A Poisson Stochastic Frontier Model with Finite Mixture Structure

Kyriakos Drivas and Claire Economidou and Efthymios G. Tsionas

Department of Economics, University of Piraeus, Piraeus,
Department of Economics, University of Piraeus, Piraeus,
Department of Economics, Athens University of Economics and Business

20. July 2014

A Poisson Stochastic Frontier Model with Finite Mixture Structure[☆]

Kyriakos Drivas[a], Claire Economidou[a,*], Efthymios G. Tsionas[b]

[a]*Department of Economics, University of Piraeus, Piraeus 184 35, Greece*
[b]*Department of Economics, Athens University of Economics and Business, Athens 104 34, Greece*

Abstract

Standard stochastic frontier models estimate log-linear specifications of production technology, represented mostly by production, cost, profit, revenue, and distance frontiers. We develop a methodology for stochastic frontier models of count data allowing for technological and inefficiency induced heterogeneity in the data and endogenous regressors. We derive the corresponding log-likelihood function and conditional mean of inefficiency to estimate technology regime-specific inefficiency. We further provide empirical evidence that demonstrates the applicability of the proposed model.

Keywords: efficiency, Poisson stochastic frontier, mixture, innovation, states
JEL: C13, C24, O33, O51

1. Introduction

Estimation of productive efficiency is performed in the framework of frontier methodologies, which have been now used extensively in economics. Rather than fitting functions that intersect data, frontier methodologies are concerned with the construction of frontiers that envelop the data and the benchmarking of performance of a decision making unit (e.g. country, region, firm) to the best practice, the frontier.[1] Units perform better than others when they use their inputs more optimally than others to produce countable outcome. The most optimal units form the efficient frontier, while less performing are situated below the frontier and their distance from the frontier represents their productive inefficiency.

The stochastic frontier (SF) methodology, in particular, constructs the efficient frontier by estimating the underlying production technology (represented either by production, cost, profit, revenue, or distance functions) across all units in the sample and specifies a two-part error term that accounts for both random error and the degree of technical inefficiency. Since its inception (Aigner et al., 1977; Meeusen and van den Broeck, 1977), the traditional SF model has been modified in many ways to confront various issues. Some recent developments have proposed modifications with respect to the distributional assumptions concerning inefficiency imposed by standard stochastic frontier models to account for perfectly efficient units (Sickles and Qian, 2009; Kumbhakar et al., 2013)[2] Others, have suggested augmentations of the traditional stochastic frontier model with Markov-switching structure (Tsionas and Kumbhakar, 2004) or finite

[☆]We thank Joanna Pagoni and Zoi Georgiopoulou for excellent research assistance. We also thank Jaap W.B. Bos for useful insights and suggestions. Kyriakos Drivas gratefully acknowledges financial support from the *National Strategic Reference Framework* No: SH1_4083. The usual disclaimer applies.

[*]Corresponding author.
Email addresses: `dribask@unipi.gr` (Kyriakos Drivas), `economidou@unipi.gr` (Claire Economidou), `tsionas@aueb.gr` (Efthymios G. Tsionas)

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

[2]For example, to account for fully-efficient units, Sickles and Qian (2009) suggest right (upper bound) truncation, whereas Kumbhakar et al. (2013) censoring of the distribution of inefficiency to overcome the typical assumption of continuous distribution of inefficiency imposed by conventional stochastic frontier modeling, and consequently, perfectly efficient units could be deemed inefficient as their full-efficiency probability is zero.

mixture (latent class) structure (Greene, 2002a,b, 2005; Orea and Kumbhakar, 2004; Bos et al., 2010a,b) to allow for various degrees of technological and inefficiency induced heterogeneity among units. In all basic or modified versions of the SF, the stochastic frontier is constructed by using continuous data analysis.

In social science literature, however, there are plenty of cases where the dependent variable is a count, taking non-negative integer values. For instance, applications of count data models are widespread in economics in modeling the relationship between number of patents granted and R&D expenditures of firms (Hausman et al., 1984) and in finance in modeling, for example, bank failures (Davutyan, 1989), unpaid installments by creditors of a bank (Dionne et al., 1996), among other applications.

To this date, there were a handful attempts in the literature that developed count data stochastic frontier techniques. The studies of Fè-Rodríguez (2007) are the first attempts in the SF literature to estimate production frontiers and calculate efficiency for discrete conditional distributions when output is an economic bad and the work of Hofler and Scrogin (2008) for economic goods. Both studies, however, lack of generality, as neither of these works can analyze both types of commodities in a single model. A more flexible count data stochastic frontier model that overcomes the restrictions of the past studies is introduced by Fè-Rodríguez and Hofler (2013). The authors evaluate its applicability and estimate a knowledge production function for a number of patents awarded to 70 pharmaceuticals in the US for the year 1976 given their expenditures on R&D. The proposed model, however, does not address dynamics and heterogeneity in the data.

This paper purports to enrich the current menu of approaches within the SF paradigm. Specifically, our contribution lies in introducing stochastic frontier estimation techniques appropriate for count models accounting for potential endogeneity of regressors and technological and efficiency induced heterogeneity in the data. To allow for different technological regimes across units, a finite mixture structure is employed to allocate regime membership. To this end, we develop a Poisson stochastic frontier model for count data augmented with finite mixture structure.

We apply our modeling approach to the economics of innovation and growth to assess the efficiency in the production of innovation. The standard approach in the economics literature, so far, is 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.[3] Only recently, a cognate strand of research employs frontier analyses to the production of innovation (Rousseau and Rousseau, 1997, 1998; Wang, 2007; Wang and Huang, 2007; Cullmann et al., 2012) and, therefore, consider the existence of (in)efficiency.[4] In estimating innovation efficiency, however, these studies have overlooked the "appropriateness" of technology, stressed by recent contributions in the economics literature (Basu and Weil, 1998; Acemoglu and Zilibotti, 2001; Jones, 2005), as economic units choose the best technology available to them, given their input mix. The latter implies the possible existence of multiple technology regimes and not just one, described by a single frontier as it is the case with past innovation efficiency studies.

In this study, we estimate a stochastic frontier of innovation production in a panel of fifty US states over the period 1993-2006 applying novel stochastic frontier techniques for count data in a dynamic and heterogenous setup. To our knowledge, empirical evidence based at disaggregated level analysis of innovation efficiency in the US has been extremely thin.[5] As states belong in the same country and, therefore, share common institutions, among other things, an interesting issue that arises is whether small differences across regions, for instance, in fiscal or employment policies have different innovation implications. According to our findings, they do. Our results support the existence of two distinct innovation classes: a very efficient one, which contains the majority of the states and exhibits increasing returns to the produc-

[3]The empirical testing of growth models has typically examined the effect of R&D on productivity or output growth ignoring any waste in the use of innovation resources. See, for instance, Jones (1995), Coe and Helpman (1995), Aghion and Howitt (1998), Griffith et al. (2004), Zachariadis (2003), Bottazzi and Peri (2003), Bottazzi and Peri (2007), and Mancusi (2008) among others.

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

[5]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 only 14 out of 51 states show 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.

tion of innovation, positive technical growth, and a less efficient class, which experience constant returns and technical regress.

The reminder of the paper proceeds as follows. Section 2 introduces a Poisson stochastic frontier model with finite mixture structure appropriate for count data, allowing for technological and efficiency induced heterogeneity and endogenous regressors. Section 3 provides empirical evidence, which demonstrates the applicability of the method proposed. Finally, Section 4 summarizes and concludes.

2. Methodology

In this section, we briefly sketch the basic idea of a standard stochastic frontier model with the use of a production function. Similar analysis could be performed, for instance, with cost, profit, and revenue functions. We then modify the traditional stochastic frontier modeling to allow for discrete conditional distributions and, therefore, introduce a Poisson stochastic frontier model. To account for potential endogeneity in the regressors as swell as technological and inefficiency induced heterogeneity, we enhance the Poisson stochastic frontier model with finite mixture structure.

2.1. A Brief Sketch of Stochastic Frontier Analysis

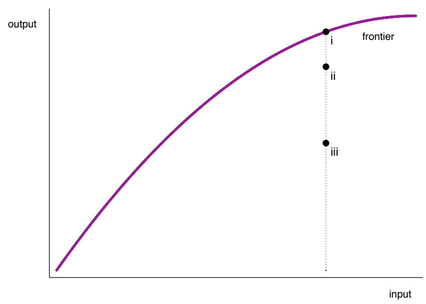
Assume a production function of a good or idea is described by the following equation:

$$Q_{it}^{*} = f(X_{it}, t; \beta) \exp\{v_{it}\} \tag{1}$$

where Q^* is the maximum (frontier) attainable output produced of unit i at time t given available vector of inputs, X, f and parameter vector β characterize the production technology, t is a time trend variable that captures neutral technical change (Solow, 1957), and v_{it} is an i.i.d. error term distributed as $N(0, \sigma_v^2)$.

Some units, however, may employ existing production resources less efficiently and, therefore, produce less than the frontier output. As Figure 1 below shows, for a given technology and set of inputs, there are units that produce at points *ii* and *iii*, in other words, their actual output is less than the frontier output.

Figure 1: A Stochastic Frontier Model of Production



To also allow for such cases, we model the performance of units' production by means of stochastic frontier production model as follows:

$$Q_{it} = Q_{it}^{*} \exp\{v_{it}\} \exp\{-u_{it}\} \tag{2}$$

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}.[6] Technical efficiency is measured as the ratio of actual over

[6]The residual in equation (2) is decomposed as $\exp\{\varepsilon_{it}\} = \exp\{v_{it}\}\exp\{-u_{it}\}$ and one can 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 (see Kumbhakar and Lovell, 2000).

maximum output, $\frac{Q_{it}}{Q^*_{it}}$, such that $0 \leq \frac{Q_{it}}{Q^*_{it}} \leq 1$ and $\frac{Q_{it}}{Q^*_{it}} = 1$ implies full efficiency. The standard way to calculate technical efficiency is to define a functional form of the efficient frontier and then log-linearize equation (2).

In the presence of count data, however, one cannot apply the log-linear transformation to equation (2).[7] An additional problem in count data sets is the existence of zero output in the set. As log of zero is not defined, a high proportion of the data could be discarded. To circumvent, however, the discrete nature of the data one can approximate the discrete random variable by a continuous one. In doing so, there is a possible loss of efficiency and even more so it could be a source of model misspecification (Cameron and Trivedi, 2013).[8]

In this paper, we focus on count data frontier models and, therefore, we approach equation (1) as Poisson process. This is the task of the section below.

A Poisson Stochastic Frontier

Suppose that actual output, Q_i, has a Poisson distribution, conditional on input vector X, with the conditional mean of the distribution $Q_i \mid \lambda_i \sim$ Poisson(λ_i), that is:

$$p(Q_i|\lambda_i) = exp(-\lambda_i)\frac{\lambda_i^{Q_i}}{Q_i!} \tag{3}$$

where $Q_i \in 0, 1, 2, ...$ are non-negative integers (counts), i = 1, ..., n are state-year observations, and λ is the mean of Poisson process and defined as:

$$log\lambda_i = x_i^{'}\beta + v_i - u_i \tag{4}$$

where x is log of the inputs vector X, β a vector of parameters, $v_i \sim$ iid $N(0, \sigma_v^2)$ and $u_i \sim$ iid $N(0, \sigma_u^2)$ a half-normal distribution.

The distribution of Q_i has density given by:

$$p(Q_i|\theta) = (2\pi\sigma_v^2)^{-1/2}(\frac{\pi}{2}\sigma_u^2)^{-1/2}\int_0^{+\infty}\int_0^{+\infty} exp(-\lambda_i)\frac{\lambda_i^{Q_i-1}}{Q_i!}exp[-\frac{(log\lambda_i - x_i^{'}\beta + u_i)^2}{2\sigma_v^2} - \frac{u_i^2}{2\sigma^2}]d\lambda_i du_i \tag{5}$$

where $\theta = (\beta', \sigma_v, \sigma_u)' \in \Theta \subset \Re^{k+2}$.

The outer integral is available in closed form and one gets:

$$p(Q_i|\theta) = \frac{2}{\sigma}\int_0^{+\infty} exp(-\lambda_i)\frac{\lambda_i^{Q_i} - 1}{Q_i!}\varphi(\frac{log\lambda_i - x_i^{'}\beta}{\sigma})\Phi(-\lambda\frac{log\lambda_i - x_i^{'}\beta}{\sigma})d\lambda_i \tag{6}$$

where $\sigma^2 = \sigma_v^2 + \sigma_u^2$, $\lambda = \frac{\sigma_u}{\sigma_v}$ and φ, Φ denote the density and the distribution function respectively of the standard error.

Technical Efficiency

We now turn into calculating (in)efficiency. From equation (3) the distribution of u_i conditional on λ_i has the well-known Jondrow et al. (1982) (JLMS) density:

$$p(u_i|\lambda, Q_i) = (2\pi\sigma_*^2)exp(-\frac{(u-\mu_*)^2}{2\sigma_*^2})\Phi(-\frac{\mu_*}{\sigma_*})^{-1} \tag{7}$$

[7] For example, let Q=4 and Q^*=11, then there is no integer value of efficiency solving the equation.

[8] Fè-Rodríguez and Hofler (2013) notes that discrete distributions often violate the third moment restrictions imposed by a continuous data stochastic frontier model. In case that output is discrete, then the log-linear transformation of output may exhibit skewness of wrong sign. The latter would result to zero inefficiency in the model even when there is substantial one.

where $\mu_* = -(log\lambda_i - x^{'}\beta)\frac{\sigma_u^2}{\sigma^2}$ and $\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma^2}$.
The mean of distribution is:

$$E(u_i|\lambda_i, Q_i) = \sigma_*[\frac{\varphi(\varepsilon_i(log\lambda_i)\lambda/\sigma)}{\Phi(-\epsilon_i(log\lambda_i)\lambda/\sigma)} - \varepsilon_i(log\lambda_i)\lambda/\sigma] \tag{8}$$

where $\varepsilon_i(log\lambda_i)$ is equal to $(log\lambda_i - x_i^{'}\beta)\lambda/\sigma$.[9]

As λ_i is unobserved, equation (6) cannot be used directly. Instead, we can use:

$$E(u_i|\lambda_i, Q_i) = \sigma_* \int_0^{+\infty} [\frac{\varphi(\varepsilon_i(log\lambda_i)\lambda/\sigma)}{\Phi(-\epsilon_i(log\lambda_i)\lambda/\sigma)} - \varepsilon_i(log\lambda_i)\lambda/\sigma](-\lambda_i)exp(-\lambda_i)\frac{\lambda_i^{Q_i}}{Q_i!}d\lambda_i \tag{9}$$

Using the change of variables $log\lambda_i = \zeta_i$ the integral can be transformed as follows:

$$E(u_i|\lambda_i) = \sigma_* \int_{-\infty}^{+\infty} [\frac{\varphi(\varepsilon_i(\zeta_i)\lambda/\sigma)}{\Phi(-\epsilon_i(\zeta_i)\lambda/\sigma)} - \varepsilon_i(\zeta_i)\lambda/\sigma]\frac{exp(\zeta_i + 1 - exp(\zeta_i))}{Q_i!}d\zeta_i \tag{10}$$

The integral is evaluated numerically.[10] Equation (10) is the analogue of the JLMS measure and provides the mean efficiency for count data frontier models.

2.2. Endogenous Regressors

The usual concern with estimating production functions is the endogeneity of regressors. In conjunction with equations (3) and (4) we assume:

$$x_i = \Gamma z_i + v_{i1} \tag{11}$$

where z_i is an $mx1$ vector of covariates, v_i=$[v_{i0}, v_{i1}^{'}]^{'} \in N_{k+1}(0, \Sigma)$, and Γ is a kxm matrix of parameters.
The distribution of Q_i has now density given by:

$$\begin{gathered} p(Q_i|\theta) = (2\pi)^{-(k+1)/2}|\Sigma|^{-1/2}\frac{1}{Q_i}(\frac{\pi}{2}\sigma_u^2)^{-1/2} \\ \int_0^{\infty}\int_0^{\infty} exp(-\lambda_i)\lambda_i^{Q_i}exp\left[-1/2(\tilde{u_{i0}}x_i - \Gamma z_i)^{'}\Sigma^{-1}(\tilde{u_{i0}}x_i - \Gamma z_i) - \frac{u_i^2}{2\sigma_u^2}\right]du_i d\lambda_i \end{gathered} \tag{12}$$

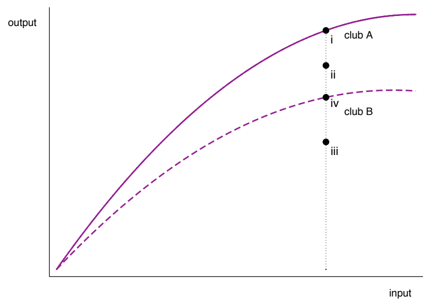
where $\tilde{u_{i0}} = log\lambda_i + u_i - x_i^{'}\beta$ is a function of latent variables.
We can formulate the likelihood based on the density in equation (12) and maximize with respect to the parameters using standard conjugate-gradient algorithms (Terza et al., 2008).
Having controlled for potential endogeneity in the regressors, we now turn to modeling different technology classes.

[9]See Kumbhakar and Lovell (2000, p 77-78).
[10]We use adaptive twenty-point Gaussian quadrature. Relative to a ten-point rule, the results were virtually the same. We truncate the infinite range of integration to an interval [a, b]. Since $\lambda_i \geq 1$ in our sample and artificial data that we examine later, we set $\alpha = 0$ and the upper bound b is determined so that the integral changes by less than 10^{-6}. A value of $b = 10$ was found more than adequate.

2.3. Technology Classes

Units use the best available technology given their input mix. Accordingly, they belong in different technological regimes. As Figure 2 below shows, for a given technology and set of inputs, there are units that produce output at point *ii* and, therefore, are inefficient compared to their own frontier and class (A) as their actual output is less than the maximum (frontier) attainable output, while other states produce at *iii* and exhibit some inefficiency compared to their own frontier and class (B).

Figure 2: Different Stochastic Frontier Models of Production



Regime membership, however, is unobserved. We account for the existence of different technology classes, where technology class membership is a function of covariates, z, and therefore advocate for a stochastic frontier model augmented with mixture structure (Geweke, 2007). The focus is on technological and inefficiency induced heterogeneity. The finite mixture approach assumes that units have access to a finite number of technologies and the inefficiency of a unit is associated with a particular technology regime. The rationale behind the mixture approach is to probabilistically identify which unit is using what technology and then measure inefficiency as a probability weighted average computed using each of these technologies as the benchmark technology.

Consequently, given the Poisson stochastic frontier model, equation (4) is modified as follows:

$$log\lambda_i = x_i^{'}\beta_{|c} + v_{i|c} - u_{i|c} \tag{13}$$

where $c \in \{1, ..., G\}$ and G is the number of distinct groups (classes), $v_{i|c} \sim \ i.i.d. N(0, \sigma^2_{v|c})$, and $u_{i|c} \sim i.i.d. N(0, \sigma^2_{u|c})$.

We assume:

$$P(c = g|z_i, \delta) = \frac{exp(z_i^{'}\delta_g)}{\sum_{g'=1}^{G} exp(z_i^{'}\delta_{g'})} \ g = 1, ..., G \tag{14}$$

The covariates z_i determine directly the probability of classification in class g and they are assumed to be the same with the covariates that we have used in equation (11). For normalization purposes, we assume: $\delta_G = 0_{(mx1)}$. In obvious notation $\delta = [\delta_i^{'}, ..., \delta_G^{'}]'$.

For the mixture model given by the Poisson model, and equations (13), (14), to account for endogeneity of regressors we also use (11).

If we denote the density in equation (12) by $p(Q_i|\theta_g)$, the density of the mixture model is:

$$p(Q_i|\theta) = \sum_{g=1}^{G} p(Q_i|\theta_g)P(c = g|z_i, \delta) \tag{15}$$

The probability of classification of observation i into group g given the data can be computed from (14) using Bayes's theorem as follows:

$$P(c = g|Q_i, z_i, \delta) = \frac{p(Q_i|\theta_g)P(c = g|z_i, \delta)}{\sum_{g'=1}^{G} p(Q_i|\theta_{g'})P(c = g'|z_i, \delta)}, \; g = 1, ..., G. \tag{16}$$

where maximum likelihood parameter estimates are substituted for the unknowns.

The resulting system of equations (13) and (16) can be estimated by maximizing iteratively, back and forth between posterior group probabilities from equation (16) and the log-likelihood function used to estimate equation (13). The likelihood maximization in equation (13) depends not only on inputs and outputs per region, but also on efficiency (λ and σ). In contrast to *a priori* clustering on the basis of some individual proxy, both the parameters β and efficiency u can be determined endogenously through latent sorting into G classes.

3. Empirical Application

In this section, we provide an empirical application of our proposed methodology. The aim is to examine whether (i) states in the US, one of the most technologically advanced country and innovation leader in the world, belong in the same (or not) technology regime, and (ii) are (in)efficient in producing new knowledge (innovation).

Our empirical specification builds on knowledge (innovation) production function, which has been first introduced in the seminal work of Griliches (1979) and empirically implemented by many studies in the literature (Pakes and Griliches, 1984; Jaffe, 1986; Hall and Ziedonis, 2001). The production of new knowledge, the innovative output or the creation of new designs in the R&D sector as in Romer (1990), is the product of knowledge generating inputs similar to the production of physical goods. Some observable measures of inputs, such as R&D expenditures and researchers, are invested in the knowledge production process and directed toward producing economically valuable knowledge, usually proxied by patents.[11] Therefore, the production of new knowledge can be described as follows:

$$Q_{it} = f(A_{it}, H_{it}) \tag{17}$$

where Q is counts of patents, A the stock of knowledge proxied by R&D stock, and H is human capital devoted to development of new knowledge proxied by the number of researchers. The units of each observation is state, i, and time, t.

In this paper, we model the performance of states' knowledge production by means of stochastic frontier production to account for inefficient use of knowledge resources.[12] To further account for different technologies employed in the knowledge production, we augment the model with a finite mixture structure. Class membership is estimated conditional on a set of covariates, included in the vector z and are state tax policies and labor mobility strictness.

More specifically, state tax policies and laws can affect the level of technology via their impact on R&D stock and human capital. As the knowledge resources (e.g. R&D and researchers) are limited, a state can remain competitive in the innovation terrain by offering motives to stimulate existed knowledge resources or to attract more innovative firms from other states. One way achieving that is to set low corporate and income tax rates and/or high R&D tax credits (Mamuneas and Nadiri, 1996; Bloom et al., 2002; Wu, 2005; Palazzi, 2011). Furthermore, labor laws can also influence innovation and technological progress by

[11] The idea of using patents counts 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). Since then, a number of papers have employed patent counts as innovation output to measure R&D (in)efficiency (e.g. Wang and Huang (2007), Sharma and Thomas (2008), Fu and Yang (2009), and Cullmann et al. (2012).

[12] States 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 (state level) data we use in our analysis. Therefore, in this paper, the term 'efficiency' refers purely to technical efficiency.

restricting or enhancing scientific labor mobility. As technological know-how acquired through research experience is embedded in the scientist's human capital, this knowledge becomes available to a competitor when the employee switches jobs. Non-competition contracts - more commonly called 'noncompetes' - is an employment agreement that limits employees' job options after leaving a company. Although the legitimate reason to enforce noncompetes is to encourage employer investment in training and information that otherwise would never take place if employees are free to depart, the literature has documented that noncompetes could limit or even impede innovation (Saxenian, 1994; Marx et al., 2009; Belenzon and Schankerman, 2012).

We redefine the production frontier as a latent class frontier, which can be characterized by a system of equations: G stochastic production frontiers and a multinomial logit model with conditioning variables in the vector z (R&D tax credits, corporate tax rate, personal tax rate, and noncompetes) for the sorting (of states) into each of the G regimes. For a translog specification of production function, where the production output, Q_i, has a Poisson distribution with the conditional mean of the distribution $Q_i \mid \lambda_i \sim \text{Poisson}(\lambda_i)$ and with a general index of technical change specified by means of time dummies t (see Baltagi and Griffin, 1988) and regimes c $(= 1, .., G)$, we can write a latent class stochastic frontier as:

$$\begin{aligned} \ln Q_{it} = {} & \beta_{o|c} + \beta_{1|c} \ln A_{it} + \beta_{2|c} \ln H_{it} + \tfrac{1}{2}\beta_{11|c} \ln A_{it}^2 + \tfrac{1}{2}\beta_{22|c} \ln H_{it}^2 + \beta_{12|c} \ln A_{it} \ln H_{it} \\ & + \gamma_{At|c} \ln A_{it} * t + \gamma_{Ht|c} \ln H_{it} * t + \delta_{1t|c} * t + \frac{1}{2}\delta_{2t|c} * t^2 + v_{it|c} - u_{it|c} \end{aligned} \tag{18}$$

Equation (18) together with equation (16) can be estimated as a system. The parameters of the system, β and efficiency u, are determined endogenously through latent sorting into G classes. Consequently, each class, c, is characterized by its own elasticities of capital and labor and level of efficiency.

An advantage of our modeling approach compared to previous latent class studies (Greene, 2002a,c, 2005; Orea and Kumbhakar, 2004) is that we allow states to switch technology regimes over time. *Within* each period, observations of a single state are not independent because the state must fall within one of the regimes during that period, and the probability of being in a regime depends on the average of the variables used to estimate regime membership. However, *across* periods, observations on a single state are treated as independent. For example, in moving from $t = t_1$ to $t = t_2$, a region is treated as a different i in the panel dimension it, and it can switch regimes. This flexibility adds an important dimension to our analysis as one can study regime migrations.

3.1. *Data*

Our empirical analysis is based on a sample of 50 US states over the period 1993-2006. Data are retrieved from various sources.

The innovative output, the result of knowledge production, is hard to capture. As new designs are usually patented, we measure the innovative output as number of patents, which are materialized innovations of business value and are actively traded in intellectual property markets. We count patents by the location of the assignee (the patent owner) whether it is individual, firm or university. Data on patent counts by assignee at grant date as well as information on the geographic location of the assignees is extracted from the NBER Data Project.[13]. We classified, during the years of our analysis, 1 million (to be precise: 1,057,301) patents assigned to US located entities. In case there are patents with more than one patent assignees, we count these patents only once based on the first assignee.[14]

Information on the two inputs of knowledge production function, R&D expenditure (for constructing R&D capital stocks) and doctoral scientists and engineers devoted to research (for human capital) is extracted from the National Science Foundation Science and Engineering State Profiles. To calculate R&D (in

[13]https://sites.google.com/site/patentdataproject/Home

[14]A mere 1.5% of patents in our sample is co-assigned.

million 2000 US dollars) stock, we use the perpetual inventory method as in Guellec and van Pottelsberghe de la Potterie (2004).[15]

Finally, information on states' tax variables and labor mobility strictness comes from a variety of sources. State top income tax rate is obtained from the National Bureau of Economic Research (NBER)[16], top corporate tax rate from the University Michigan Ross School of Business[17], and statutory R&D tax credits from Wilson (2009) for 32 states that have enacted tax credits at some time. Data on noncompete scores were obtained from Garmaise (2009) who made use of Malsberger (2004) twelve question scheme where, based on states' overall responses, a value was assigned for each state, ranging from 0 (low) to 9 (high), depending on the enforceability of noncompetes.

Annul data is used in our analysis with exemption of variables extracted from the National Science Foundation database, which are provided biannually. We use STATA's interpolation methods to fill in the gaps. All monetary variables are expressed in million of constant (2000) US dollars.

Summary statistics of the variables considered in our analysis for each state and for the period under investigation are shown in Table A.1 in the Appendix. States such as California (CA), New York state (NY) and Texas (TX) are among the top patent producers and have the highest accumulated technological knowledge (R&D stock) and human capital (scientists). In the opposite side of the spectrum are the states of Alaska (AK), Wyoming (WY) and South Dakota (SD). In terms of the policy variables, there is little variation across states. Corporate and income taxes both range between zero and eleven percent across states, with California (CA) to report the highest personal tax rate and one of the highest corporate tax rates. In terms of R&D tax credit, 32 states have some sort of R&D tax credit and the highest values, on average, for the years under consideration are observed in California (CA) (15%) and Rhode Island (RI) (15.7%). Finally, the state of California (CA), for example, has completely disregarded noncompete agreements during our sample period, whereas Florida (FL) has the most vigorous enforcement of noncompetes.

3.2. Results

Before embarking on exploring whether states in the US belong in different (or same) innovation class and whether states migrate across classes, we perform a small-scale Monte Carlo experiment in order to explore the consequences of using a log-normal distribution when the data have been generated from a Poisson distribution.

Monte Carlo Experiment

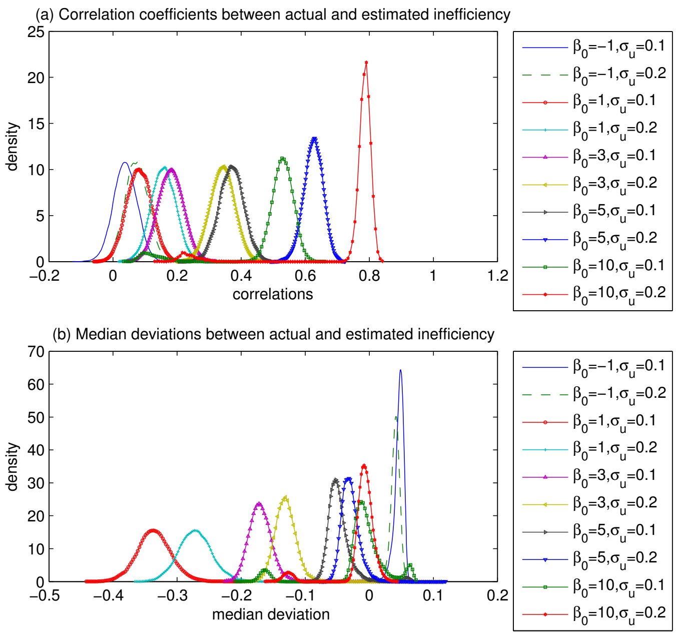
To investigate the consequences of using a log-normal distribution when the data have been generated from a Poisson we conduct a small-scale Monte Carlo experiment. The number of observations is n = 700 and the two regressors are generated as: $x_{i1} \sim$ i.i.d. N(0,1), and $x_{i2} = x_{i1} + 0.5\xi_i$, where $\xi_i \sim$ i.i.d. N(0,1). The data generating process is: $log\lambda_i = \beta_0 + 0.7x_{i1} + 0.3x_{i2} + v_i - u_i$ where $v_i \sim$ i.i.d. N(0, σ_v^2) and u_i $\sim$ i.i.d. $N^+(0, \sigma_u^2)$. We fix σ_v = 0.1 and we examine various combinations of β_0 (whose true value was proved to be critical) with σ_u = 0.1 or σ_u = 0.2. We use 10,000 simulations where the regressors x_{i1} and x_{i2} vary randomly according to the assumptions we have made. For each repetition, we compute estimated inefficiency from equation (10) and compare it to actual inefficiency using their correlation coefficient and their median deviation in the sample. Figure 3 below demonstartes the Monte Carlo results.

[15]Following the literature, we have tried different depreciation percentages, e.g., 15%, and 20%. The resulted R&D stocks are highly correlated.

[16]http://users.nber.org/ taxsim/

[17]http://www.bus.umich.edu/otpr/otpr/default.asp

Figure 3: Monte Carlo Results



As the upper panel of the Figure 3 shows the correlation between actual and estimated efficiency calculated from lognormal distribution is rather small. Similar findings hold for the median deviation (lower panel) between actual and estimated efficiency. Consequently, the log-normal distribution is not always good approximation to the Poisson distribution.

Do States Belong in Different Innovation Regimes?

We first investigate whether states in the US can be described by a common innovation production function. In estimating the mixture model specified in equation (18), we first need to determine the number of classes, G. There is little guidance as to the appropriate number of groups based on economic growth theory. Multiple regime endogenous growth models such as the ones developed in Azariadis and Drazen (1990), Easterly and Levine (2001) and Kejak (2003) corroborate the possibility of multiple steady states or growth regimes without, however, being explicit about the exact number of regimes.

We, therefore, rely upon statistical methods for determining the number of classes. The main computational problem is that a number of numerical integrations are needed with respect to λ_i and u_i. Empirically, the determination of the number of classes is specified here using the Bayesian Information Criterion (BIC) defined as $BIC = 2L - p\,log(n)/n$, where p is the total number of parameters and L the average log-likelihood, primarily because it provides consistent estimators of the model order, G. The preferred specification has the highest BIC value.

We find strong evidence in favor of two classes.[18] Accordingly, we classify states in our sample as belonging to classes A or B, respectively.

Table 1 reports the estimated parameters for the translog production function with a time trend (top panel), efficiency parameters (middle panel) and membership probability parameters (bottom panel) for

[18] Classes (G): $G = 1$ with BIC equal to -575.12; $G = 2$ with BIC equal to -568.29; and $G = 3$ with BIC equal to -569.41. In the latter class, however, parameters are jointly not significantly different from zero, and the number of observations is very small.

Table 1: Mixture Model Results

	A class		*B class*	
	Frontier			
	coeff.	st.dev.	coeff.	st.dev.
$lnR\&D$	0.230	0.023	0.117	0.015
$lnHC$	0.770	0.016	0.932	0.015
t	-0.013	0.007	0.028	0.007
$1/2lnR\&D^2$	0.054	0.003	-0.128	0.002
$lnHC * lnR\&D$	-0.315	0.002	0.454	0.031
$t * lnR\&D$	-0.043	0.003	0.127	0.002
$1/2lnHC^2$	0.073	0.002	0.002	0.001
$t * lnHC$	0.001	0.001	-0.032	0.004
$1/2t^2$	-0.001	0.001	-0.029	0.002
Constant	6.322	0.001	1.256	0.001
σ	0.303	0.017	0.335	0.025
λ	1.172	0.213	0.032	0.001
	Efficiency Estimates			
	mean	st.dev.	mean	st.dev.
	0.887	0.210	0.914	0.150
	Finite Mixture Model Coefficients			
R&D tax credit	reference group		-0.112	0.022
Corporate tax rate	reference group		-0.040	0.018
Personal tax rate	reference group		-0.031	0.022
Noncompetes	reference group		-0.662	0.034
Constant	reference group		-0.019	0.007
	Prior class probabilities at data means			
	0.821		0.179	

Note: 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=-568.29.

every regime, A and B. To examine whether parameter estimates differ significantly across regimes, we perform Wald tests for joint equality across regimes.[19]

The middle panel of Table 1 shows that inefficiency matters, too. For regime A, the efficiency parameter, λ $(= \sigma_u/\sigma_v$, the ratio of the standard deviation of efficiency over the standard deviation of the noise term), is 1.172 and significant at the 1% level. That is inefficiency is approximately 1.172 as great as noise in this innovation class. In regime B, however, the production process is quite efficient, exemplified by the small but statistically significant at 1% value of λ (0.032).

The bottom panel of Table 1 demonstrates the importance of the conditioning variables. The use of finite mixture specification implies an estimation of membership likelihood relative to the reference group, which is group A, here. For example, an increase in the R&D tax credit of 1% decreases the probability of belonging to regime A by 0.89%.[20]

The prior class probabilities (at the data means) at the bottom panel of Table 1 show that technology class A contains 23% of our sample, whereas technology class B contains 77%. The allocation of the states into the two innovation classes, A and B, is shown in Table A.1 in the Appendix. The same state can be classified as belonging to class A for some years, but also as belonging to class B for some others. However, the majority of states fall in one class (B, in this case). We observe that the smaller class, A, contains eight states, namely

[19]Wald tests available upon request, demonstrate that parameters are jointly significantly different across the two regimes. We further test whether parameters of the mixture model variables are jointly significantly different across the two regimes and find evidence in favor.

[20]We calculate probabilities by taking the exponent of the logit coefficients from the bottom panel of Table 1.

Alaska (AK), Maine (ME), Mississippi (MS), Montana (MT), North Dakota (ND), South Dakota (SD), West Virginia (WV), and Wyoming (WY), which, on average, are not high innovation performers in terms of patents production, R&D, and scientists, according to the state summary statistics. Few states, namely Arkansas (AR), Hawaii (HI), Nebraska (NE), and Rhode Island (RI) move back and forth between the two classes, with Nebraska to be the state with the most transitions between the two innovation groups.

The performance of the two classes when it comes to the sorting variables - the state policies - included in the vector z is the following: In class A, the mean R&D tax credit is, on average, 2.59%, whereas it is 4.01% in technology class B (p-value=0.001). Corporate tax rate is 6.62% for class A and 6.56% for class B (p-value=0.820), personal income tax rate is 5.29% for class A and 5.23% for class B (p-value=0.83) and, finally, noncompetes mean score in class A is 3.27 and 4.51 in class B (p-value=0.000). However, statistical difference between classes A and B are significant at 1% only for the variables R&D tax credit and noncompetes. Although policy variables are pretty much time invariant and do not differ between the two groups, still group B is the class where states, on average, provide more innovation-friendly environment to firms.

The marginal product, at the data means, of capital (R&D) stock is 0.230 in class A and in class B is about half of that in class A.[21] It appears that states in class B benefit from the fact that the marginal productivity of a unit of their researchers is higher than that of class A. Marginal products are estimated here conditional on four innovation policy variables. Although these variables do not greatly vary between the two classes, they may enhance the productivity of labor (researchers) in class B, but not necessarily the productivity of capital in the same class. Our capital and labor estimates are in line with existing empirical literature (Barro and Sala-i-Martin, 1995; Koop, 2001; Bos et al., 2010a; Wang et al., 1998). Further, the marginal rate of technical substitution (MRTS) is 0.299 (0.126) for states in class A (B), demonstrating that the rate at which researchers can be substituted for capital while holding innovative output constant is much higher for states in class A. Put differently, states in class B may use relatively cheap capital. States in class A produce at constant returns to scale, as is often reported in the literature (Barro and Sala-i-Martin, 1995; Mankiw et al., 1992), while states in class B produce at increasing returns to scale as it is supported by innovation production studied (Wang et al., 1998; Fè-Rodríguez and Hofler, 2013).

Finally, members of class A appear to be quite efficient (88.7%), when it comes to their own best practice, in producing innovation. However, the performance of states' innovation efficiency in class B is remarkable as these states produce innovation without much slack (91.4%). Including a time trend t for each class allows us to measure technical change. Interestingly, for the states that consider the frontier of technology class B their benchmark, we find that technical growth is 2.8% per year, whereas states in class A experience technical regress of 1.3%.

In sum, we find support of two innovation classes in the US, with different implications for their members' innovation growth. A highly efficient and large class that experience technical growth and with human capital to be a crucial factor for innovation and a less efficient smaller class with technical regress where accumulation of R&D capital is the main contributor.

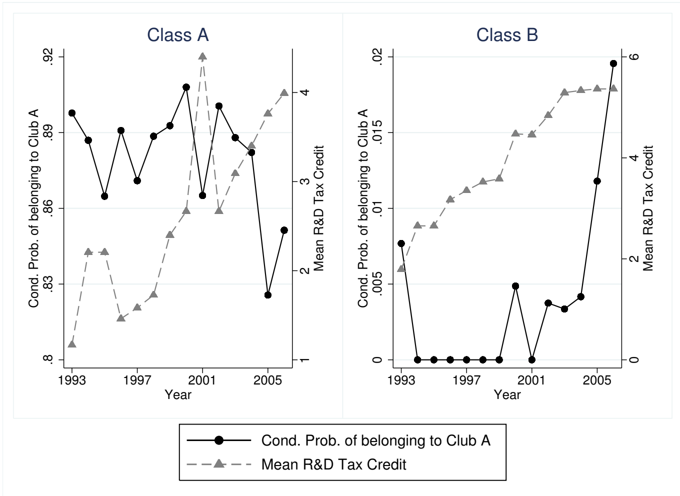
Do States Change Innovation Class Membership?

In our finite mixture model, states are not restricted to one class. In principle, a state in class A can migrate to become a member of class B (and vice versa). One of the key assumptions in our modeling strategy is that classes in our mixture model are conditional on a set of innovation policy variables, i.e., taxes and enforcement of noncompetes. As mentioned earlier, only the R&D tax credit and noncompetes differences between the two classes are statistically significant and, therefore, in this section we further explore only on these policy variables.

Figure 4 plots the conditional probability from equation (16) and R&D tax credit of states in groups A (left panel) and B (right panel).

[21] Our data are transformed, i.e., inputs are measured relative to their means, and therefore translog elasticities at means with respect to R&D stock and researchers are equal to the coefficients of R&D stock and researchers, respectively.

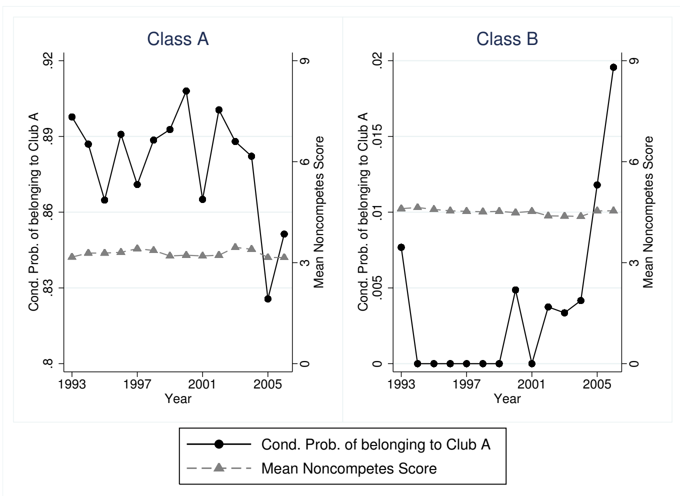
Figure 4: R&D tax credit and conditional probability of belonging to class A



The conditional probability of being a member of class A shows a volatile pattern, with some sharp changes in the middle (around 2001) and at the end (around 2005) of our sample. The mean of R&D tax credit exhibits, overall, an upward trend with some sudden changes around 2001, as some states around that year have inacted favorable R&D tax credits. Overall, the link between the development of the mean of R&D tax credit and the conditional probability of belonging to class A is, rather, weak. Only some states in class A may manage to capitalize on their high R&D tax credits they offer to their firms and become eventually a member of class B. For the states in class B, the story appears to be slightly different. Ex-ante 2001 there seems to be no relation at all; ex-post 2001, a positive relationship emerges. In other words, post 2001, the higher the R&D tax credit of the states, the higher the conditional probability of belonging in class B.

Next, Figure 5 displays the conditional probability of being a member of class A and the development of the mean of noncompetes.

Figure 5: Noncompetes and conditional probability of belonging to class A



The enforcement of noncompetes does not vary much across states and over time, as Table A.1 in the Appendix indicates. Therefore, the mean of noncompetes score is stable. The conditional probability of being member of group A does not vary that much either. Overall, the link between the development of the mean of R&D tax credit and the conditional probability of belonging to class A is, on average, absent.

We now turn to investigate the transition probabilities of states switching innovation classes. Table 2 reports these probabilities. As we saw, some states and for some time in class A have been rather successful in increasing R&D tax credits (or lowering noncompetes). These states may try to make the shift from class A to class B. In Table 2, we observe that over the sample period (less than) 5% of the states in class A manage to shift to class B.[22] For instance, the states of New Mexico (NM) and Vermont (VT), former members of class A, now join class B with Vermont to show an increase in R&D tax credit from 0% to 10% and New Mexico a slight reduction in personal tax rate. We also find that 1.7% of the states in class B, namely Alabama (AL) and Louisiana (LA), make the opposite move.

Table 2: Migrations Between Classes

	To		
From	A	B	totals
A	95 %	5 %	160
B	1.7 %	98.3 %	540
totals	161	539	700

Given that the dispersion of efficiency levels is greater in class A than B, as is shown in Table 1 (middle panel), some states are better of being efficient in the more efficient class B than inefficient in the class A, meanwhile enjoying technological progress in the former group.

4. Conclusion

Applications of count models have been ample in various disciples. In many contexts, the measuring of the technical efficiency is of central importance. Only very recently, there have been some attempts in the stochastic frontier paradigm to model efficiency when the dependent variable has discrete conditional distribution.

This paper develops a methodology appropriate for count data stochastic frontier models allowing for technological and inefficiency induced heterogeneity in the data and controlling for endogenous regressors. We, therefore, extend and generalise important aspects of past related studies in the field. The proposed model is a Poisson stochastic frontier model augmented with finite mixture structure. We derive the corresponding log-likelihood function and conditional mean of inefficiency to estimate technology regime-specific inefficiency. We feel that the methodology we proposed could be useful for applied researchers conducting efficiency studies using count data in various scientific fields, in particular in economics and finance.

To demonstrate the applicability of the proposed model, we estimate a knowledge production function, where the dependent variable is counts of patents, for the states of the US. In particular, we examine whether states in the US, one of the most technologically advanced country and innovation leader in the world, (i) belong in the same (or not) innovation regime, and (ii) are (in)efficient in producing innovation. Relevant past studies typically assume that knowledge resources are used efficiently, whereas the few which account for the latter, consider the underlying knowledge production technology to be identical for all units.

Our empirical results show that the Poisson stochastic frontier model augmented with latent class structure can be implemented in studying the innovation performance of the states of the US offering useful

[22] As we allow states to move back and forth between classes A and B, 5% also includes these cases.

insights. Our results support the existence of two distinct innovation classes with different implications for their members' innovation growth.

References

Acemoglu, D., Zilibotti, F., 2001. Productivity differences. Quarterly Journal of Economics 116(2), 563–606.

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.

Azariadis, C., Drazen, A., 1990. Threshold externalities in economic development. The Quarterly Journal of Economics 105 (2), 501–526.

Baltagi, B. H., Griffin, J. M., 1988. A general index of technical change. Journal of Political Economy 96 (1), 20–41.

Barro, R. J., Sala-i-Martin, X. X., 1995. Economic Growth. McGraw-Hill, New York.

Basu, S., Weil, D., 1998. Appropriate technology and growth. Quarterly Journal of Economics 113(4), 1025–1054.

Belenzon, S., Schankerman, M., 2012. Spreading the word: Geography, policy and knowledge spillovers. Review of Economics and Statistics, forthcoming.

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.

Cameron, C. A., Trivedi, P. K., 2013. Regression Analysis of Count Data, 2nd Edition. Econometric Society Monograph No. 53. Cambridge University Press, 1998.

Coe, D., Helpman, E., 1995. International R&D spillovers. European Economic Review 39(5), 859–887.

Coelli, T., Rao, D. P., Battese, G. E., 2005. An Introduction to Efficiency Analysis, 2nd Edition. Springer, New York.

Cruz-Cázaresa, 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., Zloczysti, 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.

Davutyan, N., 1989. Bank failures as poisson variates. Economics Letters 29(4), 333–338.

Dionne, G., Artis, M., Guillen, M., 1996. Count data models for a credit scoring system. Journal of Empirical Finance 3(3), 303–325.

Easterly, W., Levine, R., 2001. It's not factor accumulation: Stylized facts and growth models. The World Bank Economic Review 15(2), 177–219.

Fè-Rodríguez, E., 2007. Exploring a stochastic frontier model when the dependent variable is a count. University of Manchester, School of Economics Discussion Paper Series No. 0725.

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.

Fu, X., Yang, Q., 2009. Exploring the cross-country gap in patenting: A stochastic frontier approach. Research Policy 38(7), 1203–1213.

Garmaise, M. J., 2009. Ties that truly bind: Non-competition agreements, executive compensation and firm investment. Journal of Law, Economics, and Organization 27(2), 376–425.

Geweke, J., 2007. Interpretation and inference in mixture models: Simple mcmc works. Computational Statistics & Data Analysis 51(7), 3529–3550.

Greene, W. H., 2002a. Alternative panel data estimators for stochastic frontier models. Mimeo (available at http://pages.stern.nyu.edu/ wgreene/).

Greene, W. H., 2002b. Econometric Modeling Guide. New York: Econometric Software, Inc.

Greene, W. H., 2002c. Econometric Modeling Guide. New York: Econometric Software, Inc.

Greene, W. H., 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126(2), 269–303.

Griffith, R., Redding, S., van Reenen, J., 2004. 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.

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. H., Ziedonis, R. H., 2001. The patent paradox revisited: An empirical study of patenting in the u.s. semiconductor industry, 1979-1995. RAND Journal of Economics 32(1), 101–128.

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.

Hofler, R., Scrogin, D., 2008. A count data stochastic frontier. Discussion Paper, niversity of Central Florida.

Jaffe, A. B., 1986. Technology opportunity and spillovers of r&d: Evidence from firms' patents, profits, and market value. American Economic Review 76(5), 984–1001.

Jondrow, J., Lovell, C. K., Materov, I., Schmidt, P., 1982. On the estimation of technical inefficiency in the stochastic frontier production function models. Journal of Econometrics 19(2-3), 233–238.

Jones, C. I., 1995. Time series test of endogenous growth models. Quarterly Journal of Economics 110(2), 495–525.
Jones, C. I., 2005. The shape of production functions and the direction of technical change. Quarterly Journal of Economics 120(2), 517–549.
Kejak, M., 2003. Stages of growth in economic development. Journal of Economic Dynamics and Control 27 (5), 771–800.
Koop, G., 2001. Cross-sectoral patterns of efficiency and technical change in manufacturing. International Economic Review 42(1), 73–103.
Kumbhakar, S. C., Lovell, K. C., 2000. Stochastic Frontier Analysis. Cambridge: Cambridge University Press.
Kumbhakar, S. C., Parmeter, C., Tsionas, E., 2013. A zero inefficiency stochastic frontier model. Journal of Econometrics 172(1), 66–76.
Malsberger, B., 2004. Covenants Not to Compete: A State-by-State Survey. A Bloomberg BNA Publication, Washington, D.C.
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 Internationa Economics 76, 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.
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.
Orea, L., Kumbhakar, S. C., 2004. Efficiency measurement using a latent class stochastic frontier model. Empirical Economics 29(1), 169–183.
Pakes, A., Griliches, Z., 1984. Patents and R&D at the firm level: A first look, in Z. Griliches (ed): R&D, Patents and Productivity, University of Chicago Press, Chicago and London. University of Chicago Press, Chicago and London.
Palazzi, P., 2011. Taxation and innovation. OECD Taxation Working Papers, No. 9, OECD, Paris.Http://dx.doi.org/10.1787/5kg3h0sf1336-en.
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 nd Route 128. Harvard University Press, Cambridge Mass.
Sharma, S., Thomas, V., 2008. Inter-country r&d efficiency analysis: An application of data envelopment analysis. Scientometrics 76(3), 483–501.
Sickles, R., Qian, J., 2009. Stochastic frontiers with bounded inefficiency. Mimeo, Rice University.
Solow, R. M., 1957. Technical change and the aggregate production function. Review of Economics and Statistics 39(3), 312–320.
Terza, J., Basu, A., Rathouz, P., 2008. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. Journal of Health Economics 27 (3), 531–543.
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.
Tsionas, E., Kumbhakar, S., 2004. Markov switching stochastic frontier model. Econometrics Journal 7(2), 398–425.
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.
Wang, P., Cockburn, I., Puterman, M., 1998. Analysis of patent data: A mixed-poisson-regression-model approach. Journal of Business & Economic Statistics 16, 27–41.
Wilson, D. J., 2009. Beggar thy neighbor? the in-state, out-of-state, and aggregate effects of r&d tax credits. Review of Economics and Statistics 91(2), 431–436.
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.
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

Table A.1: Summary Statistics and States' Allocation per Innovation Classes

State	Patents mean	Patents st. dev.	R&D stock mean	R&D stock st. dev.	Scientists mean	Scientists st. dev.	R&D tax credit mean	R&D tax credit st. dev.	Corporate tax mean	Corporate tax st. dev.	Personal tax mean	Personal tax st. dev.	Noncompetes mean	Noncompetes st. dev.	**Class A**	**Class B**
AK	35.64	13.65	827.91	190.43	1295.64	88.33	0.00	0.00	9.40	0.00	0.00	0.00	3.00	0.00	14	0
AL	251.50	63.21	10232.51	562.85	7105.21	1191.18	0.00	0.00	6.50	0.00	3.09	0.11	5.00	0.00	2	12
AR	139.00	43.65	1777.62	229.41	3128.36	431.41	0.00	0.00	6.50	0.00	7.24	0.08	5.00	0.00	9	5
AZ	650.21	158.10	12157.40	3984.27	7739.43	1199.92	10.21	2.94	6.90	0.00	5.26	0.64	3.00	0.00	0	14
CA	16174.07	4777.37	210655.50	33458.35	84556.93	10096.99	15.00	0.00	8.97	0.22	10.25	0.77	0.00	0.00	0	14
CO	987.29	214.68	17906.91	3067.81	13117.21	1664.53	0.00	0.00	4.80	0.20	4.94	0.19	2.00	0.00	0	14
CT	2294.64	268.80	19884.92	6064.61	10257.21	1146.03	6.00	0.00	8.61	1.21	4.64	0.23	3.00	0.00	0	14
DE	1994.71	245.34	7207.65	698.14	3917.57	435.27	5.00	5.19	8.70	0.00	6.77	0.76	6.00	0.00	0	14
FL	1917.64	515.61	21802.29	2361.26	17484.29	2119.59	0.00	0.00	5.50	0.00	0.00	0.00	8.43	0.94	0	14
GA	833.71	173.30	11361.72	3128.72	12178.50	1714.06	6.43	4.97	6.00	0.00	5.83	0.02	5.00	0.00	0	14
HI	60.64	18.23	1934.93	135.79	2794.07	275.21	10.00	10.38	6.40	0.00	8.57	0.73	3.00	0.00	8	6
IA	494.07	130.32	5437.08	644.61	4919.21	292.95	6.50	0.00	10.86	1.03	5.99	0.28	6.00	0.00	0	14
ID	1183.86	737.02	4208.47	1308.76	2499.86	359.05	2.14	2.57	7.86	0.20	8.10	0.20	6.00	0.00	0	14
IL	4358.00	579.03	42800.89	6765.41	23691.57	1448.58	6.50	0.00	5.34	1.06	3.00	0.00	5.00	0.00	0	14
IN	958.43	219.77	15489.39	2549.40	9662.21	853.13	5.00	0.00	4.49	2.17	3.40	0.00	5.00	0.00	0	14
KS	276.57	50.19	4968.84	2394.31	4306.29	366.55	6.50	0.00	4.28	0.90	6.49	0.01	6.00	0.00	0	14
KY	311.79	82.68	3199.03	973.58	4914.43	520.27	0.00	0.00	8.16	0.33	6.18	0.02	6.00	0.00	0	14
LA	322.21	106.29	2863.49	538.01	5943.57	231.35	2.29	3.75	8.00	0.00	3.70	0.15	2.57	1.99	2	12
MA	3043.21	572.64	57521.14	8130.20	29144.21	3809.21	10.00	0.00	8.66	0.55	5.69	0.31	6.00	0.00	0	14
MD	854.71	208.22	41447.00	4840.35	25256.86	3298.45	5.00	5.19	7.00	0.00	5.08	0.43	5.00	0.00	0	14
ME	101.50	26.10	1032.15	427.52	2458.71	131.69	3.93	2.13	8.93	0.00	8.74	0.02	4.00	0.00	14	0
MI	3567.36	583.01	69229.04	9629.45	17645.43	1536.94	0.00	0.00	2.13	0.18	4.24	0.23	5.00	0.00	0	14
MN	2267.00	376.67	18438.57	3819.77	11415.07	1363.55	2.50	0.00	9.80	0.00	8.38	0.34	5.00	0.00	0	14
MO	777.79	143.90	10705.79	1360.25	9797.79	534.37	6.04	1.74	6.27	0.07	5.90	0.62	7.00	0.00	0	14
MS	105.79	33.82	2160.86	760.69	3375.64	227.73	0.00	0.00	5.00	0.00	5.07	0.01	4.00	0.00	14	0
MT	98.71	20.60	691.84	251.87	1979.57	188.29	2.86	2.57	6.75	0.00	7.25	0.20	2.00	0.00	14	0
NC	1087.50	298.51	19402.82	5125.77	17308.93	2469.09	3.93	2.13	7.23	0.38	8.20	0.26	4.00	0.00	0	14
ND	45.50	14.48	716.71	357.47	1629.43	462.49	4.00	0.00	10.00	1.27	5.41	0.02	0.00	0.00	14	0
NE	161.36	31.47	1918.63	534.77	2969.93	113.23	0.21	0.80	7.81	0.00	6.98	0.13	4.00	0.00	8	6
NH	330.00	63.92	4000.73	1719.61	2658.14	364.82	0.00	0.00	7.95	0.89	0.00	0.00	2.00	0.00	0	14
NJ	4072.07	691.12	53508.25	4670.01	23378.36	1919.00	9.29	2.67	9.00	0.00	7.01	1.08	4.00	0.00	0	14
NM	219.93	66.86	16344.35	2429.10	8444.43	786.31	0.00	0.00	7.60	0.00	7.43	0.97	2.00	0.00	4	10
NV	391.43	160.98	1789.66	565.60	2106.57	325.07	0.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0	14
NY	8081.14	1326.58	63772.80	2054.90	45883.79	2404.17	0.00	0.00	8.39	0.68	7.29	0.45	3.00	0.00	0	14
OH	3028.50	549.58	35754.15	1632.93	21681.64	1757.48	1.50	2.98	8.46	0.65	7.27	0.27	5.00	0.00	0	14
OK	428.57	83.24	3077.24	317.61	4956.64	242.98	0.00	0.00	6.00	0.00	5.99	0.19	1.00	0.00	0	14
OR	590.21	96.57	7514.46	3715.70	8121.00	1125.80	5.00	0.00	6.60	0.00	9.07	0.02	6.00	0.00	0	14
PA	2571.43	349.99	45865.43	2184.12	28078.14	2477.25	7.14	4.69	10.24	0.89	2.86	0.11	6.00	0.00	0	14
RI	185.71	25.86	5034.06	2133.85	2865.21	323.32	15.69	4.52	9.00	0.00	9.69	0.39	3.00	0.00	4	10
SC	387.57	75.42	4832.02	1306.03	5502.14	477.64	2.14	2.57	5.00	0.00	7.08	0.01	5.00	0.00	0	14
SD	51.79	25.14	357.06	107.10	1144.29	61.17	0.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	14	0
TN	568.71	130.03	8839.92	2402.95	9607.86	690.70	0.00	0.00	6.11	0.21	0.00	0.00	7.00	0.00	0	14
TX	4777.79	1386.10	47634.33	10747.15	34721.00	3513.42	2.14	2.57	0.32	1.20	0.00	0.00	3.29	0.73	0	14
UT	485.93	104.44	5536.01	1345.80	5325.14	451.99	3.43	3.08	5.00	0.00	6.07	0.08	6.00	0.00	0	14
VA	896.64	124.27	20974.31	6079.52	19467.71	2852.81	0.00	0.00	6.00	0.00	5.82	0.01	3.00	0.00	0	14
VT	75.57	15.46	1785.01	183.03	1913.00	203.29	2.86	4.69	9.32	0.70	9.10	0.99	5.00	0.00	12	2
WA	1546.50	516.79	36153.02	8424.19	15551.71	2205.27	0.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0	14
WI	1377.14	238.50	11873.46	2208.67	9228.29	873.28	5.00	0.00	7.90	0.00	6.82	0.09	3.00	0.00	0	14
WV	57.86	18.77	1923.62	345.04	2279.29	188.73	10.00	0.00	9.00	0.00	6.50	0.00	2.00	0.00	14	0
WY	42.57	14.92	359.22	42.49	889.93	87.05	0.00	0.00	0.00	0.00	0.00	0.00	4.00	0.00	14	0

State's two-letter abbreviation reported in first column; *R&D stock* in millions of constant (2000) US dollars; *Scientists* (science, engineering, and health researchers) are in thousands; *R&D tax credit*, *Corporate tax* and *Personal tax* are percentages (%); and *Noncompetes* range from 0 (low enforceability) to 12 (high enforceability). Classes *A* and *B* are technological regimes (classes).