

Estimating individual effects and their spatial spillovers in linear panel data models: Public capital spillovers after all?*

, 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 not generally 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 and derive appropriate GLS and IV estimators for the resulting correlated random effects spatial panel data model. We also 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, public capital does play a role in the positive “outwards” spatial contagion of the individual effects.

Keywords: correlated random effects, spatial spillovers, panel data

JEL Classification: C23, R11

*We thank P. Elhorst, A. Pérez-Laborda, Jesus Mur and participants at the ERSA Summer School 2016, 18th INFER annual conference, IAAE 2016 annual conference, 4th WIPE workshop, 40th Simposio de Análisis Económico, and 8th Jean Paelinck seminar on Spatial Econometrics 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

Does public capital have an effect on private output? And if it does, does this effect spill over nearby geographical areas? Using US state (and/or county) panel data, production function estimates have consistently concluded that public capital and its spatially weighted counterpart are not statistically significant.¹ In contrast, studies using alternative methodologies (e.g., VAR models), seem to suggest otherwise ([Pereira and Andraz, 2013](#)). In this paper we provide production function estimates supporting the existence of public capital spillovers. To be precise, we find no evidence of a direct positive effect of public capital on private output. However, we find evidence of a relation between public capital and the unobserved productivity of the states (i.e., the individual specific effect of the production function) and its spatial spillover.²

To obtain these results, this paper introduces a correlated-random effects model ([Mundlak, 1978](#); [Chamberlain, 1982](#)) that presents spatial correlation in the individual effects. To our knowledge, only the random effects model of [Kapoor et al. \(2007\)](#) accounts for this spatial correlation. In the fixed effects case, [Beer and Riedl \(2012\)](#) advocate using an extension of the Spatial Durbin Model for panel data that includes 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 lags], 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 (see e.g. [Halleck Vega and Elhorst, 2015](#)). However, this practice rules out obtaining an estimate of the individual-specific effects (net of the spatially weighted effects).³

¹See e.g. [Munnell \(1990\)](#); [Baltagi and Pinnoi \(1995\)](#); [Holtz-Eakin and Schwartz \(1995\)](#); [Garcia-Mila et al. \(1996\)](#); and [Kelejian and Robinson \(1997\)](#).

²As [Boarnet \(1998, p. 381-382\)](#) points out, “[p]ublic capital is provided at a particular place, and if such capital is productive, it enhances the comparative advantage of that location relative to other places”. Also, “productive public capital might shift economic activity from one location to another”.

³This is a critical issue, for example, in two-step models that use this 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 (e.g., for assessing their geographical distribution, which is what we do in our empirical application).

This raises the question of whether both individual effects and their spatial spillovers can indeed be identified 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 from the constant term.

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. We also prove that, regardless of the structure of the variance-covariance matrix of the correlation functions shocks, this estimator coincides with the within (fixed effects) estimator when all the explanatory variables are used to construct the correlation functions. 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 used to approximate the correlation functions. We also advocate using the backward means of these variables (i.e., the means taken, for each period, over only current and past values) as instruments.

Lastly, we use these estimators and a (correlated random effects) production function specification to address the existence of capital spillovers. Using the data and (a spatially weighted variant of) the specification used by [Munnell \(1990\)](#), we find that, under strict exogeneity, our FGLS 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](#)).⁴ However, when we explore the possibility that (some of) the explanatory variables are not exogenous, we find evidence of predeterminedness in the public capital. We then estimate the model by IV to

⁴The data set we employ 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).

find that, under a sequential exogeneity assumption, states with a larger/smaller estimated individual effect tend to have larger/smaller negative spatial spillovers. In particular, we consider both “spill-in” and “spill-out” effects (LeSage and Chih, 2016), although only the spill-out effects turn out to be statistically significant. 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 present empirical evidence based on the work of Munnell (1990). Section 5 concludes.

2 Specification and identification of the model

2.1 Spatial spillovers and the identification problem

Let us consider the spatial– (\mathbf{X}, Ψ) panel data model, that is, the spatial (lag of) \mathbf{X} model for panel data with spatially weighted fixed effects:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \Psi\boldsymbol{\mu} + \mathbf{W}\Psi\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad (2.1)$$

where $\mathbf{y} = (y_{11}, \dots, y_{1T}, \dots, y_{N1}, \dots, y_{NT})'$ is the dependent variable (as usual, $i = 1, \dots, N$ denotes cross-sectional, geographical units and $t = 1, \dots, T$ denotes the time dimension), $\mathbf{X} = (\mathbf{x}_{11}, \dots, \mathbf{x}_{1T}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2T}, \dots, \mathbf{x}_{N1}, \dots, \mathbf{x}_{NT})'$ is the $NT \times K$ matrix of explanatory variables (i.e., \mathbf{x}_{it} is a row vector of order K) is the $NT \times K$ matrix of explanatory variables, and $\boldsymbol{\varepsilon}$ is a zero-mean idiosyncratic error term with assumed variance-covariance matrix $\sigma_\varepsilon^2 \mathbf{I}_{NT}$,

with \mathbf{I}_{NT} being the $NT \times NT$ identity matrix.⁵

We assume that neighbourhood relations do not change over time, so the spatial matrix is $\mathbf{W} = \mathbf{w} \otimes \mathbf{I}_T$, with \mathbf{I}_T denoting the $T \times T$ identity matrix and $\mathbf{w} = [w_{ij}]$ being the $N \times N$ spatial weight matrix that describes the spatial arrangement of the units in the sample. Also, unobservable individual-specific effects are collected in $\Psi\boldsymbol{\mu}$, with $\Psi = \mathbf{I}_N \otimes \boldsymbol{\nu}_T$, \mathbf{I}_N being the $N \times N$ identity matrix and $\boldsymbol{\nu}_T$ a vector of ones of order T . Notice that, in contrast to the so-called SLX model (Halleck Vega and Elhorst, 2015), this model specification accounts for the spatial weights of the individual effects through the term $\mathbf{W}\Psi\boldsymbol{\alpha}$. Thus, the parameters of the model are $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ and σ_ε^2 . This means that $2(K + N) + 1$ parameters need to be estimated.

The main motivation behind the use of this model specification is the estimation of the individual effects and their spatial spillovers, since these often have a meaningful interpretation (Combes and Gobillon, 2015). In particular, following LeSage and Pace (2009), we define the spatial spillovers of the individual-specific effects in terms of the partial derivative of the (conditional expectation of the) dependent variable,

$$\frac{\partial E[\mathbf{y} | \mathbf{X}, \Psi]}{\partial \Psi_j} = (\mathbf{I}_N \mu_j + \mathbf{w}\boldsymbol{\alpha}_j) \otimes \mathbf{I}_T \quad (2.2)$$

where Ψ_j is the j -th column of Ψ and $j = 1, \dots, N$.

The off-diagonal elements of this matrix of partial derivatives represent the *spillovers* or *indirect effects* of unit j , whereas the diagonal elements of the matrix represent the *direct effects* of unit j . Notice, however, that since these effects are time-invariant, we can, without loss of generality, concentrate on the term in brackets, $\mathbf{I}_N \mu_j + \mathbf{w}\boldsymbol{\alpha}_j$. Thus, a generic off-diagonal element of this matrix, $w_{il}\alpha_j$ with $i \neq l$, measures the effect of unit l having the unobservable characteristics of unit j on the dependent variable of unit i . Similarly, a generic element of the diagonal of this matrix, $\mu_j + w_{ii}\alpha_j$, measures the effect of unit i having the unobservable characteristics of unit j on the dependent variable of unit i . This direct effect reduces to μ_j for

⁵Throughout the paper, we assume that a balanced (complete) panel data is available. However, results can easily be extended to incomplete panels.

all i when $w_{ii} = 0$, the standard case.⁶

The vector of parameters $\boldsymbol{\mu}$ has thus a neat interpretation. However, the role the spatial weight matrix \mathbf{w} and the vector of parameters $\boldsymbol{\alpha}$ play in the spill-over effects deserves some further attention. On the one hand, the structure of \mathbf{w} provides the definition of neighbourhood. That is, which units are affected (the “spill-out” effects) and which are affecting (the “spill-in” effects) by the spillover and, in distance-based matrices, how much will be affected/affecting each unit. Thus, different spatial weight matrices yield different spatial spillovers. On the other hand, the parameter α_j provides a measure of the spatial contagion of the individual effect of unit j irrespective of the number of neighbours it has and how close/distant they are. That is, α_j provides a measure of the “potentiality of the spatial contagion” associated with the individual effect of unit j . We thus refer to $\boldsymbol{\alpha}$ as the “potential” of the spatial spillovers of the individual effects.

We illustrate the calculation and interpretation of these direct and spillover effects in the empirical application of section 4. In any case, what is clear is that the estimation of these effects requires that of the parameter vectors $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ (since \mathbf{w} is assumed to be a fixed and/or known matrix). It is therefore critical to determine whether these (and the rest of) parameters of the model are identified. Proposition 1 shows that, in general, this is not the case.

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

Proof. See appendix. □

Notice that, as Beer and Riedl (2012) argue, the omission of $\mathbf{W}\boldsymbol{\Psi}\boldsymbol{\alpha}$ 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 to

⁶However, LeSage and Pace (2009) do not recommend reporting unit-level effects but scalar summary measures. Namely, the average of the main diagonal elements (direct effects) and the cumulative sum of the off-diagonal elements from each row, averaged over all rows (indirect effects). In particular, if the spatial weight matrix \mathbf{w} is row-standardised and has zeros in the diagonal, these scalar summary measures correspond to $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$, respectively. We discuss the use of alternative unit-level indirect effects (analogous to the ones proposed by LeSage and Chih 2016) in section 4.

remove one of the components – i.e., either Ψ or $\mathbf{W}\Psi$ – is perfectly sensible. This is because the model

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \Psi\boldsymbol{\mu}^* + \varepsilon, \quad (2.3)$$

with $\Psi\boldsymbol{\mu}^* = \Psi\boldsymbol{\mu} + \mathbf{W}\Psi\boldsymbol{\alpha}$, 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 discussed. To this end, we go on to propose using a spatial– (\mathbf{X}, Ψ) 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– (\mathbf{X}, Ψ) Panel Data Model

Fixed effects models implicitly assume that the individual effects are correlated with the covariates. But they somehow ignore this 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 for obtaining 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.

Our modelling approach is related to that of [Debarsy \(2012\)](#), who uses a correlated random effects specification to construct an LR test on “the relevance of the random effects approach” (p. 112).⁷ Notice, however, that although we both deal with the correlation between individual effects and covariates, our purposes differ markedly: while he seeks to correctly specify this correlation, we use it as a means to identify the spatial contagion in the individual effects. We

⁷Although this is not the aim of this paper, an analogous Wald test could be developed using our model specification and estimation procedures.

also differ in the model specification which, although similar, treats the spatial contagion of the individual effects differently. [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 (i.e., both the individual effects and their spatial spillovers are included in the specification) and assume that the individual effects and the “potential” of 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 (the additional component being a spatially weighted element).⁸

Thus, we assume the following relation between $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ and the explanatory variables:

$$\begin{aligned}\boldsymbol{\mu} &= \frac{1}{T} \boldsymbol{\Psi}' \mathbf{X}^* \boldsymbol{\Pi}_\mu + \mathbf{v}_\mu \\ \boldsymbol{\alpha} &= \frac{1}{T} \boldsymbol{\Psi}' \mathbf{X} \boldsymbol{\Pi}_\alpha + \mathbf{v}_\alpha,\end{aligned}\tag{2.4}$$

where $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$ are $(K+1) \times 1$ and $K \times 1$ parameter vectors to be estimated, respectively, and $\mathbf{X}^* = \left(\begin{array}{c} \boldsymbol{\iota}_{NT} \\ \mathbf{X} \end{array} \right)$ contains the covariates and a vector of ones. Notice that a constant term needs to be included in one of the equations in (2.4) 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 from the constant term. Also, the error terms \mathbf{v}_μ and \mathbf{v}_α are assumed to be random vectors of dimension N with $\mathbf{v}_\mu \sim (\mathbf{0}, \sigma_\mu^2 \mathbf{I}_N)$ and $\mathbf{v}_\alpha \sim (\mathbf{0}, \sigma_\alpha^2 \mathbf{I}_N)$. However, \mathbf{v}_μ and \mathbf{v}_α are not assumed to be independent, the covariance parameter, $\sigma_{\mu\alpha}$, being such that $E(\mathbf{v}_\mu \mathbf{v}_\alpha') = \sigma_{\mu\alpha} \mathbf{I}_N$ with E denoting the mathematical expectation.

⁸Another 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 $-(\mathbf{X}, \boldsymbol{\Psi})$ 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”; see, however, [Lee and Yu 2016](#)).

Plugging equations in (2.4) into the model (2.1) we obtain

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \frac{1}{T}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X}^*\boldsymbol{\Pi}_\mu + \frac{1}{T}\mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X}\boldsymbol{\Pi}_\alpha + \boldsymbol{\eta}, \quad (2.5)$$

where $\boldsymbol{\eta} = \boldsymbol{\Psi}\mathbf{v}_\mu + \mathbf{W}\boldsymbol{\Psi}\mathbf{v}_\alpha + \boldsymbol{\varepsilon}$. 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. It is different in that while we assume that the idiosyncratic term is not spatially correlated (and propose an identification strategy that takes into consideration that the individual effects may have spatial effects), Kapoor et al. (2007) assume that the idiosyncratic term is spatially autocorrelated.

It is also interesting to note that our model specification does not impose the existence of spatial contagion in the individual effects. In fact, there is no contagion at all if both $\boldsymbol{\Pi}_\alpha$ and σ_α are zero (while there would still be some “random contagion” if $\boldsymbol{\Pi}_\alpha$ is zero but σ_α is not). Similarly, the model does not impose correlation between the individual effects and the covariates. In fact, the specification becomes that of a pure random effects model if both $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$ are zero (still with spatial contagion if σ_α is not zero). Thus, it is the statistical significance of these (sets of) parameters what ultimately determines the existence of spatial contagion in the individual effects and correlation between the individual effects and the covariates.

Finally, in contrast to the spatial $-(\mathbf{X}, \boldsymbol{\Psi})$ model for panel data in (2.1), Proposition 2 shows that the correlated random effects spatial panel data model in (2.5) is generally identified.

Proposition 2. *The correlated random effects spatial panel data model in (2.5) is identified if the matrix $\tilde{\mathbf{X}} = \left(\mathbf{X} \mid \mathbf{W}\mathbf{X} \mid \frac{1}{T}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X}^* \mid \frac{1}{T}\mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X} \right)$ has full column rank.*

Proof. See appendix. □

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 $\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X}^*$ is a $NT \times K + 1$ matrix and $\mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X}$ is an $NT \times K$ matrix, both with row rank equal to N , $N \geq 2K + 1$ is an additional necessary identification condition. Further, \mathbf{X} , $\mathbf{W}\mathbf{X}$, $\boldsymbol{\Psi}'\mathbf{X}^*$ and $\mathbf{W}\boldsymbol{\Psi}'\mathbf{X}$ must have full rank. Lastly, time-invariant regressors must be included in either

the vector of explanatory variables (\mathbf{X} and/or $\mathbf{W}\mathbf{X}$) or in the vector of determinants of the individual effects and their spatial spillovers ($\Psi\Psi'\mathbf{X}^*$ and/or $\mathbf{W}\Psi\Psi'\mathbf{X}$). 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 the structure of the error term $\boldsymbol{\eta}$ is. In particular, assuming that the covariates are strictly exogenous (meaning here that $E(\varepsilon_{it}|\mathbf{x}_{j1}, \mathbf{x}_{j2}, \dots, \mathbf{x}_{jT}) = 0$ for all i, j), Ordinary Least Squares (OLS) estimates of (2.5) are consistent.⁹ Yet a more efficient Generalized Least Squares (GLS) estimator can be derived by accounting for the error components structure of the model.

To be precise, the GLS estimator does not provide efficiency gains in the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ parameters of the correlated random effects model in (2.5). In fact, as shown below (see also [Mundlak 1978](#)), the GLS estimator of these parameters coincides with the OLS estimator in (2.5) and the within or fixed-effects estimator of the (observationally equivalent) model in (2.3). This means that, if these parameters are the only ones of interest, efficiency considerations do not justify the use of the GLS estimator. In contrast, the GLS estimator of (2.5) may provide efficiency gains (with respect to OLS) in the $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$ parameters of the correlation functions. This is why, under the above strict exogeneity assumption, we propose using 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 (\mathbf{X} and $\mathbf{W}\mathbf{X}$) makes such variables endogenous when they are included among the variables that compose the correlation functions in (2.4). 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.

⁹Notice that the spatial structure of our model requires an 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 $\mathbf{W}\mathbf{X}$ and $\frac{1}{T}\mathbf{W}\Psi\Psi'\mathbf{X}$.

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 \mathbf{v}_μ and \mathbf{v}_α , the error-component $\boldsymbol{\eta} = \boldsymbol{\Psi}\mathbf{v}_\mu + \mathbf{W}\boldsymbol{\Psi}\mathbf{v}_\alpha + \boldsymbol{\varepsilon}$ has zero mean and variance-covariance matrix given by

$$\begin{aligned}\boldsymbol{\Omega} &= \sigma_\mu^2 \boldsymbol{\Psi}\boldsymbol{\Psi}' + \sigma_\alpha^2 \mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{W}' + \sigma_{\mu\alpha} \boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{W}' + \sigma_{\mu\alpha} \mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}' + \sigma_\varepsilon^2 \mathbf{I}_{NT} \\ &= \boldsymbol{\Psi} \left(\sigma_\mu^2 \mathbf{I}_N + \sigma_\alpha^2 \mathbf{w}\mathbf{w}' + \sigma_{\mu\alpha} \mathbf{w} + \sigma_{\mu\alpha} \mathbf{w}' \right) \boldsymbol{\Psi}' + \sigma_\varepsilon^2 \mathbf{I}_{NT}\end{aligned}\quad (3.1)$$

Knowledge of this matrix suffices to derive the GLS estimator (see e.g. [Wooldridge 2002](#)) of the parameters of the correlated random effects model in (2.5). In particular, [Mundlak \(1978\)](#) proves that, when all the explanatory variables are used to construct the correlation functions and $\boldsymbol{\eta} = \boldsymbol{\Psi}\mathbf{v}_\mu + \boldsymbol{\varepsilon}$, this coincides with the within (fixed effects) estimator of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in (2.3). Next we generalise this result to any (non-spherical) variance-covariance matrix of \mathbf{v}_μ .

Proposition 3. *Consider the following correlated random effects model:*

$$\mathbf{y} = \mathbb{X}\boldsymbol{\lambda} + \frac{1}{T} \boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbb{X}^*\boldsymbol{\Pi} + \boldsymbol{\eta}$$

with \mathbb{X}^* being the matrix \mathbb{X} plus a column of ones, $\boldsymbol{\lambda}$ and $\boldsymbol{\Pi}$ vectors of parameters with the appropriate dimension, and $E(\boldsymbol{\eta}\boldsymbol{\eta}') = \boldsymbol{\Omega} = \boldsymbol{\Psi}\boldsymbol{\Sigma}_v\boldsymbol{\Psi}' + \sigma_\varepsilon^2 \mathbf{I}_{NT}$, where $\boldsymbol{\Sigma}_v$ is any variance-covariance matrix. The GLS, OLS and within (fixed effects) estimators of $\boldsymbol{\lambda}$ are the same.

Proof. See appendix. □

Notice that our model corresponds to $\mathbb{X} = \left(\mathbf{X} \left| \begin{array}{c} \mathbf{W}\mathbf{X} \\ \mathbf{1} \end{array} \right. \right)$, $\frac{1}{T} \boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbb{X} = \left(\frac{1}{T} \boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X} \left| \begin{array}{c} \frac{1}{T} \mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}'\mathbf{X} \\ \frac{1}{T} \mathbf{1} \end{array} \right. \right)$, $\boldsymbol{\lambda} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$ and $\boldsymbol{\Sigma}_v = \sigma_\mu^2 \mathbf{I}_N + \sigma_\alpha^2 \mathbf{w}\mathbf{w}' + \sigma_{\mu\alpha} \mathbf{w} + \sigma_{\mu\alpha} \mathbf{w}'$. Thus, Proposition 3 fully applies. More generally, the previous proof shows that, regardless of the structure of the variance-covariance

matrix of \mathbf{v}_μ and \mathbf{v}_α , the GLS (and OLS) estimator of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ coincides with the within (fixed effects) estimator.

Next we consider the derivation of the feasible version of this GLS estimator. This basically requires a consistent 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 $\boldsymbol{\sigma}$:

$$E[\eta_{it}\eta_{ls}] = \boldsymbol{\sigma}\mathbf{M}_{ilts}, \quad (3.2)$$

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

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

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

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

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

$$\hat{\eta}_{it}\hat{\eta}_{ls} = \boldsymbol{\sigma}\mathbf{M}_{ilts} + u_{ilts} \quad (3.6)$$

where $\hat{\boldsymbol{\eta}}$ is obtained as the residual term of a consistent estimation of the model in (2.5). 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.¹⁰ In particular, $\hat{\boldsymbol{\sigma}}$ allows us to obtain $\hat{\boldsymbol{\Omega}}_{GLS}$ using (3.1) and, by Cholesky Decomposition of its

¹⁰Alternatively, 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 are non-linear in the parameters of interest and they therefore have to resort to a Generalized Moment estimator (which can also be used here).

inverse, $\widehat{\boldsymbol{\Omega}}_{GLS}^{-1} = \mathbf{D}'_{GLS} \mathbf{D}'_{GLS}$, the transformation matrix \mathbf{D}'_{GLS} . Finally, OLS estimation of the transformed model

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

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

Interestingly, we can 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 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 can account for the serial correlation in the model by including among the regressors lags of the dependent variable (and possibly of the explanatory variables in \mathbf{X} and $\mathbf{W} \mathbf{X}$). Notice, however, that the presence of the time-invariant components \mathbf{v}_μ and \mathbf{v}_α in the error term makes the lagged dependent variable endogenous. We thus propose instrumenting this variable using lags of the explanatory variables \mathbf{X} (and possibly $\mathbf{W} \mathbf{X}$) 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 these 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, $E(\varepsilon_{it} | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{it}) = 0$,

seems more appropriate, since it implies that present values of y_{it} do not affect present and past values of \mathbf{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 mathematical terms, $E(\varepsilon_{it}|\mathbf{x}_{js}) = 0$ for all $\forall i, j$ and $s \leq t$. Notice also that if (expected) future values of \mathbf{x}_{it} depend on y_{it} (i.e., present values of y_{it} affect the expected value of \mathbf{x}_{it+1}), then the explanatory variables used to construct the correlation functions in (2.4) are endogenous by construction. In other words, the presence of predetermined variables in \mathbf{X} and $\mathbf{W}\mathbf{X}$ means that $\frac{1}{T}\Psi\Psi'\mathbf{X}^*$ and $\frac{1}{T}\mathbf{W}\Psi\Psi'\mathbf{X}$ are correlated with the idiosyncratic term ε . 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).¹¹ Let

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

be the row-standardised lower triangular matrix of ones and $\mathbf{\Gamma} = \mathbf{I}_N \otimes \mathbf{L}_T$ be the transformation matrix that yields the backward-up-to- t mean of the variable (i.e., $\mathbf{\Gamma}\mathbf{X}$, 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 $\mathbf{Z}_1 = \left(\mathbf{\Gamma}\mathbf{X} \ \middle| \ \mathbf{\Gamma}\mathbf{W}\mathbf{X} \right)$

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

¹¹In 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..

with the endogenous explanatory variables $\frac{1}{T}\Psi\Psi'\mathbf{X}^*$ and $\frac{1}{T}\mathbf{W}\Psi\Psi'\mathbf{X}$. 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 used to construct the instruments but not all the explanatory variables are used 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 as with respect to the individual effect) that boils down to an initial GLS transformation of the model (using a consistent estimate of the variance-covariance matrix of the error term) and then an estimation of 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 consisting of the exogenous variables and the instruments. That is, we multiply the model by the projection matrix $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, with $\mathbf{Z} = \left(\begin{array}{c|c|c} \mathbf{X} & \mathbf{W}\mathbf{X} & \mathbf{Z}_1 \end{array} \right)$. This addresses the endogeneity problem and makes it possible 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 $\hat{\boldsymbol{\eta}}$ and, after estimating (3.6), obtain $\hat{\boldsymbol{\Omega}}_{IV}$ and \mathbf{D}_{IV} in the same way as we did for $\hat{\boldsymbol{\Omega}}_{GLS}$ and \mathbf{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 \mathbf{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 $\widehat{\boldsymbol{\Omega}}_{IV}$.¹² This guarantees that the transformed errors in period t contain elements of η_{is} for $s \geq t$ and hence the exogeneity of our instruments in the transformed model.¹³

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 \mathbf{P}_Z , except that now the sorting of the data requires using the matrix $\boldsymbol{\Gamma} = \mathbf{L}_T \otimes \mathbf{I}_N$ to construct \mathbf{Z}_1 . The transformed model,

$$\mathbf{P}_Z \mathbf{D}'_{IV} \mathbf{y} = \mathbf{P}_Z \mathbf{D}'_{IV} \mathbf{X} \boldsymbol{\beta} + \mathbf{P}_Z \mathbf{D}'_{IV} \mathbf{W} \mathbf{X} \boldsymbol{\gamma} + \frac{1}{T} \mathbf{P}_Z \mathbf{D}'_{IV} \boldsymbol{\Psi} \boldsymbol{\Psi}' \mathbf{X}^* \boldsymbol{\Pi}_\mu + \frac{1}{T} \mathbf{P}_Z \mathbf{D}'_{IV} \mathbf{W} \boldsymbol{\Psi} \boldsymbol{\Psi}' \mathbf{X} \boldsymbol{\Pi}_\alpha + \mathbf{P}_Z \mathbf{D}'_{IV} \boldsymbol{\eta}, \quad (3.8)$$

is then estimated by OLS. The IV estimates of $\widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\Pi}}_\mu$ and $\widehat{\boldsymbol{\Pi}}_\alpha$ we obtain are not only consistent but also more efficient than the initial IV estimates obtained in the first step of the procedure.

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 any predetermined variable in the model endogenous (not only those in the correlation functions). This is, however, not a major issue if the number of lags is small (the order of the moving average is low) and the time dimension

¹²Notice 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 e.g. the dependent variable was defined in Section 2 as $\mathbf{y} = (y_{11}, \dots, y_{1T}, \dots, y_{N1}, \dots, y_{NT})'$. Here we require that $\mathbf{y} = (y_{11}, \dots, y_{N1}, y_{12}, \dots, y_{N2}, \dots, y_{1T}, \dots, y_{NT})'$ and $\mathbf{X} = (\mathbf{x}_{11}, \dots, \mathbf{x}_{N1}, \mathbf{x}_{12}, \dots, \mathbf{x}_{N2}, \dots, \mathbf{x}_{1T}, \dots, \mathbf{x}_{NT})'$. In particular, notice that this sorting requires using $\mathbf{W} = \mathbf{I}_T \otimes \mathbf{w}$ and $\boldsymbol{\Psi} = \boldsymbol{\nu}_T \otimes \mathbf{I}_N$.

¹³Notice that, in a model without spatial dependence, Keane and Runkle (1992) assume that $E(\eta_{it} | \mathbf{z}_{is}) = 0$ for $s \leq t$. However, the presence of spatially weighted covariates in the model means 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, $\mathbf{D}'_{IV} \boldsymbol{\eta}$, contains the original error terms $\eta_{j,t}$ for $j = i, i+1, i+2, \dots, N$ 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 $E(\eta_{jt} | \mathbf{z}_{is}) = 0$ for all j and $s \leq t$.

of the panel is large. If the predetermined variable is among the regressors (\mathbf{X} and \mathbf{WX}), we propose using lags of the variable as instruments; and if the predetermined variable is in the correlation function, we similarly propose adjusting the periods used to compute the backward means. Thus, we lose (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 any order).

Alternatively, in applications in which strict exogeneity does not hold and there is serial correlation in the model, we may include among the regressors lags of the dependent variable (and possibly of the explanatory variables in \mathbf{X} and \mathbf{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 \mathbf{X} (and possibly \mathbf{WX}) to control for the endogeneity of the lagged dependent variable.

4 Public capital spillovers in a production function specification: empirical evidence

In this section we use our correlated random effects specification and the proposed FGLS and IV estimators to empirically address the existence of public capital spillovers. To this end, we use a Cobb-Douglas production function specification and yearly data from (Munnell, 1990, p. 77) on 48 US contiguous states over the period 1970 to 1986. The output variable is the gross state 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 unemployment are in logs. This data set 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 used spatial econometrics techniques (e.g. [Holtz-Eakin and Schwartz, 1995](#); [Kelejian and Robinson, 1997](#)).

This accumulated evidence provides an excellent benchmark for our estimates, which were obtained for a model specification that uses all the inputs and their spatial counterparts to construct the correlation functions ($\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$). Before proceeding with the estimation, however, we considered the identification of the proposed model. We thus computed $\det(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})$ to find that it was indeed positive, which means that our identification condition holds.

We report FGLS estimates of the model in [Table 1](#) (coefficients and variance components). 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 since our model specification uses all regressors (\mathbf{X} and $\mathbf{W}\mathbf{X}$) to construct the correlation functions ($\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$), estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ reported in [Table 1](#) correspond to the within estimates of the model. As for the coefficient estimates of the variables that compose the correlation functions, $\boldsymbol{\Pi}_\mu$ and $\boldsymbol{\Pi}_\alpha$, all tend to be statistically significant (both individually and jointly). Lastly, all the variance components $\boldsymbol{\sigma}$ are statistically significant and have reasonable values. This supports our correlated random effects model specification. In particular, given that we reject that $\boldsymbol{\Pi}_\alpha$ is zero, there is evidence of contagion in the individual effects and, given that we reject that $\boldsymbol{\Pi}_\mu$ is zero, there is evidence of correlation between the individual effects and the covariates (fixed effects).

[Insert [Table 1](#)]

The FGLS estimates of the $\boldsymbol{\beta}$ -coefficients associated with the main regressors (\mathbf{X}) are in line with those reported in previous studies. More precisely, 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 it 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 0.2.¹⁴ We further concur with the lack of statistical significance of public capital (see

¹⁴These estimates tend to be smaller than those reported by [Baltagi and Pinnoi \(1995\)](#) and [Garcia-Mila et al.](#)

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, γ). Lastly, the statistical significance of public capital in the correlation function of the individual effects (see the third column of Table 1, Π_μ) is consistent with evidence reported by Baltagi and Pinnoi (1995, p. 396) that rejects 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 mixed results on the endogeneity tests (Baltagi and Pinnoi, 1995; Holtz-Eakin and Schwartz, 1995; Garcia-Mila et al., 1996) and implausible results on the coefficient 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 defining μ and α 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, Z_1 contains the Γ -transformations of public capital, private capital, labour and unemployment as well as of their spatially weighted counterparts.¹⁵ At first sight, the IV estimates of the β - and γ -coefficients are not substantially different from those obtained by FGLS (perhaps with the exception of the public capital in γ). In contrast, IV and FGLS estimates of the coefficients associated with the variables that compose the correlation functions, Π_μ and Π_α , differ substantially. Indeed, a Hausman test between these two estimators strongly rejects the null hypothesis of strict exogeneity (the statistic is 159.42).

(1996), which may suggest that ignoring spatial dependence results in overestimation of the coefficients.

¹⁵We experimented with other set of instruments (e.g., without considering unemployment and its spatial weight) and found that coefficients estimates were barely altered.

This supports our tenet that public capital is actually a predetermined variable.

[Insert Table 2]

We then obtained an estimate of the direct individual effects (i.e., $\boldsymbol{\mu}$) and the “potential” of their spatial spillovers (i.e., $\boldsymbol{\alpha}$) using these IV estimates. However, we found that while the estimated direct effects were generally statistically significant (both using FGLS and IV estimates), the “potential” of the spatial spillovers of the individual effects were only statistically significant under the strict exogeneity assumption. Under the assumption that capital is a predetermined variable, the estimated α 's were not statistically significant.¹⁶

Seeking for alternative, more efficient specifications that yielded statistically significant α 's, we used a variables-selection procedure that resulted in the specification reported in Table 3.¹⁷ Notice that the (common) estimated model coefficients are not that different from those reported in Table 2. Also, the correlation between the estimated $\boldsymbol{\mu}$'s and $\boldsymbol{\alpha}$'s (obtained from the estimates reported in Table 2 and Table 3) is 0.99 and 0.90, respectively. However, 14 out of the 48 estimated α 's are now statistically significant at standard confidence levels. Alternatively, we considered an ad-hoc model selection procedure in which we explored different specifications in terms of instruments (e.g., using squared terms) and/or variables (e.g., dropping the unemployment, as in Holtz-Eakin and Schwartz 1995). This also produced

¹⁶We use t-statistics to test the statistical significance of the individual effects and their spatial spillovers. In particular, the standard errors were obtained from $\widehat{Var}(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}) = \frac{1}{T^2} \boldsymbol{\Psi}' \mathbf{X}^* \widehat{\boldsymbol{\Sigma}}_{\Pi_\mu} \mathbf{X}^{*'} \boldsymbol{\Psi} + \widehat{\sigma}_\mu^2 \mathbf{I}_N - 2 \frac{1}{T} \boldsymbol{\Psi}' \mathbf{X}^* M_\mu E[\widehat{\boldsymbol{\eta} \mathbf{v}'_\mu}]$, with $E[\cdot]$ denoting the mathematical expectation, $\boldsymbol{\Sigma}_{\Pi_\mu} = Var(\widehat{\Pi}_\mu)$ and $\widehat{\Pi}_\mu - \Pi_\mu = M_\mu \boldsymbol{\eta}$. The expression for $\widehat{Var}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})$ is analogous, only differing in the subindices (i.e., using $\widehat{\boldsymbol{\Sigma}}_{\Pi_\alpha}$ instead of $\widehat{\boldsymbol{\Sigma}}_{\Pi_\mu}$, $\widehat{\sigma}_\alpha^2$ instead of $\widehat{\sigma}_\mu^2$, M_α instead of M_μ , and \mathbf{v}'_α instead of \mathbf{v}'_μ). Under standard assumptions, these are consistent and (approximately) normally-distributed estimators of $Var(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})$ and $Var(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})$, respectively. Notice that these estimated variances allow us to also test whether the direct effects of two states are statistically equal (rather than whether the direct effect of one state is statistically equal to zero, as we did).

¹⁷In particular, we proceeded as follows. We dropped the most non-significant variable (i.e., that with the higher p-value) in the original specification (that reported in Table 2), reestimated the model and determined which was the most non-significant variable in the new specification. We then tested whether these two variables were jointly statistically significant in the original specification. If the null hypothesis of this Wald test was not rejected, we dropped the two variables and considered again the most non-significant variable in the resulting specification. Then, we constructed a new Wald test for the null that the three variables were not jointly statistically significant in the original specification. We went on dropping variables and testing their joint significance until either the jointly statistically non-significant hypothesis was rejected or all the variables in the model were statistically significant.

specifications in which the α 's were statistically significant and highly correlated with those obtained using the results reported in [Table 2](#), while the coefficients of the model and their statistical significance remained generally unaltered. This was the case, for example, when dropping public capital from \mathbf{WX} , the unemployment from $\mathbf{\Pi}_\mu$, and the unemployment and its spatially weighted counterpart from the set of instruments.

Given the illustrative aim of this empirical exercise, determining which is the best model specification is clearly beyond the scope of this paper. It is important to stress, however, that little differences were observed among the alternative specifications we considered when we plotted on a map of the US states the estimated values of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ (available from the authors upon request). Bearing this in mind, we have used the IV estimates reported in [Table 3](#) to analyse the geographical distribution of the estimated direct individual effects and their spatial spillovers. In particular, our spillover effects are based on [LeSage and Chih \(2016\)](#), who define “spill-in” and “spill-out” effects as the cumulating off-diagonal elements from row- and column-sums of the matrix of marginal effects in [\(2.2\)](#), respectively.

[Insert [Table 3](#)]

To be precise, [LeSage and Chih \(2016\)](#) interpret the row-sums of $\mathbf{w}\alpha_j$ (i.e., $\sum_{l=1}^N w_{il}\alpha_j$, assuming that $w_{ii} = 0$) as the “spill-in effects” on the outcome of unit i . Thus, this spill-in effect captures the impact on the outcome of unit i of all the units neighbouring i having the unobserved characteristics of j . However, we find more interesting to report $\sum_{l=1}^N w_{il}\alpha_l$, which is the impact on the outcome of unit i of all the units neighbouring i having their unobserved characteristics (i.e., the impact on the outcome of unit i of the individual effects of the units neighbouring i). This is reported in [Figure 1a](#).

Similarly, [LeSage and Chih \(2016\)](#) interpret the column-sums of $\mathbf{w}\alpha_j$ (i.e., $\alpha_j \sum_{l=1}^N w_{li}$, assuming that $w_{ii} = 0$) as “spill-out effects” of unit i . In particular, this spill-out captures the impact of unit i having the unobserved characteristics of j on the outcome of the neighbours of unit i . Again, we find more interesting to report $\alpha_i \sum_{l=1}^N w_{li}$, which is the impact on the outcome of the units neighbouring i of the individual effect of unit i . Notice, however, that our proposed spill-out effect is the product of α_i (the “potential” of the spatial spillovers of the individual effects)

and $\sum_{l=1}^N w_{li}$ (which in essence determines the spatial contagion, i.e., which units are affected by the spillover). Since our spill-in effect already reflects the spatial contagion of the individual effects, it seems more interesting to report α rather than $\alpha_i \sum_{l=1}^N w_{li}$. This is consequently what we do in [Figure 1b](#).

[Insert [Figure 1](#)]

Results indicate that the geographical distribution of the spatial contagion, as measured by the spill-in ([Figure 1a](#)) and spill-out ([Figure 1b](#)) effects, follows very much the same pattern (with some notable exceptions, such as Nebraska and Nevada). This means that most states show spill-in and spill-out effects that are of analogous magnitude. However, the “inwards” spatial contagion of the spill-in effects is generally not statistically significant (Colorado, Montana, Texas and Utah being the exceptions). We thus concentrate on the analysis of the spill-out effects which, as previously pointed out, tend to be statistically significant.

[Figure 1b](#) shows that, in absolute values, West-Central (from Texas to Kansas, but also North-Dakota and Indiana) and West-Mountain (New Mexico and Wyoming) states stand out as the areas with the highest “outwards” spatial contagion. These are thus states with individual effects that strongly and negatively spill over the neighbouring states. Interestingly, these spill out effects are generally statistically significant at conventional confidence levels. On the other hand, there are two areas of low spill-out effects (in absolute values): the West-Pacific (California and Washington) and the North-East (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 (and, in fact, not statistically significant).

Further, [Figure 1c](#) reveals that the states that have the highest estimated values of the direct 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) North Central, the mid-Atlantic (New York and Pennsylvania), the West-Mountain (Montana and Wyoming) and West-Pacific (California and Washington) regions. [Figure 1c](#)

also shows that the states with the lowest estimated values of the direct individual effects concentrate in New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont), although we also find some states in the West-Mountain (i.e., Idaho, Utah and Nevada) and the South-Atlantic (North and South Carolina) regions.

Lastly, it is interesting to note the overlap between [Figure 1a](#), [Figure 1b](#) and [Figure 1c](#). In fact, since the estimated values of the individual effects and their spatial spillovers show different sign (see the signs of the Π -coefficients in [Table 1](#), [Table 2](#) and [Table 3](#)), [Figure 1](#) points to a negative relation between $\hat{\mu}$ and $\hat{\alpha}$ ($\mathbf{w}\hat{\alpha}$). Importantly, this “proportionality” between the direct and indirect effects is not a feature of the model specification ([LeSage and Chih, 2016](#)). Rather, it arises as a genuine characteristic of the data.¹⁸

Therefore, states with larger/smaller estimated direct individual effects tend to have larger/smaller negative spatial spillovers (both spill-in and spill-out). Notice, however, that this is mostly driven by New England states (with small direct individual effects and small negative spatial spillovers) and the central and southern states (with large direct individual effects and large negative spatial spillovers, except for Nebraska, which shows large direct individual effects and small negative spatial spillovers). The West and North-East states, on the other hand, tend to have large estimated direct individual effects and small negative spatial spillovers.

To conclude, it is worth noting 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 \(1997\)](#), results reported in [Table 2](#) show that the spatially weighted public capital has a negative coefficient, but is not statistically significant (while private capital

¹⁸A simple regression between $\hat{\mu}$ and $\hat{\alpha}$, for example, indeed has a negative and statistically significant slope (-0.34 , with a p-value of 0.03), whereas a simple regression between $\hat{\alpha}$ and $\mathbf{w}\hat{\alpha}$ has a positive and statistically significant slope (0.86 with a p-value of 0.00).

is, and has 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 significant 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- \mathbf{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- $(\mathbf{X}, \mathbf{\Psi})$ 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 (at least more efficient with respect to the coefficients of the variables that compose the correlation functions) and propose an IV estimator to tackle situations in which 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 used 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.

The geographical distribution of the IV-estimated direct individual effects and their spatial spillovers reveals the existence of three major regions: *i)* Central and South states, where both direct individual effects and negative spatial spillovers tend to be large; *ii)* New England states, where both the direct individual effects and negative spatial spillovers tend to be small; and *iii)* West and North-East states, where the estimated direct individual effects tend to be large and the negative spatial spillovers tend to be small. In addition, both the “inwards” and “outwards” spatial contagion of the individual effects (i.e., the spill-in and spill-out effects) involve negative spillovers, although this sign is mostly associated with the private capital (and labour) and is only statistically significant for the spill-out effects. Public capital, on the other hand, is behind the positive spatial contagion of the individual effects. Consistent with previous literature, however, public capital itself does not seem to convey statistically significant spatial spillovers.

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.
- Debarys, N. (2012). The Mundlak Approach in the Spatial Durbin Panel Data Model. *Spatial Economic Analysis*, 7(1):109–131.
- 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.
- Halleck Vega, S. and Elhorst, J. P. (2015). The slx model. *Journal of Regional Science*, 55(3):339–363.
- 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.
- Henderson, H. V. and Searle, S. R. (1981). On deriving the inverse of a sum of matrices. *SIAM Review*, 23(1):53–60.

- 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.
- Lee, L.-F. and Yu, J. (2016). Identification of spatial durbin panel models. *Journal of Applied Econometrics*, 31(1):133–162.
- LeSage, J. and Chih, Y.-Y. (2016). Interpreting heterogeneous coefficient spatial autoregressive panel models. *Economics Letters*, 142:1 – 5.
- LeSage, J. and Pace, R. (2009). *Introduction to Spatial Econometrics*. Statistics: A Series of Textbooks and Monographs. CRC Press.
- 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.
- Pereira, A. and Andraz, J. (2013). On the economic effects of public infrastructure investment: A survey of the international evidence. *Journal of Economic Development*, 4(38):1–37.

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.

Appendix: Proofs of Propositions

Proof of Proposition 1. The model in (2.1) is not identified for any spatial weight matrix \mathbf{w} because Ψ and $\mathbf{W}\Psi$ are perfectly collinear. We prove this by showing that $\det \left[\left(\Psi \mid \mathbf{W}\Psi \right)' \left(\Psi \mid \mathbf{W}\Psi \right) \right]$ is zero for any spatial weight matrix \mathbf{w} . Let

$$\mathbf{A} = \left(\Psi \mid \mathbf{W}\Psi \right)' \left(\Psi \mid \mathbf{W}\Psi \right) = T \left(\begin{array}{c|c} \mathbf{I}_N & \mathbf{w} \\ \hline \mathbf{w}' & \mathbf{w}'\mathbf{w} \end{array} \right) \quad (5.1)$$

Then, by Schur complement,

$$\det(\mathbf{A}) = T^{2N} \det(\mathbf{I}_N) \det(\mathbf{w}'\mathbf{w} - \mathbf{w}'(\mathbf{I}_N)^{-1}\mathbf{w}) = T^{2N} \det(\mathbf{w}'\mathbf{w} - \mathbf{w}'\mathbf{w}) = 0 \quad (5.2)$$

□

Proof of Proposition 2. Since the correlated random effects spatial- $(\mathbf{X}, \boldsymbol{\mu})$ panel data model is linear in parameters, it is identified *iff* $\det(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}) \neq 0$. If $\tilde{\mathbf{X}}$ has full rank, it is easy to show that $\det(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}) > 0$ (see e.g. Corollary 14.2.14 and Theorem 14.9.4 in [Harville 2008](#)). □

Proof of Proposition 3. Let $\hat{\boldsymbol{\lambda}}_{OLS}$, $\hat{\boldsymbol{\lambda}}_w$ and $\hat{\boldsymbol{\lambda}}_{GLS}$ be the OLS, within and GLS estimators of $\boldsymbol{\lambda}$. We prove first that $\hat{\boldsymbol{\lambda}}_{OLS} = \hat{\boldsymbol{\lambda}}_w$. To this end, we start by noting that, by the Frisch-Waugh-Lovell theorem and given that $\Psi\Psi'\Psi\Psi' = T\Psi\Psi'$,

$$\hat{\boldsymbol{\lambda}}_{OLS} = (\mathbb{X}'\mathbf{M}_1\mathbb{X})^{-1}(\mathbb{X}'\mathbf{M}_1\mathbf{y})$$

with $\mathbf{M}_1 = \mathbf{I}_{NT} - \frac{1}{T}\Psi\Psi'\mathbb{X}^*(\mathbb{X}^{*'}\Psi\Psi'\mathbb{X}^*)^{-1}\mathbb{X}^{*'}\Psi\Psi'$. Also, let $\mathbf{Q} = \mathbf{I}_{NT} - \frac{1}{T}\Psi\Psi'$, which satisfies that $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}$ and $\mathbf{Q}\Psi = 0$. Lastly, since $\mathbb{X}' = \left(\mathbf{0}_{K \times 1} \mid \mathbf{I}_{K \times K} \right) \mathbb{X}^{*}$, it can be proved that $\mathbb{X}'\mathbf{M}_1 = \mathbb{X}'\mathbf{Q}$. Consequently,

$$\hat{\boldsymbol{\lambda}}_{OLS} = (\mathbb{X}'\mathbf{Q}\mathbb{X})^{-1}\mathbb{X}'\mathbf{Q}\mathbf{y} = \hat{\boldsymbol{\lambda}}_w$$

This concludes the first part of the proof. Next we prove that $\widehat{\boldsymbol{\lambda}}_{GLS} = \widehat{\boldsymbol{\lambda}}_{OLS} = \widehat{\boldsymbol{\lambda}}_w$. To this end, we start by noting that $\widehat{\boldsymbol{\lambda}}_{GLS}$ corresponds to the OLS estimator of the considered correlated random effects model transformed using the (upper triangular part of the) Cholesky decomposition of the inverse of the variance covariance matrix of $\boldsymbol{\eta}$, $\boldsymbol{\Omega}^{-1} = \boldsymbol{D}\boldsymbol{D}'$. Therefore, by the Frisch-Waugh-Lovell theorem,

$$\widehat{\boldsymbol{\lambda}}_{GLS} = ((\boldsymbol{D}'\boldsymbol{X})' \boldsymbol{M}_2 \boldsymbol{D}'\boldsymbol{X})^{-1} ((\boldsymbol{D}'\boldsymbol{X})' \boldsymbol{M}_2 \boldsymbol{D}'\boldsymbol{y})$$

with $\boldsymbol{M}_2 = \boldsymbol{I}_{NT} - \boldsymbol{D}'\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^* (\boldsymbol{X}^{*'}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^*)^{-1} \boldsymbol{X}^{*'}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{D}$. Also, from equation (19) in [Henderson and Searle \(1981\)](#),

$$\begin{aligned} \boldsymbol{\Omega}^{-1} &= (\sigma_\varepsilon^2 \boldsymbol{I}_{NT} + \boldsymbol{\Psi}\boldsymbol{\Sigma}_v\boldsymbol{\Psi}')^{-1} = \frac{1}{\sigma_\varepsilon^2} \boldsymbol{I}_{NT} - \frac{1}{\sigma_\varepsilon^2} \boldsymbol{\Psi}\boldsymbol{\Sigma}_v\boldsymbol{\Psi}'\boldsymbol{\Omega}^{-1} \\ &= \frac{1}{\sigma_\varepsilon^2} \boldsymbol{I}_{NT} - \frac{1}{\sigma_\varepsilon^2} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}\boldsymbol{\Sigma}_v\boldsymbol{\Psi}', \end{aligned}$$

which implies that $\boldsymbol{\Psi}\boldsymbol{\Sigma}_v\boldsymbol{\Psi}'\boldsymbol{\Omega}^{-1} = \boldsymbol{\Omega}^{-1}\boldsymbol{\Psi}\boldsymbol{\Sigma}_v\boldsymbol{\Psi}'$ and, using again that $\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{\Psi}\boldsymbol{\Psi}' = T\boldsymbol{\Psi}\boldsymbol{\Psi}'$, it can be proved that

$$\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{\Omega}^{-1} = \boldsymbol{\Omega}^{-1}\boldsymbol{\Psi}\boldsymbol{\Psi}' \tag{5.3a}$$

$$\boldsymbol{Q}\boldsymbol{\Omega}^{-1} = \boldsymbol{\Omega}^{-1}\boldsymbol{Q} = \frac{1}{\sigma_\varepsilon^2} \boldsymbol{Q} \tag{5.3b}$$

Lastly, by (5.3a) matrix \boldsymbol{M}_2 can be rewritten as

$$\boldsymbol{M}_2 = \boldsymbol{I}_{NT} - \frac{1}{T} \boldsymbol{D}'\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^* (\boldsymbol{X}^{*'}\boldsymbol{\Omega}^{-1}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{X}^*)^{-1} \boldsymbol{X}^{*'}\boldsymbol{\Psi}\boldsymbol{\Psi}'\boldsymbol{D},$$

and so $\boldsymbol{X}'\boldsymbol{D}\boldsymbol{M}_2 = \boldsymbol{X}'\boldsymbol{Q}\boldsymbol{D}$. Consequently,

$$\widehat{\boldsymbol{\lambda}}_{GLS} = (\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} (\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{\Omega}^{-1}\boldsymbol{y}) = (\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X})^{-1} (\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{y}) = \widehat{\boldsymbol{\lambda}}_{OLS} = \widehat{\boldsymbol{\lambda}}_w,$$

where the second equality holds by (5.3b). □

Table 1: **FGLS estimates.**

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

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

Table 2: IV estimates.

Coefficients	β	γ	Π_μ	Π_α
Private capital	0.255*** (0.037)	0.259*** (0.055)	0.351*** (0.081)	-0.601*** (0.135)
Labour	0.676*** (0.059)	-0.045 (0.078)	-0.666*** (0.132)	-0.100 (0.217)
Unemployment rate	-0.003 (0.002)	-0.009*** (0.003)	0.009 (0.015)	0.067** (0.029)
Public capital	-0.029 (0.125)	-0.100 (0.163)	0.541** (0.230)	0.661* (0.347)
Joint LM-test	168.57***	12.40***	12.77***	6.32***
Variance Components	σ_μ^2	σ_α^2	$\sigma_{\mu\alpha}$	σ_ε^2
	0.0046*** (0.0001)	0.0008*** (0.0003)	0.0019*** (0.0001)	0.0019*** (0.0002)

Note: *p-value<0.1; **p-value<0.05; ***p-value<0.01. The dependent variable is the gross state 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. Variance components were estimated by OLS.

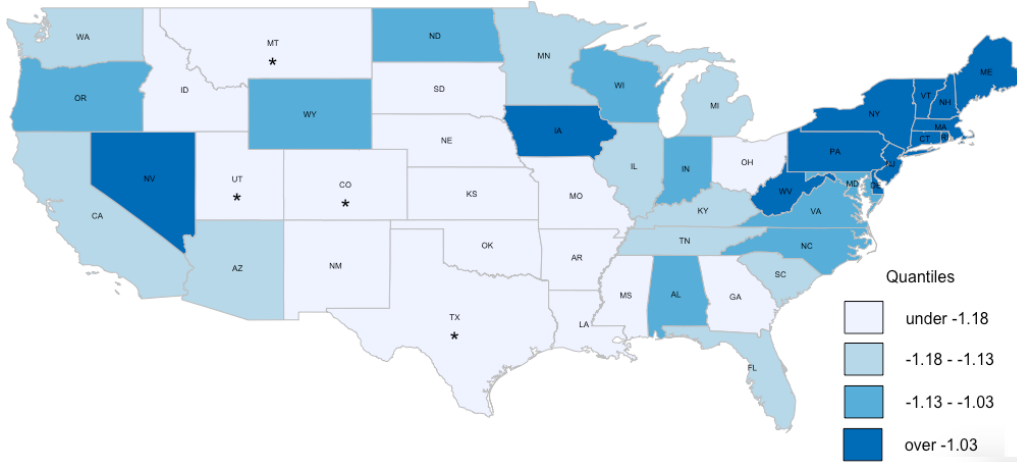
Table 3: IV estimates.

Coefficients	β	γ	Π_μ	Π_α
Private Capital	0.252*** (0.040)	0.419*** (0.068)	0.342*** (0.084)	-0.909*** (0.180)
Labour	0.666*** (0.050)	-0.279*** (0.084)	-0.776*** (0.118)	
Unemployment rate	-0.011*** (0.002)	-0.008*** (0.003)		
Public Capital			0.660*** (0.132)	0.877*** (0.193)
Variance Components	σ_μ^2	σ_α^2	$\sigma_{\mu\alpha}$	σ_ε^2
	0.0044*** (0.0001)	0.0026*** (0.0003)	0.0018*** (0.0001)	0.0015*** (0.0002)

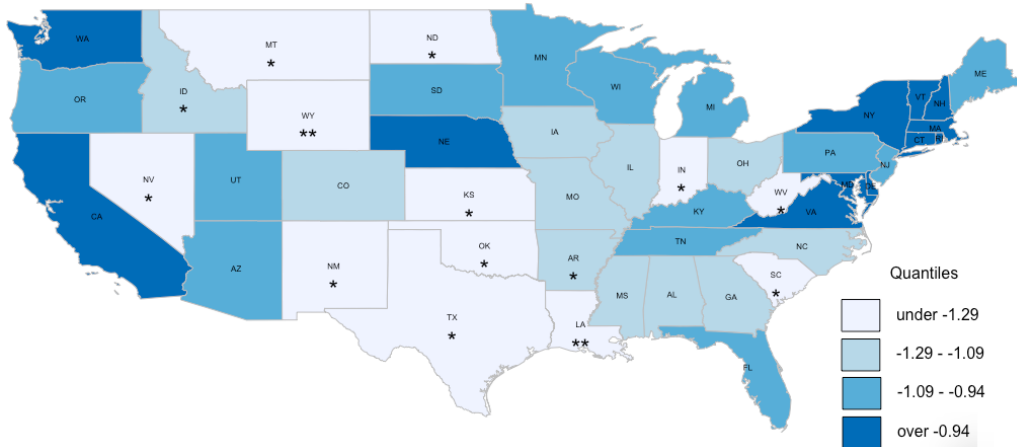
Note: *p-value<0.1; **p-value<0.05; ***p-value<0.01. The dependent variable is the gross state product. All the variables are in logs, except for the unemployment. The matrix of instruments consist of backward-up-to- t means of public capital, private capital, labour and unemployment as well as of their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

Figure 1: Estimated individual effects and their spatial spillovers.

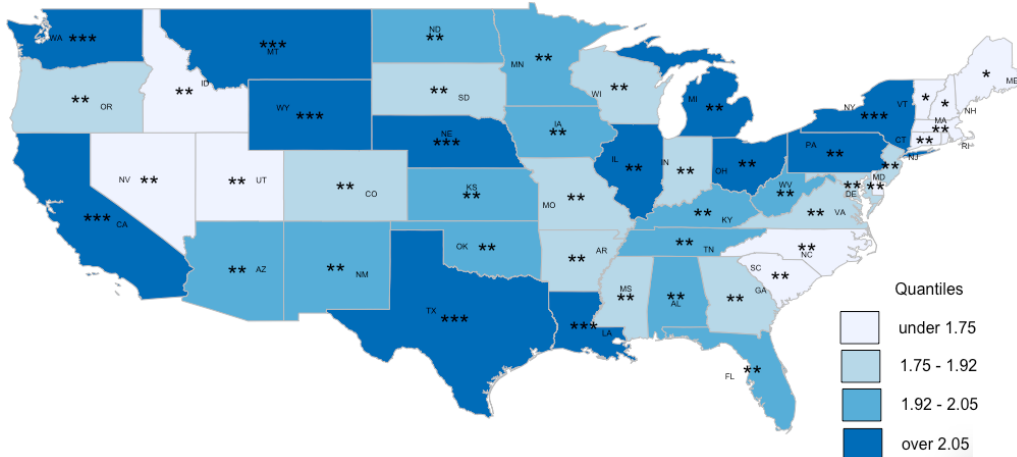
(a) Geographical distribution of $w\hat{\alpha}$



(b) Geographical distribution of $\hat{\alpha}$



(c) Geographical distribution of $\hat{\mu}$



Note: *p-value<0.1; **p-value<0.05; ***p-value<0.01.