelihood is maximized (see Kumbhakar and Lovell, 2000).

⁸States may also be inefficient if they use an input mix at which marginal returns to inputs do not equalize with true factor market prices. We do not consider this 'allocative' efficiency, because input prices are not available for the disaggregated data we use in our analysis. Therefore, in this paper, the term 'efficiency' refers purely to technical efficiency.

the observations from a number of states and not just from one. A state is inefficient, if it fails to absorb the best-practice technology. Second, our modeling approach treats the frontier as stochastic through the inclusion of the error term v_{it} , which accommodates noise in the data and, therefore, allows for statistical inference. In this respect, it fundamentally differs from other (non-parametric) frontier analyses (Wang, 2007; Cullmann et al., 2012) that do not allow for random shocks around the frontier.⁹

To operationalize equation (2), one needs to specify the functional form of the production frontier. We employ a translog production function and, therefore, equation (2) can be expressed as:¹⁰

$$\begin{aligned} \ln Y_{it} = & \beta_0 + \beta_1 \ln Capital_{it} + \beta_2 \ln Labor_{it} + \frac{1}{2}\beta_{11}(\ln Capital_{it})^2 + \frac{1}{2}\beta_{22}(\ln Labor_{it})^2 \\ & + \beta_{12} \ln Capital_{it} \ln Labor_{it} + \gamma_t t + \delta_{Capital\ t} \ln Capital_{it} + \delta_{Labor\ t} \ln Labor_{it} + v_{it} - u_{it}, \end{aligned} \quad (3)$$

where, t a time trend that captures neutral technical change.

The generic equation (2), and consequently its functional expression in (3), can describe the process of either production of goods and services or the evolution of technological knowledge creation (ideas). The production of goods and services is more straight-forwardly defined. In the R&D-based endogenous growth models, the production of knowledge is a function of labor force in the research sector and of the available stock of knowledge (Romer, 1990; Kortum, 1993; Grossman and Helpman, 1994; Abdih and Joutz, 2006). As innovation outcome, we use patents, which is a good proxy for economically profitable ideas (Bottazzi and Peri, 2003).¹¹ Knowledge and innovation, however, are a much broader concept than the count of patented blueprints. In spite of their obvious caveats - not all inventions are patented and neither do they have the same economic impact (Griliches, 1990), equally not all patents are commercially exploitable innovations, they do present the minimal standards of novelty, originality and potential economic profits.¹²

Therefore, depending on how output and inputs are defined each time, equation (3) can describe either the production of output or knowledge.

In this paper, we want to capture "jointness in production" of output and ideas in a statistical way. The reason is that the statistical error terms are unlikely to be orthogonal as decisions about knowledge and

⁹For comprehensive reviews of frontier methodologies, see Kumbhakar and Lovell (2000) and Coelli et al. (2005).

¹⁰We prefer to use a translog production function, which is a general specification, as the Cobb-Douglas (C-D) and CES are nested in it. Specification tests (not presented in the paper) rejected C-D or CES.

¹¹The idea of using patents as a metric for innovation output to examine R&D productivity dates at least back to Hausman et al. (1984). For a more extensive review of early work of using patent counts consult Hall et al. (2001).

¹²Bibliographic data could be also used as an alternative indicator of innovative output. In this paper, we restrict ourselves only to patents, which are a proxy of industrially applicable innovation, and leave alternative aspects of innovative output for future research.

goods productions, in most of firms, are simultaneous. Therefore, the two production functions, output and knowledge, are jointly estimated.

Therefore, our model consists of the following two production functions:

The stochastic frontier production of output,

$$\begin{aligned} \ln Output_{it} = & \alpha_i + \alpha_1 \ln K_{it} + \alpha_2 \ln N_{it} + \frac{1}{2}\alpha_{11}(\ln K_{it})^2 + \frac{1}{2}\alpha_{22}(\ln N_{it})^2 \\ & + \alpha_{12} \ln K_{it} \ln N_{it} + \gamma_{t,1}t + \delta_{K_t} \ln K_{it} + \delta_{N_t} \ln N_{it} + v_{it,1} - u_{it,1}, \end{aligned} \quad (4)$$

where, *Output* is gross state product (GSP) in state *i*, *K*, state's *i* level of gross capital stock, and *N* number of workers in state *i*.

The stochastic frontier production of knowledge (ideas):

$$\begin{aligned} \ln Patents_{it} = & \beta_i + \beta_1 \ln R\&D_{it} + \beta_2 \ln S_{it} + \frac{1}{2}\beta_{11}(\ln R\&D_{it})^2 + \frac{1}{2}\beta_{22}(\ln S_{it})^2 \\ & + \beta_{12} \ln R\&D_{it} \ln S_{it} + \gamma_{t,2}t + \delta_{R\&D_t} \ln R\&D_{it} + \delta_{S_t} \ln S_{it} + v_{it,2} - u_{it,2}, \end{aligned} \quad (5)$$

where, *Patents* is patents weighted by their citations in state *i*, *R&D* is state's *i* level of business (industry) R&D stock, and *S* number of scientists and engineers in state *i*.

Now, the inputs in the two production functions can be endogenous and bias our estimates. To account for potential endogeneity of inputs in the production of output, we rely on the innovation effort of states and, therefore, express capital stock (*K*) and number of workers (*N*) as functions of state level business R&D stock. The literature has identified various channels through which R&D affects technology and, therefore, the production of output. It has been argued that R&D embodies the factor productivity of labor and capital (Gordon et al., 1987; Hall and Mairesse, 1995; Griliches, 1998; Hercowitz, 1998; Link and Siegel, 2003, 2007) and, thereby, affects the technology parameters β_k and β_l .¹³ Consequently, R&D affects the marginal productivities of the production function inputs, the state of art of technology as well as the production efficiency.¹⁴

¹³There is a large literature on the 'dual' role of R&D spending or 'old knowledge' (captured by the R&D stock). According to innovation-driven growth models (Grossman and Helpman, 1991; Aghion and Howitt, 1998), R&D has two roles or 'faces'. The first role, stimulating innovation, has received the most attention in the existing empirical literature. The second role is in facilitating the imitation of others' discoveries. Rigorous econometric work assessing the statistical significance and quantitative importance of the 'two faces of R&D' has been provided by (among others) Hall and Mairesse (1995), Griliches (1998), Griffith et al. (2004b), Cameron et al. (2005), and Kneller and Stevens (2006) who find evidence that R&D is statistically and economically important in the catch up process as well as in directly stimulating innovation.

¹⁴Physical and human capital could also be functions of other variables (e.g., past output), but the intention is to identify the

Similarly, we instrument for the inputs in the knowledge production by relying on innovation efforts of the academic institutions. The latter amounts to about 13% of total R&D for the US and the period under examination, 1993-2006. Furthermore, up to 70% of the US academic R&D is sponsored by US government agencies (National Science Board, 2012), rendering thus academic R&D an interesting policy related instrument. Most importantly, studies have shown that academic research has significant positive effects in regional industry R&D and economic growth (Jaffe, 1989; Acosta et al., 2009). Further, Adams (1990) links academic knowledge with productivity growth and finds a significant positive contribution and Mansfield (1991), in an influential survey, finds that approximately 10% of industry innovations could not have happened without academic research.¹⁵ Therefore, we instrument for the inputs in the knowledge production by relying on innovation efforts of the academic institutions.

Following the literature, we express the inputs in both production processes as follows:

$$\ln K_{it} = g_1(\text{totalR\&D}_{it-1}, t) + v_{it,3}, \quad \ln N_{it} = g_2(\text{totalR\&D}_{it-1}, t) + v_{it,4} \quad (6)$$

and

$$R\&D_{it} = h_1(\text{academicR\&D}_{it-1}, t) + v_{it,5}, \quad S_{it} = h_2(\text{academicR\&D}_{it-1}, t) + v_{it,6} \quad (7)$$

where, *totalR&D* and *academicR&D* are total and academic R&D stocks in state *i*, respectively, functions *g* and *h* are linear and include state-fixed effects, and $v_{it,k}$ where $k=3, 4, 5,$ and 6 are error terms.

An important question of this paper is whether the productions of output and ideas have been efficient in the US. We model technical efficiency, u_{it} , in both production functions by following Cornwell et al. (1990) (1990, CSS hereafter), who introduced a flexible (quadratic) function of time, with coefficients varying over time and units, without invoking strong distributional assumptions about technical efficiency or random noise. We, therefore, estimate time-varying efficiency levels for individual states as described below:

underlying *exogenous* factors (output is endogenous) to keep our model ergonomic and efficient, avoiding meanwhile endogeneity issues. Such factors could be culture, endowments, type of government, natural resources but are not examined in our study; as we explore different regions within the same country these factors could vary, but the variations would be far more greater in case we studies different countries. The literature has also identify a number of policy related factors (e.g., state tax policies, labor laws, R&D tax credits, non-competition contracts ('non-competes'), which may affect the level of technology via their impact on R&D stock and human capital, see (Mamuneas and Nadiri, 1996; Bloom et al., 2002; Wu, 2005; Palazzi, 2011; Saxenian, 1994; Marx et al., 2009; Belenzon and Schankerman, 2013)). Therefore, we have also instrumented physical capital and human capital as function of state marginal income tax, marginal corporate tax, R&D tax credits, and non-competes, but instrumentation results were weak and volatile. Estimates of regressions have also shown that all these factors affect the R&D (stock).

¹⁵For a recent discussion on the contribution of academic R&D in the economy-wide innovation, see Berman (2012).

$$u_{it,Output} = a_{i0} + a_{i1}t + a_{i2}t^2 \quad (8)$$

where, $u_{it,Output}$ is technical efficiency in the production of output, and

$$u_{it,Knowledge} = b_{i0} + b_{i1}t + b_{i2}t^2 \quad (9)$$

where, $u_{it,Knowledge}$ is technical efficiency in the knowledge production function.

Both equations (8) and (9) are linear in the elements of a_{ij} and b_{ij} ($j= 1, 2, 3$) and allow for state-specific and state effects. State-specific and time-varying estimates of efficiency can be recovered as follows:

$$\hat{u}_{it,Output} = \max_j (\tilde{u}_{jt,Output}) - \tilde{u}_{it,Output} \quad (10)$$

$$\hat{u}_{it,Knowledge} = \max_j (\tilde{u}_{jt,Knowledge}) - \tilde{u}_{it,Knowledge} \quad (11)$$

where, $\tilde{u}_{it,Output}$ and $\tilde{u}_{it,Knowledge}$ represent the fitted values from (8) and (9).¹⁶

The system of equations (4)-(7) is estimated by a Continuously-Updated-Estimation (CUE) version of Generalized Method of Moments (GMM), where the instruments include lagged values of K_{it} , N_{it} , $R\&D_{it}$, S_{it} , $academicR\&D_{it}$, and $industryR\&D$ as well as their squares and cross-products.¹⁷ A more technical presentation of our estimation strategy is presented in the Appendix (section A2).

2.2. Decomposing Output and Knowledge Change

Next, we decompose output (knowledge) change into three components: input change, which represents movements along the frontier; technical change, which reflects shifts of the frontier; and efficiency change, which captures movements towards (or away from) the frontier as states absorb and implement best practice technologies and reduce (or increase) technical inefficiencies.

In doing so, we take logs and totally differentiate the deterministic part of equation (2) with respect to time, which yields a convenient expression of output change:

¹⁶In CSS this is accomplished by considering a production function: $y_{it} = f(z_{it}) + \alpha_{it}$ where, $\alpha_{it} = \theta_{i0} + \theta_{i1}t_{it} + \theta_{i2}t_{it}^2$. Suppose, $\tilde{\alpha}_{it} = \tilde{\theta}_{i0} + \tilde{\theta}_{i1}t_{it} + \tilde{\theta}_{i2}t_{it}^2$ represents the fitted values. Then, technical inefficiency is estimated by $\hat{u}_{it} = \max_j (\tilde{\alpha}_{jt}) - \tilde{\alpha}_{it}$.

¹⁷We have experimented with various versions; lagged values of the inputs of the two productions, in (4) and (5) and present-time values of instruments appear in (6) and (7), which gives rise to a triangular simultaneous equations model, but the results were qualitatively and quantitatively very close to the results reported here.

$$\frac{\dot{Y}_{it}}{Y_{it}} = \frac{\partial \ln f_{it}}{\partial t} - \frac{\partial u_{it}}{\partial t} + \epsilon_{it}^{Capital} \frac{Capital_{it}}{Capital_{it}} + \epsilon_{it}^{Labor} \frac{Labor_{it}}{Labor_{it}} \quad (12)$$

where $\epsilon_{Capital}$ and ϵ_{Labor} denote the partial elasticity of stochastic frontier output with respect to the inputs, capital and labor, respectively and dotted variables refer to time derivatives.

Equation (12) indicates that output (knowledge) growth can be broken down into three components. The first term, $\frac{\partial \ln f_{it}}{\partial t}$, corresponds to technical change (TC), where $\frac{\partial \ln f_{it}}{\partial t} > 0$, represents an upward shift of the frontier (technical progress). The second term corresponds to efficiency change (EC), $-\frac{\partial u_{it}}{\partial t}$, where $-\frac{\partial u_{it}}{\partial t} > 0$ represents a reduction of inefficiency. Finally, the last two terms capture the scale changes (SC), $\epsilon_{it}^{Capital} \frac{Capital_{it}}{Capital_{it}}$ and $\epsilon_{it}^{Labor} \frac{Labor_{it}}{Labor_{it}}$ due to input accumulation in capital ($\beta_{Capital} Capital_{it}$) and labor ($\beta_{Labor} Labor_{it}$), respectively.

Table 1 below summarizes the decomposition of output and knowledge changes.

Table 1: Decomposition of Output (Ideas) Change

$\frac{\dot{Y}_{it}}{Y_{it}}$	=	$TC_{it} + EC_{it} + SC_{it}$
TC_{it}	=	$\frac{\partial \ln f(Capital, Labor, t; \beta)}{\partial t}$
EC_{it}	=	$\frac{\partial v_{it}}{\partial t}$
SC_{it}	=	$(\epsilon_{it}^{Capital} \frac{Capital_{it}}{Capital_{it}} + \epsilon_{it}^{Labor} \frac{Labor_{it}}{Labor_{it}})$

Decomposition of output (knowledge) changes into their various components, allows one to assess the individual contribution of each of the components to the growth process of either output or knowledge.

2.3. Data

Our empirical analysis is based on a sample of 50 US states over the period 1993-2006, the longest period for which data is available for the variables employed in our analysis.¹⁸ Data are retrieved from various sources. Analytical description of the variables and sources can be found in the Appendix (see section A2.2: Variables and Sources).

Table 2 below provides summary statistics of the variables in our analysis.

¹⁸The state of D.C. is excluded from the analysis as it produces almost zero patents.

Table 2: Descriptive Statistics

variables	Observations	Mean	St. Dev.	Min	Max
Output	700	188.3523	223.614	14.71788	1538.082
<i>K</i>	700	196233.5	244716.1	13826.23	1557929
<i>N</i>	700	2657644	2835773	230580	1.73E+07
<i>Patents</i>	700	576.7667	1261.345	0.800	10080.2
<i>R&D</i>	700	14998.27	27046.76	24.18379	210846.8
<i>S</i>	700	11925.95	14426.04	730	99060
<i>academicR&D</i>	700	2549.487	3112.389	95.76921	23453
<i>totalR&D</i>	700	19898.2	33393.62	250.6079	263371.4

Note: Information on all variables is provided at the state level. *Output* is real gross state product (GSP), *K* is gross fixed capital, *N* are thousands of employees (fte), *Patents* is patents filled per state weighted by their citations, *R&D* is industry (business) R&D stock, *S* is number of doctoral scientists and engineers, *academicR&D* is the R&D stock of universities and research institutions, and *totalR&D* stock is total R&D stock in a state. All monetary variables are expressed in million (2000 US) dollars.

Table 2, in combination with Table A.1 in the Appendix, shows that states, on average, produce annually real GSP of 188 million US dollars, with California, New York State, Texas, Illinois and Florida to report the highest GSP in the US, whereas the Dakotas, Wyoming, Vermont and Minnesota the lowest. Further, the average state appears to be quite productive in terms of patents weighted by their citations. However, as the standard deviation shows, the distribution of patent production is skewed with a few states to account for a disproportionate large share. Particularly, intense patenting activity is concentrated mainly in East (New York State, New Jersey) and West (California) Coasts along with some states around the Great Lakes (e.g. Illinois) and in Texas. The lowest patent production is recorded in Alaska, Hawaii, the Dakotas (ND & SD), Wyoming, Montana and Vermont. States have developed, on average, a stock of knowledge in their industry sector of about 15 billion (2000 US) dollars. Accumulated knowledge, proxied by business R&D stock, as well as scientists are also concentrated in the top performer states (including Massachusetts and Pensilvania), while Alaska, Wyoming and the Dakotas score low in both. Finally, academic and total stock of knowledge per state worths of approximately 2.5 and 20 billion (2000 US) dollars. Industry R&D stock accounts for three quarters of total R&D stock, while academic R&D stock accounts for, roughly, 12%.

3. Empirical Results

3.1. How Efficient is the Productions of Output and Ideas?

Table 3 reports estimates of the system of equations (4)-(7), where the technology parameters, i.e., the coefficients for capital, labor and time, and level of efficiency for both production functions, i.e., output and knowledge/ideas, are estimated simultaneously.¹⁹

¹⁹The estimates of the instrumented inputs, in both production processes, are reported in Table A.2 in the Appendix.

Table 3: System Estimates

Frontier Estimates				
	Output		Knowledge	
	coeff.	st.dev.	coeff.	st.dev.
<i>lnCapital</i>	0.425***	0.035	0.658***	0.061
<i>lnLabor</i>	0.577***	0.052	0.424**	0.190
<i>t</i>	0.045***	0.002	0.104***	0.018
$1/2\ln\text{Capital}^2$	0.282**	0.126	0.056	0.048
$1/2\ln\text{Labor}^2$	0.210	0.145	0.120	0.265
<i>lnLabor * lnCapital</i>	-0.242*	0.132	-0.090	0.096
<i>t * lnCapital</i>	-0.002	0.002	-0.006	0.005
<i>t * lnLabor</i>	-0.001	0.002	-0.007	0.009
$1/2t^2$	0.000	0.001	0.002	0.022
<i>Constant</i>	5.203***	0.089	1.886***	0.363
σ	0.300***	0.020	0.315***	0.022
λ	1.182***	0.256	0.021***	0.002

ρ (correlation between output and knowledge functions): 0.751

Efficiency scores

	mean	st.dev.	mean	st.dev.
	0.905	0.024	0.851	0.027

Note: (*/**/***): significance at the 10/5/1 % level; the generic term *Capital* denotes tangible (physical) capital stock (*K*) and R&D stock (*R&D*) in the production of output and knowledge/ideas, respectively; the generic term *Labor* denotes workers (*L*) and scientists (*S*) in the production of output and knowledge/ideas, respectively; λ and σ are efficiency parameters, where λ ($= \sigma_u/\sigma_v$), the ratio of the standard deviation of efficiency over the standard deviation of the noise term, and σ ($= (\sigma_u^2 + \sigma_v^2)^{1/2}$), the composite standard deviation; $BIC=-4.234$; ρ : correlation between production of output and knowledge is computed as $\rho = \sigma_{12}/\sqrt{\sigma_{11} * \sigma_{22}}$ and is derived from the 2x2 covariance matrix for the errors of the output and knowledge functions.

Beginning with the production of output (GSP), our estimation of marginal products of capital and labor of 0.226 and 0.762, respectively, is in line with existing empirical literature (Barro and Sala-i-Martin, 1995; Koop, 2001; Bos et al., 2010a). States produce output at about constant returns to scale, as is often reported in the literature (Barro and Sala-i-Martin, 1995; Mankiw et al., 1992). The inclusion of a time trend, *t*, allows us to measure the technical growth. We find that technical growth is positive and significant at approximately 0.8% per year. However, the latter finding does not necessarily imply that *all* states benefit from 0.8% technical growth, as technical growth is measured at the frontier. We also need to consider the efficiency of the states.

States appear to be very efficient in the production of their output and operate very close to their frontier, as the average efficiency in the states is almost 91% and is not that dispersed. This is also confirmed by the parameters σ and λ , which measure the total variance and the relative magnitude of variance that is attributed to inefficiency, respectively. A positive and significant λ shows that much of this variance indeed

consists of inefficiency.²⁰ In Table A.2 in the Appendix, one can see that there is minimum waste in the production of output for the period of our analysis in the states of Iowa, North Carolina and Delaware, while the states of Hawaii, Alaska and Vermont are the most wasteful states in the use of their production inputs. The top contributor states to the US GDP, California, New York State and Texas, exhibit an average performance in terms of efficiency and rank at 39, 18 and 24 position, respectively. Based on efficiency scores, there is, for example, more waste in producing output for every dollar spent, for instance in California, than in New York State.

Proceeding with the production of ideas, we find the marginal products of R&D stock and scientists to be 0.287 and 0.687, respectively. Although cross-study comparisons are even harder in knowledge production literature, due to different level of aggregation, input definition, and often omission of efficiency of the innovation inputs, we can still recover some findings that can be compared with ours; for instance, the effect of knowledge stock on innovative output, ideas (Porter and Stern, 2000; Ulku, 2004). Our estimates confirm that the stock of existing knowledge (proxied by R&D stock) has a positive effect on innovation. This supports the "standing on shoulders" assumption, to the extent that the accumulation of past knowledge increases the creation of new knowledge. Compared to the production of output, the marginal productivity of a unit of knowledge stock in the production of ideas is higher than that of physical capital stock in the production of output, while the opposite holds for the marginal productivity of labor. The marginal rate of technical substitution (MRTS) in the knowledge (output) production is 0.418 (0.297), demonstrating the rate at which labor can be substituted for capital, while holding output constant, is lower for the production of output. Further, the production of knowledge in the states exhibits about constant returns.

Technological knowledge expands, on average, every year by approximately 0.6% as the estimate of technical change indicates. States are not fully efficient in the production of output or knowledge, but mean efficiency in both productions is pretty high. To examine the presence of inefficiency in our data, we decompose the composite residual of equation (2), $\exp\{\varepsilon_{it}\} = \exp\{v_{it}\} \exp\{-u_{it}\}$, and identify its components, $\exp\{v_{it}\}$ and $\exp\{-u_{it}\}$, by re-parameterizing λ in the maximum likelihood procedure, where λ ($= \sigma_u/\sigma_v$), the ratio of the standard deviation of efficiency over the standard deviation of the noise term, and σ ($= (\sigma_u^2 + \sigma_v^2)^{1/2}$) is the composite standard deviation. The frontier is identified by the λ for which the log-likelihood is maximized (Kumbhakar2000). Consequently, the parameters σ and λ measure the total variance and the relative magnitude of variance that is attributed to inefficiency, respectively. A positive

²⁰This result reflects an important advantage of our stochastic frontier approach to a comparable data envelopment analysis. If λ - the efficiency to noise ratio - is statistically insignificant, then almost all of the unexplained variance is indeed noise, while a statistically significant λ is an indication of inefficiency rather than noise. This is clearly something that a DEA model would fail to capture.

and significant λ shows that much of this variance indeed consists of inefficiency. This result reflects an important advantage of our stochastic frontier approach to a comparable data envelopment analysis. If λ - the efficiency to noise ratio - is statistically insignificant, then almost all of the unexplained variance is indeed noise, while a statistically significant λ is an indication of inefficiency rather than noise.²¹

Consequently states are about 85% efficient in producing new knowledge and 91% in producing output; therefore, there is more waste in the production of ideas. This should not come as a surprise. Although innovation is one of the most crucial elements in economic growth is by nature uncertain and risky, which relies upon scarcely available resources. Also, best practices may derive from the learning associated with research failures and successes, or arise with the entry into the market of high-tech start-up. We further find that there is no large dispersion across states' efficiencies as the small standard deviation of knowledge efficiency indicates. Among the most efficient states in the production of new knowledge, as Table A.2 in the Appendix shows, are the states of Maine, Oklahoma, Hawaii and Mississippi. These states are at the bottom of the ranking in terms of business R&D and patent production; however, they appear to manage their innovation inputs, i.e., R&D stock and scientists, in a rather efficient manner, so they are able to produce more patents for every dollar spent. The study of Thomas et al. (2011) also finds Mississippi and Oklahoma (among other twelve) states to be the ones that show improvements in innovation efficiency during 2004-2008. As they authors argue, certain high tech firms located in these states (e.g., IBM, Intel, Honeywell, AT&T, Lexmark, Halliburton, Microsoft, Boeing and Micron Technologies) have experienced moderate to sharp increases in patent granted during 2004-2008. Among the least efficient states are Colorado, New Jersey, Ohio and Massachusetts. As before, California together with New York State, Illinois and Texas, which spend the largest sums in R&D and produce the largest number of patents, perform fair in terms of knowledge efficiency. In particular, California ranks at 26, New York State at 44, Illinois at 31 and Texas at 10 place among the 50 states in terms of efficiency in producing new technological knowledge.²²

As to *why* these efficiency differentials exist and, in particularly, why wealthier states, on average, appear to be less efficient compared to the rest of the states, a possible explanation could be that wealthier state can 'afford' to go for riskier innovation projects that not all of them eventually turn into commercialised ideas or attracts more research sources (e.g., researchers and R&D) in their production of innovation process that their quality might considerably range compared to a less wealthier state. A number

²¹This is clearly something that an alternative frontier methodology, for instance Data Envelop Analysis (DEA) would fail to capture. Therefore, the values of λ and σ in particular confirm the presence of inefficiency in our data.

²²As an exercise, we calculate efficiency scores following the approach of Thomas et al. (2011). That is, we calculate the ratio of the average number of patents to industry R&D stock (or spending) per period for each state. Efficiency scores (mildly) vary, but the main conclusion holds: large and wealthy states do not score at the top of the efficiency score ranking. For example, California ranks 16th and Texas 17th. With our SFA approach these two big states rank 26th and 10th, respectively.

of mechanisms could be responsible for that. State-specific policies such as personal, corporate and R&D taxes, labor and market regulations and institutions could make the production of innovation in a state more (or less) efficient by attracting specific capital (human or physical) and affecting their productivities. Delving deeper into regional micro data, Fritsch and Slavtchevc (2011) and Broekel (2015), in studying the innovation efficiency in Germany, have examined a number of regional factors such as knowledge networkings and collaboration among regional organizations. For example, Fritsch and Slavtchevc (2011) consider a number of factors, namely the density and industrial composition of regional actors, region's accessibility, the technological, industrial and institutional infrastructure (networks) of a region that may shape the innovation efficiency differentials across regions. The study concludes that spillovers from private as well as from universities and other public institutions have a positive effect on the innovation efficiency, and regions that are dominated by large establishments tend to be less efficient compared to regions with a lower average establishment size. Further, Broekel (2015) uses a rich panel dataset, covering 270 German labour market regions and four industries to examine whether subsidies for R&D cooperation are a suitable policy measure for stimulating the innovation efficiency of regions. The study shows that cooperation among regional firms and subsidized links to non-regional public research institutes benefit regions with low innovation capacities while subsidization of cooperation with non-regional universities is more important for regions with large innovation capacities.

Therefore, these factors, if coupled with certain institutional settings, may influence conditions under which research inputs are managed. As a result, very different innovation outcomes may be achieved even when a similar amount of resources is devoted to R&D. Consequently, identifying the 'right' mechanism requires thorough and comprehensive analysis with detailed regional data. Due to lack of state level information on some regional factors, such investigation is left for future research.

Thus far, our analysis was based on patent counts by assignees and each patent was assigned to the first assignee in case of multiple assignees. For robustness purposes, we re-estimate our system considering patent counts by inventor, controlling for patents with multiple inventors. We have also removed outliers (i.e., states with exceptional high and low patent production or gross state product) and applied various weighting schemes (e.g. weighting by the size of population in a state). Results, overall, have remained robust and do not alter in any significant way.²³

To visualize our findings, Figure 2 below plots mean efficiency in the production of output versus mean efficiency in the production of knowledge.

necessarily the most efficient ones in terms of knowledge production. In fact, according to the global innovation index (GII), the US ranks at the 6th position in terms of quality of universities, availability of micro-finance and venture capital, to gauge innovation capabilities and measurable results, but it is not rank among the top tier countries in terms of innovation efficiency. Instead, middle-(or even low-) income countries with low - compared to the US - R&D spending show a higher efficiency score compared to the US.²⁴ Apparently, the limited available resources to the knowledge production, for instance, were used with great efficiency in these countries.

3.2. What are the Components of Output and Ideas Growth?

We now can examine how states grow in terms output and ideas. In doing so, we decompose output and knowledge growth into three components, input change (*SC*), technical change (*TC*) and efficiency change (*EC*), as shown in Table 4, below.

Table 4: Components of Output and Knowledge Growth

	total change $\frac{\dot{Y}}{Y}$	input change <i>SC</i>	technical change <i>TC</i>	efficiency change <i>EC</i>	Obs
<i>Output</i>	0.032 (0.029)	0.025 (0.018)	0.008 (0.002)	-0.001 (0.025)	600
<i>Knowledge</i>	0.028 (0.306)	0.024 (0.282)	0.006 (0.004)	-0.002 (0.030)	600

During our sample period, gross state product grew by 3.2%; a rate which is consistent to the figures reported from the Bureau of Economic Analysis about the growth in the US during 90s and mid 2000. This growth was mainly attributed to the accumulation of inputs, 2.5%, and to technical change, 0.8%. Efficiency, in the use of production inputs, remained pretty much unaltered, exhibiting a small decline of 0.1%.

Similar picture emerges for the production of new ideas, as patents weighted by their citations, on average, grew by 2.8% and this growth was primarily driven from the growth in the innovation inputs (R&D and scientists), which was 2.4% and to a lesser extent from the technical change, which contributed by 0.6%. During the same period, there was no improvement in the innovation efficiency, which actually decreased by 0.2%.

²⁴The GII measures 142 countries, using 84 indicators, which include the quality of universities, availability of micro-finance and venture capital, to gauge innovation capabilities and measurable results. Read more at <http://knowledge.insead.edu/entrepreneurship-innovation/global-innovation-index-2930#5CwzK0lz9oqX6jy8.99>. About countries' rankings, see https://www.globalinnovationindex.org/userfiles/file/blog/GII_DB_Bali_03.pdf

The important role of input growth in the output and innovation growth process is also consistent with findings in the related literature (Koop et al., 1999; Koop, 2001; Kumar and Russell, 2002).

At this point, it is interesting to consider how changes in the inputs of both production processes relate to changes in the instruments. More specifically, physical capital and workers are related to the changes in total R&D and similarly the changes in business R&D stock and scientists to the changes in academic R&D. Table 5 below, reports these correlations.

Table 5: Correlations

	$\text{corr}(\text{growth}K, \text{growth totalR\&D})$	$\text{corr}(\text{growth}N, \text{growth totalR\&D})$
<i>Output</i>	0.220***	0.120***
	$\text{corr}(\text{growthR\&D}, \text{growth academicR\&D})$	$\text{corr}(\text{growth}S, \text{growth academicR\&D})$
<i>Knowledge</i>	0.040	-0.030

Significance at the 10/5/1 % level (*/**/***).

It is possible that physical capital (R&D stock) and worker (scientist) accumulation may reflect the use of total (academic) R&D in both functions. In general, we expect the correlations to be positive. However, a negative association could also be the case. For example, capital with new vintages may be replacing outdated vintages, and less but more skilled labor may be replacing more but less skilled labor. For the case of the production of output, inputs' growth is positively and statistically related with total R&D growth. As for the knowledge production, although the growth of business R&D stock positively relates with academic R&D, the growth of scientists is negatively associated with the growth of academic R&D. In the latter case, this suggests that the remaining scientific force is more skilful, as states may use more academic R&D per unit of scientist in order to increase the production of patents.

4. Conclusion

This paper examines efficiency and growth patterns in the production of output and ideas in the US. Employing frontier techniques we jointly estimate the production of output and the production of ideas controlling meanwhile for the endogeneity of the inputs in both production functions. With our modeling approach we are able to unfold the roots of output and ideas growth and discuss important questions in the literature, namely whether there is waste in the production of output and ideas and efficient in production of output states are also efficient in producing ideas.

Our results show that the states of the US are particularly efficient in the use of their inputs in producing output and less efficient in the use of their innovation resources for producing new knowledge. States appear to vary less in terms of output efficiency and more in knowledge efficiency, as the latter shows higher dispersion. The level of mean efficiency in either output or ideas production, however, remains pretty much unaltered throughout our sample period with an even small tendency to decline over time. Our findings do not support the common perception that richer (most innovative) states are more efficient than less richer (less innovative) states. A policy implication of our findings is that government efforts should not only be directed toward states that produce high R&D, but also to states with modest R&D that are able to produce more output or patents for every dollar spent. Furthermore, an efficient in producing knowledge state is not necessarily efficient in producing goods and services. Finally, a consistent finding is that input accumulation is an important source of growth both for output and for the knowledge.

Future research could provide more guidance to regional factors, economic and institutional among other, responsible for (in)efficiency differentials across regions, both in output and knowledge production.

References

- Abdih, Y., Joutz, F., 2006. Relating the knowledge production function to total factor productivity. *International Monetary Fund Staff Papers* 53(2), 242–271.
- Acosta, M., Coronado, D., Leon, M. D., Martinez, M., 2009. Production of university technological knowledge in european regions: Evidence from patent data. *Regional Studies* 43(9), 1167–1181.
- Adams, J., 1990. Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy* 98(4), 673–702.
- Aghion, P., Howitt, P., 1998. *Endogenous Growth Theory*. MIT Press, Cambridge, MA.
- Aigner, D. J., Lovell, K. C., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6(1), 21–37.
- Altvater-Mackensen, N., Balicki, G., Bestakowa, L., Bocatus, B., Braun, J., Brehmer, L., Verena Brune, K. E., Erdem, F., Fritscher, R., Jacobs, A., Klingsporn, B., Konsinski, M., Kuntze, J., Lee, J.-R., Osterhage, A., Probst, M., Risch, T., Schmitt, T., Stock, W. G., Sturm, A., Weller, K., Werner, K., 2005. Science and technology in the region: The output of regional science and technology, its strengths and its leading institutions. *Scientometrics* 63(3), 463–529.
- Asheim, B., Gertler, M., 2005. The geography of innovation: Regional innovation systems. In: Fagerberg J., M. D. C., (Eds), N. R. R. (Eds.), *Handbook of Innovation*. Oxford University Press, Oxford, pp. 291–317.
- Bai, J., 2013. On regional innovation efficiency: Evidence from panel data of china's different provinces. *Regional Studies* 47(5), 773–788.
- Barra, C., Zotti, R., 2015. Regional innovation system (in)efficiency and its determinants: an empirical evidence from italian regions. Munich Personal RePEc Archive (MPRA). Available at: <https://mpra.ub.uni-muenchen.de/67067/>.
- Barro, R. J., Sala-i-Martin, X. X., 1995. *Economic Growth*. McGraw-Hill, New York.
- Battese, G. E., Corra, G. S., 1977. Estimation of a production frontier model, with application to the pastoral zone of eastern australia. *Australian Journal of Agricultural Economics* 21(3), 169–179.
- Belenzon, S., Schankerman, M., 2013. Spreading the word: Geography, policy and knowledge spillovers. *Review of Economics and Statistics* 95(3), 884–903.
- Berman, E. P., 2012. *Creating the market university: How academic science became an economic engine*. Princeton University Press.
- Bloom, N., Griffith, R., Reenen, J. V., 2002. Do r&d tax credits work? evidence from a panel of oecd countries 1979-1997. *Journal of Public Economics* 85(1), 1–31.
- Bos, J., Economidou, C., Koetter, M., 2010a. Technology clubs, r&d and growth patterns: Evidence from eu manufacturing. *European Economic Review* 54(1), 60–79.
- Bos, J., Economidou, C., Koetter, M., Kolari, J., 2010b. Do all countries grow alike? *Journal of Development Economics* 91(1), 113–127.
- Bottazzi, L., Peri, G., 2003. Innovation and spillovers in regions: Evidence from european patent data. *European Economic Review* 47(4), 687–717.
- Bottazzi, L., Peri, G., 2007. The international dynamics of r&d and innovation in the long run and in the short run. *Economic Journal* 117(518), 486–511.
- Brenner, T., Broekel, T., 2011. Methodological issues in measuring innovation performance of spatial units. *Industry and Innovation* 18(1) (1), 7–37.
- Broekel, T., 2012. Collaboration intensity and regional innovation efficiency in germany - a conditional efficiency approach. *Industry and Innovation* 19(2), 155–179.
- Broekel, T., 2015. Do cooperative research and development (r&d) subsidies stimulate regional innovation efficiency? evidence from germany. *Regional Studies* 49(7), 1087–1110.
- Broekel, T., Rogge, N., Brenner, T., 2013. The innovation efficiency of german regions - a shared-input dea approach. Working Papers on Innovation and Space No 2013-08, Philipps University Marburg, Department of Geography. Available at: <ftp://137.248.191.199/RePEc/pum/wpaper/wp0813.pdf>.
- Broekel, T., Schlump, C., 2009. The importance of r&d subsidies and technological infrastructure for regional innovation performance - a conditional efficiency approach. *Papers in Evolutionary Economic Geography*. Available at: <http://econ.geo.uu.nl/peeg/peeg0921.pdf>.
- Cameron, G., Proudman, J., Redding, S., 2005. Technological convergence, R&D, trade and productivity growth. *European Economic Review* 49(3), 775–807.
- Chen, K., Guan, J., 2012. Measuring the efficiency of china's regional innovation systems: Application of network data envelopment analysis (dea). *Regional Studies* 46 (3), 355–377.
- Chen, Y., Zhiping, Y., Shu, F., Zhengyin, H., Meyer, M., Bhattacharya, S., 2009. A patent based evaluation of technological innovation capability in eight economic regions in pr china. *World Patent Information* 31(2), 104–110.
- Coe, D., Helpman, E., 1995. International R&D spillovers. *European Economic Review* 39(5), 859–887.
- Coelli, T., Rao, D., Donnell, C. O., Battese, G., 2005. *An Introduction to Efficiency Analysis*, 2nd Edition. Springer, New York.
- Cooke, 2001. Regional innovation systems, clusters, and the knowledge economy. *Industrial and Corporate Change* 10(4) (4), 945–974.
- Cooke, P., 2004. The role of research in regional innovation systems: new models meeting knowledge economy demands. *International Journal of Technology Management* 28(3/4/5/6), 507–533.
- Cornwell, C., Schmidt, P., Sickles, R. C., 1990. Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics* 46(1-2), 185–200.
- Cruz-Cázares, C., Bayona-Sáezb, C., García-Marco, T., 2013. You can't manage right what you can't measure well: Technological innovation efficiency. *Research Policy* 42(6-7), 1239–1250.
- Cullmann, A., Schmidt-Ehmcke, J., Zloczyski, P., 2012. Innovation, r&d efficiency and the impact of the regulatory environment: A two-stage semi-parametric dea approach. *Oxford Economic Papers* 64(1), 176–196.

- Färe, R., Grosskopf, S., Norris, M., Zhang, Z., 1994. Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review* 84(1), 66–83.
- Fè-Rodríguez, E., Hofler, R., 2013. Count data stochastic frontier models, with an application to the patents-R&D relationship. *Journal of Productivity Analysis* 39(3), 271–284.
- Feldman, M., 1994. *The Geography of Innovation*. Boston: Kluwer Academic Publishers.
- Fosfuri, A., Helmers, C., Roux, C., 2012. Are joint patents collusive? evidence from the us and europe. TILEC Discussion Paper No. 2012035, University of St. Gallen Law & Economics Working Paper No. 201315. Available at <http://dx.doi.org/10.2139/ssrn.2164222>.
- Franco, C., Leoncini, R., 2013. Measuring china's innovative capacity: a stochastic frontier exercise. *Economics of Innovation and New Technology* 22(2), 199–217.
- Franco, C., Pieri, F., Venturini, F., 2016. Product market regulation and innovation efficiency. *Journal of Productivity Analysis* 45(3), 299–315.
- Fritsch, M., 2002. Measuring the quality of regional innovation systems - a knowledge production function approach. *International Regional Science Review* 25(1), 86–101.
- Fritsch, M., Slavtchev, V., 2010. How does industry specialization affect the efficiency of regional innovation systems? *The Annals of Regional Science* 45(1), 87–108.
- Fritsch, M., Slavtchev, V., 2011. Determinants of the efficiency of regional innovation systems. *Regional Studies* 45(7), 905–918.
- Fu, X., Yang, Q., 2009. Exploring the cross-country gap in patenting: A stochastic frontier approach. *Research Policy* 38(7), 1203–1213.
- Gordon, R. H., Schankerman, M. A., Spady, R. H., 1987. Estimating the effects of r&d on bell system productivity: a model of embodied technical change. NBER Working Paper Series No. 1607.
- Griffith, R., Redding, S., van Reenen, J., 2004a. Mapping the two faces of r&d: Productivity growth in a panel of oecd industries. *Review of Economics and Statistics* 86(4), 883–895.
- Griffith, R., Redding, S., van Reenen, J., 2004b. Mapping the two faces of r&d: Productivity growth in a panel of oecd industries. *Review of Economics and Statistics* 86(4), 883–895.
- Griliches, Z., 1979. Issues in assessing the contribution of r&d to productivity growth. *Bell Journal of Economics* 10(1), 92–116.
- Griliches, Z., 1990. Patent statistics as economic indicators: A survey. *Journal of Economic Literature* 28(4), 1661–1707.
- Griliches, Z., 1992. The search for r&d spillovers. *Scandinavian Journal of Economics* 94(0), S29–47, Supplement.
- Griliches, Z., 1998. *R&D and Productivity: The Econometric Evidence*. Chicago: University of Chicago Press, Chicago, USA.
- Grossman, G., Helpman, E., 1991. *Innovation and Growth in the Global Economy*. MIT Press, Massachusetts, USA.
- Grossman, G. M., Helpman, E., 1994. *Technology and Trade*. NBER Working Papers 4926, National Bureau of Economic Research, Inc.
- Guellec, D., van Pottelsberghe de la Potterie, B., 2004. From r&d to productivity growth: Do the institutional settings and the source of funds of r&d matter? *Oxford Bulletin of Economics and Statistics* 66(3), 353–378.
- Hall, B., Jaffe, A., Trajtenberg, M., 2001. The nber patent citation data file: Lessons, insights and methodological tools. NBER Working Paper No. 8498.
- Hall, B., Mairesse, J., 1995. Exploring the relationship between r&d and productivity in french manufacturing firms. *Journal of Econometrics* 65(1), 263–293.
- Hall, B. H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. *RAND Journal of Economics* 36(1), 16–38.
- Hausman, J. A., Hall, B., Griliches, Z., 1984. Econometric models for count data with an application to the patents-r&d relationship. NBER Working Paper No. 0017.
- Henderson, D. J., Zelenyuk, V., 2007. Testing for (efficiency) catching-up. *Southern Economic Journal* 73(4), 1003–1019.
- Hercowitz, Z., 1998. The 'embodiment' controversy: A review essay. *Journal of Monetary Economics* 41(1), 217–224.
- Jaffe, A., 1989. Real effects of academic research. *American Economic Review* 79(5), 957–970.
- Jones, C. I., 1995. Time series test of endogenous growth models. *Quarterly Journal of Economics* 110(2), 495–525.
- Kneller, R., Stevens, P., 2006. Frontier technology and absorptive capacity: Evidence from oecd manufacturing industries. *Oxford Bulletin of Economics and Statistics* 68(1), 1–21.
- Koop, G., 2001. Cross-sectoral patterns of efficiency and technical change in manufacturing. *International Economic Review* 42(1), 73–103.
- Koop, G., Osiewalski, J., Steel, M. F., 1999. The components of output growth: A stochastic frontier analysis. *Oxford Bulletin of Economics and Statistics* 61(4), 455–487.
- Koop, G., Osiewalski, J., Steel, M. F., 2000. Modeling the sources of output growth in a panel of countries. *Journal of Business and Economic Statistics* 18(3), 284–299.
- Kortum, S., 1993. Equilibrium r&d and the patent-r&d ratio: U.s. evidence. *American Economic Review* 83(2), 450–457.
- Kumar, S., Russell, R. R., 2002. Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *American Economic Review* 92(3), 527–548.
- Kumbhakar, S. C., Lovell, K. C., 2000. *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.
- Lai, R., D'Amour, A., Yu, A., Sun, Y., Fleming, L., 2011. Disambiguation and co-authorship networks of the u.s. patent inventor database (1975–2010). Harvard Business School, The Institute for Quantitative Social Science.
URL <http://hdl.handle.net/1902.1/15705> UNF:5:RqsI3LsQEYLHkkg5jG/jRg== The Harvard Dataverse Network [Distributor] v5 [Version]
- Li, X., 2009. China's regional innovation capacity in transition: An empirical approach. *Research Policy* 38 (2), 338–357.
- Limam, Y. R., Miller, S. M., 2004. Explaining economic growth: Factor accumulation, total factor productivity growth, and production efficiency improvement. *Quarterly Review of Economics and Finance* 44 (2), 296–320.
- Link, A. N., Siegel, D. S., 2003. *Technological Change and Economic Performance*. Routledge, New York, USA.
- Link, A. N., Siegel, D. S., 2007. *Innovation, Entrepreneurship, and Technological Change*. Oxford University Press, Oxford, United Kingdom.

- Mamuneas, T., Nadiri, M., 1996. Public r&d policies and cost of behavior of the us manufacturing industries. *Journal of Public Economics* 63(1), 57–81.
- Mancusi, M. L., 2008. International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations. *Journal of International Economics* 76(2), 155–165.
- Mankiw, G. N., Romer, D., Weil, D. N., 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107(2), 407–437.
- Mansfield, E., 1991. Academic research and industrial innovation. *Research Policy* 20(1) (1), 1–12.
- Marx, M., Strumsky, D., Fleming, L., 2009. Mobility, skills, and the michigan non-compete experiment. *Management Science* 55(6), 875–889.
- Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from cobb-douglas production functions with composed error. *International Economic Review* 18(2), 435–444.
- National Science Board, 2012. National Science Board: Science and Engineering Indicators. National Science Board.
- Palazzi, P., 2011. Taxation and innovation. OECD Taxation Working Papers, No. 9, OECD, Paris. [Http://dx.doi.org/10.1787/5kg3h0sf1336-en](http://dx.doi.org/10.1787/5kg3h0sf1336-en).
- Porter, M. E., Stern, S., 2000. Measuring the "ideas" production function: Evidence from international patent output. Working Paper Series No. 7891, NBER, Cambridge, MA.
- Romer, P. M., 1990. Endogenous technological change. *Journal of Political Economy* 98(5), 71–102.
- Rousseau, S., Rousseau, R., 1997. Data envelopment analysis as a tool for constructing scientometric indicators. *Scientometrics* 40(1), 45–56.
- Rousseau, S., Rousseau, R., 1998. The scientific wealth of european nations: Taking effectiveness into account. *Scientometrics* 42(1), 75–87.
- Saxenian, A., 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge Mass.
- Thomas, V., Sharma, S., Jain, S., 2011. Using patents and publications to assess r&d efficiency in the states of the usa. *World Patent Information* 33(1), 4–10.
- Trajtenberg, M., 2002. Government support for commercial r&d: Lessons from the israeli experience. In: Jaffe, A. B., Lerner, J., Stern, S. (Eds.), *Innovation Policy and the Economy*, Volume 2. NBER Books, National Bureau of Economic Research, Inc., Ch. Chapter 3, pp. 79–134.
- Ulku, H., 2004. R&d, innovation, and economic growth: An empirical analysis. IMF Working Paper, WP/04/185.
- USPTO, 2001. Manual of Patent Examining Procedure (MPEP), Eighth Edition, August 2001. United States Patent and Trademark Office.
- Wang, E. C., 2007. R&d efficiency and economic performance: A cross-country analysis using the stochastic frontier approach. *Journal of Policy Modeling* 29(2), 345–360.
- Wang, E. C., Huang, W., 2007. Relative efficiency of r&d activities: A cross-country study accounting for environmental factors in the dea approach. *Research Policy* 36(2), 260–273.
- Whalley, A., Hicks, J., 2014. Spending wisely? how resources affect knowledge production in universities. *Economic Inquiry* 52(1), 35–55.
- Wu, Y., 2005. The effects of state r&d tax credits in stimulating private r&d expenditure: A cross-state empirical analysis. *Journal of Policy Analysis and Management* 24(4), 785–802.
- Yamarik, S., 2013. State-level capital and investment: Updates and implications. *Contemporary Economic Policy* 31(1), 62–72.
- Zachariadis, M., 2003. R&d, innovation, and technological progress: a test of the schumpeterian framework without scale effects. *Canadian Journal of Economics* 36(3), 566–586.

APPENDIX

A.1. Brief Literature Review

The conventional empirical testing of growth models has examined output (or productivity) growth assuming no waste (full efficiency) in the use of resources.²⁵ With the development of frontier methodologies, a growing body of empirical literature has conducted efficiency analyses using a variety of modeling approaches. Most of the studies, however, focused on assessing the (in)efficiency of the production process and factors that potentially affect it. So far, the attention has largely been at decomposing aggregate (country-level) output. For instance, Färe et al. (1994) use data envelopment analysis (DEA), while Koop et al. (1999, 2000), Limam and Miller (2004) and Bos et al. (2010b) use stochastic frontier analysis (SFA) to examine country-specific inefficiency in a number of developed and developing countries and factors that shape it. Recently, a number of studies have investigated the role of efficiency in explaining growth differentials for a panel of manufacturing industries in the OECD countries. The study of Koop (2001) explores the forces of output growth in six manufacturing industries during the 1970s and 1980s, while Kneller and Stevens (2006) investigate the sources of inefficiency in nine industries over the same period. In similar vein, Bos et al. (2010a) unfold the drivers of output change in a panel of EU manufacturing industries allowing for inefficient use of resources and differences in the production technology across industries and discuss output growth, technology spillovers and catch-up issues across industries and countries.

While it is well documented that a large portion of productivity differentials can be ascribed to how countries (or industries) manage production inputs, and which factors are behind international efficiency gaps, it remains largely unexplored how countries (or industries) efficiently perform innovation activity. As acknowledged by R&D-based endogenous growth theories, innovation is one of the most crucial elements in fostering economic growth. Innovation, however, is by nature an uncertain and risky activity, which relies upon scarcely available resources. Moreover, best practices may derive from the learning associated with research failures and successes, or arise with the entry into the market of high-tech start-up. These factors, if coupled with certain institutional settings, may influence conditions under which research inputs are managed. As a result, very different innovation outcomes may be achieved even when a similar amount of resources is devoted to R&D. Consequently, if research is performed inefficiently, policies aimed at fostering investment in an area are unlikely to achieve the expected outcome of encouraging economic growth over the long run.

Thus far, the standard approach in the economics of innovation literature has been the use of a knowledge (innovation) production function where the innovative output, the counts of patents, is produced analogously to the production of real output, employing existed knowledge and human capital allowing no waste in their use (Griliches, 1979, 1992; Jaffe, 1989; Feldman, 1994).

A rising body of studies has examined at the efficiency performance in research and patenting. The studies of Rousseau and Rousseau (1997, 1998) and Franco et al. (2016) are among the few attempts, which consider efficiency in innovation process. These studies evaluate innovation performance of a number of, mainly OECD, countries and rank them according to their efficiency scores.²⁶ A number of subsequent studies, perform similar analysis controlling also for socio-economic factors, institutions, and regulations. Among them, Wang (2007) examines technical efficiency in research across 23 OECD and seven non-OECD countries finding large disparities in R&D efficiency scores. The latter, are largely explained by cross-country differentials in technology endowments and open-market institutional settings. In a similar vein, Fu and Yang (2009) find that countries may rank differently in terms of innovation capacity (patenting frontier) and innovation (in)efficiency. The former is positively correlated with the share of high-tech industries while innovation efficiency is affected positively by the degree of economic development, the share of business research and the strength of intellectual property protection. By means of a two-stage non-parametric DEA approach, Cullmann et al. (2012) study R&D efficiency differences among OECD

²⁵See, for instance, Jones (1995), Coe and Helpman (1995), Aghion and Howitt (1998), Griffith et al. (2004a), Zachariadis (2003), Bottazzi and Peri (2003), Bottazzi and Peri (2007), and Mancusi (2008) among others.

²⁶See Cruz-Cázares et al. (2013) for an updated review on innovation efficiency studies.

countries focusing on the role of regulatory environment. A more recent study that of Franco et al. (2016) examines cross-industry (cross-country) efficiency levels in knowledge production focusing on the role played by the institutional framework regulating the functioning of upstream markets, i.e. how regulation in upstream sectors (services) influences innovation efficiency of downstream (manufacturing) firms that purchase factor inputs from the former. The key idea under exploration is that if intermediate input markets are imperfect, due to administrative barriers, licensing, among others, innovating firms are likely to reach sub-optimal efficiency levels, falling behind the innovation frontier. To test this hypothesis, the authors carry out a stochastic frontier analysis (SFA) on innovation data for fifteen manufacturing industries of ten OECD economies between 1990 and 2002 which, as widely documented, was period of intense market reforms and widespread pro-competition initiatives. Their results show that upstream regulation does not influence directly innovation efficiency; rather it has second-order effects, i.e., by interacting with the functioning of the technology, labour and the financial market. Regulation in the product market is associated with lower levels of efficiency in contexts where intellectual property protection is strong, whilst seems beneficial in the presence of less regulated financial markets and stringent employment protection.

Global economic forces have raised the profile of regions and the prominence of regional and local business clusters as vehicles for global and national economic competitiveness. Regions, especially when they have developed clusters and appropriate administrative machinery for supporting innovative enterprise, represent more meaningful communities of economic interest, define genuine flows of economic activities and can take advantage of true linkages and synergies among economic actors. A newly developed strand of literature has shifted its attention to regional innovation systems.²⁷ As production becomes more science-based, developed research infrastructure, a highly qualified workforce and an innovative culture are becoming more important than natural resources, which means that a supportive environment for innovative companies can deliberately be created. Locations differ with regard to the quality or the efficiency of regional innovation systems (RIS), leading to different levels of innovative output even if the inputs are identical in quantitative as well as in qualitative terms. However, the available empirical evidence for regional innovation efficiency is rather thin. We still know very little about the conditions that are conducive or unfavourable for innovation activity and how policy could help improve the functioning of regional innovation systems.

A handful of regional studies though managed to shed some light. Rich and detailed data on German regions have allowed the study of regional innovation performance and the evaluation of various policy measures.²⁸ The studies of Broekel and Schlump (2009), Brenner and Broekel (2011), Fritsch and Slavtchev (2011), Broekel (2012) and Broekel et al. (2013), in particular, have studied the efficiency of regional (and industry) specific innovativeness and the role of regional factors, i.e., knowledge networkings and collaboration among regional organizations using (mainly non-parametric) frontier methodologies in measuring innovation efficiency. For example, Fritsch and Slavtchev (2011) estimate a knowledge production function to assess innovation efficiency for 97 planning German regions, documenting considerable differences in technical efficiency across regions, which are specially divided into different regimes with divergent level of performance. Overall, regions with similar efficiency scores are found to cluster together. The top performer regions are found in the south, west and central Germany (particularly in large and densely populated areas), while the laggards in innovation efficiency in the north, east and at the borders. Additionally, the authors consider a number of factors, namely the density and industrial composition of regional actors, region's accessibility, the technological, industrial and institutional infrastructure (networks) of a region that may shape the innovation efficiency differentials across regions. The study concludes that spillovers from private as well as from universities and other public institutions have a positive effect on the innovation efficiency. Furthermore, regions that are dominated by large establishments tend to be less efficient

²⁷For a systematic account of the idea and content of regional innovation systems, see Cooke (2001), Cooke (2004) and Asheim and Gertler (2005) among others

²⁸The works of Michael Fritsch, Tom Broekel, Michael Binder and their co-authors, among others, provide a rich line of research on regional of dynamic aspects of German regional innovativeness. The study of Fritsch (2002) expands to European regional level analysis to assess the quality of regional innovation systems. Building on data for eleven European regions, estimates, a knowledge production function employing conventional, non-frontier techniques.

compared to regions with a lower average establishment size.²⁹ A more recent study of Broekel (2015) builds on a rich panel dataset, covering 270 German labour market regions and four industries, examines whether subsidies for R&D cooperation are a suitable policy measure for stimulating the innovation efficiency of regions. Results show that cooperation among regional firms and subsidized links to non-regional public research institutes benefit regions with low innovation capacities while subsidization of cooperation with non-regional universities is more important for regions with large innovation capacities.³⁰

To our knowledge, empirical evidence, based at aggregated or disaggregated level analysis of either production or innovation efficiency in the US, has been extremely thin. Only a handful of studies that have performed comparative cross-country studies of either production or knowledge efficiency and include the US (Koop et al., 1999; Henderson and Zelenyuk, 2007; Wang and Huang, 2007; Cullmann et al., 2012). The study of Thomas et al. (2011) is among the very few attempts that examines innovation efficiency of the US at the state level for the period 2004-2008. The study, however, measures innovation efficiency based on the ratio of R&D outputs (e.g. patents granted or scientific publications) to R&D inputs (e.g. R&D expenditure), concluding that between 2004 and 2008 there has been a marked decrease in R&D efficiency in 37 of the 51 geographic regions of the US and only 14 (i.e., Alaska, Arizona, Georgia, Idaho, Illinois, Kentucky, Michigan, Mississippi, New York, Ohio, Oklahoma, Vermont, Washington and West Virginia) states shown modest improvements in innovation efficiency. A closer to ours study is that of Fè-Rodríguez and Hofler (2013), which proposes a stochastic frontier count model and performs a cross-section analysis to study innovation efficiency in a number of pharmaceutical firms in the US. As states belong in the same country and, therefore, share common institutions, among other things, an interesting issue that arises is whether small institutional and networking differences across regions have different innovation implications.

²⁹A related study of Fritsch and Slavtchev (2010) has also analyzed the relationship between the specialization of a region in a panel of German industries and the efficiency of the region in generating new knowledge. Externalities resulting from high R&D intensity of the local private sector as well as knowledge from local public research institutions are identified to positively influence regional innovation efficiency.

³⁰A number of other recent studies such as those of Li (2009), Chen and Guan (2012), Bai (2013) and Barra and Zotti (2015) have investigated the innovation performance of Chinese and Italian (the latter study) regions applying various frontier analyses.

A.2. Methodology

A.2.1. Estimation

Input endogeneity is captured by making K , N , $R\&D$ and S functions of $TotalRD_{i,t-1}$ or $AcademicRD_{i,t-1}$ and, most importantly, by assuming that $v_{it} = [v_{it,j}, j = 1, \dots, 5]'$ are not necessarily independent. Technical and knowledge inefficiency are also not assumed necessarily orthogonal to these error terms.

The CUE-GMM criterion that we minimize is the following:

$$\min_{\theta \in \Theta} : \mathbf{F}(\theta)' \Omega(\theta)^{-1} \mathbf{F}(\theta) \quad (\text{A.1})$$

where, $\Omega(\theta)$ is the usual GMM weighting matrix used for the CUE version of GMM. This is given by,

$$\Omega(\theta) = (nT)^{-1} \sum_{i,t} \left[\mathbf{f}(X_{it}, \theta) \mathbf{f}(D_{it}, \theta)' \right] \quad (\text{A.2})$$

and the sample equivalent of the moment conditions is,

$$\mathbf{f}(X_{it}, \theta) = \left[(nT)^{-1} \sum_{i,t} u_{it,j}(X_{it}, \theta) z_{it,m} = 0, j = 1, \dots, 5, m = 1, \dots, M \right]' \quad (\text{A.3})$$

where, X_{it} denotes the data, $\theta \in \Theta$ is the parameter vector, $z_{it,m}, m = 1, \dots, M$ denotes the instruments and $u_{it,j}(X_{it}, \theta)$ denotes deviations of data for dependent variables from the translog specifications in equations (3) to (7).

For the system as a whole the p-value of Hansen's J -statistic to test the over-identifying restrictions is 0.26.

An R^2 for each equation is computed by considering the square of the correlation coefficient between actual and predicted values.

To account for heterogeneity we introduce state fixed effects in all equations. CUE-GMM is implemented in two "stages". First, the fixed effects are wiped out by taking deviations from the means. Suppose the typical equation in our system is,

$$y_{it,j} = a_{i,j} + g_j(X_{it}, \theta) + \varepsilon_{it,j} \quad (\text{A.4})$$

where, $\varepsilon_{it,j} = v_{it,j} - u_{it,j}$ for $j = 1, 2$ and $\varepsilon_{it,j} = v_{it,j}$ otherwise.

The fixed effects are wiped out via the transformation,

$$y_{it,j} - \bar{y}_{t,j} = g_j(X_{it}, \theta) - T^{-1} \sum_t g_j(X_{it}, \theta) + \varepsilon_{it,j} - \bar{\varepsilon}_{t,j} \quad (\text{A.5})$$

We ignore any econometric issues introduced by the averaging as our purpose is to obtain good starting values for θ . Specifically, we impose the restrictions that (i) the production function is monotone and concave, and (ii) the knowledge function is monotone, at each observed data point. In the second stage, we implement CUE-GMM in the estimation of (A.4) using as starting values for all parameters (including the fixed effects $a_{i,j}$) the final estimates from the first stage.³¹

The monotonicity and concavity restrictions in (i) and (ii) can be written as,

$$\theta \in \mathcal{R} \Leftrightarrow \mathbf{q}(\mathbf{X}, \theta) \leq \mathbf{0} \quad (\text{A.6})$$

³¹All computations are performed in Fortran 77 making extensive use of IMSL libraries. Numerical optimization for second-stage CUE-GMM is performed using subroutine DUMIDH, a quasi-Newton algorithm with a finite-difference Hessian. Numerical optimization for first-stage CUE-GMM is performed using subroutine DUMPOL, a Nelder-Mead simplex algorithm, to avoid non-convergence problems due to bad starting values and / or the penalty function. The final results of both optimizations are checked using a simulated annealing global optimization procedure. The reported results are based on a HAC covariance matrix robust to heteroskedasticity and autocorrelation of order two.

for a vector function $\mathbf{q}(\mathbf{X}, \boldsymbol{\theta}) \leq \mathbf{0}$, $q : \mathfrak{R}^{d_x} \times \mathfrak{R}^k \rightarrow \mathfrak{R}^{d_q}$, where \mathcal{R} denotes the feasible region for the parameter and $\mathbf{X} \equiv [X_{it}]$ vector.

For example, first-order derivatives at each observed point should be positive, and concavity for the production function imposes restrictions on the Hessian sub-matrices.

In both stages of CUE-GMM we impose the restrictions by modifying the criterion function in (A.1) using a penalty function, as follows:

$$\min_{\boldsymbol{\theta} \in \Theta} : \mathbf{F}(\boldsymbol{\theta})' \boldsymbol{\Omega}(\boldsymbol{\theta})^{-1} \mathbf{F}(\boldsymbol{\theta}) + \lambda \Pi(\boldsymbol{\theta}, \mathcal{R}) \quad (\text{A.7})$$

where, $\lambda > 0$ is a fixed parameter and the penalty function is,

$$\Pi(\boldsymbol{\theta}, \mathcal{R}) = \max \left\{ \mathbf{q}(\mathbf{X}, \boldsymbol{\theta}), \mathbf{0}_{d_q} \right\} \quad (\text{A.8})$$

Parameter λ is increased from 10 to 10^p for $p = 2, 3, \dots$ until we can satisfy 90% of the restrictions in (A.6). In the first stage for $p = 5$ we can satisfy 90% of the restrictions but not all of them when $p > 5$. In the second stage all restrictions can be satisfied for $p = 8$. In the second stage, fixed effects $a_{i,j}$ are jointly estimated with all other parameters. In both stages, when $\boldsymbol{\Omega}(\boldsymbol{\theta})$ does not appear to be numerically invertible we use regularization: $\boldsymbol{\Omega}(\boldsymbol{\theta}) := \boldsymbol{\Omega}(\boldsymbol{\theta}) + 10^{-7} \mathbf{I}$. Our criterion for non-invertibility is that the minimum eigenvalue is below machine precision.

A.2.2. Variables and Sources

Data are retrieved from various sources. Information on real gross state product, GSP (*Output*), which is the dependent variable of the production of output function, is extracted from the *Regional Accounts* of the Bureau of Economic Analysis.³² Data on gross fixed capital formation to construct physical capital stock (K) and number of workers (N) - the two inputs of the production function - are obtained from Yamarik (2013) and the Bureau of Labor Statistics, respectively.

We use patents weighted by their citations (*Patents*), as measure of innovative output for the knowledge function.³³ We only take into account patents assigned to corporations. Every patent is assigned to an issue year and technology field. We have 14 years and 37 technology groups; therefore, every patent is classified in one out of $14 \times 37 = 518$ groups. Each patent in every group is then weighted by the number of citations it has in the group's distribution. The weighting scheme is the following one: $w_1 = 0.1$, if citation belongs to the first quantile, $w_2 = 0.2$ for the second, $w_3 = 0.3$ for the third, and $w_4 = 0.4$ for the fourth. We then sum these values up for every state and for every year t and calculate our weighted measure of innovation output.

There is a number of issues raised with patents. In the US only a person can file for a patent application. Subsequently, when the *inventor* is an employee to a firm or to an institution, she re-assigns the patent to the employer (the so-called *assignee*). Thus, there are two main entities that are involved with a patent. First, it is the inventor who is the driving force behind the creation and manifestation of the idea into an invention and second, the firm or institution which supply the framework and resources for inventors to operate, and reaps the benefits of the innovation and use patents as a means to protect its inventions. Therefore, one can address the efficiency in the production of patents either way, i.e., patents by inventors or patents by assignees. In our sample, almost all patents, about 98.5%, have a single assignee. So, if we focus just on the first (and single, in the vast majority of the cases) assignee and allocate each patent to the (first) assignee's state, then we have a very small measurement error; merely 1.5% of the patents disclose more than one assignee and even in these cases, we miss information only if assignees are located in different states. In

³²See http://www.bea.gov/iTable/index_regional.cfm

³³A patent grants the patent holder a monopoly right for a given period of time. In the US prior to 1995, a patent was potentially active for 17 years from issue date. After June 8, 1995 where the the P.L. 103-465 (also referred to as "Uruguay Round Agreements Act") patent life changed to 20 years from first application date (USPTO, 2001). Note that in the case of Pharmaceutical patents, there is increase in patent life by an additional 5 years due to delays of FDA approvals per Hatch-Waxman Act of 1984.

our sample, 57% of patents have more than one inventor. Very frequently, inventors are located in different states. In case a patent has more than inventor, then patent allocation issues are raised. For example, one could divide the patent by the number of its inventor, say by $1/n$ (n is the inventors), and allocate $1/n$ to each state. Another way is to n -count that specific patent, i.e., equally allocate it to all (inventors') states. However, every weighing scheme has its own pros and cons and there is no theoretical justification as to what scheme is the optimal one. Therefore, we opted to count patents based on assignee's location rather than inventor's location based on three reasons: first, counting patents based on inventor's location is a less ambiguous task.³⁴ Second, the inventor could reside outside the US and we are mainly interested in the amount of patents corresponding in each state. Third, in all states, the percentage of patents that disclosed the same state for the lead inventor and assignee was quite high.³⁵ Finally, as our goal is to weight patents by their citations, we opted to weight them by the citations the owner (assignee) receives rather than the inventor. In other words, our goal is to measure the value generated to the firm and not to the inventor.

Data on patent counts by assignee (at grant date) is extracted from Hall et al. (2001).³⁶ This database also provides location information on the patent assignees (individual, firm or government). By extracting information on the location for the patent assignees, we are able to count the number of patents issued to applicants within a state within a given year. We only count a patent once; i.e. to the first assignee, in case there is more than one assignee.³⁷ However, it could well be the case that the headquarters (as well as the actual research unit) of a firm to be located in a different state and in such a case the assignee address could be the firm's headquarters. For robustness purposes, we also count patents by state of inventor.³⁸ We extract this information from the Patent Network Dataverse constructed by Lai et al. (2011).³⁹ As in the case of patents counts by assignee, we also rely only on the first inventor's state. As additional robustness, we double count patents with more than one inventors.

As inputs of the knowledge production function we consider industry level R&D stock, as a proxy of a knowledge stock, and the number of doctoral scientists and engineers as a proxy of human capital. Information at the state level on business R&D expenditure for constructing R&D capital stock (*R&D*), and doctoral scientists and engineers (*S*) is extracted from the National Science Foundation *Science and Engineering State Profiles*. From the latter, we also derived state level data on academic R&D and total R&D expenditures to construct academic R&D stock (*academicR&D*) and total R&D stock (*totalR&D*), respectively. To calculate physical capital stock and all R&D stocks, we use the perpetual inventory method as in Guellec and van Pottelsberghe de la Potterie (2004) and Hall et al. (2005).⁴⁰

Finally, merchandise trade flows at the state level are extracted from the Bureau of Transportation Services *Commodity Flow Surveys*. Data are available only for the years 1993, 1997, and 2002.⁴¹

³⁴For robustness purposes, however, we replicated our analysis by counting our dependent variable (patents) by lead inventor's location and results did not change in any significant way.

³⁵For instance, for the case of California the percentage was over 90%, for Texas approximately 90%, for Missouri more than 90% and so forth. Therefore, focusing either on the lead inventor or assignee would make little difference, if any, in terms of *ex post* analysis.

³⁶The time span of the Hall *et al.* (2001) stretches till 2001. A recent update of the data (till year 2006) can be found at <https://sites.google.com/site/patentdataport/>. Assignee names are cleaned and each unique assignee is assigned a unique id number.

³⁷In our data, only 1.5% of patents are co-assigned, i.e., disclose more than one patent assignee; a result consistent with findings both from the United States Patent and Trademark Office (USPTO) and European Patent Office (EPO) (Fosfuri et al., 2012). Therefore, relying solely on the first assignee is not likely to cause any measurement issues. For robustness purposes, however, we also allow for double counting of patents that are co-assigned. For instance, a patent that has three assignees, for example, two in California and one in New York, is counted twice for California and once for New York.

³⁸In our data, approximately 200,000 patents have their first inventor in the US, while the first assignee is not located in the US. Similarly, 80,000 patents have the first assignee located in the US and the first inventor located outside the US. Finally, of the approximately 1,000,000 patents where both the first assignee and inventor are located in the US, for 27% of the cases the first inventor and first assignee are located in different states.

³⁹<http://hdl.handle.net/1902.1/15705> UNF:5:9kQaFvALs6qcuoy9Yd8uOw== V1

⁴⁰Following the literature, we have tried different depreciation rates, e.g., 15%, and 20%. The resulted stocks, using different depreciation rates, are highly correlated (around 0.998).

⁴¹What is only available is the value of state's total imports; there is no information, however, by type of (imported) good.

A.3. Tables

Table A.1: Summary Statistics and Efficiency Scores per State

State	Summary Statistics					Efficiency Scores		
	Output	Patents	K	N	R&D	S	Output Efficiency	Knowledge Efficiency
AK	30.99116	2.271429	30575.57	299391.4	79.2154	1295.643	0.901	0.854
AL	116.1839	40.64286	119278.1	2017008	4020.303	7105.214	0.902	0.851
AR	67.31779	17.78572	70791.03	1194604	1345.823	3128.357	0.906	0.852
AZ	152.8057	144.7429	146869.9	2313891	8895.276	7739.429	0.904	0.848
CA	1223.555	7288.136	1276883	1.56E+07	167332.2	84556.93	0.904	0.849
CO	158.2229	262.3429	180745.4	2190529	13131.24	13117.21	0.906	0.848
CT	154.4251	865.8786	156307.9	1677994	17402.83	10257.21	0.907	0.852
DE	39.26646	1217.279	29493.8	388794.4	6987.615	3917.572	0.913	0.848
FL	474.5455	378.2786	432871.6	7264079	18523.29	17484.29	0.906	0.849
GA	271.5889	219.3857	281044.7	3907527	6379.468	12178.5	0.905	0.849
HI	44.20113	4.085714	37907.46	574688.4	659.1832	2794.072	0.899	0.859
IA	90.46287	119.8571	87206.04	1545363	3219.471	4919.214	0.909	0.849
ID	33.79185	458.1071	36005.27	627782.5	3542.033	2499.857	0.908	0.865
IL	452.2365	1453.514	467716.6	5948563	33293.98	23691.57	0.906	0.849
IN	187.1401	314.1214	194584.9	2985235	13012.71	9662.214	0.906	0.850
KS	82.9617	61.20714	89181.02	1338993	3840.226	4306.286	0.905	0.848
KY	111.4007	81.8	116472.8	1826911	2118.31	4914.429	0.905	0.849
LA	146.8658	40.64286	144642.6	1895359	795.5873	5943.571	0.909	0.852
MA	251.3861	960.7214	263638.6	3142044	42785.72	29144.21	0.905	0.848
MD	182.5274	183.5643	165455.6	2695068	10753.89	25256.86	0.903	0.850
ME	35.4286	20.86429	33643.86	638719.9	623.7079	2458.714	0.900	0.855
MI	309.9446	1288.536	333901.6	4698547	63932.98	17645.43	0.903	0.849
MN	177.4107	806.0214	187425.3	2645793	15733.58	11415.07	0.907	0.849
MO	172.6266	159.2429	175887.2	2763049	7896.597	9797.786	0.905	0.849
MS	65.42122	34.36429	63947.28	1208321	565.1187	3375.643	0.906	0.862
MT	22.48844	16.81429	27143.64	443969.7	154.8606	1979.571	0.903	0.853
NC	260.4055	268.9	233290.6	3788597	13803.4	17308.93	0.910	0.849
ND	18.74678	3.914286	21019.01	333285.4	249.7103	1629.429	0.908	0.855
NE	56.16074	39.82143	59832.4	902453	840.4363	2969.928	0.904	0.852
NH	40.62439	92.25714	42185.21	650766	3054.036	2658.143	0.905	0.850
NJ	334.3741	1978.236	333216.7	4005004	47925.43	23378.36	0.903	0.848
NM	51.6931	46.14286	45539.72	793734.5	5201.694	8444.429	0.906	0.851
NV	74.18292	100.6071	73647.97	946756.1	1048.772	2106.572	0.904	0.852
NY	742.3105	4336.814	810810.6	8473489	50725.27	45883.79	0.906	0.848
OH	359.757	1150.6	353494.6	5453330	31919.88	21681.64	0.905	0.848
OK	93.74434	107.3929	119483.3	1567856	2339.058	4956.643	0.904	0.856
OR	105.2072	154.3143	111446.3	1667749	6308.506	8121	0.910	0.849
PA	385.0078	946.6643	418319.9	5735257	36553.5	28078.14	0.904	0.848
RI	33.33106	47.04286	30382.81	500194.2	3204.849	2865.214	0.905	0.849
SC	110.0137	107.3143	101976.7	1843972	3293.413	5502.143	0.904	0.851
SD	23.30956	6.507143	20770.21	391422.7	128.97	1144.286	0.906	0.852
TN	171.7585	150.6929	170128	2677278	5794.982	9607.857	0.905	0.849
TX	720.4307	1582.636	945561.2	9738478	35367.85	34721	0.905	0.853
UT	65.58952	115.8071	65501	1065516	3616.472	5325.143	0.913	0.849
VA	254.6196	239.8786	222346.4	3490540	9605.989	19467.71	0.906	0.852
VT	17.38005	13.86429	18238.95	324263.8	1392.172	1913	0.902	0.852
WA	211.4904	482.9572	209758.7	2849684	30859.67	15551.71	0.902	0.852
WI	169.6502	411.9	172697.4	2819648	8755.62	9228.286	0.905	0.848
WV	42.60871	8.228572	55126.92	753086.8	830.2287	2279.286	0.905	0.854
WY	20.02302	5.635715	27281.87	254870.1	68.30658	889.9286	0.907	0.855

State's two-letter abbreviation reported in first column; *Output* is real gross state product, *K* is gross fixed capital, *N* are thousands of employees (fte), *Patents* is patents filled per state weighted by their citations, *R&D* is industry (business) R&D stock, *S* is number of doctoral scientists and engineers, *Output Efficiency* is the production of output efficiency and ranges between zero and one, and *Knowledge Efficiency* is the production of knowledge efficiency and ranges between zero and one. All monetary variables are expressed in million (2000 US) dollars.

Table A.2: Input Instrumentation

	Output		Knowledge	
	$\ln K$	$\ln N$	$\ln R\&D$	$\ln S$
$\ln RD_{total\ it-1}$	0.017*** (0.003)	0.135*** (0.002)		
$\ln RD_{academic\ it-1}$			0.472*** (0.021)	0.145*** (0.042)
D_{it}	0.012*** (0.001)	0.011*** (0.002)	0.212*** (0.007)	0.017*** (0.001)
<i>Constant</i>	-2.335*** (0.014)	-1.175*** (0.017)	0.717*** (0.046)	5.345*** (0.135)
Observations	650	650	650	650
R^2	0.88	0.72	0.85	0.83

All regressions include state fixed effects and time dummies. Standard errors reported in parentheses; respectively; (***) (**), and (*): significance at 1%, 5%, and 10% level, respectively.