



## Documents de Travail

Laboratoire d'Économie de Dijon  
Université de Bourgogne –  
CNRS UMR 6307  
Inserm U1200

### Working Papers

#### **Technology Spillover and TFP Growth: a Spatial Durbin Model**

Aligui Tientao, Diègo Legros and Marie-Claude Pichery

[2015 – 03]

« Economie des Territoires et de l'Environnement »



Pôle d'Économie et de gestion - 2 boulevard Gabriel - BP 26611 - F 21066 DIJON CEDEX  
[Secretariat.ledi@u-bourgogne.fr](mailto:Secretariat.ledi@u-bourgogne.fr) - Tél : 03 80 39 54 41 – Fax : 03 80 39 54 43

[ledi.u-bourgogne.fr](http://ledi.u-bourgogne.fr)

# Technology Spillover and TFP Growth: a Spatial Durbin Model

A. Tientao<sup>a,\*</sup>, D. Legros<sup>a,†</sup> and M.C. Pichery<sup>a,‡</sup>

<sup>a</sup>Université de Bourgogne, Laboratoire d'Économie de Dijon

March 18, 2015

## Abstract

This paper contributes to the extensive literature on technological spillovers by examining the major determinants of total-factor productivity (TFP) evolution. Beginning with a model in which technological progress is reflected by product variety, we decompose TFP into quality and variety components. We address the quality component by introducing a country's distance to the technological frontier. Quality is assumed to be a negative function of the technological gap of country  $i$  with respect to its own technological frontier. This technological frontier is defined as the geometric means of knowledge levels in all countries. We deal with the variety component by using R&D expenditure combined with human capital stocks. We obtain a spatial Durbin structure in TFP growth that can be estimated using spatial econometrics tools. Our TFP growth model is estimated from a sample of 107 countries for the period 2000-2011. The main focus is on the role played by technological spillovers. They impact productivity growth substantially, as do traditional factors such as R&D and human capital stock. Technological spillovers are captured by the spatial autocorrelation coefficient and the indirect impact of R&D.

**KEYWORDS:** Diffusion; Productivity; R&D; Spatial Auto-correlation.

**JEL:** R12; E23; O32; C21.

---

\*Université de Bourgogne, Laboratoire d'Économie de Dijon UMR 6307 - CNRS, 2 boulevard Gabriel, BP 26611, 21066 DIJON CEDEX, Phone: +33 (0)3 8039 5430, E-mail: [aligui.tientao@u-bourgogne.fr](mailto:aligui.tientao@u-bourgogne.fr) corresponding author

†Université de Bourgogne, Laboratoire d'Économie de Dijon UMR 6307 - CNRS, 2 boulevard Gabriel, BP 26611, 21066 DIJON CEDEX, Phone: +33 (0)3 8039 3520, E-mail: [diego.legros@u-bourgogne.fr](mailto:diego.legros@u-bourgogne.fr)

‡Université de Bourgogne, Laboratoire d'Économie de Dijon UMR 6307 - CNRS, 2 boulevard Gabriel, BP 26611, 21066 DIJON CEDEX, Phone: +33 (0)3 8039 5434, E-mail: [marie-claude.pichery@u-bourgogne.fr](mailto:marie-claude.pichery@u-bourgogne.fr)

# 1 Introduction

Continued economic growth depends on our ability to maintain and increase current levels of innovation. Governments implement a wide range of policies to promote innovation including in R&D, intellectual property rights, education, labour markets, financial markets and product market regulations. Improving the business environment in order to encourage innovation is an especially important policy area and open trade is conducive to the free flow of technologies across borders, enhanced competitive pressure and the opening up of new markets. International trade provides a way for global firms to exploit innovations, and it is also a major source of innovation (Grossman and Helpman, 1991).

There is a mass of theoretical research showing that international openness impacts growth and productivity positively in various ways (Aghion and Howitt, 2009 Chap. 15). Trade can boost productivity because producers gain access to new imported varieties of inputs. This can reduce the cost of innovation engendering more variety creation in the future. The effect of increased product variety on productivity is thought to depend upon the elasticity of substitution among different varieties of a good, and/or upon shifts in the apportionment of expenditure among new, remaining, and disappearing goods. Increasing the number of varieties does not appear to affect productivity much if new varieties are close substitutes for existing varieties or if the proportion of new varieties is small relative to existing ones (Broda *et al.*, 2006).

Since the seminal paper of Coe and Helpman (1995), several empirical studies have documented that R&D cross-country spillovers, through the channel of trade flows, have been an important engine of TFP growth in the industrialized countries (Coe *et al.*, 1997; Bayoumi *et al.*, 1999; Crespo *et al.*, 2004; Coe *et al.*, 2009). Coe and Helpman (1995) test the prediction of the trade and growth models of Grossman and Helpman (1991) and Rivera-Batiz and Romer (1991) in which foreign R&D creates new intermediate inputs and perhaps spillovers that the home country can access through imports. Subsequent studies reveal that productivity spillovers arising from international openness are largely determined by the host country's capability to absorb and innovate. True, a large technology gap between local and foreign firms may signal considerable "catch-up" potential; however, it may also indicate the very poor absorptive capabilities of the local partners (Blomström and Sjöholm, 1999). The availability of adequate human capital and basic infrastructure facilities is crucial for the adoption and development of advanced technologies (Borensztein *et al.*, 1998). Empirical studies have reported that trade enhances competitive pressure. For instance, fierce competition arising from the entry of multinational corporations (MNCs) is found to be detrimental to the economy because it crowds out the least efficient domestic firms (Kokko, 1996).

Other channels of international technology diffusion have been examined. For example, Keller and Yeaple (2009) examine R&D spillover by substituting bilateral measures

of Foreign Direct Investment (FDI) for imports. Lee (2006) uses bilateral technological proximity and patent citations between countries while Keller (2002) and Ertur and Musolesi (2013) use geographic location.

While trade and FDI are known to be important for the performance of innovation systems, not enough is known about how trade affects the innovation process. This paper aims to contribute to the extensive literature on technological spillovers by analysing the major determinants of TFP evolution. The main contribution of this paper to the existing literature is twofold. First, we propose an alternative method for estimating technology spillovers based on a model in which technological progress shows up as an expansion of the number of varieties of products. To this end, we focus on the impact of international openness in terms of foreign trade and geographical proximity on productivity. Second, we decompose TFP into two components: quality and variety. We address the quality component by determining a country's distance from the technological frontier. A country that is far from the technology frontier derives a certain advantage from this deficit, because it can grow rapidly simply by adopting technologies that have already been developed in more advanced countries. Technology transfer will stabilize the gap between rich and poor countries, allowing the poor countries to grow as fast as the rich. We assume that quality is a negative function of the technological gap of country  $i$  with respect to its own technological frontier. This technological frontier is defined as the geometric means of knowledge levels in all countries.

Technological knowledge is often tacit and circumstantially specific. It cannot simply be copied and transplanted to another country. Instead, the receiving country must have a certain capacity in order to master the technology and adapt it to local conditions. We use R&D expenditure combined with human capital stocks to deal with the variety component and obtain a spatial Durbin model that can be estimated using spatial econometrics tools. This enables us to capture both the direct and indirect effects of R&D through trade and geographical proximity on TFP growth.

It is worth noting that our first contribution uses a novel method for estimating technology spillovers. This method allows us to account for the quality and variety components of TFP simultaneously.

We use a sample of 107 countries over the period 2000-2011 to estimate our TFP growth model. The role played by technological spillovers is the main focus of this analysis. The main results from our estimations are that, in addition to traditional factors such as R&D and human capital stock, technological spillovers have a strong impact on productivity growth. These spillovers are captured by the spatial autocorrelation coefficient and by the indirect impact of R&D and human capital.

The remainder of the paper has the following structure. In Section 2 we lay out the theoretical model. Section 3 presents the estimation procedure. Section 4 is about data

and estimation results and Section 5 concludes.

## 2 The theoretical model

We consider models in which technological progress shows up as an expansion of the number of varieties of products (Romer 1987, 1990). We think of a change in this number as a basic innovation, akin to opening up a new industry. Of course, the identification of the state of technology with the number of varieties of products should be viewed as a metaphor; it selects one aspect of technical advance and thereby provides a tractable framework in which to study long-term growth.

Another metaphor has been developed in which progress shows up as quality improvements for an array of existing kinds of products (Grossman and Helpman, 1991a; Aghion and Howitt, 1992). These quality enhancements represent the more or less continuous process of upgrading that occurs within an established industry. The two metaphors should be viewed as complementary and not opposing approaches (Barro and Sala-i-Martin, 2003).

### 2.1 Production relations

A common point of departure in the literature is to start from a stylized regional production function to model the transmission channels of trade and FDI activity as well as additional private and public inputs to economic growth. A spatially extended version of the production function approach is presented in Ertur and Koch (2007), for instance. We consider a single country in a world economy with  $n$  different countries. There is a fixed number  $L$  of people, each of whom lives forever and has a constant flow of one unit of labour that can be used in manufacturing. For simplicity we suppose that no one has a demand for leisure time, so each person offers her one unit of labour for sale inelastically (that is, no matter what the wage rate). Her utility in each period depends only on consumption, according to the same isoelastic function (Aghion and Howitt, 2009 Chap. 3).

$$u(c) = \frac{c^{1-\varepsilon}}{1-\varepsilon}, \quad \varepsilon > 0$$

and she discounts utility using a constant rate of time preference  $\rho$ . This means that in the steady state the growth rate  $g$  and the interest rate  $r$  must obey the Euler equation, which can be written as:

$$g = \frac{r - \rho}{\varepsilon}$$

There is one final good  $Y_i(t)$ , produced under perfect competition by labour  $L_i(t)$  and a continuum of intermediate products, indexed by  $v$  in the interval  $[0, M_i(t)]$ .  $M_i(t)$  is

our measure of product variety. We follow Broda *et al.* (2006) by writing the production function as

$$Y_i(t) = (A_i(t)L_i(t))^{1-\alpha} \left[ \int_0^{M_i(t)} x_{i,v}^\nu(t) dv \right]^{\frac{\alpha}{\nu}} \quad (1)$$

where  $A_i(t)$  is a productivity parameter,  $\alpha \in [0, 1]$  is one minus the share of labour in output and  $\nu \in [0, 1]$  measures the elasticity of substitution between varieties of input goods  $x_{i,\nu}(t)$ , with a higher  $\nu$  corresponding to more substitutable inputs.

All intermediates enter symmetrically into the production function, and all command the same price. At equilibrium, each intermediate is demanded to the same extent  $x_i(t) = x_{i,\nu}(t)$  (Grossman and Helpman 1991). Using this fact, (1) can be simplified to

$$Y_i(t) = (A_i(t)L_i(t))^{1-\alpha} M_i(t)^{\frac{\alpha}{\nu}} x_i^\alpha(t) \quad (2)$$

Each intermediate product is produced using the final good as input, one for one. That is, each unit of intermediate product  $v$  produced requires the input of one unit of final good (Aghion and Howitt, 2009 Chap. 3). According to this one-for-one technology, the aggregate capital stock is given by  $K_i(t) = M_i(t)x_i(t)$ . Using this fact, we can rewrite (2) as

$$Y_i(t) = A_i(t)^{1-\alpha} L_i(t)^{1-\alpha} M_i(t)^{\left(\frac{1-\nu}{\nu}\right)\alpha} K_i(t)^\alpha \quad (3)$$

From equation (3) we can specify the total factor productivity (TFP) as follows:<sup>1</sup>

$$Z_i(t) = \frac{Y_i(t)}{L_i(t)^{1-\alpha} K_i(t)^\alpha} \quad (4)$$

Plugging (3) into (4) yields:

$$Z_i(t) = A_i(t)^{1-\alpha} M_i(t)^{\left(\frac{1-\nu}{\nu}\right)\alpha} \quad (5)$$

Unlike in Coe and Helpman (1995, 2009) and Keller (1998), this measure of TFP has two components: a product-variety component captured by the term in  $M_i(t)^{\left(\frac{1-\nu}{\nu}\right)\alpha}$  and a quality component embodied in the term in  $A_i(t)^{1-\alpha}$ .

## 2.2 Quality of knowledge and expanding variety

Equation (5) shows that TFP depends on quality of innovation and product-variety. In order to capture the quality component, we draw on Ertur and Koch, 2011 by defining  $A_i(t)^{1-\alpha}$  as:

$$A_i(t)^{1-\alpha} = \zeta \prod_{j=1}^n \left( \frac{Z_j(t)}{Z_i(t)} \right)^{\gamma w_{ij}} \quad (6)$$

---

<sup>1</sup>See Coe and Helpman (1995, 2009) and Keller (1998).

where  $\gamma \in [-1, 1]$  captures the degree of technology diffusion. We assume that quality is a negative function of the technological gap of country  $i$  with respect to its own technological frontier. This technological frontier is defined as the geometric mean of knowledge levels in all countries denoted by  $Z_j(t)$ , for  $j = 1, 2, \dots, n$ . We can also define the technology frontier as the world or global technological leader. But this approach requires a normalization of the technology frontier to a reference country, so each computed technology frontier must be interpreted relative to a particular country that has to be chosen in advance. Moreover, the specification proposed in this paper encompasses the particular case of the world or global technological leader. We assume that the interaction terms  $w_{ij}$  are non negative, finite and non stochastic. The gap with respect to the technological frontier determines the quality of productivity. Indeed, the closer a country is to its own technological frontier the higher is its productivity quality.

Plugging (6) into (5) yields:

$$Z_i(t) = \zeta \prod_{j=1}^n \left( \frac{Z_j(t)}{Z_i(t)} \right)^{\gamma w_{ij}} M_i(t)^{\left(\frac{1-\nu}{\nu}\right)\alpha} \quad (7)$$

For the product variety component, we follow Grossman and Helpman (1991) in assuming that in a world with international trade in goods and services, foreign direct investment, and an international exchange of information and dissemination of knowledge, a country's productivity depends on its own R&D as well as on the R&D efforts of its trading partners. In another world, a country's level of productivity will be related to the number of contacts that local agents have with their counterparts in the international and business communities. Explicitly, we have:

$$M_i(t)^{\left(\frac{1-\nu}{\nu}\right)\alpha} = R_i^\theta(t) H_i^\psi(t) \prod_{j=1}^n \left( R_j^\theta(t) H_j^\psi(t) \right)^{\gamma w_{ij}} \quad (8)$$

where  $\theta > 0$  and  $\psi > 0$  are the elasticities of R&D and human capital stock. We therefore suppose that country  $i$ 's product variety depends on its own R&D expenditure  $R_i(t)$  and on R&D of all countries, denoted by  $R_j(t)$ ,  $j = 1, 2, \dots, n$ . The term  $H_i$  captures country's  $i$  ability and absorption capacity which is measured by the human capital stock.

Plugging (8) into (7) yields:

$$Z_i(t) = \zeta \prod_{j=1}^n \left( \frac{Z_j(t)}{Z_i(t)} \right)^{\gamma w_{ij}} R_i^\theta(t) H_i^\psi(t) \prod_{j=1}^n \left( R_j^\theta(t) H_j^\psi(t) \right)^{\gamma w_{ij}} \quad (9)$$

Taking (9) in logarithm form yields:

$$\ln Z_i(t) = \ln \zeta - \ln Z_i(t) + \gamma \sum_{j=1}^n w_{ij} \ln Z_j(t) + \theta \ln R_i(t) + \psi \ln H_i(t) + \gamma \theta \sum_{j=1}^n w_{ij} \ln R_j(t) + \gamma \psi \sum_{j=1}^n w_{ij} \ln H_j(t)$$

Arranging the terms, we obtain:

$$\begin{aligned} \ln Z_i(t) = & \ln \zeta/2 + \gamma/ \sum_{j=1}^n w_{ij} \ln Z_j(t) + \theta/2 \ln R_i(t) + \psi/2 \ln H_i(t) + \\ & \gamma\theta/2 \sum_{j=1}^n w_{ij} \ln R_j(t) + \gamma\psi/2 \sum_{j=1}^n w_{ij} \ln H_j(t) \end{aligned} \quad (10)$$

Equation (10) can be rewritten as:

$$\begin{aligned} \ln Z_i(t) = & \beta_0 + \beta_1 \ln R_i(t) + \beta_2 \ln H_i(t) + \rho \sum_{j=1}^n w_{ij} \ln Z_j(t) + \lambda_1 \sum_{j=1}^n w_{ij} \ln R_j(t) \\ & + \frac{\gamma\psi}{2} \sum_{j=1}^n w_{ij} \ln H_j(t) \end{aligned} \quad (11)$$

where  $\beta_0 \equiv \frac{\ln \zeta}{2} > 0$  is the constant term;  $\beta_1 \equiv \frac{\theta}{2} > 0$  is the coefficient that captures the impact of country  $i$ 's R&D;  $\beta_2 \equiv \frac{\psi}{2} > 0$  is the coefficient associated with human capital;  $\rho \equiv \frac{\gamma}{2} > 0$  is the spatial autocorrelation coefficient that captures knowledge diffusion from neighbouring countries;  $\lambda_1 \equiv \frac{\gamma\theta}{2} > 0$  measures the average impact of neighbouring countries' R&D and  $\lambda_2 \equiv \frac{\gamma\psi}{2} > 0$  is the coefficient that captures the average impact of neighbouring countries' human capital.

In matrix form we obtain:

$$\mathbf{Z} = \mathbb{I}\beta_0 + \mathbf{R}\beta_1 + \mathbf{H}\beta_2 + \lambda_1 \mathbf{W}\mathbf{R} + \lambda_2 \mathbf{W}\mathbf{H} + \rho \mathbf{W}\mathbf{Z} \quad (12)$$

where  $\mathbf{Z} = \ln Z_i(t)$ , a matrix ( $n \times 1$ ) of TFP growth;  $\mathbb{I}$ , a matrix ( $n \times 1$ ) of 1;  $\mathbf{R} = \ln R_i(t)$  a matrix ( $n \times 1$ ) of R&D;  $\mathbf{H} = \ln h_i(t)$ , a matrix ( $n \times 1$ ) of human capital stock;  $\mathbf{W} = \sum_{j=1}^n w_{ij}$  is our interaction matrix ( $n \times n$ ). Equation (12) is a version of the well known specification in the spatial econometric literature referred to as the Spatial Durbin model (SDM). This kind of econometric specification includes spatial lags of all the exogenous variables in addition to the spatial lag of the endogenous variable.

## 3 Estimation procedure

### 3.1 SDM model estimation

When the spatial autocorrelation is modelled, ordinary least squares regression (OLS) is no longer appropriate: the estimates obtained by this method are not convergent if there is a lagged endogenous variable and they are inefficient in the presence of spatial autocorrelation. Other estimation methods are then necessary to find convergent and efficient estimates. The method widely used is that of maximum likelihood (Lee, 2004;



LeSage and Pace, 2009).<sup>2</sup> Consider the following SDM model:

$$y = \alpha \iota_n + X\beta + WX\theta + \rho Wy + \varepsilon \quad (13)$$

$$y = (I_n - \rho W)^{-1}(\alpha \iota_n + X\beta + WX\theta + \rho Wy + \varepsilon)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

where 0 represents an  $n \times 1$  vector of zeros and  $\iota_n$  an  $n \times 1$  vector of ones associated with the constant term parameter  $\alpha$ . This model can be written as a SAR<sup>3</sup> model by defining:  $Z = [\iota_n \ X \ WX]$  and  $\delta = [\alpha \ \beta \ \theta]'$ . This means that the likelihood function for SAR and SDM models can be written in the same form. The log-likelihood function for the SDM and SAR models takes the following form:

$$\ln L = -(n/2) \ln(2\pi\sigma^2) + \ln |I_n - \rho W| - \frac{e' e}{2\sigma^2}$$

$$e = y - \rho Wy - Z\delta$$

$$\rho \in (\min(\omega)^{-1}, \max(\omega)^{-1})$$

where:  $Z = [\iota_n \ X]$  for the SAR model and  $Z = [\iota_n \ X \ WX]$  for the SDM model;  $\omega$  is the  $n \times 1$  vector of eigenvalues of the matrix  $W$  (LeSage and Pace, 2009). If  $\omega$  contains only real eigenvalues, a positive definite variance-covariance matrix is ensured by the condition:  $\rho \in (\min(\omega)^{-1}, \max(\omega)^{-1})$  (Ord, 1975).

Maximizing the log-likelihood for the SAR model would involve setting the first derivatives with respect to the parameters  $\beta$ ,  $\sigma^2$  and  $\rho$  equal to zero and simultaneously solving these first-order conditions for all parameters. In contrast, the equivalent maximum likelihood estimates could be found using the log-likelihood function concentrated with respect to the parameters  $\beta$  and  $\sigma^2$ . This involves substituting closed-form solutions from the first order conditions for the parameters  $\beta$  and  $\sigma^2$  to yield a log-likelihood that is said to be concentrated log-likelihood function with respect to these parameters.

$$y = Z\delta + \rho Wy + \varepsilon \quad (14)$$

From the model statement (14), if the true value of the parameter  $\rho$  was known to be say  $\rho^*$ , we could rearrange the model statement in (14) as:

$$y - \rho^* Wy = Z\delta + \varepsilon \quad (15)$$

This suggests an estimate for  $\delta$  of  $\hat{\delta} = (Z'Z)^{-1}Z'(I_n - \rho^* W)y$ . In this case we could also find an estimate for the noise variance parameter  $\hat{\sigma}^2 = n^{-1}e(\rho^*)'e(\rho^*)$ , where  $e(\rho^*) =$

<sup>2</sup>For estimation we used James LESAGE's Econometrics Toolbox which is available at <http://www.spatial-econometrics.com/>.

<sup>3</sup>Spatial Auto-Regressive model:  $y = \alpha \iota_n + X\beta + \rho Wy + \varepsilon$

$y - \rho^*Wy - Z\widehat{\delta}$ . These ideas mean that we can concentrate the full (log) likelihood with respect to the parameter  $\beta$ ,  $\sigma^2$  and reduce the maximum likelihood to a univariate optimization problem in the parameter  $\rho$ .

Working with the concentrated log-likelihood yields exactly the same maximum likelihood estimates  $\widehat{\beta}$ ,  $\widehat{\sigma}$ , and  $\widehat{\rho}$  as would arise from maximizing the full log-likelihood (Davidson and Mackinnon, 1993).

As noted, the log-likelihood can be concentrated with respect to the coefficient vector  $\delta$  and the noise variance parameter  $\sigma^2$ . Pace and Barry (1997) suggest a convenient approach for concentrating out the parameters  $\delta$  and  $\sigma^2$ .

$$\ln L(\rho) = \kappa + \ln |I_n - \rho W| - (n/2) \ln(S(\rho)) \quad (16)$$

$$S(\rho) = e(\rho)'e(\rho) = e_0'e_0 - 2\rho e_0'e_d + \rho^2 e_d'e_d$$

$$e(\rho) = e_0 - \rho e_d$$

$$e_0 = y - Z\delta_0$$

$$e_d = Wy - Z\delta_d$$

$$\delta_0 = (ZZ)^{-1}Z'y$$

$$\delta_d = (ZZ)^{-1}Z'Wy$$

The term  $\kappa$  is a constant that does not depend on the parameter  $\rho$ , and  $|I_n - \rho W|$  is the determinant of the  $n \times n$  matrix. We use the notation  $e(\rho)$  to indicate that this vector depends on the values taken by the parameter  $\rho$ , as does the scalar concentrated log-likelihood function value  $\ln L(\rho)$ .

To simplify optimization of the log-likelihood with respect to the parameter  $\rho$ , Pace and Barry (1997) propose evaluating the log-likelihood using a  $q \times 1$  vector of values for  $\rho$  in the interval  $[\rho_{\min}, \rho_{\max}]$ , labelled  $\rho_1, \rho_2, \rho_3, \dots, \rho_q$ .

$$\begin{pmatrix} \ln L(\rho_1) \\ \ln L(\rho_2) \\ \vdots \\ \ln L(\rho_q) \end{pmatrix} = \kappa + \begin{pmatrix} \ln |I_n - \rho_1 W| \\ \ln |I_n - \rho_2 W| \\ \vdots \\ \ln |I_n - \rho_q W| \end{pmatrix} - (n/2) \begin{pmatrix} \ln(S(\rho_1)) \\ \ln(S(\rho_2)) \\ \vdots \\ \ln(S(\rho_q)) \end{pmatrix} \quad (17)$$

Given a sufficiently fine grid of  $q$  values for the log-likelihood, interpolation can supply intervening points to any desired precision (which follows from the smoothness of the log-likelihood function). Note, the scalar moments  $e_0'e_0$ ,  $e_0'e_d$ , and  $e_d'e_d$ , and the  $k \times 1$  vectors  $Z\delta_0$ ,  $Z\delta_d$  are computed prior to optimization and so, given a value for  $\rho$ , calculating  $S(\rho)$  simply involves weighting three numbers. Given the optimum value of  $\rho$ , this becomes the maximum likelihood estimate of  $\rho$  denoted as  $\widehat{\rho}$ . Therefore, it requires very little computation to arrive at the vector of concentrated log-likelihood values.

Given the maximum likelihood estimate  $\hat{\rho}$ , the maximum likelihood estimates for the coefficients  $\hat{\delta}$ , the noise variance parameter  $\hat{\sigma}^2$ , and associated variance-covariance matrix for the disturbances are:

$$\hat{\delta} = \delta_0 - \hat{\rho}\delta_d \quad (18)$$

$$\hat{\sigma}^2 = n^{-1}S(\hat{\rho}) \quad (19)$$

$$\hat{\Omega} = \hat{\sigma}^2[(I_n - \rho W)'(I_n - \rho W)]^{-1} \quad (20)$$

Maximum likelihood estimation could proceed using a variety of univariate optimization techniques. These could include the vectorized approach just discussed based on a fine grid of values of  $\rho$  (large  $q$ ), non-derivative search methods such as the Nelder-Mead simplex or bisection search scheme, or derivative-based optimization techniques (Press *et al.*, 1996). Some form of Newton's method with numerical derivatives has the advantage of providing the optimum as well as the second derivative of the concentrated log-likelihood at the optimum  $\hat{\rho}$ . This numerical estimate of the second derivative in conjunction with other information can be useful in producing a numerical estimate of the variance-covariance matrix for the parameter.

### 3.2 Interpreting parameter estimates

Linear regression parameters can be interpreted simply as the partial derivation of the dependent variable with respect to the explanatory variable. This arises from linearity and the assumed independence of observation in the model:  $y = \sum_{r=1}^k x_r \beta_r + \varepsilon$ . The partial derivatives of  $y_i$  with respect to  $x_{ir}$  have a simple form:  $\partial y_i / \partial x_{ir} = \beta_r$  for all  $i, r$ ; and  $\partial y_i / \partial x_{jr} = 0$ , for  $j \neq i$  and all variables  $r$ .

In models containing spatial lags of explanatory or dependent variables, interpretation of the parameters becomes richer and more complicated. A number of researchers have noted that models containing spatial lags of the dependent variable require special interpretation of the parameters (Anselin and LeGallo, 2006; Kelejian *et al.*, 2006; Kim *et al.*, 2003; LeGallo *et al.*, 2003).

Spatial regression models expand the information set to include information from neighbouring regions/observations. Consider the SDM model which we have re-written as:

$$\begin{aligned} (I_n - \rho W)y &= X\beta + WX\theta + \iota_n + \varepsilon \\ y &= \sum_{r=1}^k S_r(W)x_r + V(W)\iota_n\alpha + V(W)\varepsilon \\ S_r(W) &= V(W)(I_n\beta_r + W\theta_r) \\ V(W) &= (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots \end{aligned} \quad (21)$$

Equation (21) can be re-written as:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{r=1}^k \begin{pmatrix} S_r(W)_{11} & S_r(W)_{12} & \dots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & & \\ \vdots & \vdots & \ddots & \\ S_r(W)_{n1} & S_r(W)_{n2} & \dots & S_r(W)_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{pmatrix} \quad (22)$$

$$+ V(W)_{\ell n} \alpha + V(W) \varepsilon$$

It follows from (22) that the derivative of  $y_i$  with respect to  $x_{jr}$  is potentially non-zero, taking a value determined by the  $i, j$ th element of the matrix  $S_r(W)$ .

$$\partial y_i / \partial x_{jr} = S_r(W)_{ij}$$

An implication of this is that a change in the explanatory variable for a single region can potentially affect the dependent variable in all other regions. It is also the case that the derivative of  $y_i$  with respect to  $x_{ir}$  does not usually equal  $\beta_r$  as in least-squares.

$$\partial y_i / \partial x_{ir} = S_r(W)_{ii}$$

The own derivative for the  $i$ th region measures the impact on the dependent variable observation  $i$  from a change in  $x_{ir}$ . This impact includes the effect of feedback loops where observation  $i$  affects observation  $j$  and observation  $j$  also affects observation  $i$  as well as longer paths which might go from observation  $i$  to  $j$  to  $k$  and back to  $i$ .

## 4 Empirical Implementation

### 4.1 Data

Our study uses a sample of 107 countries for the period 2000-2011. The sample contains 25 African countries, 21 American countries, 23 Asian countries, 36 European countries and 2 Oceanic countries. We extract our basic data from the Feenstra *et al.* (2013) Penn World Table (PWT version 8.0). This database contains information on TFP growth, and the index of human capital per person (among many other variables) for a large number of countries. We measure all variables for  $i = 1, \dots, n$  as the average over the period 2000-2011. Our index of human capital per person is based on years of schooling (Barro and Lee, 2012) and returns to education (Psacharopoulos, 1994).

R&D is often said to have two faces: the first is innovation, while the second is to facilitate the understanding and imitation of others' discoveries. The latter is related to absorptive capacity and provides for efficient technology transfer. R&D is likely to take place at the firm or industry level, but will ultimately promote overall economic development through enhanced productivity. R&D has two sources, domestic (as already

described), or it can be generated from international spillovers. The literature seems to suggest that both channels are important for TFP growth. R&D expenditure data are from the United Nations Educational, Scientific and Cultural Organization.

The interaction matrix  $\mathbf{W}$  corresponds to the so-called spatial weights matrix commonly used in spatial econometrics to model spatial interdependence between observations (LeSage and Pace, 2009). More precisely, each country is connected to a set of the neighbouring countries by means of a purely spatial pattern introduced exogenously in  $\mathbf{W}$ . Elements  $w_{ii}$  on the main diagonal are set to zero by convention, whereas elements  $w_{ij}$  indicate the way country  $i$  is spatially connected to country  $j$ . In order to normalize the outside influence upon each country, the weight matrix is standardized such that the elements of a row sum up to one. For the variable  $x$ , this transformation means that the expression  $Wx$ , called the spatial lag variable, is simply the weighted average of the neighbouring observations. It is important to stress that the friction terms  $w_{ij}$  should be exogenous to the model (Ertur and Koch, 2007). Traditionally, connectivity has been understood as geographical proximity, and various weight matrices based on geographical space have thus been used in the spatial econometrics literature, such as contiguity, nearest neighbours and geographical distance-based matrices. However, the definition is in fact much broader and can be generalized to any network structure to reflect any kind of interactions between observations. This is why we prefer to use the term interaction matrix for  $\mathbf{W}$ .

We specify two different interaction matrices. The first,  $\mathbf{Wt}$ , is defined as follows:

$$w_{ij} = \begin{cases} 0 & \text{if } i = j \\ M_{ij} & \text{if } i \neq j \end{cases}$$

where  $M_{ij}$  is defined as the average imports of country  $i$  from country  $j$  over the 1990-1999 period to prevent endogeneity problems that might arise. We design the second interaction matrix  $\mathbf{Wd}$  using a decreasing function of pure geographical distance. This interaction matrix is defined as follows:

$$w_{ij} = \begin{cases} 0 & \text{if } i = j \\ \frac{1}{d_{ij}^2} & \text{if } i \neq j \end{cases}$$

Trade flows are from the UN COMTRADE database.<sup>4</sup> We use bilateral distances (in kilometres) between capital cities taken from the CEPII database.<sup>5</sup> They are computed using the great circle distance formula applied to the capitals' geographic coordinates.

In order to visualize potential interaction patterns between countries, we consider four interdependent countries. The structure of interaction is represented in the following  $(4 \times 4)$  matrix:

---

<sup>4</sup><http://comtrade.un.org/>

<sup>5</sup><http://www.cepii.fr/>

$$W = \begin{pmatrix} 0 & M_{12} & M_{13} & M_{14} \\ M_{21} & 0 & M_{23} & M_{24} \\ M_{31} & M_{32} & 0 & M_{34} \\ M_{41} & M_{42} & M_{43} & 0 \end{pmatrix}$$

The flows of technology between countries go from country  $j$  to country  $i$  (for instance  $M_{23}$  represents the flow from country 3 to country 2). In other words, each row represents the receiving country and each column represents the emitting country.

Since the Moran test for spatial autocorrelation among residuals shows that there is no spatial correlation in error terms (See Table 1 columns 2 and 3),<sup>6</sup> we do not need to use a spatial econometric specification that takes into account the spatial lag of the error term, such as the spatial error model (SEM), general spatial model (GSM) or spatial Durbin error model (SDEM)(see Anselin, 1988a).

In the next section, we perform several estimation procedures depending on the spatial interdependencies. First, we estimate the model (12) without spatial interdependencies (i.e.  $\rho = 0$ ) using OLS. Second, we estimate two versions of spatial models using our two interaction matrices  $\mathbf{Wt}$  and  $\mathbf{Wd}$ : a SAR model and an SDM model.

## 4.2 Estimation results

As explanatory variables in our SDM regression model we use R&D, a constant term, and the human capital index for the period 2000-2011. Since this is a Spatial Durbin model, the explanatory variables also include the average of these variables from neighbouring countries, which we label as W-R&D and W-hc. Table 1 displays the full results:

---

Table 1 around here

---

Table 1 presents least-squares, SAR and SDM model estimates based on our two interaction matrices. The results show that the coefficient  $\rho$  associated with spatial autocorrelation is positive and significantly different from zero for all spatial estimations (Table 1 columns 2-5). Since the estimate for the parameter  $\rho$  is significantly different from zero, least-squares estimates are biased and inconsistent. This coefficient captures technological diffusion from neighbouring countries' TFP growth. We note that the estimate for the parameter R&D is positive and significant for all regressions, but that the human capital turns out to be insignificant using SAR.

As regards the average impacts from neighbouring countries, the estimate for the parameter W-R&D turns out to be significantly different from zero. Whereas the coefficient

---

<sup>6</sup>The tests have been made on the OLS SAR and SDM models (See Anselin, 1988b).

related to the parameter  $W_{hc}$  is non-significant. This suggests that the SAR models suffer from a bias of omitted variables. Therefore, in what follows we focus on the estimates of the SDM models alone.

The SDM model estimates cannot be interpreted as partial derivatives in the typical regression model fashion. Spatial regression models exploit the complicated dependence structure between observations representing countries, regions, etc. Because of this, parameter estimates contain a wealth of information about relationships among the observations or regions. A change in a single observation (region) associated with any given explanatory variable will affect the region itself (a direct impact) and potentially affect all other regions indirectly (an indirect impact). To assess the signs and magnitudes of impacts arising from changes in the two explanatory variables, we turn to the summary measures of direct, indirect and total impacts presented in Table 2

---

Table 2 around here

---

Let us consider the direct impacts of R&D. We see that these are close to the SDM model coefficient estimates associated with the variable R&D reported in Table (1). The difference between the coefficient estimate of 0.005 (0.007) and the direct effect estimate of 0.006 (0.007) in Table (2) of 0.001 (0.000) represents feedback effects that arise as a result of impacts passing through neighbouring countries and back to the country itself. The discrepancy is positive (null) since the impact estimate exceeds the coefficient estimate, reflecting some positive (null) feedback. Since the difference between the SDM coefficient and the direct impact estimate is very small, we could conclude that feedback effects are small and not likely to be of economic significance. The feedback effects of human capital are very small (see Tables 1 and 2)

From the table we see that the direct impact of R&D is positive and significant, suggesting a positive impact on TFP growth. The indirect effect of R&D is positive and significant. This suggests that R&D in neighbouring countries has a positive impact on TFP growth, which seems intuitively plausible. The total effect from R&D is positive and comprised mostly of the indirect impact, a large R&D spillover.

The direct impact of human capital is positive and significant, suggesting a positive impact on TFP growth. However, the indirect effect of human capital is not significant. This means that we do not have an impact of human capital from neighbouring countries. The total effect of human capital is positive and composed mostly of the indirect impact.

We can interpret the total impact estimates as elasticities since the model is specified using logged growth of TFP, R&D and human capital. Based on the positive 0.043 estimates for the total impact of R&D, we can conclude that a 10 percent increase in R&D would result in 0.43 percent growth in TFP.

## 5 Conclusions

We have proposed an alternative method for estimating technology spillovers based on a model in which technological progress shows up as an expansion in the number of varieties of products. We have analysed the major determinants of TFP evolution by decomposing it into quality and variety components. To deal with both components, we have introduced a country's distance from the technological frontier and a variable that captures international R&D spillover. In doing so, we have obtained a Spatial Durbin model.

We have performed several estimation procedures depending on the spatial interdependencies. First, we have estimated the model (12) without spatial interdependencies (i.e.  $\rho = 0$ ) using OLS. Secondly, we have estimated two versions of spatial models using our two interaction matrices  $\mathbf{Wt}$  and  $\mathbf{Wd}$ : a SAR model and an SDM model. The empirical results have shown the presence of spatial autocorrelation suggesting international technological diffusion between countries. Moreover, when the spatial autocorrelation is modelled, OLS is no longer appropriate: the estimates obtained by this method are not convergent if there is a lagged endogenous variable and they are inefficient in the presence of spatial autocorrelation of errors.

The results of our estimation also show a positive direct impact of R&D and a positive indirect impact suggesting R&D spillover from neighbouring countries. The total impact of R&D shows that a 10 percent increase in R&D results in 0.43 percent growth in TFP. The results also highlight a positive impact of human capital but no spillover from the neighbouring countries' human capital.

The role played by technological spillovers has constituted the main focus of this analysis. We show that in addition to traditional factors such as R&D and human capital stock, technological spillovers strongly impact productivity growth. The technological spillovers are captured by both the spatial autocorrelation coefficient and the indirect impact of R&D.

## References

- [1] Aghion, P. and P. Howitt (1992) A model of growth through creative destruction. *Econometrica*, 60(2): 323–351.
- [2] Aghion, P. and P. Howitt (2009) *The Economics of Growth*. Number 9780262012638 in MIT Press Books. The MIT Press.
- [3] Anselin, L. (1988a) *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers: Dordrecht.



- [4] Anselin, L. (1988b) ‘A test for spatial autocorrelation in seemingly unrelated regressions’, *Economics Letters*, 28 (4): 335–341.
- [5] Anselin, L. and J. Le Gallo (2006) Interpolation of air quality measures in hedonic house price models: spatial aspects. *Spatial Economic Analysis*, 1 (1): 31–52.
- [6] Barro, R. J. and J. W. Lee (2012) A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*.
- [7] Barro, R. J. and X. Sala-I-Martin (2003) *Economic Growth, 2nd Edition*, volume 1 of *MIT Press Books*. The MIT Press.
- [8] Bayoumi, T., D. T. Coe and E. Helpman (1999) R&D spillovers and global growth. *Journal of International Economics*, 47(2): 399–428.
- [9] Blomström, M. and F. Sjöholm (1999) Technology transfer and spillovers: Does local participation with multinationals matter? *European Economic Review*, 43(4): 915–923.
- [10] Borensztein, E., J. De Gregorio and J. W. Lee (1998) How does foreign direct investment affect economic growth? *Journal of International Economics*, 45 (1): 115–135.
- [11] Broda, C., J. Greenfield and D. Weinstein (2006) From groundnuts to globalization: A structural estimate of trade and growth. NBER Working Papers 12512, National Bureau of Economic Research, Inc.
- [12] Coe, D. T. and E. Helpman (1995) International r&d spillovers. *European Economic Review*, 39(5): 859–887.
- [13] Coe, D. T., E. Helpman, and A. W. Hoffmaister (1997) North-South R&D spillovers. *The Economic Journal*, 107(40): 134–149.
- [14] Coe, D. T., E. Helpman, and A. W. Hoffmaister (2009) International r&d spillovers and institutions. *European Economic Review*, 53(7): 723–741.
- [15] Crespo, G., C. Martin and F. J. Velázquez (2004) The role of international technology spillovers in the economic growth of the OECD countries. *Global Economy Journal*, 4(3): 1–18.
- [16] Davidson, R. and J. Mackinnom (1993) *Estimation and Inference in Econometrics*. New York: Oxford University Press.

- [17] Ertur, C., and W. Koch (2007) Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22(6): 1033–1062.
- [18] Ertur, C., and W. Koch (2011) A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions. *Journal of Economic Growth*, 16(3): 215–255.
- [19] Ertur, C. and A. Musolesi (2013) Weak and strong cross-sectional dependence: a panel data analysis of international technology diffusion. *Grenoble Applied Economics Laboratory (GAEL), Working Papers*, 2013-09.
- [20] Feenstra, R.C., R. Inklaar and M. Timmer (2013) The next generation of the Penn World Table. *National Bureau of Economic Research, Working Paper*, 19255.
- [21] Grossman, G. M and E. Helpman (1991a) Quality ladders in the theory of growth. *Review of Economic Studies*, 58(1): 43–61.
- [22] Grossman, G. M and E. Helpman (1991b) Trade, knowledge spillovers, and growth. *European Economic Review*, 35(2-3): 517–526.
- [23] Kelejian, H.H., G.S. Tavlas and G. Hondronyiannis (2006) A spatial modelling approach to contagion among emerging economies. *Open Economies Review*, 17 (4/5): 423–442.
- [24] Keller, W. (1998) Are international R&D spillovers trade-related?: Analysing spillovers among randomly matched trade partners. *European Economic Review*, 42(8): 1469–1481.
- [25] Keller, W. (2002) Geographic localization of international technology diffusion. *American Economic Review*, 92): 120-142.
- [26] Keller, W. and S.R. Yeaple (2009) Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States. *The Review of Economics and Statistics*, 91(4): 821–831.
- [27] Kim, C.W., T.T.Phipps and L. Anselin (2003) Measuring the benefits of air quality improvement: a spatial hedonic approach. *Journal of Environmental Economics and Management*, 45: 24–39.
- [28] Kokko, A. (1996) Productivity spillovers from competition between local firms and foreign affiliates. *Journal of International Development*, 8 (4): 517-530.

- [29] Lee, L.F. (2004) Asymptotic Distributions of Quasi-maximum Likelihood Estimators for Spatial Autoregressive Models, *Econometrica*, 72 (6): 1899-1925.
- [30] G. Lee (2006) The effectiveness of international knowledge spillover channels. *European Economic Review*, 50 (8): 2075–2088.
- [31] LeGallo J., C. Ertur and C. Baumont (2003) A spatial econometric analysis of convergence across European regions, 1980–1995, in B. Fingleton (ed.). *European Regional Growth*.
- [32] LeSage, J.P. and R.K. Pace (2009) *Introduction to Spatial Econometrics*, Statistics: A Series of Textbooks and Monographs. Chapman and Hall/CRC.
- [33] Ord, J.K (1975) Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70, 120–126.
- [34] Pace, R.K. and R.P. Barry (1997) Quick computation of spatial autoregressive estimators. *Geographical Analysis*, 29, 232–246.
- [35] Press, W., S. Teukolsky, W. Vetterling and B. Flannery (1996) *Numerical Recipes in Fortran 77*, second edition, New York: Cambridge University Press.
- [36] Psacharopoulos, G. (1994) Returns to investment in education: A global update. *World Development*, 22 (9): 1325–1343.
- [37] Rivera-Batiz, L.A. and P. M. Romer (1991) Economic Integration and Endogenous Growth. *The Quarterly Journal of Economics*, 106 (2): 531–55.
- [38] Romer, P. M. (1987) Growth based on increasing returns due to specialization. *American Economic Review*, 77(2): 56–62.
- [39] Romer, P. M. (1990) Endogenous technological change. *Journal of Political Economy*, 98(5): 71–102.

Table 1: Estimation results

Model	OLS	SAR		SDM	
		<b>W-d</b>	<b>W-t</b>	<b>W-d</b>	<b>W-t</b>
Dependent V.	TFPG	TFPG	TFPG	TFPG	TFPG
R&D	0.007** (0.003)	0.006* (0.003)	0.006* (0.003)	0.005* (0.002)	0.007** (0.003)
hc	0.024* (0.014)	0.023 (0.016)	0.020 (0.016)	0.029* (0.015)	0.034** (0.017)
$\rho$		0.436*** (0.114)	0.406*** (0.106)	0.386*** (0.120)	0.184*** (0.037)
w-R&D		Moran I-stat 3.641	Moran I-stat 1.993	0.009* (0.004)	0.026* (0.014)
w-HC		Prob <0.001	Prob 0.046	-0.019 (0.034)	0.089 (0.066)
$R^2$	0.042				
AIC		-2.626	-2.639	-2.608	-2.815
BIC		-2.552	-2.564	-2.484	-2.690

*Notes:* Standard errors are given in parentheses. \*\*\* significant at 1%; \*\* significant at 5% and \* significant at 10%. AIC and BIC stand for the Akaike and the Schwarz information criteria, respectively.

Table 2: Cumulative effects

Models	Variables	Direct	Indirect	Total
SAR-Wd	R&D	0.006*	0.004	0.011*
		(0.003)	(0.002)	(0.006)
	hc	0.023	0.017	0.041
		(0.017)	(0.015)	(0.032)
SAR-Wt	R&D	0.006*	0.004	0.010*
		(0.003)	(0.003)	(0.005)
	hc	0.021	0.013	0.035
		(0.017)	(0.013)	(0.029)
SDM-Wd	R&D	0.006**	0.019*	0.026**
		(0.002)	(0.011)	(0.012)
	hc	0.029*	-0.016	0.012
		(0.015)	(0.473)	(0.043)
SDM-Wt	R&D	0.007**	0.036**	0.043**
		(0.003)	(0.018)	(0.022)
	hc	0.033*	0.098	0.132
		(0.018)	(0.080)	(0.156)

*Notes:* Standard errors are given in parentheses. \*\*\* significant at 1%; \*\* significant at 5% and \* significant at 10%.