ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 06/2006

Modeling Regional Heterogeneity with Panel Data: Application to Spanish Provinces

Antonio Álvarez y Julio del Corral

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm

UNIVERSIDAD DE OVIEDO

DEPARTAMENTO DE ECONOMÍA

PERMANENT SEMINAR ON EFFICIENCY AND PRODUCTIVITY

MODELLING REGIONAL HETEROGENEITY WITH PANEL DATA: WITH APPLICATION TO SPANISH PROVINCES

Antonio Álvarez♠ and Julio del Corral♣

Efficiency Series Paper 06/2006

Abstract: The estimation of aggregate production functions is common in regional economics. Regional data are usually characterized by a high level of heterogeneity among observations which are not frequently reflected in the data. If this unobserved heterogeneity, is not treated adequately can arise serious econometric problems. In this paper we review the different approaches the literature has used to deal with this problem. Moreover, in order to compare different models we estimate aggregate production functions using data from Spanish regions. First we estimate the Cornwell, Schmidt and Sickles (1990) model. Furthermore, we estimate a Battese and Coelli (1995) model. We include a time trend as an efficiency explanatory variable so the efficiency is time varying. We compare both estimates with the traditional fixed effects model.

Key words: aggregate production function, time-variant efficiency, regional unobserved heterogeneity

♠ Departamento de Economía, Universidad de Oviedo, Spain.

♣ Corresponding author. Departamento de Economía, Universidad de Oviedo, Spain.
e-mail: corraljulio@uniovi.es.

1. Introduction

The estimation of aggregate production functions is common in regional economics. Regional production functions have been used to study different topics, including the evolution of productivity, the influence of infrastructures or human capital on the private sector and the existence of catching-up to the technological frontier, among many others. The traditional framework is to estimate a Cobb-Douglas aggregate production function using as explanatory variables private capital and labor. Likewise other inputs are often also included such as human capital (e.g., de la Fuente, 1995) or public capital (e.g., Puig-Junoy, 2001; Munnell, 1990). Some studies used control variables. For instance, Evans and Karras (1994) used the composition of public capital, García-Milà and McGuire (1992), Munnel (1990) used the business cycle, Álvarez, Arias and Orea (2006) used a specialization index.

Regional data are usually characterized by a high level of heterogeneity among observations since regions have usually important differences (i.e., climate, orography, natural resources…). The existence of regional heterogeneity, which can be unobservable for the analyst, can lead to biased estimates. Hence it is important to account for the heterogeneity embedded in regional data. To mitigate this problem two kinds of models can be used. First, unobserved heterogeneity can be modelled as an individual effect. The individual effects can be time variant or time invariant. Alternatively, models that estimate different technologies in the sample (i.e., random parameters models, latent class models, kernel estimation) can also be used.

In this paper we review some of the different models used in the literature to mitigate unobserved regional heterogeneity. Especially we focus on models where the unobserved heterogeneity is modelled through individual effects. In the empirical sections we compare different models using Spanish regional data from 1986 to 2003.

The rest of the paper is organized as follows. Next we review some alternatives to account for unobserved regional heterogeneity. After that, we describe the methodologies used in the empirical analysis. The next section contains the data and the empirical models. Then the results are discussed and finally some conclusions are drawn.

2. Unobserved heterogeneity in regional production functions

In most empirical situations there are differences across observations that are not reflected in the data. This information is referred to as unobserved heterogeneity. When this information is not important, it can be accommodated in the error term. However, if these differences are important and the unobserved heterogeneity is correlated with the explanatory variables, the estimated parameters will be biased (Griliches, 1957). A frequent case occurs when the information not included in the model can be considered as invariant over time, e.g., location or orography. If panel data are available, the solution to this problem is to model the heterogeneity as an individual effect (see Mundlak, 1961). Several papers have used the Within or random effects estimators in order to estimate aggregate production functions using regional data. For instance, García-Milà, McGuire and Porter (1996), Evans and Karras (1994) or Holtz-Eakin (1994) used data from the U.S. states. Mas et al. (1996), Moreno et al. (1997) used data from Spanish regions.

However, traditional panel data techniques assume that individual effects are time invariant. Schmidt and Sickles (1984) indicated that this assumption is less valid as the panel becomes longer. For instance, some regional characteristics, such as its economic structure, abundance of natural resources[1] or productive efficiency, may vary over time. Therefore models that are able to control not only the time invariant unobserved heterogeneity but also the time variant unobserved heterogeneity should be used. Cornwell, Schmidt and Sickles (1990) developed an extension for the traditional fixed effects in which individual effects are allowed to vary over time. Wu (1995) used this model to study total factor productivity growth, technological progress and technical efficiency change in postreform China.

Stochastic frontier models[2] (Aigner, Lovell and Schmidt, 1977) are other approach capable of modelling time-varying unobserved heterogeneity. The main characteristic of these models is that they have a composed error term. Specifically, an asymmetric error term is added to the traditional symmetric error term. The former is supposed to represent technical inefficiency[3] while the latter is supposed to engage random shocks

[1] The natural resources are more or less invariant but the exploitation can be very different.

[2] See Kumbhakar and Lovell (2000) for an extensive survey and Dorfman and Koop (2005) for an overview of recent developments.

[3] Technical efficiency is the ability to obtain maximum output from a given input vector (Kumbhakar and Lovell, 2000; p. 42).

and measurement errors. There are several papers that have used stochastic frontier models to estimate aggregate production functions. For instance Puig-Junoy (2001) estimated technical efficiency indexes for the 48 contiguous U.S. states. More recently Mastromarco and Woitek (2006) studied the link between public infrastructure investment and efficiency in the Italian regions. Delgado and Álvarez (2004) and Arias and Rodriguez-Vález (2004) estimated aggregate production functions using Spanish datasets. Moreover, Greene (2005) proposed a "true fixed effects model" which is able to estimate not only individual effects but also an inefficiency term. This model was used recently by Álvarez (2007) in order to decompose the productivity growth of Spanish regions.

3. Time-Varying individual effects models

Our basic framework is an aggregate production function, such as the following:

$$Y = Af\left(K, L, G, H\right) \tag{1}$$

where *Y* is aggregate production, *A* is a technical efficiency index which reflects the state of technology and omitted factors, *K* is private capital, *L* is labor, *G* is a measure of public capital and *H* is a human capital index. We estimate two models that consider *A* as time-varying. First, we estimate the model developed by Cornwell, Schmidt and Sickles (1990) where individual effects are time variant. Then we use the Battese and Coelli (1995) model in which a time trend is used as an efficiency explanatory variable.

a) Cornwell, Schmidt and Sickles (1990), (CSS)

This model is an extension of the traditional fixed effects model in which the intercept varies across individuals. In particular in this model not only does the intercept vary across individuals but also over time. The model can be written as follows:

$$y_{it} = \beta' x_{it} + \delta_i' w_{it} + v_{it} \quad ; \qquad w_{it} = \left[1, t, t^2\right] \tag{2}$$

where *y* represents the output, *x* is a M-dimensional vector of explanatory variables, *w* is a J-dimensional vector that contains a constant, a time trend and a squared time trend, *β* and *δ* are matrices of coefficients, *v* is a typical random disturbance. Subscript *i* indicates region while subscript *t* denotes time. The important feature in this model is that w_{it} has specific coefficients for each region[4]. Thus individual effects are allowed to vary over time.

Cornwell, Schmidt and Sickles (1990) developed three alternative ways to estimate model (2). In particular, we are going to use a generalization of the within estimator[5]. This within estimator is based on some results for the partitioned regression where the estimation is done in two steps. In the first step the coefficients of the x vector are obtained in the following way:

$$\hat{\beta} = \left(x'Mx\right)^{-1}\left(x'My\right) \tag{3}$$

where matrix M is defined as: $M = \left(I - w\left(w'w\right)^{-1} w'\right)$ in which I is an identity matrix. The above expression can be viewed as a residual maker (Greene, 2003; p. 24). Then, once the coefficients of the x vector have been estimated, the coefficients δ_i can be estimated regressing the residuals obtained in the first step on w_{it} for each region $\left(\hat{\delta}_i = \left(ww'\right)^{-1} w'\hat{u}_{it}\right)$, where $\hat{u}_{it}$ are the residuals from the estimated coefficients in (3).

Consequently the individual effects are calculated as:

$$\hat{\alpha}_{it} = w'_{it}\hat{\delta}_i \tag{4}$$

b) Battese and Coelli (1995), (BC95)

In this model technical inefficiency effects are assumed to be a function of some exogenous variables. Hence, the model can be expressed in the following way:

$$y_{it} = \beta' x_{it} + v_{it} - u_{it} \quad ; \qquad u_{it} = z_{it}\delta + W_{it} \tag{5}$$

where u is a random term which reflects the technical inefficiency of the regions, while v is a typical random disturbance term which is assumed to follow a normal distribution centered at zero with σ_v standard deviation, z are the explanatory variables associated with technical inefficiency and W_{it} is defined by a normal distribution truncated at $-z_{it}\delta$

[4] Actually, if w_{it} just contains a constant, then the model in (2) is just the standard fixed effects model where the intercept only varies across individuals (regions).

[5] We have also estimated the GLS estimator proposed in Cornwell, Schmidt and Sickles (1990), but the hypothesis that the effects are not correlated with the variables was rejected.

with zero mean and variance σ^2. It is also assumed that v and u are independent. The parameters of the stochastic frontier and the model for the technical inefficiency effects are estimated simultaneously by maximum likelihood.

4. Data and empirical model

The data consists of 18 annual observations for the 50 provinces of Spain over the period 1986-2003. Spain is a country where there is a huge heterogeneity among provinces due to different weather conditions, orography and geographical location, among many other sources of heterogeneity. Hence, these data seem appropriate to compare different ways to control unobserved heterogeneity.

We use Value Added in thousands of euros of 1986 as the measure of output. Four inputs are included: private capital, labor, road infrastructures and a human capital index. The source of private capital is Mas et al. (2005). Since this variable is only available till 2000 we have extrapolated the data for the years 2001, 2002 and 2003[6]. In order to consider productive capital, we subtracted residential capital. Labor, measured as number of workers, has been taken from *Instituto Valenciano de Investigaciones Economicas* (IVIE)[7]. Road infrastructure is measured as the number of kilometers of highways (Source: *Ministerio de Fomento*). The human capital index has been constructed as the number of skilled workers (i.e., superior studies and preceding to superior studies) from IVIE. Furthermore, in order to account for the changing economic structure we use the number of agricultural workers as a control variable (Source: IVIE)[8].

We use the Cobb-Douglas functional form which is the most frequently used in the literature. We also include Hicks neutral technical change. Therefore, the models to be estimated are:

a) Fixed effects model

[6] It does not seem a major drawback since the changes in private capital are highly stable.

[7] "Capital Humano en España y su distribución provincial (1964-2004)".

[8] Note that it is empirically analogous to include the number of agricultural workers or the percentage of agricultural employment. The only difference is that in the former case, the labor coefficient is itself the labor coefficient. In contrast, in the latter case in order to obtain the labor coefficient it should be added the coefficient of the percentage of agricultural employment to the labor coefficient obtained in the regression. The same applies to the human capital index.

$$ln\ y_{it} = \beta_i + \sum_{m=1}^{M} \beta_m\ ln\ x_{imt} + \beta_z\ ln\ z_{it} + \lambda_t \cdot t + v_{it} \qquad (6)$$

where *y* is the output, *x* are the inputs, *z* is the control variable, *t* is a time trend and *v* is random noise.

b) Cornwell, Schmidt and Sickles (1990)

$$ln\ y_{it} = \beta_{it} + \sum_{m=1}^{M} \beta_m\ ln\ x_{imt} + \beta_z\ ln\ z_{it} + v_{it} \qquad (7)$$

c) Battese and Coelli (1995)

$$ln\ y_{it} = \beta + \sum_{m=1}^{M} \beta_m\ ln\ x_{imt} + \beta_z\ ln\ z_{it} + \lambda_t \cdot t + v_{it} - u_{it}; \qquad u_{it} = \delta + \delta_t \cdot t + W_{it} \qquad (8)$$

where *u* is the inefficiency term.

5. Results

The model in (6) was estimated by OLS using the Within estimator (see Greene, 2003; p. 234), while models (7) and (8) were estimated using the estimators described above[9]. First we show the estimation of the traditional fixed effects model.

Table 1. Estimation of the fixed effects model

Variable	Coefficients
Private capital	0.306***
Labor	0.300***
Infrastructures	0.009***
Human capital	-0.007
Agricultural employment	-0.048***
Time trend	0.010***
R^2	0.99

* , **, *** indicate significance at the 10%, 5% and 1% level, respectively.

We can see that the estimated coefficient for labor is smaller than expected, since this parameter is supposed to be equal to the labor share. In turn, this implies that the scale elasticity is very low; so this model is likely to face some problems. On the other hand,

[9] Models (6) and (8) were estimated using Limdep 8.0. while model (7) was estimated using Matlab 6.0.

infrastructures have a positive and significant estimated coefficient. However, its contribution to productivity growth is rather low but is similar to other elasticities obtained using similar data (Álvarez, Arias and Orea, 2006). The estimated human capital coefficient is negative but not significant. As expected, the control variable has a negative and significant coefficient. Finally, there is technical change.

In Table 2 we present the estimation of equation (7).

Table 2. CSS estimation

Variable	Coefficients
Private capital	0.273***
Labor	0.493***
Infrastructures	-0.004
Human capital	-0.038***
Agricultural employment	-0.033***
R^2	0.99

*, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

In this model not only the scale elasticity is higher than in the fixed effects model but also the output elasticity of labor seems to be closer to the input share. In addition, the results show that in this model highway infrastructures do not appear to have any impact on the output. Moreover the results suggest that there is a negative impact of human capital on productivity.

Finally, Table 3 presents the results for the Battese and Coelli (2005) model.

Table 3. Battese and Coelli (1995) estimation

Variable	Coefficients
Constant	6.244***
Private capital	0.389***
Labor	0.503***
Infrastructures	0.033***
Human capital	0.143***
Agricultural employment	-0.070***
Time trend	-0.014***
Squared time trend	0.000
Inefficiency variables	
Constant	0.161***
Time Trend	-0.016***
$\lambda = \sigma_u / \sigma_v$	2.14

$\sigma = [\sigma_v^2 + \sigma_u^2]^{1/2}$	0.111
Likelihood Function	1,035

*, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

The main difference between this model and both fixed effects and the CSS models is that the coefficients of human capital and infrastructures are positive and significant. With respect to the inefficiency term, we can see that the time trend has a negative and significant effect. This means that provinces reduce their inefficiency over time. In contrast, there is negative technological change. It seems that this model is not able to separate the effect of technical change from efficiency change.

Next we are going to analyze technical efficiency. The procedure to calculate technical efficiency is different in the Battese and Coelli model than in the fixed effects models. The difference is that in the fixed effects models the technical efficiency indexes are obtained comparing each individual effect with the largest individual effect. In particular, the technical efficiency indexes in the fixed effect model are calculated as suggested by Schmidt and Sickles (1984):

$$TEFE_i = exp\left(\alpha_i - max(\alpha_i)\right) \quad (9)$$

where *α* are the individual effects. However, in the CSS model there is not only variation across individual but also across time. Hence the technical efficiency indexes are calculated as:

$$TECSS_{it} = exp\left(\alpha_{it} - max(\alpha_{it})\right) \quad (10)$$

On the other hand, the BC95 model calculates the technical efficiency index using the following expression:

$$TEBC_{it} = exp(-\hat{u}_{it}) \quad (11)$$

The average technical efficiency of each province is presented in Table 4.

Table 4. Average technical efficiency of the Spanish provinces

Province	Fixed effects	CSS	BC95
Álava	0.27	0.40	0.92
Albacete	0.22	0.30	0.83
Alicante	0.43	0.50	0.88
Almería	0.30	0.40	0.96
Asturias	0.40	0.48	0.89

Ávila	0.18	0.27	0.90
Badajoz	0.27	0.34	0.85
Baleares	0.38	0.48	0.92
Barcelona	0.84	0.84	0.86
Burgos	0.28	0.39	0.92
Cáceres	0.22	0.30	0.81
Cádiz	0.40	0.49	0.95
Cantabria	0.30	0.40	0.89
Castellón	0.31	0.41	0.93
Ciudad Real	0.26	0.35	0.85
Córdoba	0.32	0.41	0.93
Coruña	0.39	0.46	0.88
Cuenca	0.17	0.26	0.87
Gerona	0.35	0.45	0.93
Granada	0.34	0.44	0.92
Guadalajara	0.17	0.28	0.83
Guipúzcoa	0.39	0.50	0.92
Huelva	0.28	0.38	0.96
Huesca	0.21	0.31	0.89
Jaén	0.31	0.40	0.95
Las Palmas	0.37	0.47	0.92
León	0.27	0.36	0.84
Lérida	0.29	0.40	0.94
Lugo	0.22	0.28	0.87
Madrid	1.00	1.00	0.93
Málaga	0.41	0.49	0.91
Murcia	0.42	0.50	0.94
Navarra	0.36	0.48	0.94
Orense	0.21	0.27	0.81
Palencia	0.18	0.29	0.87
Pontevedra	0.34	0.40	0.87
Rioja	0.25	0.37	0.94
Salamanca	0.23	0.33	0.87
Segovia	0.18	0.28	0.89
Sevilla	0.49	0.57	0.93
Soria	0.15	0.25	0.89
Sta. Cruz de Tenerife	0.37	0.47	0.92
Tarragona	0.37	0.49	0.95
Teruel	0.17	0.26	0.93
Toledo	0.27	0.35	0.87
Valencia	0.58	0.64	0.91
Valladolid	0.31	0.43	0.92
Vizcaya	0.45	0.56	0.86
Zamora	0.18	0.27	0.89
Zaragoza	0.41	0.51	0.90
Total	0.33	0.42	0.90

Mean technical efficiency is 0.33 in the Fixed Effects, 0.42 in the CSS model and 0.90 in the BC95 model. Thus, there are rough differences across models in the technical efficiency indexes. The biggest difference arises between the BC95 and the fixed effects models. As it was explained above the technical efficiency indexes are calculated in a very different way. Actually the correlation coefficient between technical efficiency indexes of the fixed effect model and the CSS model is 0.99 while the correlation coefficient between the BC95 and the fixed effects model and CSS model are respectively 0.23 and 0.31. These huge differences arise not only because of the different way in which technical efficiency is calculated but also from the different treatment of unobserved heterogeneity.

Even though the main purpose of the paper is to study the different treatment to deal with unobserved heterogeneity, the most important feature of the CSS and BC95 models is that the efficiency is allowed to vary over time. Hence, it is possible to know whether the provinces improved their efficiency levels. Table 5 presents the average of the technical efficiency for every year in the CSS model and in the Battese and Coelli model.

Table 5. Average technical efficiencies for the years

Year	CSS	BC95
1986	0.43	0.86
1987	0.43	0.86
1988	0.43	0.88
1989	0.43	0.88
1990	0.43	0.88
1991	0.42	0.88
1992	0.42	0.88
1993	0.42	0.89
1994	0.42	0.90
1995	0.42	0.91
1996	0.42	0.91
1997	0.42	0.92
1998	0.41	0.92
1999	0.41	0.91
2000	0.41	0.91
2001	0.41	0.92
2002	0.41	0.93
2003	0.40	0.92

We can see in the CSS model that there is a decreasing trend in the technical efficiency. On the contrary, in the Battese and Coelli model the technical efficiency is increasing.

It is also interesting to know how technical efficiency has evolved over time in the provinces. Hence, the next table presents the efficiency changes in the Spanish provinces through time in both models. Concretely we show the efficiency for each province in 1986, the difference between the technical efficiency 2003 and the efficiency in 1986, and the ratio between the improvement in the efficiency in these two years and 1986.

Table 6. Changes in the efficiency

	CSS			BC95		
Province	TE_0	DTE	GRTE	TE_0	DTE	GRTE
Álava	0.44	-0.05	-0.11	0.94	-0.02	-0.02
Albacete	0.29	0.00	0.01	0.78	0.05	0.07
Alicante	0.53	-0.06	-0.12	0.87	0.00	0.00
Almería	0.39	0.00	-0.01	0.95	0.01	0.01
Asturias	0.51	-0.05	-0.11	0.84	0.08	0.10
Ávila	0.26	-0.01	-0.03	0.81	0.14	0.17
Badajoz	0.34	0.00	0.01	0.79	0.08	0.10
Baleares	0.50	-0.08	-0.16	0.90	0.01	0.01
Barcelona	0.85	-0.05	-0.06	0.80	0.10	0.13
Burgos	0.41	-0.03	-0.07	0.88	0.04	0.05
Cáceres	0.31	-0.01	-0.04	0.78	0.09	0.11
Cádiz	0.54	-0.07	-0.13	0.94	0.02	0.02
Cantabria	0.40	-0.02	-0.05	0.79	0.13	0.17
Castellón	0.44	-0.05	-0.12	0.94	0.00	0.00
Ciudad Real	0.34	0.01	0.02	0.78	0.14	0.17
Córdoba	0.39	0.00	0.00	0.84	0.11	0.13
Coruña	0.43	0.02	0.05	0.78	0.12	0.15
Cuenca	0.24	0.02	0.09	0.81	0.11	0.14
Gerona	0.48	-0.05	-0.11	0.90	0.06	0.06
Granada	0.43	0.01	0.02	0.85	0.10	0.12
Guadalajara	0.30	-0.03	-0.11	0.78	0.15	0.20
Guipúzcoa	0.54	-0.07	-0.13	0.87	0.04	0.04
Huelva	0.44	-0.06	-0.14	0.96	0.01	0.02
Huesca	0.33	-0.03	-0.10	0.85	0.09	0.10
Jaén	0.40	-0.04	-0.10	0.97	-0.03	-0.03
Las Palmas	0.49	-0.04	-0.08	0.91	0.01	0.01
León	0.34	0.06	0.17	0.78	0.15	0.19
Lérida	0.42	-0.05	-0.11	0.88	0.10	0.11
Lugo	0.26	0.02	0.09	0.78	0.13	0.17
Madrid	1.00	0.00	0.00	0.87	0.09	0.11
Málaga	0.52	-0.03	-0.06	0.90	0.05	0.05
Murcia	0.55	-0.10	-0.17	0.96	-0.05	-0.05
Navarra	0.51	-0.04	-0.07	0.90	0.06	0.07
Orense	0.24	0.05	0.22	0.78	0.09	0.11

Palencia	0.29	0.00	-0.01	0.78	0.16	0.20
Pontevedra	0.40	-0.01	-0.01	0.83	0.04	0.04
Rioja	0.39	-0.04	-0.09	0.92	0.04	0.04
Salamanca	0.32	0.00	0.00	0.78	0.15	0.20
Segovia	0.27	0.04	0.16	0.80	0.17	0.22
Sevilla	0.57	-0.02	-0.04	0.87	0.06	0.07
Soria	0.26	-0.02	-0.07	0.83	0.12	0.14
Sta. Cruz de Tenerife	0.53	-0.13	-0.24	0.93	-0.03	-0.04
Tarragona	0.56	-0.12	-0.21	0.95	0.01	0.01
Teruel	0.26	-0.01	-0.04	0.82	0.13	0.16
Toledo	0.35	-0.01	-0.04	0.78	0.11	0.15
Valencia	0.69	-0.10	-0.14	0.90	-0.01	-0.01
Valladolid	0.44	-0.04	-0.08	0.93	-0.03	-0.03
Vizcaya	0.57	0.00	-0.01	0.81	0.14	0.17
Zamora	0.27	0.00	0.00	0.82	0.10	0.12
Zaragoza	0.54	-0.07	-0.13	0.85	0.05	0.06

TE_0- Technical efficiency in 1986.
DTE- Difference between technical efficiency in 2003 and technical efficiency in 1986.
GRTE- Ratio between DTE and TE_0.

The results are similar in both models since the correlation between the differences is 0.71 and the correlation in the growth rates is 0.66.

6. Conclusions

In this paper we have reviewed the different approaches put forward in the literature to mitigate the problem that can arise from regional unobserved heterogeneity. Subsequently, we have estimated an aggregate production function using data from Spanish regions from 1986 to 2003 using two different models. First, we have estimated the model suggested by Cornwell, Schmidt and Sickles (1990). On the other hand, we have estimated a Battese and Coelli (1995) model, including a time trend as an exogenous variable for the efficiency term. The results show rough differences among models not only in estimated coefficients but also in technical efficiency indexes.

Moreover, the empirical results obtained show that there is not a strong link between the public capital, measured as the kilometres of highways, and human capital and productivity.

References

Aigner, D., K. Lovell and P. Schmidt, (1977), “Formulation and Estimation of Stochastic Frontier Production Function Models”, *Journal of Econometrics* 6, 21-37.

Álvarez, A., (2007), “Decomposing Regional Productivity Growth using an Aggregate Production Frontier”, *Annals of Regional Science*, forthcoming.

Álvarez, A., C. Arias and L. Orea, (2006), “Econometric Testing of Spatial Spillovers from Public Capital”, *Hacienda Pública Española*, 178, 9-21.

Arias, C., and J. Rodríguez-Vález, (2004), “Spatial Spillover of Infrastructure: An Application with Stochastic Frontiers”, *Estudios de Economía Aplicada*, 22, 657-673.

Battese, G. and T. Coelli, (1995), “A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data”, *Empirical Economics*, 20, 325-332.

Cornwell, C., P. Schmidt and R. Sickles, (1990), “Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels”, *Journal of Econometrics*, 46, 185-200.

Delgado, M.J. and I. Álvarez, (2004), “Infraestructuras y eficiencia técnica: un análisis a partir de técnicas frontera”, *Revista de Economía Aplicada*, 35, 65-82.

de la Fuente, A., (1995), “Infrastructure and Education as Instruments of Regional Policy: Evidence from Spain”, *Economic Policy*, 10, 11-51.

Dorfman, J. and G. Koop, (2005), “Current Developments in productivity and efficiency measurement”, *Journal of Econometrics*, 126, 233-240.

Evans, P. and G. Karras, (1994), “Are government activities productive? Evidence from a Panel of U.S. States”, *The Review of Economics and Statistics*, 76 (1), 1-11.

Garcia-Milà, T. and T. McGuire, (1992), “The Contribution of Publicly Provided Inputs to States’ Economies”, *Regional Science and Urban Economics*, 22, 229-241.

Garcia-Milà, T., T. McGuire, and H. Porter, (1996), “The Effect of Public Capital in State-Level Production Functions Reconsidered”, *The Review of Economics and Statistics*, 78(1), 177-180.

Greene, W. (2003). *Econometric Analysis*. Fith Edition, New Jersey, Prentice Hall.

Greene, W., (2005), “Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model”, *Journal of Econometrics*, 126, 269-303.

Griliches, Z., 1957, “Specification Bias in Estimates of Production Function”*, Journal of Farm Economics*, 39, 8-20.

Holtz-Eakin, D., (1994), “Public-Sector Capital and the Productivity Puzzle”, *The Review of Economics and Statistics*, 76 (1), 12-21.

Kumbhakar, S. and K. Lovell, (2000), *Stochastic Frontier Analysis*. Cambridge University Press.

Mas, M., J. Maudos, F. Pérez, and E. Uriel, (1996), "Infrastructures and Productivity in the Spanish Regions", *Regional Studies,* 30(7), 641-649.

Mas, M., F. Pérez and E. Uriel, (2005), "El stock de capital y su distribución territorial (1964-2002)", *Fundación BBVA*.

Mastromarco, C. and U. Woitek, (2006), "Public Infrastructure Investment and Efficiency in Italian Regions", *Journal of Productivity Analysis*, 25, 57-95.

Moreno, R., M. Artís, E. López-Bazo and J. Suriñach, (1997), "Evidence on the complex link between infrastructure and regional growth", *International Journal of Development Planning Literature,* 20, 81-108.

Mundlak, Y., (1961), "Empirical Production Function Free of Management Bias", *Journal of Farm Economics,* 43, 44-56.

Munnell, A., (1990), "How Does Public Infrastructure Affect Regional Economic Performance", *New England Economic Review,* Federal Reserve Bank of Boston, sept-oct, 3-22.

Puig-Junoy, J., (2001), "Technical Inefficiency and Public Capital in U.S. States: A Stochastic Frontier Approach", *Journal of Regional Science*, 41(1), 75-96.

Schmidt, P. and R. Sickles, (1984), "Production Frontiers and Panel Data", *Journal of Business and Economic Statistics*, 2, 367-374.

Wu, Y., (1995), "Productivity Growth, Technological Progress, and Technical Efficiency Change in China: a Three Sector Analysis", *Journal of Comparative Economics*, 21, 207-229.