

## ePub<sup>WU</sup> Institutional Repository

Manfred M. Fischer and Monika Bartkowska and Aleksandra Riedl and Sascha Sardadvar and Andrea Kunnert

The impact of human capital on regional labor productivity in Europe

Paper

*Original Citation:*

Fischer, Manfred M. and Bartkowska, Monika and Riedl, Aleksandra and Sardadvar, Sascha and Kunnert, Andrea (2008) The impact of human capital on regional labor productivity in Europe. WU Vienna University of Economics and Business, Vienna.

This version is available at: <http://epub.wu.ac.at/3963/>

Available in ePub<sup>WU</sup>: July 2014

ePub<sup>WU</sup>, the institutional repository of the WU Vienna University of Economics and Business, is provided by the University Library and the IT-Services. The aim is to enable open access to the scholarly output of the WU.

# The impact of human capital on regional labor productivity in Europe

*Manfred M. Fischer, Monika Bartkowska, Aleksandra Riedl,  
Sascha Sardadvar and Andrea Kunnert*

Institute for Economic Geography and GIScience, Department for Social Sciences,  
Vienna University of Economics and Business

November, 2008

## Abstract

This paper employs a spatial Durbin model for analyzing the impact of human capital on regional productivity using for 198 NUTS-2 European regions for the sample period from 1995 to 2004. The study provides evidence for the existence of spatial externalities and interactions of the sort as emphasized by new growth theory. To interpret results meaningfully, we calculate summary measures that account for the simultaneous feedback nature of the underlying model. By sampling from the parameter distribution we present measures of dispersion, revealing that it is *relative* regional advantages in human capital that matter most for productivity growth.

## 1 Introduction

Education is often considered a key determinant in economic growth. It is viewed as one of the most important potential policy instruments for raising both productivity growth and economic growth in general. Education has also been the subject of intensive policy discussion in Europe, as evidenced, for example, by the emphasis on education and the information society in recent years.

A traditional way of studying the role of education in economic growth is to allow for human capital as an explicit determinant of economic growth<sup>1</sup>. The

---

<sup>1</sup> Studies of economic growth often seek to explain differences in economic growth rates across countries or regions in terms of levels and changes in human capital, among other variables. However, these estimates are plagued by measurement errors and specification problems (see Brock and Durlauf 2001), and may suffer due to omitted spatial dependence.

objective of this paper is to provide evidence from a cross-sectional point of view and to use spatial econometric methods to account for spatial externalities of the sort emphasized in endogenous growth theory. By human capital we mean, for the purpose of this paper, simply the skills of the workforce as given by the level of educational attainment of the population. We use gross value added [GVA] per worker as metric of economic level to examine cross-regional evidence of the importance of human capital.

The observation units are NUTS-2 regions<sup>2</sup>. The GVA data were calculated on the basis of the 1995 European System of Accounts [ESA 95] and refer to the time years 1995 and 2004, the latest year for which data is available. The time period is relatively short due to a lack of reliable data in Central and Eastern Europe. This comes partly from the substantial change in accounting conventions now used in these countries. But more important, even if estimates of the change in the volume of output did exist, these would be impossible to interpret meaningfully because of the fundamental change of production from a centrally planned to a market system (Fischer and Stirböck 2006).

The remainder of this paper consists of two sections and a conclusion. Section 2 describes the spatial regression framework, along with the relevant methodology to estimate the impact of human capital on regional productivity levels. Section 3 applies the methodology to a sample of 198 NUTS-2 regions that covers 22 countries in Europe, presents the estimation results and quantifies the impact of human capital, drawing on recent work by LeSage and Pace (2008). Unlike most previous research in regional growth analysis, with the notable exception of LeSage and Fischer (2008), this quantification accounts for the simultaneous feedback nature of the underlying spatial regression model.

The paper shows that the unrestricted spatial Durbin model is an appropriate model specification for estimating the impact of human capital on regional productivity in a cross-sectional regression framework. The analysis provides evidence for the existence of spatial externalities and interactions of the sort as emphasized by new growth theory. Since the model involves spatial lags of the dependent and independent variables, the traditional least-squares *ceteris paribus* interpretation of the regression parameters does not hold any longer.

---

<sup>2</sup> NUTS is an acronym for Nomenclature of Units of Territorial Statistics. For details of the definition of NUTS-2 regions see [http://ec.europa.eu/eurostat/ramon/nuts/basicnuts\\_regions\\_en.html](http://ec.europa.eu/eurostat/ramon/nuts/basicnuts_regions_en.html).

## 2 The regression framework

### (i) The spatial Durbin model

In this study we employ a spatial Durbin model (SDM) given by Eq. (1) as some sort of framework for analyzing the impact of human capital on regional productivity. This model is a generalization of the conventional least-squares regression that attempts to explain cross-regional differences in terms of labour productivity and human capital levels<sup>3</sup>, and incorporates three types of spatial externalities into the cross-sectional regression specification:

- (i) spatial effects working through the dependent variable, labour productivity at the end of the sample period (2004),
- (ii) spatial effects working through the level of labour productivity and
- (iii) spatial effects working through the level of human capital, both at the beginning of the sample period (1995).

This spatial regression model takes the form

$$\mathbf{y} = \mathbf{1}_n \alpha + \mathbf{X}\boldsymbol{\beta} + \rho \mathbf{W} \mathbf{y} + \mathbf{W} \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (1)$$

where all variables are in log form.  $\mathbf{y}$  represents an  $n$ -by-1 vector of observations of the labor productivity level at the end of the sample period, with  $n$  being the number of observations (regions).  $\mathbf{1}_n$  is an  $n$ -by-1 vector of ones with the associated scalar parameter  $\alpha$ .  $\mathbf{X}$  is an  $n$ -by-2 matrix of observations on the two (non-constant) explanatory variables: labor productivity and human capital at the beginning of the sample period, while  $\boldsymbol{\beta}$  is the associated 2-by-1 parameter vector.

$\mathbf{W}$  is an  $n$ -by- $n$  non-stochastic, non-negative spatial weight matrix that specifies the spatial dependence among observations (regions), or in other words expresses for each row (observation/region) those regions (columns) which belong to its neighborhood<sup>4</sup> set as non-zero elements. Formally, we define  $W_{ij} = 1$  when region  $j$  is neighbor of region  $i$ , and  $W_{ij} = 0$  otherwise. By convention, the diagonal elements of  $\mathbf{W}$  are set to zero. The matrix  $\mathbf{W}$  is row-standardized, which guarantees that all weights are between zero and one. This facilitates the

<sup>3</sup> Note that both explanatory variables are measured at the beginning of the sample period to avoid endogeneity.

<sup>4</sup> Note that the term neighborhood is used here in a more general sense of spatial relatedness, despite that we will use it later in the more restricted sense of map-based first-order contiguity relations (see Section 3).

interpretation of operations with the spatial weight matrix as an averaging of neighboring values.

The  $n$ -by-1 vector  $W\mathbf{y}$  is the spatial lag of  $\mathbf{y}$  that captures the first type of spatial dependence mentioned above [see (i)].  $\rho$  is the scalar parameter of the first order spatial autoregressive (SAR) process, and is typically referred to as the spatial autoregressive parameter assumed to lie  $(-1,1)$ . This parameter reflects spatial dependence, which is expected to be positive in our model, indicating that regional productivity levels are positively related to a linear combination of neighboring regions' productivity. The presence of the spatial lag variable  $W\mathbf{y}$  on the right side of the equation will induce a non-zero correlation with  $\boldsymbol{\varepsilon}$  that represents an  $n$ -by-1 normally distributed, constant variance disturbance term,  $\boldsymbol{\varepsilon} \sim (0, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_n)$ . The spatial lag for an observation (region)  $i$  is not only correlated with the error term at  $i$ , but also with the error terms at  $j \neq i$ . Thus, an ordinary least-squares estimator will not be consistent for this model.

$WX$  is the  $n$ -by- $q$  matrix of the spatially lagged non-constant explanatory variables. The  $q$  element vector  $\boldsymbol{\gamma}$  contains the regression parameters associated with these variables. The coefficient estimate on the spatial lag of the labour productivity variable captures the second type of spatial externalities [see (ii)] and that on the spatial lag of human capital reflects the third type of spatial externalities [see (iii)], mentioned above.

This SDM occupies an interesting position in the field of spatial growth regression analysis because it nests many of the models widely used in the literature (see Abreu et al. 2005; Fingleton and López-Bazo 2006):

- (i) Imposing the restriction  $\boldsymbol{\gamma} = \mathbf{0}$  leads to the spatial autoregressive (SAR) growth regression model<sup>5</sup> that includes a spatial lag of labor productivity from neighboring regions, but excludes the influence of the spatially lagged explanatory variables.
- (ii) The so-called common factor parameter restriction  $\boldsymbol{\gamma} = -\rho\boldsymbol{\beta}$  yields the spatial error growth regression model (SEM) specification that assumes that externalities across regions are mostly a nuisance spatial dependence problem caused by the regional transmission of random shocks.
- (iii) The restriction  $\rho = 0$  results in a least-squares spatially lagged X growth regression model (labeled SLX by LeSage and Pace 2008) that assumes independence between regional productivity levels, but includes characteristics from neighboring regions in the form of spatially lagged explanatory variables.

---

<sup>5</sup> The term spatial *autoregressive* is used since the dependent variable is regressed on a spatial lag of itself.

- (iv) Finally, imposing the restriction  $\rho=0$  and  $\boldsymbol{\gamma}=\mathbf{0}$  yields the standard least-squares growth regression model (LSM).

Testing whether the restrictions hold or not implies not much effort. Of particular importance are common factor tests that discriminate between the unrestricted SDM and the SEM, or in other words between substantive and residual dependence in the analysis. The likelihood ratio test proposed by Burridge (1981) is the most popular test in this context (see Mur and Angulo 2006 for alternative tests and a comparison based on Monte Carlo evidence).

### (ii) Model estimation

The spatial Durbin model cannot be estimated by the least-squares approach due to endogeneity problems stemming from the dependence of the regressor  $W\mathbf{y}$  and the error term  $\boldsymbol{\varepsilon}$ . An alternative is to use maximum likelihood methods estimation, which requires solving a univariate optimization problem that involves the spatial autoregressive parameter  $\rho$ . This is achieved by concentrating the likelihood with respect to the parameters  $\boldsymbol{\beta}$ ,  $\boldsymbol{\gamma}$ ,  $\alpha$  and  $\sigma^2$ :

$$\ln \mathcal{L}(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha, \sigma^2) = C - \frac{n}{2} \ln(\mathbf{e}'\mathbf{e}) + \ln |\mathbf{I}_n - \rho\mathbf{W}| \quad (2)$$

with

$$\mathbf{e} = \mathbf{e}_o - \rho \mathbf{e}_d \quad (3)$$

where

$$\mathbf{e}_o = \mathbf{y} - \bar{X} \boldsymbol{\beta}_o \quad (4)$$

$$\mathbf{e}_d = \mathbf{W} \mathbf{y} - \bar{X} \boldsymbol{\beta}_d \quad (5)$$

$$\boldsymbol{\beta}_o = (\bar{X}' \bar{X})^{-1} \bar{X}' \mathbf{y} \quad (6)$$

$$\beta_d = (\bar{X}' \bar{X})^{-1} \bar{X}' W y. \quad (7)$$

with  $\bar{X} = [X, W X]$ , and where  $C$  represents a constant not involving the parameters. The computationally troublesome part in the numerical optimization is the need to compute the log-determinant of the  $n$ -by- $n$  matrix  $|I_n - \rho W|$ . Operation counts for computing this determinant via eigenvalues<sup>6</sup> increase with the cube of  $n$  for a dense matrix  $W$ . While  $W$  is an  $n$ -by- $n$  matrix, it is sparse by construction and becomes more sparse with increases in sample size. Thus, direct sparse matrix algorithms such as Cholesky or LU decompositions to compute the log-determinant might be used (see Pace and Barry 1997).

### (iii) Interpretation of estimated parameters

To attain the objective of this paper, we explicitly consider spatial effects in estimating the impact of human capital on regional productivity. Note that the least-squares *ceteris paribus* interpretation of regression parameters does not hold any longer in a spatial regression context. For models that contain spatial lags of dependent and/or explanatory variables, interpretation of the parameter estimates becomes more complicated (see Anselin 2003; Kim et al. 2003; LeSage and Pace 2008).

In our spatial Durbin regression setting the labor productivity of region  $i$  (that we denote by  $y_i$ ) depends on

- *first*, labor productivity from regions neighboring  $i$ , captured by the spatial lag variable  $W_i y$  where  $W_i$  represents the  $i$ th row of the spatial weight matrix  $W$ ,
- *second*, the own-region initial period level of productivity, represented by  $x_{i1}$ , the first column of the  $n$ -by-2 matrix  $X$ ,
- *third*, the initial period levels of productivity in the neighboring regions, represented by the spatially lagged productivity variable  $W_i x_1$ ,
- *fourth*, the own-region initial period level of human capital, represented by  $x_{i2}$ , the second column of  $X$ ,
- *fifth*, the initial period levels of human capital in the neighboring regions, represented by the spatially lagged human capital variable  $W_i x_2$ .

Thus, a change in the human capital level in region (observation)  $i$  will not only exert a *direct effect* on the productivity level of this region, but also an *indirect*

---

<sup>6</sup> Note that  $\ln|I_n - \rho W| = \sum_{r=1}^n \ln(1 - \rho \lambda_r)$  where  $\{\lambda_r, r = 1, \dots, n\}$  being the eigenvalues of the spatial weight matrix  $W$ .

*effect* on other regions  $j \neq i$ . This is a result of the spatial connectivity relationships incorporated in the model (LeSage and Fischer 2008). To arrive at a correct interpretation of the impact of human capital on productivity growth, we draw on recent work by LeSage and Pace (2008) and quantify the impact using their computationally feasible means of calculating scalar summary measures of direct and indirect impacts that arise from changes in the human capital variable in our general spatial Durbin model.

The data generating process for this model is given as

$$y = \sum_{r=1}^q \mathbf{S}_r(\mathbf{W})x_r + (\mathbf{I}_n - \rho\mathbf{W})^{-1} \mathbf{1}_n \alpha + (\mathbf{I}_n - \rho\mathbf{W})^{-1} \boldsymbol{\varepsilon} \quad (8)$$

with

$$\mathbf{S}_r(\mathbf{W}) = (\mathbf{I}_n - \rho\mathbf{W})^{-1} (\mathbf{I}_n \beta_r + \mathbf{W} \gamma_r) \quad (9)$$

where the index  $r$  runs from 1 to  $q$ , and  $x_r$  is the  $r$ th explanatory variable ( $r$ th column of  $\mathbf{X}$ ). There are  $2q+1$  explanatory variables. The  $q$ -by-1 vector  $\boldsymbol{\beta}$  contains the regression parameters associated with the explanatory variables in  $\mathbf{X}$ , and the  $q$ -by-1 vector  $\boldsymbol{\gamma}$  the regression parameters associated with the spatially lagged variables  $\mathbf{WX}$  (in this study  $q=2$ ).

In the case of standard least-squares growth analysis, where  $\rho=0$  and  $\boldsymbol{\gamma}=\mathbf{0}$ , the partial derivatives of  $y_i$  with respect to  $x_{ir}$  [the  $(i, r)$ th element of  $\mathbf{X}$ ] have a simple form

$$\frac{\partial y_i}{\partial x_{ir}} = \beta_r \quad \text{for all } i \text{ and } r \quad (10)$$

and

$$\frac{\partial y_i}{\partial x_{jr}} = 0 \quad \text{for all } j \neq i \text{ and for all } r. \quad (11)$$

The spatial error growth regression model (given by the restriction  $\boldsymbol{\gamma} = -\rho\boldsymbol{\beta}$ ) inherits this characteristic from the least-squares model, since  $\mathbf{S}_r(\mathbf{W}) = \mathbf{I}_n \beta_r$ .

To interpret estimated spatial regression models that contain spatial lags of the dependent variable, one has to examine up to  $qn^2$  partial derivatives, as LeSage and Pace (2008) point out. The derivative of  $y_i$  with respect to  $x_{ir}$  usually does not equal  $\beta_r$  and the derivative of  $y_i$  with respect to  $x_{jr}$  for  $j \neq i$  does not equal



zero. From Eq. (8) it follows that changes in the  $r$ th explanatory variable in a spatial regression model have a partial derivative impact on  $y_i$  equal to

$$\frac{\partial y_i}{\partial x_{jr}} = \mathbf{S}_r(\mathbf{W})_{ij} \quad \text{for } j \neq i, \text{ and for all } r \quad (12)$$

where  $\mathbf{S}_r(\mathbf{W})_{ij}$  refers to the  $(i, j)$ th element of the  $n$ -by- $n$  matrix  $\mathbf{S}_r$  given by Eq. (9), and

$$\frac{\partial y_i}{\partial x_{ir}} = \mathbf{S}_r(\mathbf{W})_{ii} \quad \text{for all } i \text{ and } r \quad (13)$$

LeSage and Pace (2008) label  $\mathbf{S}_r(\mathbf{W})_{ii}$  the *direct* effect that is measured by the  $(i, i)$ th element of  $\mathbf{S}_r$ . This includes feedback influences that arise as a result of impacts passing through neighbors, and back to the observation (region) itself<sup>7</sup>. The *indirect* effects that arise from changes in all observations  $j=1, \dots, n$  of an explanatory variable  $x_r$  are found as the sum of the off-diagonal elements of row  $i$  from the matrix  $\mathbf{S}_r$ , for each observation  $i$ . Direct plus indirect effects equal the *total* effect from ceteris paribus changes in variable  $x_r$ .

Since the impact of changes in an explanatory variable differs over all observations, LeSage and Pace (2008) suggest the following scalar summary measures:

- (i) the *average direct* effect constructed as an average of the diagonal elements of  $\mathbf{S}_r(\mathbf{W})$ ,
- (ii) the *average indirect* effect constructed as an average of the off-diagonal elements of  $\mathbf{S}_r(\mathbf{W})$ , where the off-diagonal row elements are summed up first, and then an average of these sums is taken,
- (iii) the *average total* effect is the sum of the direct and indirect impacts.

We will use these scalar summary measures to draw inferences regarding the statistical significance of the direct, indirect and total impacts that arise from changes in the human capital variable. For inference, we need the distribution of

---

<sup>7</sup> Despite the fact that the main diagonal of the spatial weight matrix  $\mathbf{W}$  contains zeros, the main diagonal of higher order matrices  $\mathbf{W}^m$  ( $m$  integer) that arise in the infinite series expansion representation of the matrix inverse are non-zero.  $W_{ii}^2$ , for example, is non-zero to reflect the fact that region  $i$  is a second-order neighbor to itself, that is a neighbor to its neighbor. This accounts for the feedback effects.

the scalar summary measures. To produce measures of dispersion, we produce samples of parameters  $\beta$ ,  $\gamma$  and  $\rho$  that obey the distribution implied by the maximum likelihood estimates. This is a multivariate normal distribution with means equal to the maximum likelihood estimates and a variance-covariance matrix based on the numerical Hessian that comes from the maximum likelihood procedure.

### 3 Application of the methodology

#### (i) Variables, sample data and the spatial weight matrix

Our sample includes 198 NUTS-2 regions in continental Europe including 159 regions<sup>8</sup> located in *Western Europe* covering Austria (nine regions), Belgium (11 regions), Denmark (one region), Finland (four regions), France (21 regions), Germany (40 regions), Italy (19 regions), Luxembourg (one region), the Netherlands (12 regions), Norway (seven regions), Portugal (five regions), Spain (15 regions), Sweden (eight regions) and Switzerland (seven regions), and 39 regions in *Central Eastern Europe* covering the Baltic States (three regions), Czech Republic (eight regions), Hungary (seven regions), Poland (16 regions), Slovakia (four regions) and Slovenia (one region).

There are shortcomings of the NUTS definition of regions, which can raise a form of the modifiable areal unit problem, illustrated, for example, by Openshaw and Taylor (1979). Ideally, the definition of regions should be based on theoretical considerations leading to functionally defined regions, but empirical studies are typically constrained by the availability of public data. A poor boundary matching is most likely to induce nuisance spatial dependence, which is in sharp contrast to substantive spatial dependence caused by knowledge spillovers, forward and backward linkages, factor mobility and trade among regions.

We use gross value added [GVA] per worker as metric of regional growth, expressed in ECU, the former European Currency Unit, replaced by the Euro in 1999. GVA is the net result of output at basic prices less intermediate consumption valued at purchasers' prices, and measured in accordance with the European System of Accounts [ESA] 1995. Our main data source is Eurostat's REGIO database. The data for Norway and Switzerland stem from Statistics Norway (Division for National Accounts) and the Swiss Office Fédéral de la Statistique (Comptes Nationaux), respectively. GVA has the comparative

---

<sup>8</sup> We exclude the Spanish North African territories of Ceuta y Melilla, the Portuguese non-continental territories Azores and Madeira, the French Départements d'Outre-Mer Guadeloupe, Martinique, French Guayana and Réunion, and moreover, Åland (Finland), the Spanish Balearic islands, Corse, Sardegna and Sicilia.

advantage of being the direct outcome of variation in factors that determine regional competitiveness.

All variables are in log form. The dependent variable is labor productivity measured in terms of GVA per worker in 2004, and there are two (non-constant) explanatory variables, per worker labour productivity and human capital in 1995. Human capital is proxied by the skills of the workforce as given by the levels of educational attainment of the active population (aged 15 and over, with tertiary education).

The time period from 1995 to 2004 is short due to a lack of reliable figures for the regions in Central and Eastern Europe. The political changes since 1989 have resulted in the emergence of new or re-established states (the Baltic states, Czech Republic, Slovakia and Slovenia) with only a very short history as sovereign national entities. In most of these states historical series simply do not exist. Even for states such as Hungary and Poland that existed for much longer time periods in their present boundaries, the quality of data referring to the period of central planning imposes serious limitations on the analysis of regional growth. This is closely related to the change in accounting conventions, from the Material Product Balance System to the European System of Accounts 1995. Cross-region comparisons require interregionally comparable regional data which are not only statistically consistent but are also expressed in the same numéraire such as ECU/euro. The absence of market exchange rates in the planned economies is seen as a further impediment.

The definition of a spatial lag in spatial regression models depends on the choice of a spatial weight matrix that summarizes the topology of the data set. Clearly a large number of weight matrices can be derived for the same spatial layout<sup>9</sup>. In this study we employ a first-order contiguity spatial weight matrix, constructed on the basis of digital boundary files in a GIS and implemented in row-standardized form to make the parameter estimates between different models more comparable. Two regions are defined as neighbors when they show a common boundary.

## **(ii) Estimation results and interpretation of the coefficient estimates**

This section presents the estimation results of the spatial Durbin model and quantifies the impact of human capital on productivity growth, using the scalar summary measures suggested by LeSage and Pace and (2008). As previously noted, the spatial error growth regression model, in contrast, estimates only direct effects, and hence may be used as a benchmark for comparison with the direct effects from the SDM specification. Thus, Table 1 reports the parameter estimates, the associated *t*-statistics and standard errors not only of the SDM, but also of the

---

<sup>9</sup> For extensive reviews see Cliff and Ord (1981), Anselin (1988), Anselin and Bera (1998), and Griffith (1995). The latter provides some guidelines for specifying the weight matrix.

SEM specification. A likelihood ratio test rejects the common factor restriction (test statistic: 13.79,  $p = 0.001$ ) and, thus, the SEM specification. This indicates that spatial externalities are substantive phenomena rather than random shocks diffusing through space. The parameter estimate of the spatial autoregressive parameter ( $\hat{\rho} = 0.664$ ) provides evidence for the existence of significant spatial effects working through the dependent variable.

**Table 1.** Raw estimates from SEM and SDM specifications

Variables	SDM			SEM		
	Parameter	Standard error	<i>t</i> -statistic	Parameter	Standard error	<i>t</i> -statistic
Constant	1.0831	0.2098	5.1631	3.3926	0.1915	17.7190
Initial Labor prod.	0.6621	0.0260	25.4287	0.6716	0.0197	34.1043
Human capital	0.1476	0.0198	7.4576	0.1365	0.0194	7.0218
<i>W</i> -initial labor prod.	-0.4150	0.0503	-8.2482	–	–	–
<i>W</i> -human capital	-0.1691	0.0247	-6.8577	–	–	–
Spatial autoregressive parameter	0.6640	0.0598	11.1002	0.7380	0.0500	14.7480
Sigma squared		0.0064			0.0066	
Log-LIK/n		1.3974			1.3626	

*Notes:* The dependent variable is labor productivity in 2004, the independent variables are labor productivity and human capital in 1995. The dependent and the independent variables are in log form. Thus, the coefficient estimates with the independent variables can be interpreted on an elasticity scale.

Table 2 reports the summary direct, indirect and total impact measures for our SDM specification, along with inferential statistics<sup>10</sup>. A comparison of the direct impact estimates and the coefficient estimates associated with the non-spatially lagged variables presented in Table 1 shows that the two sets of estimates are rather similar in magnitude. The smaller direct impact estimate for human capital (0.1317 in comparison to 0.1476) indicates that feedback effects diminished the importance of changes in this variable on productivity growth. The same holds for the case of the initial labor productivity level where the direct impact estimate was 0.6677 whereas the coefficient estimate was 0.6621.

Turning to the indirect impact estimates in Table 2, we observe larger discrepancies between these estimates and the model coefficients on the spatially lagged explanatory variables given in Table 1. The estimates associated with the spatially lagged variables are often interpreted (incorrectly) as measures of the size and significance of indirect impacts in spatial regression models. The differences in the table indicate that this could lead to incorrect inferences about

<sup>10</sup> To produce interpretation measures of dispersion we simply simulate from the multivariate normal distribution with means equal to the maximum likelihood estimates and a variance-covariance matrix based on the numerical Hessian that comes from the maximum likelihood procedure. We insert these in the above formulas to produce a series of 10,000 matrices of the effects, and then construct scalar summary measures based on the trace (diagonal) and off-diagonal averages from these 10,000 matrices. These represent 10,000 scalar summary effects from which we calculate a mean and variance.

the true role of neighboring regions' human capital levels. The parameter estimate given in Table 1 is -0.1691, whereas the average indirect impact for this variable is larger, being -0.1968.

LeSage and Pace (2008) point out that indirect impact estimates can be interpreted in two ways. One involves the impact a region has on all other regions, and the other relates the impact of all other regions on a particular region. In terms of the impact a region has on all other regions, a one percent increase in the level of human capital in a region will on average result in all other regions collectively experiencing a 0.1968 percent drop in labor productivity. This impact is spread out over multiple regions, and thus individual regions will experience a smaller drop.

**Table 2.** Direct, indirect and total impact estimates for SDM  
(*t*-statistics in parentheses)

Variables	SDM		
	Direct impact	Indirect impact	Total impact
Labor productivity	0.6677 (27.5716)	0.0683 (1.8992)	0.7361 (26.3921)
Human capital	0.1317 (6.8644)	-0.1968 (-3.7637)	-0.0650 (-1.1847)

Note: *t*-statistics based on sampled raw parameter estimates of SDM.

The other interpretation involves the impact of all other regions on a particular region, that is, a one percent increase in human capital in all other regions will on average lead to a 0.1968 percent reduction in labor productivity for the region of interest. Although the estimated magnitude of -0.1968 is the same in both cases, it matters whether the interpretative focus is on a typical region's relation to all others (impact from an observation), or all other regions' relation to a typical region (impact on an observation).

Finally, the total impact of human capital on labor productivity may be of interest. For the SDM specification, a one percent increase in the human capital stock in all regions has an insignificant total impact<sup>11</sup> on labor productivity. Compare this with the spatial error regression model specification, where the estimated total impact would be 0.1365. This seems to be intuitively correct, since the spatial error model ignores spatial externalities and interactions across regions and indirect or feedback impacts which are incorporated in the spatial Durbin model.

## 4 Concluding remarks

This paper shows that the unrestricted spatial Durbin model is a suitable model specification for estimating the impact of human capital on regional productivity levels in a cross-sectional regression framework. The analysis provides evidence

<sup>11</sup> Note that the total impact is defined as the sum of the direct and indirect impacts.

for the existence of spatial externalities and interactions of the sort as emphasized by new growth theory, and fits well with the interpretations given by Fingleton and López-Bazo (2006). Since the model involves spatial lags of the dependent and independent variables, the traditional least-squares *ceteris paribus* interpretation of the regression parameters does not hold any longer.

A change in the human capital variable in region (observation)  $i$  has a direct impact on region  $i$  as well as an indirect impact on neighboring regions  $j \neq i$ . This is a result of the spatial connectivity relationships incorporated in the model. We use the scalar summary measures suggested by LeSage and Pace (2008), and exploit the multivariate normal distribution of the maximum likelihood parameter estimates to draw inferences regarding the statistical significance of the direct and indirect impacts that arise from changes in the human capital variable. The results obtained shed some interesting light on the role given to human capital in European growth. First, a *ceteris paribus* increase in the level of human capital has a significant and positive direct impact on regional productivity growth. This is what we expect since human capital has long been stressed as a pre-requisite for economic growth. Second, this positive direct impact is offset by a significant and negative indirect impact producing a negative total effect that is not significantly different from zero. Reflecting on this result, we note that it seems more intuitive to think in terms of what LeSage and Pace (2008) label the *average total impact on an observation view* of a change in the human capital levels during the initial period. The intuition here arises from the notion that it is relative regional advantages in human capital that matter most for productivity growth.

The inferences were made conditional on the data and the specification of the spatial weight matrix. The assumption that a particular spatial weight matrix specification is correct might be relaxed by treating spatial weight specification as an additional unknown feature, that is, by explicitly incorporating model uncertainty in the statistical analysis. To accommodate this uncertainty issue one might follow LeSage and Fischer (2008) in endorsing the use of Bayesian methods such as Bayesian model averaging in combination with Markov Chain Monte Carlo Model Composition. Another avenue for future research is to extend our framework to allow not only for geographical, but also for time dependence. This would permit us to study the impact of human capital over time.

**Acknowledgements.** The authors gratefully acknowledge the grant no. P19025-G11 provided by the Austrian Science Fund (FWF). They thank James LeSage for initiating this research and for helpful suggestions. All computations were made using his Spatial Econometrics library, <http://www.spatial-econometrics.com/>.

## References

Abreu M, de Groot HLF, Florax RJGM (2005) Space and growth: A survey of empirical evidence and methods. *Région et Développement* 21: 12-43

- Anselin L (1988) *Spatial econometrics: Methods and models*. Kluwer, Dordrecht
- Anselin L. (1990) Some robust approaches to testing and estimation in spatial econometrics. *Regional Science and Urban Economics* 20 (2): 141-163
- Anselin L (2003) Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review* 26 (2): 153-166
- Anselin L, Bera AK (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A, Giles D (eds) *Handbook of Applied Economic Statistics*. Marcel Dekker, New York, pp 237-289
- Bivand R, Brunstad R (2006) Regional growth in Western Europe: Detecting spatial misspecification using the R environment. *Papers in Regional Science* 85(2): 277-297
- Brock WA, Durlauf SN (2001) Growth empirics and reality. *The World Bank Economic Review* 15(2): 229-272
- Burridge J (1981) A note on maximum likelihood estimation for regression models using grouped data. *Journal of the Royal Statistical Society, Series B* 43(1): 41-45
- Cliff A, Ord JK (1981) *Spatial processes: Models and applications*. Pion, London
- Efron B (1982) *The jackknife, the bootstrap, and other resampling plans*. Society for Industrial and Applied Mathematics, Philadelphia
- Efron B, Tibshirani R (1993) *An introduction to the bootstrap*. Chapman & Hall, New York
- Fingleton B, López-Bazo E (2006) Empirical growth models with spatial effects. *Papers in Regional Science* 85(2): 177-198
- Fischer MM, Stürböck C (2006) Pan-European regional income growth and club-convergence. *Annals of Regional Science* 40(4): 693-721
- Griffith DA (1995) Some guidelines for specifying the geographic weights matrix contained in spatial statistics models. In: Arlinghaus SL, Griffith DA, Arlinghaus WC, Drake WD, Nystrom JD (eds) *Practical handbook of spatial analysis*. CRC Press, Boca Raton, pp 65-82
- Kim CW, Phipps TT, Anselin L (2003) Measuring the benefits of air quality improvement: A spatial hedonic approach. *Journal of Environmental Economics and Management* 45 (1): 24-39
- LeSage JP, Fischer MM (2008) *Spatial growth regressions: Model specification, estimation and interpretation*. Available at SSRN: <http://ssrn.com/abstract=980965>
- LeSage JP, Pace RK (2008) *An introduction to spatial econometrics*. Chapman & Hall/CRC (December 26, 2008) [forthcoming]
- Mooney CZ, Duval RD (1993) *Bootstrapping. A nonparametric approach to statistical inference*. Sage Publications, Newbury Park, London and New Delhi
- Mur J, Angulo A (2006) The spatial Durbin model and the common factor tests. *Spatial Economic Analysis* 1(2): 207-226
- Openshaw S, Taylor P (1979) A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In: Wrigley N (ed) *Statistical applications in the spatial science*. Pion, London, pp 127-144
- Pace RK, Barry R (1997) Quick computation of spatial autoregressive estimators. *Geographical Analysis* 29: 232-246