Estimating individual effects and their spatial spillovers in linear panel data models*

K. Miranda, O. Martínez-Ibañez, and M. Manjón-Antolín

QURE-CREIP Department of Economics, Rovira i Virgili University

Abstract

Individual-specific effects and their spatial spillovers are generally not identified in linear panel data models. In this paper we present identification conditions under the assumption that covariates are correlated with the individual-specific effects. We also derive appropriate GLS and IV estimators for the resulting correlated random effects spatial panel data model with strictly-exogenous and predetermined explanatory variables, respectively. Lastly, we illustrate the proposed estimators using a Cobb-Douglas production function specification and US state-level data from Munnell (1990). As in previous studies, we find no evidence of public capital spillovers. However, the public capital does play a role in the positive spatial contagion of the nevertheless negative spillovers that states produce in and receive from their neighbours.

Keywords: correlated random effects, spatial panel data

JEL Classification: C23

*We thank seminar participants at the 4^{th} WIPE workshop, the 40^{th} *Simposio de Análisis Económico*, the 8^{th} seminar Jean Paelinck on Spatial Econometrics (Dijon-France) and the University Rovira i Virgili for helpful comments. This research was funded by grants ECO2014-55553-P (Ministerio de Economía y Competitividad) and 2014FI_B00301 (Agaur, Generalitat de Catalunya). Usual caveats apply.

1 Introduction

Spatial linear panel data models (see e.g. Elhorst, 2010, for an overview) essentially differ in the model component that is spatially weighted. That is, to account for the existence of spatial correlations between individuals (while controlling for their unobserved heterogeneity by including a "random" or "fixed" effect), the Spatial Lag Model for example spatially weights the dependent variable; the Spatial Error Model, the error term; and the Spatial Durbin Model both the independent and the dependent variables LeSage and Pace (2009). To date, however, proposed model specifications have generally not considered the case for the individual effects, with the notable exception of the random effects model considered by Kapoor et al. (2007). In this paper we introduce a correlated-random effects model (Mundlak, 1978; Chamberlain, 1982) that also presents spatial correlation in the individual effects.

To our knowledge, there are no analogous models in the fixed effects case. Beer and Riedl (2012), for example, advocate for using an extension of the Spatial Durbin Model for panel data that does include the spatially weighted individual effects. Ultimately, however, they argue that "it is (...) advisable to remove the spatial lag of the fixed effects from the equation as the inclusion of both, [the individual effects] and [their spatial spillovers], leads to perfect multicollinearity" (p. 302). Removing the spatial lag of the fixed effects does not generally preclude the consistent estimation of the parameters of the model. However, such practice rules out obtaining an estimate of the individual-specific effects (net of the spatially weighted effects). This is a critical issue, for example, in two-steps models that use such estimate as the dependent variable (Combes and Gobillon, 2015). Similarly, obtaining an estimate of the spatial spillovers of the individual-specific effects may be of great interest in certain applications (e.g., to asses their geographical distribution, as we do in our empirical application).

This raises the question of whether it is possible to identify both the individual effects and their spatial spillovers in linear panel data models. In this paper we provide identifying conditions in a model specification that spatially weights both the independent variables and the individual effects. In particular, we show that there is no identification problem if the covariates are correlated with the individual-specific effects (and the individual effects correspond to deviations with respect to the constant term). Thus, the specification we consider is a spatial$-(\boldsymbol{X}, \boldsymbol{\mu})$ panel data model with correlated random effects.

Our work is related to that of Debarsy (2012), who uses a correlated random effects specification to construct a LR test on "the relevance of the random effects approach" (p. 112).[1] Notice, however, that although we both deal with the correlation between individual effects and the covariates, our purposes differ markedly: while he seeks to correctly specify such correlation, we use such correlation as a means to identify the spatial contagion in the individual effects. We also differ in the model specification, which although similar, involves a different treatment of the spatial contagion of the individual effects. Debarsy (2012) assumes that the individual effects depend on both the explanatory variables and the explanatory variables in their neighbourhood, but there is no spatial contagion in the individual effects. In contrast, we account for the spatial contagion in the individual effects and assume that the individual effects and their spatial spillovers depend on the (mean of the) explanatory variables, which allows us to identify both the individual effects and their spatial spillovers. These alternative assumptions yield different error component structures: a one-way error in his case, a two-way error in ours (being the additional component a spatially weighted element, as in Kapoor et al. 2007).[2]

Having proved that the model is identified, we then consider the estimation of its parameters under alternative exogeneity assumptions on the explanatory variables. Under the assumption that all the explanatory variables are strictly exogenous (with respect to the idiosyncratic term), we derive a Feasible Generalised Least Squares (FGLS) estimator. Under the assumption that the explanatory variables are predetermined, we propose an Instrumental Variables (IV) estimator to address the endogeneity of the means of the predetermined explanatory variables employed to approximate the correlation function. We also advocate for using the backward

[1] Although this is not the aim of this paper, an analogous Wald test could be developed using our model specification and estimation procedures.

[2] Another important difference with the work of Debarsy (2012) is that whereas he analyses the Spatial Durbin Model (as Beer and Riedl 2012 do), our results are derived for the Spatial (Lag of) X model. This allows us to address the identification and estimation of the model in a linear setting, whereas considering correlated random effects in a Spatial Durbin specification would result in a non-linear model in which identification and estimation are more involved (Debarsy (2012, p. 115), for example, "assume(s) that all parameters are identified"). We leave the study of this model specification for future research.

means of these variables (i.e., the means taken, for each period, over only current and past values) as instruments.

Lastly, we illustrate these estimators using the data and (a spatially weighted variant of the) specification employed by Munnell (1990).[3] We find that, under strict exogeneity, our estimates of a Cobb-Douglas production function for the US states over the period 1970 to 1986 are largely consistent with those reported in related studies (using this data set, as e.g. Baltagi and Pinnoi 1995; Kelejian and Robinson 1997; and using analogous data sets, as e.g. Holtz-Eakin and Schwartz 1995; Garcia-Mila et al. 1996). In particular, our FGLS estimates are very close to a fixed-effects estimation of the model using the error components structure proposed by Kapoor et al. (2007), which suggests a correct specification of the correlation functions. Yet we find evidence of predeterminedness in the public capital. Under sequential exogeneity, IV estimates indicate that states with a larger/smaller estimated individual effect tend to have larger/smaller negative spatial spillovers (both inwards and outwards). Also, while the part of the individual effects associated with the private capital produces negative spatial contagion, the part associated with the public capital produces positive spatial contagion. Consistent with previous literature, however, we find no significant spatial spillovers in the public capital.

The rest of the paper is organised as follows. In Section 2 we discuss the identification problem and show that, under mild rank conditions, the correlated random effects model considered is identified. In Section 3 we present appropriate (FGLS and IV) estimators. In Section 4 we discuss how to estimate models with serial correlation and present illustrative empirical evidence based on the work of Munnell (1990). Section 5 concludes.

[3] The data set is publicly available and can be downloaded, for example, from the **Ecdat** package in R (a standardised binary contiguity spatial weights matrix of the US states is also included in the package).

2 Specification and identification of the model

2.1 The identification problem

Let us consider the spatial$-(\boldsymbol{X}, \boldsymbol{\mu})$ panel data model, i.e., the spatial (lag of) $\boldsymbol{X}$ model for panel data with spatially weighted fixed effects:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{WX\gamma} + \boldsymbol{\Psi\mu} + \boldsymbol{W\Psi\alpha} + \boldsymbol{\varepsilon}, \tag{2.1}$$

where $\boldsymbol{y} = (y_{11}, \cdots, y_{1T}, \cdots, y_{N1}, \cdots, y_{NT})'$ is the dependent variable (as usually, $i = 1, \cdots, N$ denotes cross-sectional, geographical units and $t = 1, \cdots, T$ denotes the time dimension) and $\boldsymbol{X}$ is the $NT \times K$ matrix of explanatory variables.[4] We assume that neighbourhood relations do not change over time, so that the spatial matrix is $\boldsymbol{W} = \mathbf{w} \otimes \boldsymbol{I}_T$, with $\boldsymbol{I}_T$ denoting the $T \times T$ identity matrix and $\mathbf{w} = [w_{ij}]$ being the spatial weight matrix that describes the spatial arrangement of the units in the sample. Also, unobservable individual-specific effects are collected in $\boldsymbol{\Psi\mu}$, with $\boldsymbol{\Psi} = \boldsymbol{I}_N \otimes \boldsymbol{\iota}_T$ and $\boldsymbol{I}_N$ being the $N \times N$ identity matrix and $\boldsymbol{\iota}_T$ a vector of ones of order T. Thus, $\boldsymbol{W\Psi\alpha}$ captures the spatial spillovers of these individual-specific effects.[5] Lastly, $\boldsymbol{\varepsilon}$ is a zero-mean idiosyncratic error term with assumed variance-covariance matrix $\sigma_\varepsilon^2 \boldsymbol{I}_{NT}$, being $\boldsymbol{I}_{NT}$ the $NT \times NT$ identity matrix.

The parameters of the model are $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ and σ_ε^2. This means that there are $2(K+N)+1$ parameters to be estimated. In particular, our main interest is the estimation of the individual-specific effects ($\boldsymbol{\mu}$) and their spatial spillovers ($\boldsymbol{\alpha}$). Proposition 1 shows, however, that in general these parameters are not identified.

Proposition 1. *The Spatial (Lag of) $\boldsymbol{X}$ model for panel data with spatially weighted fixed effects is not identified for any spatial weight matrix* $\mathbf{w}$.

[4]Through the paper, we assume that a balanced (complete) panel data is available. However, results can easily be extended to incomplete panels.

[5]To be precise, α_j is the spillover "on the neighbours" of unit j and $[w_{\cdot j}]\boldsymbol{\alpha}$ is the spillover "from the neighbours" on unit j, where the definition of neighbourhood is given by the structure of $\mathbf{w} = [w_{ij}]$. We illustrate this in the empirical application of section 4.

Proof. The model in (2.1) is not identified for any spatial weight matrix $\mathbf{w}$ because $\boldsymbol{\Psi}$ and $\boldsymbol{W\Psi}$ are perfectly collinear. We proof this by showing that $\det \left[\left(\boldsymbol{\Psi} \,\vdots\, \boldsymbol{W\Psi} \right)' \left(\boldsymbol{\Psi} \,\vdots\, \boldsymbol{W\Psi} \right) \right]$ is zero for any spatial weight matrix $\mathbf{w}$. Let

$$\boldsymbol{A} = \left(\boldsymbol{\Psi} \,\vdots\, \boldsymbol{W\Psi} \right)' \left(\boldsymbol{\Psi} \,\vdots\, \boldsymbol{W\Psi} \right) = T \begin{pmatrix} \boldsymbol{I}_N & \mathbf{w} \\ \mathbf{w}' & \mathbf{w}'\mathbf{w} \end{pmatrix} \tag{2.2}$$

Then, by Schur complement,

$$\det(\boldsymbol{A}) = T^{2N} \det(\boldsymbol{I}_N) \det\left(\mathbf{w}'\mathbf{w} - \mathbf{w}'(\boldsymbol{I}_N)^{-1}\mathbf{w}\right) = T^{2N} \det(\mathbf{w}'\mathbf{w} - \mathbf{w}'\mathbf{w}) = 0 \tag{2.3}$$

□

Notice that, as Beer and Riedl (2012) argue, the omission of $\boldsymbol{W\Psi\alpha}$ does not preclude the consistent estimation of the parameters of the model. Thus, if the spatial spillovers of the individual-specific effects are of no interest for the application in hand, their suggestion of removing one of the components, i.e., either $\boldsymbol{\Psi}$ or $\boldsymbol{W\Psi}$, is perfectly sensible. This is because the model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{WX\gamma} + \boldsymbol{\Psi\mu}^* + \boldsymbol{\varepsilon}, \tag{2.4}$$

with $\boldsymbol{\mu}^* = \boldsymbol{\Psi\mu} + \boldsymbol{W\Psi\alpha}$, is observationally equivalent to (2.1).

On the other hand, if the individual-specific effects and/or their spatial spillovers are of some interest, then the identification and estimation of the model need to be addressed. To this end, next we propose using a spatial$-(\boldsymbol{X}, \boldsymbol{\mu})$ panel data model with correlated random effects. In particular, we show that, under mild assumptions, the model is identified. Later we present appropriate estimators under alternative exogeneity assumptions.

2.2 The Correlated Random Effects Spatial$-(\boldsymbol{X},\boldsymbol{\mu})$ Panel Data Model

Fixed effects models implicitly assume that the individual effects are correlated with the covariates. But they somehow ignore such correlation in the estimation procedure. In fact, what the within and analogous transformations do (see e.g. Beer and Riedl, 2012) is to wipe out the individual effects so that this correlation is no longer a concern for the consistent estimation of the model. An alternative procedure to obtain consistent estimates, however, is to incorporate this correlation into the model (Mundlak, 1978; Chamberlain, 1982). This is the approach followed here. In particular, we make use of the correlation between covariates and the (spatially weighted) individual effects to identify the spatial contagion in the individual effects.

Thus, we assume that $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ may be correlated with the explanatory variables:

$$\begin{aligned}\boldsymbol{\mu} &= \frac{1}{T}\boldsymbol{\Psi}'\,\boldsymbol{X}^*\,\boldsymbol{\Pi}_\mu + \boldsymbol{\upsilon}_\mu \\ \boldsymbol{\alpha} &= \frac{1}{T}\boldsymbol{\Psi}'\boldsymbol{X}\boldsymbol{\Pi}_\alpha + \boldsymbol{\upsilon}_\alpha,\end{aligned} \tag{2.5}$$

where $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$ are $K \times 1$ parameter vectors to be estimated and $\boldsymbol{X}^*$ contains the covariates and a vector of ones, i.e, $\boldsymbol{X}^* = \left(\begin{array}{c:c} \boldsymbol{\iota}_{NT} & \boldsymbol{X} \end{array}\right)$. Notice that the inclusion (or the lack) of a constant term in one of the equations in (2.5) is made to guarantee identification, since the spatial contagion of any common factor in the individual effects $\boldsymbol{\mu}$ (in particular, a constant term) is not identified. Ultimately, this means that we are implicitly assuming that the individual effects correspond to deviations with respect to the constant term. Also, the error terms $\boldsymbol{\upsilon}_\mu$ and $\boldsymbol{\upsilon}_\alpha$ are assumed to be random vectors of dimension N with $\boldsymbol{\upsilon}_\mu \sim (\boldsymbol{0}, \sigma^2_\mu \boldsymbol{I}_N)$ and $\boldsymbol{\upsilon}_\alpha \sim (\boldsymbol{0}, \sigma^2_\alpha \boldsymbol{I}_N)$. However, $\boldsymbol{\upsilon}_\mu$ and $\boldsymbol{\upsilon}_\alpha$ are not assumed to be independent, being the covariance parameter, $\sigma_{\mu\alpha}$, such that $E(\boldsymbol{\upsilon}_\mu \boldsymbol{\upsilon}'_\alpha) = \sigma_{\mu\alpha}\boldsymbol{I}_N$ (where E denotes the mathematical expectation).

Plugging equations in (2.5) into the model (2.1) we obtain

$$\boldsymbol{y} \;=\; \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\gamma} + \frac{1}{T}\boldsymbol{\Psi}\boldsymbol{\Psi}'\,\boldsymbol{X}^*\,\boldsymbol{\Pi}_\mu + \frac{1}{T}\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}\boldsymbol{\Pi}_\alpha + \boldsymbol{\eta}, \tag{2.6}$$

where $\boldsymbol{\eta} = \boldsymbol{\Psi} \boldsymbol{v}_\mu + \boldsymbol{W} \boldsymbol{\Psi} \boldsymbol{v}_\alpha + \boldsymbol{\varepsilon}$. Notice that the resulting error component is similar to the one proposed by Kapoor et al. (2007) in that both error components allow for spatial contagion in the individual (random) effects. Also, in contrast to the spatial (lag of) $\boldsymbol{X}$ model for panel data in (2.1), Proposition 2 shows that the correlated random effects spatial panel data model in (2.6) is generally identified.

Proposition 2. *The correlated random effects spatial panel data model in* (2.6) *is identified if the matrix* $\widetilde{\mathbf{X}} = \left(\boldsymbol{X} \,\vdots\, \boldsymbol{W}\boldsymbol{X} \,\vdots\, \frac{1}{T}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^* \,\vdots\, \frac{1}{T}\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X} \right)$ *has full column rank.*

Proof. Since the correlated random effects spatial$-(\boldsymbol{X}, \boldsymbol{\mu})$ panel data model is linear in parameters, it is identified *iff* $det(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}) \neq 0$. If $\widetilde{\mathbf{X}}$ has full rank, it is easy to show that $det(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}) > 0$ (see e.g. Corollary 14.2.14 and Theorem 14.9.4 in Harville 2008). □

Notice that since the number of parameters in the model is $4K + 1$ (excluding the σ's), $NT \geq 4K + 1$ is a necessary identification condition. Notice also that since $\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^*$ is a $NT \times K + 1$ matrix and $\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}$ is an $NT \times K$ matrix, both with row rank equal to N, $N \geq 2K + 1$ is an additional necessary identification condition. Further, $\boldsymbol{X}$, $\boldsymbol{W}\boldsymbol{X}$, $\boldsymbol{\Psi}'\boldsymbol{X}^*$ and $\boldsymbol{W}\boldsymbol{\Psi}'\boldsymbol{X}$ must have full rank. Lastly, time-invariant regressors must be included in either the vector of explanatory variables ($\boldsymbol{X}$ and/or $\boldsymbol{W}\boldsymbol{X}$) or in the vector of determinants of the individual effects and their spatial spillovers ($\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^*$ and/or $\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}$). Otherwise, there is exact multicollinearity between the explanatory variables.

3 Estimation

We start by noticing that consistent estimation of the parameters of the model does not depend on what is the structure of the error term $\boldsymbol{\eta}$. In particular, assuming that the covariates are strictly exogenous (meaning here that $E(\varepsilon_{it} | \boldsymbol{x}_{j1}, \boldsymbol{x}_{j2}, \ldots, \boldsymbol{x}_{jT}) = 0$ for all i, j), Ordinary Least Squares (OLS) estimates of (2.6) are consistent.[6] Yet a more efficient Generalized Least Squares

[6]Notice that the spatial structure of our model requires a orthogonality condition involving not only all the time periods (a standard assumption in applied work; see e.g. Wooldridge 2002) but also all the units. Otherwise, we cannot guarantee the exogeneity of $\boldsymbol{W}\boldsymbol{X}$ and $\frac{1}{T}\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}$.

(GLS) estimator can be derived by accounting for the error components structure of the model. Thus, we propose a FGLS estimator based on the estimates of the parameters of the variance-covariance matrix of $\boldsymbol{\eta}$. On the other hand, none of these estimators are consistent if the strict exogeneity assumption does not hold. In particular, the presence of predetermined variables among the regressors ($\boldsymbol{X}$ and $\boldsymbol{WX}$) makes such variables endogenous when they are included among the variables that compose the correlations functions in (2.5). To obtain consistent estimates, we propose an IV estimator and the means of the endogenous variables taken, for each period, over only current and past values (backward means) as instruments.

Next we discuss the derivation of the proposed estimators in detail.

3.1 GLS estimation under strict exogeneity

Given our initial assumption of spherical disturbances and the stochastic assumptions about the behaviour of $\boldsymbol{\upsilon}_\mu$ and $\boldsymbol{\upsilon}_\alpha$, the error-component $\boldsymbol{\eta} = \boldsymbol{\Psi}\boldsymbol{\upsilon}_\mu + \boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\upsilon}_\alpha + \boldsymbol{\varepsilon}$ has zero mean and variance-covariance matrix given by

$$\boldsymbol{\Omega} = \sigma_\mu^2 \boldsymbol{\Psi}\boldsymbol{\Psi}' + \sigma_\alpha^2 \boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{W}' + \sigma_{\mu\alpha}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{W}' + \sigma_{\mu\alpha}\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}' + \sigma_\varepsilon^2 \boldsymbol{I}_{NT} \tag{3.1}$$

Knowledge of this matrix suffices to derive the GLS estimator (see e.g. Wooldridge 2002). However, to derive the feasible version of the estimator we require an estimate of the vector of parameters $\boldsymbol{\sigma} = (\sigma_\mu^2, \sigma_\alpha^2, \sigma_{\mu\alpha}, \sigma_\varepsilon^2)$. To this end, we notice that each component of $\boldsymbol{\Omega}$ can be written as a linear function of $\boldsymbol{\sigma}$:

$$E\left[\eta_{it}\eta_{ls}\right] = \boldsymbol{\sigma}\boldsymbol{M}_{ilts}, \tag{3.2}$$

where $i, l = 1, \cdots, N$; $t, s = 1, \cdots, T$, $E\left[\eta_{it}\eta_{ls}\right]$ denotes the mathematical expectation of $\eta_{it}\eta_{ls}$, and $\boldsymbol{M}$ is a 4×1 vector whose rows are functions of $\mathbf{w}$. More specifically,

$$E\left[\eta_{it}^2\right] = \sigma_\mu^2 + \sigma_\alpha^2 \sum_{j=1}^{N} w_{ij}^2 + 2\sigma_{\mu\alpha} w_{ii} + \sigma_\varepsilon^2 \tag{3.3}$$

$$E\left[\eta_{it}\eta_{is}\right] = \sigma_\mu^2 + \sigma_\alpha^2 \sum_{j=1}^{N} w_{ij}^2 + 2\sigma_{\mu\alpha} w_{ii} \quad \text{for } t \neq s \tag{3.4}$$

$$E\left[\eta_{it}\eta_{ls}\right] = \sigma_\alpha^2 \sum_{j=1}^{N} w_{ij} w_{lj} + \sigma_{\mu\alpha}(w_{il} + w_{li}) \quad \text{for } i \neq l \tag{3.5}$$

This allows us to consider the following linear regression to estimate $\boldsymbol{\sigma}$:

$$\hat{\eta}_{it}\hat{\eta}_{ls} = \boldsymbol{\sigma}\boldsymbol{M}_{ilts} + u_{ilts} \tag{3.6}$$

where $\hat{\boldsymbol{\eta}}$ is obtained as the residual term of a consistent estimation of the model in (2.6). Given the assumption of strict exogeneity of the covariates, OLS may be used for this purpose.

Under mild conditions, OLS estimation of (3.6) provides consistent estimates of $\boldsymbol{\sigma}$ (denoted by $\hat{\boldsymbol{\sigma}}$) and, with these in hand, we can obtain the FGLS estimates of the model.[7] In particular, $\hat{\boldsymbol{\sigma}}$ allows us to obtain $\widehat{\boldsymbol{\Omega}}_{GLS}$ using (3.1) and, by Cholesky Decomposition of its inverse, $\widehat{\boldsymbol{\Omega}}_{GLS}^{-1} = \boldsymbol{D}_{GLS}\boldsymbol{D}'_{GLS}$, the transformation matrix $\boldsymbol{D}'_{GLS}$. Finally, OLS estimation of the transformed model

$$\boldsymbol{D}'_{\boldsymbol{GLS}}\boldsymbol{y} = \boldsymbol{D}'_{\boldsymbol{GLS}}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{D}'_{\boldsymbol{GLS}}\boldsymbol{W}\boldsymbol{X}\boldsymbol{\gamma} + \frac{1}{T}\boldsymbol{D}'_{\boldsymbol{GLS}}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^*\boldsymbol{\Pi}_\mu + \frac{1}{T}\boldsymbol{D}'_{\boldsymbol{GLS}}\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}\boldsymbol{\Pi}_\alpha + \boldsymbol{D}'_{\boldsymbol{GLS}}\boldsymbol{\eta}, \tag{3.7}$$

provides the FGLS estimates of $\widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\Pi}}_\mu$ and $\widehat{\boldsymbol{\Pi}}_\alpha$.

Interestingly, we may proceed in an analogous way to deal with error structures with idiosyncratic shocks following autoregressive and moving-average processes. In particular, the main difference with respect to the procedure proposed to deal with spherical disturbances is

[7] Alternatively, one may impose the positiveness of the variances (σ_μ^2, σ_α^2 and that σ_ε^2) and that the correlation between $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ lies in the $[-1, 1]$ interval ($-1 \leq \sigma_{\mu\alpha} \times (\sigma_\mu^2 \times \sigma_\alpha^2)^{-1/2} \leq 1$, and use e.g a Non-Linear Least Squares estimator of $\boldsymbol{\sigma}$. Notice that this differs from the approach followed by e.g. Kapoor et al. (2007) in that their estimating equations non-linear in the parameters of interest and they have thus to resort to a Generalized Moment estimator (which may anyway be used here).

that the presence of a serial correlation matrix Kronecker (post-)multiplying the last term of equation (3.1) results in additional (autoregressive and moving-average) parameters in (3.2). Alternatively, we may account for the serial correlation in the model by including among the regressor lags of the dependent variable (and possibly of the explanatory variables in $\boldsymbol{X}$ and $\boldsymbol{WX}$). Notice, however, that the presence of the time-invariant components $\boldsymbol{\upsilon}_\mu$ and $\boldsymbol{\upsilon}_\alpha$ in the error term makes the lagged dependent variable endogenous. We thus propose instrumenting this variable using lags of the explanatory variables $\boldsymbol{X}$ (and possibly $\boldsymbol{WX}$) to control for the endogeneity of the lagged dependent variable.

3.2 Instrumental Variables Estimation Under Sequential Exogeneity

The assumption of strict exogeneity of the covariates is critical to guarantee that the GLS estimators presented in the previous section provide consistent estimates of the parameters of interest. However, in applications the assumption that ε_{it} is uncorrelated with the covariates in all the time periods may not hold. If for example the values of an explanatory variable in period t are related to past values of the dependent variable (e.g., in $t-1$), then future values of such explanatory variables (e.g., in $t+1$) may depend on the values of the idiosyncratic term in t, thus breaking the strict exogeneity assumption (see e.g. Wooldridge 2002).

In such circumstances, a sequential exogeneity assumption, $E(\varepsilon_{it}|\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \cdots, \boldsymbol{x}_{it}) = 0$, seems more appropriate, since it implies that present values of y_{it} do not affect present and past values of $\boldsymbol{x}_{it}$. However, given the spatial structure of our model and following the strict exogeneity case, we instead propose using an "extended sequential exogeneity assumption" involving all the units in the sample. In maths, $E(\varepsilon_{it}|\boldsymbol{x}_{js}) = 0$ for all $\forall i, j$ and $s \leq t$. Notice also that if (expected) future values of $\boldsymbol{x}_{it}$ depend on y_{it} (i.e., present values of y_{it} affect the expect value of $\boldsymbol{x}_{it+1}$), then the explanatory variables employed to construct the correlation functions in (2.5) are endogenous by construction. In other words, the presence of predetermined variables in $\boldsymbol{X}$ and $\boldsymbol{WX}$ makes $\frac{1}{T}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^*$ and $\frac{1}{T}\boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}$ to be correlated with the idiosyncratic term $\boldsymbol{\varepsilon}$. Therefore, under sequential exogeneity, the GLS estimators presented in the previous section no longer provide consistent estimates of the parameters of interest. Rather, an IV estimator should be considered for this purpose.

The main challenge IV estimators face in practice is that it is often difficult to find good instruments. In this case, however, the structure of the model provides natural candidates. Namely, the means of the exogenous explanatory variables constructed using values up to period t (rather than using all T values).[8] In maths, let

$$\boldsymbol{L}_T = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} & \frac{1}{T} & \frac{1}{T} & \cdots & \frac{1}{T} \end{pmatrix}$$

be the row-standardised lower triangular matrix of ones and $\boldsymbol{\Gamma} = \boldsymbol{I}_N \otimes \boldsymbol{L}_T$ be the transformation matrix that yields the backward-up-to-t mean of the variable (i.e., $\boldsymbol{\Gamma X}$, for example, yields a matrix composed by the means of the exogenous explanatory variables constructed using values up to period t). The matrix of instruments can be thus written as $\boldsymbol{Z}_1 = \left(\boldsymbol{\Gamma X} \;\vdots\; \boldsymbol{\Gamma W X} \right)$

Notice that these backward means are exogenous variables under the extended sequential exogeneity assumption. But they are also relevant, since by construction they are correlated with the endogenous explanatory variables $\frac{1}{T}\boldsymbol{\Psi\Psi}'\boldsymbol{X}^*$ and $\frac{1}{T}\boldsymbol{W\Psi\Psi}'\boldsymbol{X}$. Notice also that if we use the same explanatory variables to construct both the instruments and the correlation functions (or different variables but the same number), then the model is exactly identified. However, if all the explanatory variables are employed to construct the instruments but not all the explanatory variables are employed to construct the correlation functions, then the model is overidentified.

To construct the IV estimator, we follow Hausman and Taylor (1981) and Keane and Runkle (1992). Hausman and Taylor (1981) propose a two-step procedure to estimate linear panel data models with endogenous explanatory variables (with respect to the idiosyncratic term, as well

[8] In fact, provided that the number of available periods is long enough, one may use up to lagged periods to construct such means, i.e., one may use values up to period $t-1$, $t-2$, etc..

as with respect to the individual effect) that boils down to first GLS-transform the model (using a consistent estimate of the variance-covariance matrix of the error term) and then estimate the transformed model by IV. However, Keane and Runkle (1992) show that this procedure may not yield consistent estimates when the instruments are predetermined. This is because the GLS-transformation proposed by Hausman and Taylor (1981) results in individual errors that are linear combinations of the errors of the individual in all time periods. To obtain consistent estimates, Keane and Runkle (1992) instead propose using the upper-triangular Cholesky decomposition of the serial correlation matrix (forward filtering) to GLS-transform the model. In essence, this is the procedure we follow, except that the complex structure of our error term requires a different GLS transformation and alternative orthogonality conditions between the errors and the explanatory variables.

To be precise, we obtain the IV estimates of our model in the following way. First, we transform the model using the projection matrix onto the column space of the matrix composed by the exogenous variables and the instruments. That is, we multiply the model by the projection matrix $\boldsymbol{P_Z} = \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'$, with $\boldsymbol{Z} = \left(\boldsymbol{X} \;\vdots\; \boldsymbol{WX} \;\vdots\; \boldsymbol{Z}_1 \right)$. This addresses the endogeneity problem and allow us to consistently estimate the transformed model by OLS. However, a more efficient estimation may be obtained if we transform the model to obtain spherical disturbances. To this end, we use these OLS estimates to generate the residuals $\hat{\eta}$ and, after estimating (3.6), obtain $\widehat{\boldsymbol{\Omega}}_{IV}$ and $\boldsymbol{D}_{IV}$ in an analogous way as done for $\widehat{\boldsymbol{\Omega}}_{GLS}$ and $\boldsymbol{D}_{GLS}$.

In particular, since our instruments are predetermined, we propose using the upper-triangular Cholesky decomposition of the inverse of the variance-covariance matrix (rather than that of the serial correlation matrix used by Keane and Runkle 1992) to obtain $\boldsymbol{D}_{IV}$. However, given the spatial structure of our model, we need to sort the data first by time and then by units within each time period before computing the upper-triangular Cholesky decomposition of $\widehat{\boldsymbol{\Omega}}_{IV}$.[9] This guarantees that the transformed errors in period t contain elements of η_{is} for

[9]Notice that so far we have followed the standard practice of having the data sorted first by units and then by time within each unit, so that e.g. the dependent variable was defined in Section 2 as $\boldsymbol{y} = (y_{11}, \cdots, y_{1T}, \cdots, y_{N1}, \cdots, y_{NT})'$. Here we require that $\boldsymbol{y} = (y_{11}, \cdots, y_{N1}, y_{12}, \cdots, y_{N2}, \cdots, y_{1T}, \cdots, y_{NT})'$, and analogously for the explanatory variables, which will be now $\boldsymbol{X} = (\boldsymbol{x}_{11}, \cdots, \boldsymbol{x}_{N1}, \boldsymbol{x}_{12}, \cdots, \boldsymbol{x}_{N2}, \cdots, \boldsymbol{x}_{1T}, \cdots, \boldsymbol{x}_{NT})$.

$s \geq t$ and thus the exogeneity of our instruments in the transformed model.[10]

In the second step of the procedure, we estimate the GLS-transformed model by IV. This means that we again transform the model using the projection matrix $\boldsymbol{P_Z}$, except that now the sorting of the data requires using the matrix $\boldsymbol{\Gamma} = \boldsymbol{L}_T \otimes \boldsymbol{I}_N$ to construct $\boldsymbol{Z}_1$. The transformed model,

$$\boldsymbol{P_Z D}'_{IV}\boldsymbol{y} = \boldsymbol{P_Z D}'_{IV}\boldsymbol{X\beta} + \boldsymbol{P_Z D}'_{IV}\boldsymbol{WX\gamma} + \frac{1}{T}\boldsymbol{P_Z D}'_{IV}\boldsymbol{\Psi\Psi' X}^*\Pi_\mu + \frac{1}{T}\boldsymbol{P_Z D}'_{IV}\boldsymbol{W\Psi\Psi' X}\Pi_\alpha + \boldsymbol{P_Z D}'_{IV}\boldsymbol{\eta}, \tag{3.8}$$

is then estimated by OLS. The IV estimates of $\widehat{\boldsymbol{\beta}}$, $\widehat{\gamma}$, $\widehat{\boldsymbol{\Pi}}_\mu$ and $\widehat{\boldsymbol{\Pi}}_\alpha$ we obtain are not only consistent but also (more) efficient.

Lastly, it is interesting to note that, unlike the strict exogeneity case, the treatment of serial correlation under sequential exogeneity does not follow immediately. In the present case, the presence of lags of the idiosyncratic term $\boldsymbol{\varepsilon}$ makes endogenous any predetermined variable in the model (not only those in the correlation functions). This is not a major issue as long as the number of lags is small (the order of the moving average is low) and the time dimension of the panel is large. If the predetermined variable is among the regressors ($\boldsymbol{X}$ and $\boldsymbol{WX}$), we propose using lags of the variable as instruments; if the predetermined variable is in the correlation function, we similarly propose adjusting the periods employed to compute the backward means. Thus, we loose (at least) one period for each additional lagged term in the idiosyncratic error. The problem, of course, is that if the order of the moving average process driving the idiosyncratic term is not smaller than the number of time periods minus one, then there is no room for using lags and backward means as instruments.In particular, by the Wold representation theorem, this situation arises if the idiosyncratic term follows an AR process (of

In particular, notice that this sorting requires using $\boldsymbol{W} = \boldsymbol{I}_T \otimes \mathbf{w}$ and $\boldsymbol{\Psi} = \boldsymbol{\iota}_T \otimes \boldsymbol{I}_N$.

[10] Notice that, in a model without spatial dependence, Keane and Runkle (1992) assume that $E(\eta_{it}|\boldsymbol{z}_{is}) = 0$ for $s \leq t$. However, the presence of spatially weighted covariates in the model makes that, for the proposed instruments, this only holds if the extended sequential exogeneity assumptions holds. Notice also that the observation (i, t) of the transformed error term, $\boldsymbol{D}'_{IV}\boldsymbol{\eta}$, contains the original error terms $\eta_{j,t}$ for $j = i, i+1, i+2, ..., N$ as well as $\eta_{j,s}$ for all j and $s > t$. This is why, in the presence of spatial dependence, the orthogonality condition proposed by Keane and Runkle (1992) does not suffice. In contrast, our extended sequential exogeneity assumption guarantees that the proposed instruments are exogenous to the transformed errors, since $E(\eta_{jt}|\boldsymbol{z}_{is}) = 0$ for all j and $s \leq t$.

any order).

Alternatively, in applications where the strict exogeneity does not hold and there is serial correlation in the model, we may include among the regressor lags of the dependent variable (and possibly of the explanatory variables in $\boldsymbol{X}$ and $\boldsymbol{WX}$) and then apply the two-step procedure previously described. Notice, however, that in this specification the lagged dependent variable is endogenous (for the reasons pointed out in the GLS case). We thus proposed extending our matrix of instruments to include lags of the explanatory variables $\boldsymbol{X}$ (and possibly $\boldsymbol{WX}$) to control for the endogeneity of the lagged dependent variable.

4 Empirical Application

We illustrate the proposed FGLS and IV estimators using a Cobb-Douglas production function specification and yearly data from (Munnell, 1990, p. 77) on 48 US contiguous states over the 1970 to 1986 period. The output variable is the gross social product and the inputs include public capital, private capital and labour. "The unemployment rate is also included [in the regressions] to reflect the cyclical nature of productivity". All the variables except the unemployment are in logs. This data set, which is one of the few spatial panel datasets publicly available, has the additional interest of having been partially (e.g. Garcia-Mila et al., 1996) or totally (e.g. Baltagi and Pinnoi, 1995) used in a number of studies on the relation between public capital and private output —see also Boarnet (1998) and Sloboda and Yao (2008). In particular, some of these studies have employed spatial econometrics techniques (e.g. Holtz-Eakin and Schwartz, 1995; Kelejian and Robinson, 1997). This accumulated evidence provides an excellent benchmark for our estimates.

We report FGLS estimates of the model in Table 1 (coefficients and variance components).[11] We also report the joint significance LM-tests of each subset of coefficients ($\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$). The first thing to notice is that the coefficient estimates of the variables that compose the

[11] It is worth noting that, before proceeding with the estimation, we considered the identification of the model. We thus computed $det(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})$ to find that it was indeed positive, which means that our identification condition holds.

correlation functions, $\mathbf{\Pi}_\mu$ and $\mathbf{\Pi}_\alpha$, tend to be statistically significant (both individually and jointly). Also, all the variance components $\boldsymbol{\sigma}$ are statistically significant and have reasonable values. This supports our correlated random effects model specification.

[Insert Table 1]

Also, estimates of the $\boldsymbol{\beta}$-coefficients associated with the main regressors ($\boldsymbol{X}$) are in line with those reported in previous studies. In particular, they are close to those reported by Holtz-Eakin and Schwartz (1995) and Kelejian and Robinson (1997). While our estimate of the elasticity of labour is 0.7, for example, they estimated to be between 0.6 and 0.9; similarly, our estimate of the elasticity of private capital is 0.2, while their estimates range from 0.06 to 0.2.[12] We further concur with the lack of statistical significance of the public capital (see also Baltagi and Pinnoi, 1995; Garcia-Mila et al., 1996) and the statistical significance of the spatially weighted public capital (see the second column of Table 1, $\boldsymbol{\gamma}$). Lastly, the statistical significance of the public capital in the correlation function of the individual effects (see the second column of Table 1, $\mathbf{\Pi}_\mu$) is consistent with evidence reported by Baltagi and Pinnoi (1995) rejecting the orthogonality between regressors and individual effects "only when the public capital stock (is) included in the production function".

Next we explore the possibility that the explanatory variables are not exogenous but predetermined. Previous related studies have analysed the endogeneity of (some of) the explanatory variables, with mixing results with respect to the endogeneity tests (Baltagi and Pinnoi, 1995; Holtz-Eakin and Schwartz, 1995; Garcia-Mila et al., 1996) and implausible results with respect to the coefficients' estimates (Baltagi and Pinnoi, 1995; Holtz-Eakin and Schwartz, 1995). Here we address the predeterminedness of the public capital variable (and its spatially weighted counterpart). This would be the case, for example, if the amount states spend on public capital is related to past values of private output (e.g., because more prosperous states are likely to generate higher tax revenues). Under such circumstances, our previous discussion on the FGLS estimates is flawed, since the variables that compose the correlation functions

[12]These estimates tend to be smaller that those reported by Baltagi and Pinnoi (1995) and Garcia-Mila et al. (1996), which may suggest that ignoring spatial dependence results in overestimation of the coefficients.

defining $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ become endogenous and the FGLS estimator is no longer consistent.

We thus report results from an IV estimation in Table 2. These were obtained using as instruments backward-up-to-t means of all the explanatory variables and their spatially weighted counterparts. That is, $\boldsymbol{Z}_1$ contains the $\boldsymbol{\Gamma}-$transformations of public capital, private capital, labour and unemployment as well as of their spatially weighted counterparts.[13] At first sight, the IV estimates of the $\boldsymbol{\beta}$- and $\boldsymbol{\gamma}$-coefficients are not substantially different from those obtained by FGLS (perhaps with the exception of the public capital in $\boldsymbol{\gamma}$). In contrast, IV and FGLS estimates of the coefficients associated with the variables that compose the correlation functions, $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$, differ substantially. Indeed, a Hausman test between these two estimators strongly rejects the null hypothesis of strict exogeneity (the statistic is 159.42). This supports our tenet that public capital is actually a predetermined variable.

[Insert Table 2]

We use these IV estimates to analyse the geographical distribution of individual effects and their spatial spillovers. To this end, we plot the estimated values of the individual effects (Figure 1a) and their spatial spillovers (Figure 1b and Figure 1c) in a map of the US states. Notice that we consider spillovers on and from the neighbours: while $\widehat{\boldsymbol{\alpha}}$ provides an estimate of the spillover of each state's individual effect on its (first ring) neighbours, $\mathbf{w}\widehat{\boldsymbol{\alpha}}$ provides an estimate of the spillover of the (first ring) neighbours' individual effects on each state. To facilitate the analyses, the maps group states by the quantiles values of the estimate of interest ($\widehat{\boldsymbol{\mu}}$, $\widehat{\boldsymbol{\alpha}}$ and $\mathbf{w}\widehat{\boldsymbol{\alpha}}$).

[Insert Figure 1]

Figure 1a reveals that the states that have the highest estimated values of the individual effects are mostly located in the North and West of the country (plus Texas and Louisiana in the South). More precisely, in the East (Illinois, Michigan and Ohio) and West (Nebraska)

[13] We experimented with other set of instruments (e.g., without considering unemployment and its spatial weight) and found that coefficients estimates were barely altered.

North Central, the Mid-Atlantic (New York and Pennsylvania), the West-Mountain (Montana and Wyoming) and West-Pacific (California and Washington) regions. Figure 1a also shows that the states with the lowest estimated values of the individual effects concentrate in New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont), although we also find some states in the Mountain-West (i.e., Idaho, Utah and Nevada) and South-Athlantic regions (North and South Carolina).

Also, Figure 1b shows that, in absolute values, the West-Central (from Texas to Iowa, but also North-Dakota and Indiana) and West Mountain (New Mexico and Wyoming) states stand out as the area with the highest outwards spatial contagion. These are thus states with individual effects that strongly negatively spill over the neighbouring states. On the other hand, there are two areas of low outwards spatial contagion (in absolute values): the West-Pacific (California, Oregon and Washington) and the Northeast (New York in the Mid-Atlantic and Connecticut, Rhode Island, Massachusetts and Vermont in New England). These are thus states with individual effects that negatively spill over the neighbouring states, but for which the relative magnitude of these spillovers is small.

Interestingly, Figure 1c shows that the geographical distribution of the inwards spatial contagion follows very much the same pattern (with some notable exceptions, such as Utah, Indiana and Iowa). This means that most states generate inwards and outwards spatial spillovers that are of analogous magnitude, i.e., most states have individual effects that negatively spill over the neighbouring states on a magnitude similar to the negative spillovers they receive from neighbouring states. Further, there is analogous overlap in Figure 1a and Figure 1b (and hence Figure 1c). In fact, since the estimated values of the individual effects and their spatial spillovers show different sign (see the signs of the $\mathbf{\Pi}$-coefficients in Table 1 and Table 2), Figure 1 points to a negative relation between $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\alpha}}$ ($\mathbf{w}\widehat{\boldsymbol{\alpha}}$).[14]

Therefore, states with a larger/smaller estimated individual effect tend to have a larger/smaller negative spatial spillover (both inwards and outwards). Notice, however, that

[14]A simple regression between $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\alpha}}$, for example, has indeed a negative slope (-0.17) that is close to significance (the p-value is 0.12). Also, the slope of this regression is steeper (-0.78) and significant at standard confidence levels when we drop the unemployment and its spatial weight from the list of instruments.

this is mostly driven by the Central and South states (with large individual effects and large negative spatial spillovers,) and New England states (with small individual effects and small negative spatial spillovers, as North and South Carolina, and to a lesser extent Colorado, also have). The West and Northeast states, on the other hand, tend to have large estimated individual effects and small negative spatial spillovers.

To conclude, it is interesting to note that negative spillovers that "might shift economic activity from one location to another" have previously been found by for example Boarnet (1998, p. 382) and Sloboda and Yao (2008) with respect to the stock of and the investment in public infrastructure, respectively. However, the source of this "crowding out" effect in our model are the unobservable characteristics of the states or "unobserved productivity" (which these studies cannot identify). In fact, consistent with the work of Holtz-Eakin and Schwartz (1995) and Kelejian and Robinson (1997), the spatially weighed public capital has a negative coefficient, but is not statistically significant (while the private capital is, and with a positive sign). Notice also that both private and public capital are positively related to the states' unobserved productivity (both variables show positive and statistically signs in $\mathbf{\Pi}_\mu$). However, the **role** of these variables in the spatial spillover of the productivity, $\mathbf{\Pi}_\alpha$, differs. While the investment in private capital is associated with negative spillovers, the investment in public capital is associated with positive spillovers.

5 Conclusions

In this paper we analyse the problem of estimating individual effects and their spatial spillovers in linear panel data models. In particular, we consider models in which the exogenous regressors are spatially weighted and there is no spatially lagged dependent variable (i.e., the so-called spatial-$\boldsymbol{X}$ model). We first show that in this model specification the individual effects and their spatial spillovers are not identified for any spatial weight matrix. Under mild assumptions, however, we show that they are identified in a correlated random effects specification. To be precise, we show that there is no identification problem in a spatial$-(\boldsymbol{X}, \boldsymbol{\mu})$ panel data model with correlated random effects if certain rank conditions hold and the individual effects correspond to deviations with respect to the constant term.

We then consider the estimation of the parameters of the (identified) model. Under strict exogeneity of the covariates, OLS estimates are consistent. Here, though, we provide more efficient FGLS estimators and propose an IV estimator to tackle situations where the strict exogeneity assumption may not hold and a sequential exogeneity assumption is upheld. In particular, we suggest using the means of the exogenous explanatory variables constructed using values up to period t as instruments for the endogenous explanatory variables employed to construct the correlation function (which, ultimately, are "means-up-to-T" of the exogenous variables). Also, dropping the most recent periods used to construct these instrumental variables (i.e., using "means-up-to-$(t-s)$", with s being a positive integer) may provide further instruments and/or instruments for potentially endogenous regressors.

Lastly, we present results from an empirical application: the estimation of a Cobb-Douglas production function using US state data. We find statistically significant differences between the FGLS and IV estimates, which suggest that the strict exogeneity assumption that sustains the FGLS estimates may not hold because the public capital variable is actually predetermined. Also, IV (and FGLS) estimates show that the variables that compose the correlation functions, as well as the variance components, all tend to be statistically significant. This supports our correlated random effects model specification. Lastly, the geographical distribution of the (both FGLS- and IV-) estimated individual effects and their spatial spillovers reveals the existence of three major regions: *i)* Central and South states, where both individual effects and negative spatial spillovers tend to be large; *ii)* New England states, where both the individual effects and negative spatial spillovers tend to be small; and *iii)* West and Northeast states, where estimated individual effects tend to be large and negative spatial spillovers tend to be small. In addition, both the inwards and outwards spatial contagion of the individual effects involve negative spillovers, although this sign is mostly associated with the private capital (and labour). The public capital, on the other hand, is behind the positive spatial contagion of the individual effects. Consistent with previous literature, however, the public capital by itself does not seem to convey statistically significant spatial spillovers.

As for the future research, our finding of a negative proportionality between individual effects and their spatial spillovers suggests an alternative approach to the identification of the model. Notice, however, such a proportionality assumption makes the model nonlinear, thus making the identification (and estimation) more involved. Still, since this is a restricted case of the one presented here, we should expect the model to be identified. Also, this paper deals exclusively with the spatial$-(\boldsymbol{X}, \boldsymbol{\mu})$ panel data model with correlated random effects. We hope to extend results presented here to the Durbin model, i.e., to a model specification that includes the spatially lagged dependent variable and spatially weights both the independent variables and the individual effects.

References

Baltagi, B. and Pinnoi, N. (1995). Public capital stock and state productivity growth: Further evidence from an error components model. *Empirical Economics*, 20(2):351–59.

Beer, C. and Riedl, A. (2012). Modelling spatial externalities in panel data: The spatial durbin model revisited. *Papers in Regional Science*, 91(2):299–318.

Boarnet, M. G. (1998). Spillovers and the locational effects of public infrastructure. *Journal of Regional Science*, 38(3):381–400.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, 18(1):5–46.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In Duranton, G., Henderson, V., and Strange, W., editors, *Handbook of Urban and Regional Economics, vol. 5A*, pages 247–348. Elsevier.

Debarsy, N. (2012). The Mundlak Approach in the Spatial Durbin Panel Data Model. *Spatial Economic Analysis*, 7(1):109–131.

Elhorst, J. (2010). Spatial panel data models. In Fischer, M. M. and Getis, A., editors, *Handbook of Applied Spatial Analysis*, pages 377–407. Springer Berlin Heidelberg.

Garcia-Mila, T., McGuire, T. J., and Porter, R. H. (1996). The Effect of Public Capital in State-Level Production Functions Reconsidered. *The Review of Economics and Statistics*, 78(1):177–80.

Harville, D. (2008). *Matrix Algebra From a Statistician's Perspective*. Springer.

Hausman, J. A. and Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica*, 49(6):1377 – 1398.

Holtz-Eakin, D. and Schwartz, A. (1995). Spatial productivity spillovers from public infrastructure: Evidence from state highways. *International Tax and Public Finance*, 2(3):459–468.

Kapoor, M., Kelejian, H. H., and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1):97 – 130. Analysis of spatially dependent data.

Keane, M. P. and Runkle, D. E. (1992). On the estimation of panel-data models with serial correlation when instruments are not strictly exogenous. *Journal of Business & Economic Statistics*, 10(1):1–9.

Kelejian, H. H. and Robinson, D. P. (1997). Infrastructure productivity estimation and its underlying econometric specifications: A sensitivity analysis. *Papers in Regional Science*, 76(1):115–131.

LeSage, J. P. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Chapman & Hall/CRC.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1):69–85.

Munnell, A. H. (1990). How does public infrastructure affect regional economic performance? In Munnell, A. H., editor, *Is There a Shortfall in Public Capital Investment?*, pages 69–103. Federal Reserve Bank of Boston.

Sloboda, B. W. and Yao, V. W. (2008). Interstate spillovers of private capital and public spending. *The Annals of Regional Science*, 42(3):505–518.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Table 1: **FGLS estimates**.

Coefficients	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\Pi}_\mu$	$\boldsymbol{\Pi}_\alpha$
Private capital	0.199^{***}	0.260^{***}	0.197^{***}	-0.477^{***}
	(0.030)	(0.043)	(0.052)	(0.089)
Labour	0.724^{***}	-0.027	-0.212^{***}	0.101
	(0.035)	(0.050)	(0.066)	(0.115)
Unemployment rate	-0.002	-0.007^{***}	-0.013	0.035^{*}
	(0.001)	(0.002)	(0.010)	(0.018)
Public capital	-0.023	-0.129^{**}	0.186^{***}	0.230
	(0.030)	(0.051)	(0.070)	(0.146)
Joint LM-test	250.07^{***}	17.83^{***}	13.10^{***}	8.28^{***}
Variance Components	σ^2_μ	σ^2_α	$\sigma_{\mu\alpha}$	σ^2_ε
	0.0045^{***}	0.0012^{***}	0.0017^{***}	0.0013^{***}
	(0.0001)	(0.0003)	(0.0001)	(0.0002)

Note: *p-value<0.1; **p-value<0.05; ***p-value<0.01. The dependent variable is the gross social product. All the variables are in logs, except for the unemployment. Variance components were estimated by OLS.

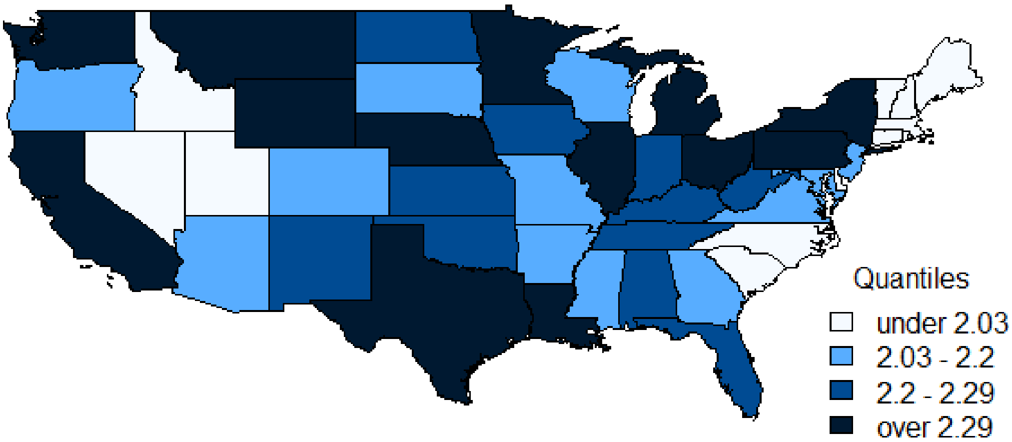
Table 2: **IV estimates**

Coefficients	$\widehat{\boldsymbol{\beta}}$	$\widehat{\boldsymbol{\gamma}}$	$\widehat{\boldsymbol{\Pi}}_\mu$	$\widehat{\boldsymbol{\Pi}}_\alpha$
Private capital	0.255^{***}	0.259^{***}	0.351^{***}	-0.601^{***}
	(0.037)	(0.055)	(0.081)	(0.135)
Labour	0.676^{***}	-0.045	-0.666^{***}	-0.100
	(0.059)	(0.078)	(0.132)	(0.217)
Unemployment rate	-0.003	-0.009^{***}	0.009	0.067^{**}
	(0.002)	(0.003)	(0.015)	(0.029)
Public capital	-0.029	-0.100	0.541^{**}	0.661^{*}
	(0.125)	(0.163)	(0.230)	(0.347)
Joint LM-test	168.57^{***}	12.40^{***}	12.77^{***}	6.32^{***}
Variance Components	$\widehat{\sigma}^2_\mu$	$\widehat{\sigma}^2_\alpha$	$\widehat{\sigma}_{\mu\alpha}$	$\widehat{\sigma}^2_\varepsilon$
	0.0046^{***}	0.0008^{***}	0.0019^{***}	0.0019^{***}
	(0.0001)	(0.0003)	(0.0001)	(0.0002)

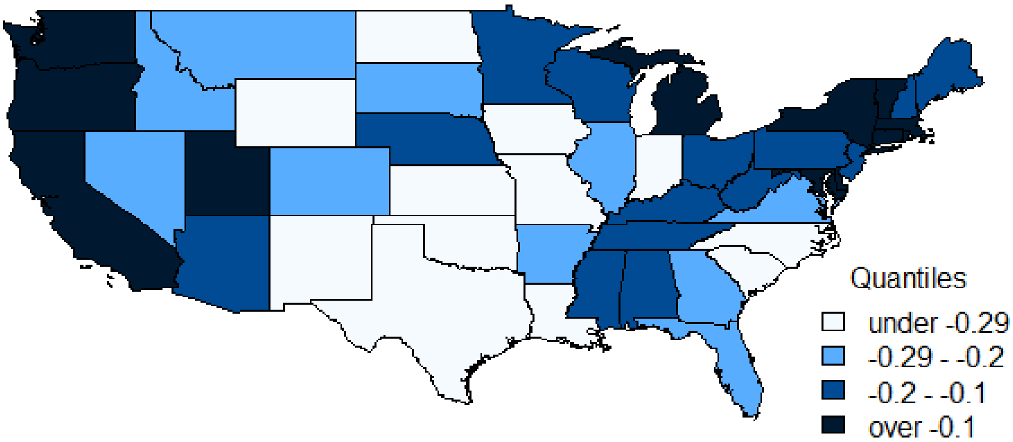
Note: *p-value<0.1; **p-value<0.05; ***p-value<0.01. The dependent variable is the gross social product. All the variables are in logs, except for the unemployment. The matrix of instruments consist of backward-up-to-t means of public capital, private capital, labour and unemployment as well as of their spatially weighted counterparts.

Figure 1: **Estimated individual effects and their spatial spillovers.**

(a) Geographical distribution of $\widehat{\boldsymbol{\mu}}$



(b) Geographical distribution of $\widehat{\boldsymbol{\alpha}}$



(c) Geographical distribution of $\mathbf{w}\widehat{\boldsymbol{\alpha}}$