Returns to scale and curvature in the presence of spillovers: evidence from European countries

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This is a pre-copyedited, author-produced version of an article accepted for publication in Oxford Economic Papers following peer review. The version of record GLASS, A.J., KENJEGALIEVA, K. and SICKLES, R.C., 2015. Returns to scale and curvature in the presence of spillovers: evidence from European countries. Oxford Economic Papers, 68 (1), pp.16-39. is available online at: http://dx.doi.org/10.1093/oep/gpv058

Metadata Record: https://dspace.lboro.ac.uk/2134/18272

Version: Accepted for publication

Publisher: Oxford University Press

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Returns to scale and curvature in the presence of spillovers: evidence from European countries

Anthony J. Glass∗, Karligash Kenjegalieva† and Robin C. Sickles‡

Abstract

Drawing on a recent development on the interpretation of spatial econometric models we extend two classic characteristics of production (returns to scale and diminishing marginal productivity of factor inputs) to the spatial case. In the context of a spatial translog production function we define internal (i.e. direct), external (i.e. indirect) and total (direct plus indirect) returns to scale. The spatial production function gives rise to direct, indirect and total productions functions so we set out empirical checks to establish if these functions are concave. The ideas of spatial returns to scale and spatial concavity/convexity can easily be applied to other technologies (e.g. cost and distance functions) and other functional forms (e.g. Cobb-Douglas). We apply these ideas to aggregate production of European countries using balanced panel data for the period 1990 – 2011.

JEL Classification: C23; D24

---

∗School of Business and Economics, Loughborough University, Leicestershire, UK, LE11 3TU; e-mail: A.J.Glass@lboro.ac.uk.
†School of Business and Economics, Loughborough University; e-mail: K.A.Kenjegalieva@lboro.ac.uk.
‡Department of Economics, Rice University, Houston, TX, and School of Business and Economics, Loughborough University; e-mail: rsickles@rice.edu
1 Introduction

In this paper we introduce the idea of spatial returns to scale by drawing on a recent development in applied spatial econometrics. LeSage and Pace (2009) demonstrate that the coefficients on the independent variables from a model with a spatial autoregressive (SAR) variable cannot be interpreted as marginal effects. This is because the marginal effect of an independent variable is a function of the SAR variable. To address this issue they propose a method to calculate direct, indirect and total (direct plus indirect) marginal effects.\(^1\) In the context of a spatial production function we use the direct, indirect and total marginal effects to propose the concepts of internal, external and total returns to scale, respectively. Internal, external and total returns to scale can easily be calculated for other technologies (cost, standard and alternative profit, revenue, and input and output distance functions) so there is considerable scope for wider application of the spatial returns to scale which we propose. Internal returns to scale has the same interpretation as returns to scale from a non-spatial production function. External returns refers to the rate of increase in a unit’s output following an increase in the factor inputs of all the other units in the sample. Total returns to scale is the rate of increase in a unit’s output following a simultaneous increase in its own factor inputs and the factor inputs of all the other units in the sample. Having estimated spatial production functions for European countries over the period 1990 – 2011 using various spatial weights matrices, we estimate internal, external and total returns to scale.

The closest relatives to this paper fall into two categories: (i) empirical applications of New Economic Geography (NEG) and (ii) empirical growth models estimated using spatial econometric techniques.\(^2\) There are a wide range of empirical applications of NEG to, for example, countries (Redding and Venables, 2004), Chinese prefectures (Roberts, et al., 2012) and U.S. counties (Hanson, 2005). For a sample of local areas in Great Britain, Fingleton (2006) combines spatial econometric methods and an artificial nesting model to test non-nested NEG and urban economic theories against one another. The urban economic theory is based on the benefits from good connections with producers in the service sector, where these linkages are better in urban areas where employment density is higher. Interestingly, Fingleton finds that the data supports urban economic theory over NEG. Solving theoretical NEG models yields a set of price and wage equations, where the vast majority of empirical NEG studies estimate the wage equation because of its tractability. The effect of geography in empirical NEG studies is in terms of the effect of

---

\(^1\)A direct elasticity is interpreted in the same way as an elasticity from a non-spatial model, although a direct elasticity takes into account feedback effects (i.e. effects which pass through first order and higher order neighbours via the spatial multiplier matrix and back to the unit which initiated the change). An indirect elasticity is a spillover elasticity and is the change in the dependent variable for one particular unit following a change in an explanatory variable in all the other units. A total elasticity is the sum of the direct and indirect elasticities.

\(^2\)See Fujita et al. (1997) for a theoretical textbook treatment of NEG.
a region’s market potential on its wages, where a region’s market potential depends on its geographical location as this affects, among other things, the access cost to its own market and other regional markets. In this paper, however, the effect of geography is explicitly related to neighbours’ factor inputs.

The equations which have been estimated using spatial econometric techniques to analyse economic growth include: (i) the standard neoclassical output per worker equation (Solow, 1956; 1957), and (ii) Verdoorn’s model, which postulates that output growth in a region will lead to an increase in the regions’s labour productivity because of increasing returns to scale. A small number of studies use spatial econometric techniques in conjunction with an artificial nesting model to test between the standard neoclassical growth model and NEG (Fingleton, 2008; Fingleton and Fischer, 2010). In both studies the standard neoclassical equation for output per worker is augmented with a spatial error term. Other studies have extended the Solow (1956; 1957) neoclassical growth set-up by endogenising technical change which involves modelling this change as a function of, among other things, capital per worker spillovers (López-Bazo et al., 2004; Egger and Pfaffermayr, 2006; Koch, 2009; Pfaffermayr, 2009). Using this approach Koch (2009) shows that a spatial neoclassical development accounting equation for output per worker is linearly approximated by the spatial Durbin model.  

Fingleton and McCombie (1998) and Fingleton and López-Bazo (2006) all observe strong empirical support for the Verdoorn specification using data for NUTS–2 European regions. Furthermore, the theoretical model proposed by Fingleton (2001) yields a reduced form of Verdoorn’s model, which is consistent with theories from the urban and regional economics literature which rely on agglomeration economies and also some endogenous growth models based on increasing returns to scale. As Pfaffermayr (2009) notes, however, it is not possible to construct an artificial nesting model to test the non-nested Verdoorn and neoclassical growth models against one another. Notwithstanding the important contributions of the above spatial empirical growth models, the authors do not calculate direct, indirect and total marginal effects. With the spatial error specification the direct, indirect and total marginal effects relate to the error term. In the context of a spatial error production function, changes in Total Factor Productivity (TFP) spillovers are therefore separable from the production factors. In this paper, however, we model TFP spillovers which are non-separable from the production factors by allowing

---

3 The spatial Durbin specification models global spatial dependence (1st order neighbour effects through to \(N - 1\)th order neighbour effects) via the SAR variable and local spatial dependence (1st order neighbour effects only) via a spatial lag of each independent variable.

4 The reason is because independent variables in the artificial nesting model would include log differences of income and population, which are used to calculate the dependent variable.

5 Using a spatial neoclassical growth model similar to Koch (2009), Ertur and Koch (2007) calculate direct and indirect marginal effects, where they refer to the latter as cross-elasticities. In contrast to our paper where the focus is on the direct, indirect and total marginal effects, the focus of their paper lies elsewhere so they only discuss the direct elasticities briefly and do not calculate the associated \(t\)–statistics.
the SAR term to shift the production technology. As a result, we can be more explicit about the spillovers which are at work because the direct, indirect and total marginal effects relate to the independent variables and thus indicate, among other things, the effect of TFP spillovers which are due to production factor spillovers.

In an empirical analysis where the focus is on internal, external and total returns to scale it is more illuminating to calculate these returns from a function with a flexible form. Accordingly, in the application to aggregate production of European countries, rather than estimate a spatial Cobb-Douglas production function which underpins the above spatial empirical growth models and where returns to scale are the same at every point in the sample, we estimate a spatial translog production function which allows returns to scale to vary over the sample. Specifically, we estimate SAR and spatial Durbin production functions. At various junctures in the remainder of this paper we explain why we have a preference for the spatial Durbin specification. To give an insight into our key empirical findings, from our preferred spatial Durbin model we find that, on average, the 2004 EU enlargement led to a marked fall (rise) in total returns to scale for EU (non-EU) countries.

The direct, indirect and total marginal effects can be used to construct three translog production functions. An issue which is related to internal, external and total returns to scale over the sample is the curvature of the direct, indirect and total translog production functions (i.e. internal, external and total curvature, respectively). This is an important issue because concavity of a direct translog production function would satisfy the curvature property of the function and would indicate diminishing marginal productivity of a unit’s own factor inputs. We conduct an empirical check of the curvature of the fitted direct translog production functions by adapting the approach which is used to check the curvature of a non-spatial translog production function. Indirect and total translog production functions have no such curvature properties but we find it informative to conduct an empirical check of the curvature of these fitted functions in a similar way.

The remainder of this paper is organised as follows. In section 2 we set out the two specifications of the spatial translog production function which we estimate and then we explain the approach to calculate internal, external and total returns to scale. Section 3 discusses how we conduct empirical checks of the monotonicity and curvature of the fitted direct, indirect and total translog production functions. Section 4 is dedicated to the application where, among other things, we test for constant internal, constant external and constant total returns to scale. Section 5 concludes.
2 Spatial Translog Production Functions and Spatial Returns to Scale

Our starting point is the spatial Durbin Cobb-Douglas production function which Ertur and Koch (2007) and Koch (2009) estimate. LeSage and Pace (2009) make a compelling econometric case for the spatial Durbin model. There are two strands to their econometric case. The first is based on their belief that the principal focus of spatial modelling should be the analysis of substantive global spillovers which relate to the SAR term, rather than the analysis of the global spillover of shocks which emanate from the spatial error term. The second strand concerns the unbiased parameter estimates which the spatial Durbin model yields even if the true data generating process is, among others, the spatial error model or the SAR model.

The origin of productivity spillovers which result in increasing total returns to scale is explicit in Romer’s (1986; 1987) seminal endogenous growth theories. In these models, a firm’s production function is characterised by the usual constant or decreasing internal returns to scale. In the Romer (1986) model, a firm’s own knowledge and aggregate knowledge in the economy are distinct inputs. Aggregate knowledge as an input indicates that the production possibilities of a firm increase as a result of knowledge spillovers. These spillovers between firms arise because new knowledge cannot be perfectly patented or kept secret. In the Romer (1987) model, however, because a firm’s own knowledge and own physical capital are assumed to be used in fixed proportions in production, the set-up follows Arrow (1962) and collectively represents own knowledge and own physical capital using a composite own capital input. It then follows that aggregate composite capital in the economy is also an input in the firm’s production function in Romer’s (1987) model. Irrespective of how a firm’s own knowledge and aggregate knowledge in the economy is modelled in the above endogenous growth theories, in both cases external returns to scale from spillovers are sufficiently large so that even if internal returns to scale are decreasing, adding external returns to the internal returns yields increasing total returns.

Notable empirical growth studies which employ spatial econometric techniques and are also explicit about the spillover process include Ertur and Koch (2007) and Koch (2009). These studies are explicit about the origin of productivity spillovers from a theoretical perspective and also an empirical perspective. The former is the case because both studies extend the standard Solow growth model by endogenising technical change and modelling this change as a function of well-defined spillovers (e.g. capital per worker spillovers). The latter is the case because the theoretical spatial Solow equation for output per worker

---

6 Arrow (1962) developed a neoclassical growth model where technical change is completely embodied in new units of the composite own capital input. Technical progress in this model results from learning by doing. In other words, each new unit of physical capital which is used in the production process gives rise to learning and thus the accumulation of knowledge.
is estimated using the spatial Durbin specification which captures local spillovers via a spatial lag of capital per worker and global spillovers via the SAR process. Our approach is similar in spirit as we use a spatial Durbin production framework. In contrast, however, rather than use the relatively restrictive Cobb-Douglas specification we use the flexible translog function. Moreover, we estimate indirect marginal effects which relate to the independent variables so we are empirically explicit about the spillover process but in a different way to Ertur and Koch (2007) and Koch (2009).\footnote{We thank an anonymous referee for guidance on the nature of the spillover process in the related literature.}

We estimate the following spatial Durbin translog production function for panel data:

$$\ln y_{it} = \kappa + \psi_t + \alpha_i + \zeta_1 t + \zeta_2 t^2 + TL(x_{it}) + z_i \theta + \sum_{j=1}^{N} w_{ij} TL(x_{jt}) + \sum_{j=1}^{N} w_{ij} z_{jt} + \lambda \sum_{j=1}^{N} w_{ij} \ln y_{jt} + \varepsilon_{it},$$

where $N$ is a cross-section of units indexed $i = 1, 2, ..., N$, $T$ is the number of time periods indexed $t = 1, 2, ..., T$, $\ln y_{it}$ is the log of output of the $i$th unit at time $t$, $\kappa$ is the intercept parameter and $\alpha_i$ is a fixed effect. We follow the spatial decomposition of aggregate TFP growth for European countries by Glass et al. (2013) and capture the effects of time in Eq. 1 by, firstly, including a time trend, $t$, and the associated quadratic term to capture average annual technical change over the study period. Secondly, we include some time period dummy variables, $\psi_t$, to capture departures from the time trend in a particular year due to, for example, common macroeconomic shocks. Time period dummy variables for some years are omitted to avoid perfect collinearity with the time trend.\footnote{In the application we estimate a SAR production function (see Eq. 2), its non-spatial counterpart, which is Eq. 2 with the $\lambda \sum_{j=1}^{N} w_{ij} \ln y_{jt}$ term omitted, and the spatial Durbin production function (Eq. 1) by demeaning in the cross-sectional dimension (i.e. the Within transformation) but not in the time dimension. This is because demeaning in the time dimension would eliminate the $t$ and $t^2$ terms from Eqs. 1 and 2 which we want to retain because their direct, indirect and total marginal effects have important economic implications in an empirical analysis of TFP. Also, as will be apparent from Eq. 2 in due course, if we demeaned in the time dimension the modelling of the interaction terms which include $t$ in Eq. 2 would be erroneous. This is because we would eliminate $t$ and retain the interaction terms which include $t$. Moreover, in our application the time period dummy variables for particular years are a relatively small consumer of degrees of freedom because $N$ is relatively large compared to $T$. Inconsistent estimates due to an incidental parameter problem associated with the time period dummy variables for particular years is not therefore a great cause for concern. Finally, we note when we estimate the non-spatial counterpart of Eq. 2 using STATA the time period dummy variables for 2000 and 2011 are dropped due to perfect collinearity with the time trend, and all the coefficients on the time period dummy variables are relative to the omitted category which is the time period dummy variable for 1990. To remain consistent when estimating Eqs. 1 and 2 using MATLAB we drop the same time period dummy variables.}

$x_{it}$ in Eq. 1 is a $(1 \times K)$ vector of inputs indexed $k = 1, ..., K$. $TL(x_{it}) = \gamma' \ln x_{it} + \frac{1}{2} \ln x_{it}' \Theta \ln x_{it}$ represents the $i$th unit’s translog production function at time $t$, where
\( \gamma' \) is a vector of parameters and \( \Theta \) is a matrix of \( \theta \) parameters. 
\[ \sum_{j=1}^{N} w_{ij} TL (x_{jt}) = \hat{\gamma}' \sum_{j=1}^{N} w_{ij} \ln x_{jt} + \frac{1}{2} \sum_{j=1}^{N} w_{ij} \Phi \sum_{j=1}^{N} w_{ij} \ln x_{jt} \]
represents the weighted sum of the first order \( j \)th units' translog production functions at time \( t \), where \( \hat{\gamma}' \) is a vector of parameters and \( \Phi \) is a matrix of \( \phi \) parameters. It follows from the properties of the non-spatial translog production function (Christensen et al., 1973) that Eq. 1 is twice differentiable with respect to the \( i \)th unit's inputs, where the associated Hessian is symmetric because of the symmetry restrictions which are imposed on \( \Theta \) i.e. \( \theta_{1K} = \theta_{K1} \). Eq. 1 is also twice differentiable with respect to the spatially weighted sum of the inputs of the first order \( j \)th units, where the associated Hessian is symmetric because of the symmetry restrictions which are imposed on \( \Phi \). The inputs of higher order \( j \)th units affect \( \ln y_{it} \) via their effect on the SAR term. \( z \) in Eq. 1 is a vector of variables, where \( \vartheta \) and \( \upsilon \) are vectors of parameters, and \( \varepsilon_{it} \) is an i.i.d. disturbance for \( i \) and \( t \) with zero mean and variance \( \sigma^2 \). \( \lambda \) is the SAR parameter and \( w_{ij} \) is a known non-negative element of the \((N \times N)\) spatial weights matrix, \( W \), where as is standard the diagonal elements of \( W \) are set to zero to rule out 'self-influence'. \( W \) captures the spatial arrangement of the cross-sectional units and also the strength of the spatial interaction in the cross-section. \( W \) must be specified prior to estimation and is usually specified according to some measure of geographical or economic proximity. In the application to aggregate production of European countries \( W \) is normalised to have row sums of unity so that the spatial lag of the dependent variable is a weighted average of observations for the dependent variable for neighbouring countries which preserves the scaling of the data.\(^9\)

Whereas with non-spatial production functions \( z \) variables shift the technology, \( z_{it} \hat{\vartheta}, \) 
\[ \sum_{j=1}^{N} w_{ij} z_{jt} \hat{\upsilon}, \sum_{j=1}^{N} w_{ij} TL (x_{jt}) \text{ and } \lambda \sum_{j=1}^{N} w_{ij} \ln y_{jt} \text{ all shift the technology in Eq. 1.} \]
as is standard we make the following assumptions. (i) \((I_N - \lambda W)\) is non-singular and the parameter space of \( \lambda \) is \( \left( \frac{1}{h_{\min}}, 1 \right) \), where \( I_N \) is the \((N \times N)\) identity matrix and \( h_{\min} \) is the most negative real characteristic root of \( W \). Since we use a row-normalised \( W \) in the application, 1 is the largest real characteristic root of \( W \) which rules out explosive growth. (ii) The row and column sums of \( W \) and \((I_N - \lambda W)^{-1} \) are bounded uniformly in absolute value before \( W \) is row-normalised. As a result of this assumption the spatial process for the dependent variable has a 'fading' memory (e.g. Kelejian and Prucha, 1998).

---

\(^9\)If we attempted to also model TFP spillovers which are separable from the production factors by introducing the spatial error term, \( \rho \sum_{j=1}^{N} w_{ij} \varepsilon_{jt} \), to Eq. 1 resulting in what is widely referred to as the Manski (1993) model, where \( \rho \) is the spatial error parameter, the parameters would not be identified. Lee (2007) shows, however, that the parameters of the Manski model are identified if the dependent variable and the independent variables are pre-multiplied by a spatial weights matrix which differs from that which is used to construct the spatial error term. This would create quite a large number of possible combinations of spatial weights matrices so this was not a line of enquiry that we pursued. There is no such identification problem if Eq. 2 is augmented with the spatial error term. Again this was not something which we pursued because, as with the SAR model, there is an implausible restriction on the ratio of the indirect and direct marginal effects of the independent variables. We discuss this restriction further in this section where we also note that we favour the spatial Durbin specification as it is free of such a restriction.
In Eq. 1 we omit the technology shifters $t$ and $t^2$ from the translog function to avoid perfect collinearity but include the $t$ and $t^2$ terms additively to account for a Hicks-neutral type of productivity change.\footnote{If in Eq. 1 we introduced $t$ and $t^2$ to the translog function and $\sum_{j=1}^{N} w_{ij} t$ and $\sum_{j=1}^{N} w_{ij} t^2$ to the local spatial translog function there would be perfect collinearity because pre-multiplying $t$ and $t^2$ by $W$ yields $t$ and $t^2$.} Strictly therefore Eq. 1 is a partial spatial Durbin model. Hicks-neutral technological change, however, restricts the factor substitution possibilities in this model. Alternatively, we could relax the assumption of Hicks-neutral technological change in Eq. 1 by retaining $\sum_{j=1}^{N} w_{ij} T(x_{it})$ and replacing $T(x_{it})$ with $T(x_{it}, t) = \gamma' \ln x_{it} + \frac{1}{2} \ln x_{it} \Theta \ln x_{it} + \zeta_1 t + \zeta_2 t^2 + \phi' \ln x_{it} t$, where $\phi'$ is a vector of parameters. This would, however, introduce a theoretical inconsistency between the functional forms of the non-spatial and the local spatial translog production functions. With the SAR translog specification, technological change need not be Hicks-neutral. This is evident because the SAR translog panel production function which we estimate is:

$$\ln y_{it} = \kappa + \psi_t + \alpha_i + T(x_{it}, t) + z_{it} \theta + \lambda \sum_{j=1}^{N} w_{ij} \ln y_{jt} + \varepsilon_{it},$$

where the variables and parameters are as described above.

Eqs. 1 and 2 can be estimated using Maximum Likelihood (ML), General Method of Moments and Bayesian MCMC. Here we follow the procedure in Elhorst (2009) and use ML to estimate Eq. 2, where Eq. 1 is estimated in similar fashion. The estimation of the models has a number of important features. Firstly, we include in the log-likelihood functions the scaled logged determinant of the Jacobian of the transformation from $\varepsilon_t$ to $\ln y_t$ (i.e. include in the log-likelihood functions $T \ln |I_N - \lambda W|$). As is standard in ML estimation of spatial econometric models, this ensures that $\lambda$ lies in its parameter space, and the transformation accounts for the endogeneity of the SAR variable and the fact that $\varepsilon_t$ is not observed. Secondly, we use a row-normalised $W$ which simplifies the computation of $\ln |I_N - \lambda W|$ to evaluate the log-likelihoods. For more details on this see Elhorst (2009). Thirdly, we use the Within transformation to estimate the non-spatial and spatial models by demeaning in the cross-sectional dimension to circumvent the incidental parameter problem associated with the fixed effects, which eliminates these effects (and the intercept). Lee and Yu (2010) show that demeaning in the cross-sectional dimension to estimate a fixed effects spatial model which contains the SAR variable results in a biased estimate of $\sigma^2$ if $N$ is large and $T$ is fixed, which we denote $\sigma^2_B$, where the bias is of the type identified in Neyman and Scott (1948). Following Lee and Yu (2010) and Elhorst (2014) we correct for this bias by replacing $\sigma^2_B$ with the bias corrected estimate of $\sigma^2$, $\sigma^2_{BC} = \frac{T \sigma^2}{(T-1)}$, which changes the standard errors.

As noted in the opening section of this paper, LeSage and Pace (2009) demonstrate that for models such as Eqs. 1 and 2 the coefficients on the independent variables cannot
be interpreted as elasticities. They therefore suggest the following approach to calculate direct, indirect and total marginal effects. We can rewrite Eqs. 1 and 2 as follows, respectively, where the subscript \(i\)'s are dropped to denote successive stacking of cross-sections.

\[
\ln y_t = (I_N - \lambda W)^{-1} \left( \begin{array}{c} \kappa \ell + \psi_t \ell + \alpha + \rho_1 \ell + \rho_2 \ell^2 + \\
\Gamma_t \beta + W \Gamma_t \eta + Z_t \vartheta + WZ_t v + \varepsilon_t \end{array} \right),
\]

\[
\ln y_t = (I_N - \lambda W)^{-1} \left( \kappa \ell + \psi_t \ell + \alpha + \Omega_t \zeta + Z_t \vartheta + \varepsilon_t \right),
\]

where \(\ell\) is an \((N \times 1)\) vector of ones, \(\alpha\) is an \((N \times 1)\) vector of fixed effects and \(Z_t\) is a matrix of stacked observations for \(z_{it}\). \(\Gamma_t\) is a \((N \times U)\) matrix of stacked observations for \(TL(x_{it}) = \gamma' \ln x_{it} + \tfrac{1}{2} \ln x_{it}' \Theta \ln x_{it}\) or in other words, a matrix which includes stacked observations for \(\ln x_{it}\) and \(\ln x_{it}' \ln x_{it}\), where the stacked components of \(TL(x_{it})\) are indexed \(u = 1, ..., U\). If there are two inputs then \(U = 5\). The \(\gamma'\) and \(\tfrac{1}{2} \Theta\) translog parameters are for simplicity collected in the vector of parameters \(\beta\) in Eq. 3. Similarly, \(W \Gamma_t\) is a \((N \times U)\) matrix of stacked observations for \(\sum_{j=1}^{N} w_{ij} TL(x_{jt}) = \varrho' \sum_{j=1}^{N} w_{ij} \ln x_{jt} + \tfrac{1}{2} \sum_{j=1}^{N} w_{ij} \ln x_{jt}' \Phi \sum_{j=1}^{N} w_{ij} \ln x_{jt}\) or put another way, a matrix of observations which includes stacked observations for \(\sum_{j=1}^{N} w_{ij} \ln x_{jt}\) and \(\sum_{j=1}^{N} w_{ij} \ln x_{jt}' \sum_{j=1}^{N} w_{ij} \ln x_{jt}\). Once again for simplicity the \(\varrho'\) and \(\tfrac{1}{2} \Phi\) local spatial translog parameters are collected in the vector of parameters \(\eta\) in Eq. 3. \(\Omega_t\) is a \((N \times V)\) matrix of stacked observations for \(TL(x_{it}, t) = \gamma' \ln x_{it} + \tfrac{1}{2} \ln x_{it}' \Theta \ln x_{it} + \varsigma_1 t + \varsigma_2 t^2 + \phi' \ln x_{it} t\) or in other words, a matrix which includes stacked observations for \(\ln x_{it}\), \(\ln x_{it}' \ln x_{it}, t, t^2\) and \(\ln x_{it} t\). If there are two inputs \(V = 9\), where \(U < V\) because of the additional four variables in \(TL(x_{it}, t)\) (i.e. \(t, t^2, \ln x_{it1} \times t\) and \(\ln x_{it2} \times t\)). The \(\gamma', \Theta, \varsigma_1, \varsigma_2\) and \(\phi'\) translog parameters are collected in the vector of parameters \(\zeta\) in Eq. 4.

We set out the approach to calculate the direct, indirect and total marginal effects for the \(uth\) component of \(\Gamma_t\) in Eq. 3 at the sample mean. Using mean adjusted data all the fitted parameters from a local spatial translog production function (i.e. Eq. 1 without the SAR variable) are elasticities because at the sample mean the own and local spatial quadratic and interaction terms are zero. Extending this to Eq. 3, the fitted \(\beta\) and \(\eta\) parameters for the \(uth\) component of \(\Gamma_t\) can be used to directly calculate the direct, indirect and total elasticities at the sample mean. Differentiating Eq. 3 with respect to the \(uth\) component of \(\Gamma_t\), which we denote \(\Gamma_{ut}\), yields the following matrix of direct and indirect elasticities for each unit, where the right-hand side of Eq. 5b is independent of the time index:
Since Eq. 5b yields different direct and indirect elasticities for each unit, to facilitate interpretation LeSage and Pace (2009) suggest reporting a mean direct elasticity (average of the diagonal elements of Eq. 5b) and a mean indirect elasticity (average row sum of the non-diagonal elements of Eq. 5b in the empirical section of this paper). The mean total elasticity is the sum of the mean direct and mean indirect elasticities.\(^{11}\)

For Eq. 2 and if we augmented Eq. 2 with the spatial error term, the direct, indirect and total marginal effects for the independent variables are also calculated using Eq. 5b but with the \(w_{ij}\eta_u\) off-diagonal elements set equal to zero by construction.\(^{12}\) The upshot is that the ratio of the indirect and direct marginal effects is the same for all the independent variables for models which contain the SAR variable but no local spatial variables. This is considered unrealistic and is a key reason why the spatial Durbin model is now widely favoured.

Returns to scale from a non-spatial production function and internal returns to scale from Eqs. 1 and 2 measure the percentage change in the \(i\)th unit’s output due to a one percent increase in the \(i\)th unit’s inputs. Unlike non-spatial returns to scale, however, internal returns to scale from Eqs. 1 and 2 also include feedback effects (i.e. when the \(i\)th unit’s inputs change, which via the spatial multiplier matrix affects neighbours’ outputs, some of this effect on neighbours’ outputs rebounds and affects the output of the \(i\)th unit). External returns to scale from Eqs. 1 and 2 refers to the percentage change in the \(i\)th unit’s output due to a one percent increase in the sum of the inputs across all

\[^{11}\]To compute the \(t\)-statistics for the mean direct, mean indirect and mean total elasticities we follow LeSage and Pace (2009) who propose conducting a Monte Carlo experiment to simulate the distribution of the mean elasticities. For more details on this see Elhorst (2014).

\[^{12}\]In the application, having fitted the partial spatial Durbin model, for some of the independent variables \((t, t^2\text{ and time period and EU membership dummy variables})\) \(w_{ij}\eta_u = 0\) is also the case in the calculation of the direct, indirect and total marginal effects. As we touched on in footnote 10, this is because we omit the spatial lags of \(t, t^2\) and the dummy variables from the specification of the partial spatial Durbin model due to perfect collinearity. To illustrate, using a time period dummy variable, \(\psi_t\): \(W\psi_t = \psi_t\).
the other \( N - 1 \) units. Total returns to scale from Eqs. 1 and 2 is the sum of internal and external returns to scale and is the percentage change in the \( i \)th unit’s output due to a one percent increase in the sum of the inputs across all \( N \) units in the sample. We can compute internal, external and total returns to scale (denoted \( RTS^{Int} \), \( RTS^{Ext} \) and \( RTS^{Tot} \), respectively) at the sample mean from Eqs. 1 and 2 as follows.

\[
\sum_{k=1}^{K} e_{k}^{Dir} + \sum_{k=1}^{K} e_{k}^{Ind} = \sum_{k=1}^{K} e_{k}^{Tot},
\]

where \( e_{k}^{Dir} \), \( e_{k}^{Ind} \) and \( e_{k}^{Tot} \) are direct, indirect and total elasticities at the sample mean for the \( k \)th input, and \( RTS^{Int} = \sum_{k=1}^{K} e_{k}^{Dir} \), \( RTS^{Ext} = \sum_{k=1}^{K} e_{k}^{Ind} \) and \( RTS^{Tot} = \sum_{k=1}^{K} e_{k}^{Tot} \).

We observe decreasing internal, external and total returns if \( RTS^{Int} < 1 \), \( RTS^{Ext} < 1 \) and \( RTS^{Tot} < 1 \), constant internal, external and total returns if \( RTS^{Int} = 1 \), \( RTS^{Ext} = 1 \) and \( RTS^{Tot} = 1 \), and increasing internal, external and total returns if \( RTS^{Int} > 1 \), \( RTS^{Ext} > 1 \) and \( RTS^{Tot} > 1 \). The classification of \( RTS^{Int} \), \( RTS^{Ext} \) and \( RTS^{Tot} \) need not of course be the same. It should be emphasised that internal returns to scale refers to the \( i \)th unit changing its inputs, external returns to scale refers to all the other \( N - 1 \) units changing their inputs and, finally, the calculation of total returns to scale is based on all \( N \) units in the sample simultaneously changing their inputs and not just the \( i \)th unit or the other \( N - 1 \) units. We test the null hypotheses of constant internal, constant external and constant total returns to scale for the sample average country using one-sided \( t \)-tests.

3 Internal, External and Total Curvature and Monotonicity

With reference to Eqs. 1 and 2 the internal, external and total curvature and monotonicity propositions are set out in Propositions 1—6.\(^\text{13}\) The internal, external and total curvature propositions are presented without loss of generality in terms of two input vectors, \( x_{A} \) and \( x_{B} \).

**Proposition 1:** Internal concavity of the effect of the \( i \)th unit’s own input vector

Any linear combination of the effects of the two input vectors for the \( i \)th unit on its output, where these effects are represented by \( e_{A,i}^{Dir} x_{A,i} \) and \( e_{B,i}^{Dir} x_{B,i} \), and \( e_{A,i}^{Dir} \) and \( e_{B,i}^{Dir} \) are vectors of the direct input elasticities, will produce direct output for the \( i \)th unit, \( y_{A_{AB},i}^{Dir} \).

\(^{13}\) Internal concavity and internal monotonicity are properties of non-spatial and direct production functions. External (total) concavity and external (total) monotonicity, however, are not properties of indirect (total) production functions. We refer to the internal concavity and internal monotonicity properties of a direct production function as propositions so that there is uniformity with the external and total concavity and monotonicity propositions.
Furthermore, it is evident from Proposition that internal concavity of the effect of the $ith$ unit’s own inputs on its output. Furthermore, it is evident from Proposition 1 that internal concavity of the effect of $x_i$ depends on the specification of the spatial weights matrix. This is because the magnitudes of $e^{Dir}_{A,i}$ and $e^{Dir}_{B,i}$ depend on the specification of the spatial weights matrix.

**Proposition 2: External concavity of the effect of the sum of the input vectors across the $jth$ units**

Any linear combination of the effects of the two summed input vectors across the $jth$ units, where these effects are represented by $e^{Ind}_{A,i} \sum_{j=1}^{N} x_{A,j}$ and $e^{Ind}_{B,i} \sum_{j=1}^{N} x_{B,j}$, and $e^{Ind}_{A,i}$ and $e^{Ind}_{B,i}$ are vectors of the indirect input elasticities, will produce indirect output for the $ith$ unit, $y^{Ind}_{A,B,i}$, that is no less than a linear combination of the indirect outputs $y^{Ind}_{A,i} = f(e^{Ind}_{A,i} \sum_{j=1}^{N} x_{A,j})$ and $y^{Ind}_{B,i} = f(e^{Ind}_{B,i} \sum_{j=1}^{N} x_{B,j})$. Formally:

$$f \left( \xi e^{Ind}_{A,i} \sum_{j=1}^{N} x_{A,j} + (1 - \xi) e^{Ind}_{B,i} \sum_{j=1}^{N} x_{B,j} \right) \geq \xi f \left( e^{Ind}_{A,i} \sum_{j=1}^{N} x_{A,j} \right) + (1 - \xi) f \left( e^{Ind}_{B,i} \sum_{j=1}^{N} x_{B,j} \right),$$

for all $0 \leq \xi \leq 1$.

If Proposition 2 holds empirically the fitted indirect production function will be characterised by a diminishing marginal effect of the sum of the input vectors across the $jth$ units on the $ith$ unit’s output. Along similar lines to Proposition 1, we can see from Proposition 2 that external concavity of the effect of $\sum_{j=1}^{N} x_j$ depends on the specification of the spatial weights matrix. This is because the magnitudes of $e^{Ind}_{A,i}$ and $e^{Ind}_{B,i}$ depend on the specification of the spatial weights matrix.

**Proposition 3: Total concavity of the effect of the sum of the input vectors across all $N$ units in the sample**

Any linear combination of the effects of the two summed input vectors across all $N$ units in the sample, where these effects are represented by $e^{Tot}_{A,i} \left( x_{A,i} + \sum_{j=1}^{N} x_{A,j} \right) = e^{Dir}_{A,i} x_{A,i} + e^{Ind}_{A,i} \sum_{j=1}^{N} x_{A,j}$ and $e^{Tot}_{B,i} \left( x_{B,i} + \sum_{j=1}^{N} x_{B,j} \right) = e^{Dir}_{B,i} x_{B,i} + e^{Ind}_{B,i} \sum_{j=1}^{N} x_{B,j}$, and $e^{Tot}_{A,i}$ and $e^{Tot}_{B,i}$ are vectors of the total input elasticities, will produce total output for the $ith$ unit, $y^{Tot}_{A,B,i}$, that is no less than a linear combination of the total outputs $y^{Tot}_{A,i} = f \left( e^{Tot}_{A,i} \left( x_{A,i} + \sum_{j=1}^{N} x_{A,j} \right) \right)$ and $y^{Tot}_{B,i} = f \left( e^{Tot}_{B,i} \left( x_{B,i} + \sum_{j=1}^{N} x_{B,j} \right) \right)$. Formally:
the magnitudes of discussion of Propositions This follows because the magnitudes of associated direct, indirect and total elasticities in a similar way. variables for ease of exposition. Having estimated Eq. recognizing that give rise to direct, indirect and total translog production functions. For a ... fitted Eq. 1, empirical checks to see if the curvature and monotonicity propositions hold involves characterised by a diminishing marginal effect of the sum of the input vectors across the N units on the ith unit’s output. Proposition 3 shows that total concavity of the effect of \( x_i + \sum_{j=1}^{N} x_{i,j} \) depends on the specification of the spatial weights matrix. This is because the magnitudes of \( e_{A,i}^{Tot} \) and \( e_{B,i}^{Tot} \) depend on the specification of the spatial weights matrix. This follows because \( e_{A,i}^{Tot} = e_{A,i}^{Dir} + e_{A,i}^{Ind} \) and \( e_{B,i}^{Tot} = e_{B,i}^{Dir} + e_{B,i}^{Ind} \), where we observed that the magnitudes of \( e_{A,i}^{Dir} \), \( e_{B,i}^{Dir} \), \( e_{A,i}^{Ind} \) and \( e_{B,i}^{Ind} \) depend on the spatial weights in the above discussion of Propositions 1 and 2.

**Proposition 4: Internal Monotonicity**

\( y_{i,t} \) is monotonically increasing in the kth input of the ith unit, \( x_{k,i,t} \), if \( \partial y_{i,t}/\partial x_{k,i,t} \equiv e_{k,i}^{Dir} \geq 0 \).

**Proposition 5: External Monotonicity**

\( y_{i,t} \) is monotonically increasing in the summed kth input across the jth units, \( \sum_{j=1}^{N} x_{k,j} \), if \( \partial y_{i,t}/\partial \sum_{j=1}^{N} x_{k,j} \equiv e_{k,j}^{Ind} \geq 0 \).

**Proposition 6: Total Monotonicity**

\( y_{i,t} \) is monotonically increasing in the summed kth input across all N units in the sample, \( x_{k,i} + \sum_{j=1}^{N} x_{k,j} \), if \( \partial y_{i,t}/\partial \left( x_{k,i} + \sum_{j=1}^{N} x_{k,j} \right) \equiv e_{k,i}^{Tot} \geq 0 \).

Having estimated Eqs. 1 and 2, the resulting direct, indirect and total elasticities give rise to direct, indirect and total translog production functions. For a fitted Eq. 1, empirical checks to see if the curvature and monotonicity propositions hold involves recognising that \( \ln y_{i,t} \) can be expressed as follows, where we omit the time period dummy variables for ease of exposition. Having estimated Eq. 2, \( \ln y_{i,t} \) can be expressed using the associated direct, indirect and total elasticities in a similar way.

\[
\ln y_{i,t} = f \left[ \begin{array}{c}
\xi_{1}^{Dir} t_{1}, \xi_{2}^{Dir} t_{2}, TL \left( x_{i} \right)^{Dir}, z_{i} \partial^{Dir}, \xi_{1}^{Ind} \sum_{j=1}^{N} t_{j}, \xi_{2}^{Ind} \sum_{j=1}^{N} t_{j}, \sum_{j=1}^{N} z_{j} \partial^{Ind} \\
\xi_{1}^{Tot} \left( t_{1} + \sum_{j=1}^{N} t_{j} \right), \xi_{2}^{Tot} \left( t_{2} + \sum_{j=1}^{N} t_{j} \right), TL \left( x_{i} + \sum_{j=1}^{N} x_{j} \right)^{Tot}, \left( z_{i} + \sum_{j=1}^{N} z_{j} \right) \partial^{Tot}
\end{array} \right], \tag{7}
\]

\[
= f \left[ \begin{array}{c}
\xi_{1}^{Dir} t_{1}, \xi_{2}^{Dir} t_{2}, TL \left( x_{i} \right)^{Dir}, z_{i} \partial^{Dir}, \xi_{1}^{Ind} \sum_{j=1}^{N} t_{j}, \xi_{2}^{Ind} \sum_{j=1}^{N} t_{j}, \sum_{j=1}^{N} z_{j} \partial^{Ind} \\
\xi_{1}^{Tot} \left( t_{1} + \sum_{j=1}^{N} t_{j} \right), \xi_{2}^{Tot} \left( t_{2} + \sum_{j=1}^{N} t_{j} \right), TL \left( x_{i} + \sum_{j=1}^{N} x_{j} \right)^{Tot}, \left( z_{i} + \sum_{j=1}^{N} z_{j} \right) \partial^{Tot}
\end{array} \right], \tag{8}
\]
where the direct translog production function is:

\[
TL(x_i)^{\text{Dir}} = \gamma^{\text{Dir}} \ln x_i + \frac{1}{2} \ln x_i' \Theta^{\text{Dir}} \ln x_i,
\]

the indirect translog production function is:

\[
TL \left( \sum_{j=1}^{N} x_j \right)^{\text{Ind}} = \gamma^{\text{Ind}} \sum_{j=1}^{N} \ln x_j + \frac{1}{2} \sum_{j=1}^{N} \ln x_j' \Theta^{\text{Ind}} \sum_{j=1}^{N} \ln x_j
\]

and the total translog production function is:

\[
TL \left( x_i + \sum_{j=1}^{N} x_j \right)^{\text{Tot}} = \gamma^{\text{Tot}} \ln \left( x_i + \sum_{j=1}^{N} x_j \right) + \frac{1}{2} \ln \left( x_i + \sum_{j=1}^{N} x_j \right)' \Theta^{\text{Tot}} \ln \left( x_i + \sum_{j=1}^{N} x_j \right).
\]

\(\gamma^{\text{Dir}}, \gamma^{\text{Ind}}\) and \(\gamma^{\text{Tot}}\) are the direct, indirect and total counterparts of the \(\gamma'\) vector of parameters in \(TL(x_i)\) (see the above discussion of Eq. 1). Likewise \(\Theta^{\text{Dir}}, \Theta^{\text{Ind}}\) and \(\Theta^{\text{Tot}}\) are matrices of direct, indirect and total \(\theta\) parameters and are the direct, indirect of total counterparts of \(\Theta\) in \(TL(x_i)\).

From the properties of the non-spatial translog production function (Christensen et al., 1973) (i)-(iv) follow. (i) Eq. 7 is twice differentiable with respect to each of the \(i\)th unit’s inputs and the summation of each input across the \(j\)th units. (ii) Eq. 8 is twice differentiable with respect to the summation of each input across the \(N\) units in the sample. (iii) The matrices of second order partial derivatives of Eq. 7 with respect to each of the \(i\)th unit’s inputs and the summation of each input across the \(j\)th units, which we refer to as the direct and indirect Hessians (denoted \(H^{\text{Dir}}\) and \(H^{\text{Ind}}\), respectively), are symmetric because of the symmetry restrictions which are imposed on \(\Theta^{\text{Dir}}\) and \(\Theta^{\text{Ind}}\). (iv) Similarly, the matrix of second order partial derivatives of Eq. 8 with respect to the summation of each input across the \(N\) units in the sample, which we refer to as the total Hessian (denoted \(H^{\text{Tot}}\)), is symmetric because of the symmetry restrictions which are imposed on \(\Theta^{\text{Tot}}\).

By applying the approach in Diewert and Wales (1987) for non-spatial translog functions, we can check the curvature of the direct, indirect and total translog production functions at the sample mean and outside the sample mean by checking the sign patterns of the principal minors of \(H^{\text{Dir}}, H^{\text{Ind}}\) and \(H^{\text{Tot}}\). This involves recognising that the sign patterns of the principal minors of \(H^{\text{Dir}}, H^{\text{Ind}}\) and \(H^{\text{Tot}}\) equal the sign patterns of the principal minors of three related matrices, \(\widehat{H}^{\text{Dir}}, \widehat{H}^{\text{Ind}}\) and \(\widehat{H}^{\text{Tot}}\), respectively. Using the indirect case to illustrate the form of \(\widehat{H}^{\text{Dir}}, \widehat{H}^{\text{Ind}}\) and \(\widehat{H}^{\text{Tot}}\):

\[
\widehat{H}^{\text{Ind}} = \Theta^{\text{Ind}} - \gamma^{\text{Ind}} + e^{\text{Ind}} e^{\text{Ind}'}.
\]
where $\Theta^{Ind}$ is as above, $\Gamma^{Ind}$ is a matrix with indirect input elasticities on the main diagonal and zeros elsewhere, and $e^{Ind}$ is a vector of indirect input elasticities. At the sample mean $e^{Ind}$ equals the vector of indirect parameters $\gamma^{Ind}$. This is because when the data is mean adjusted all the quadratic and cross terms in $TL \left( \sum_{j=1}^{N} x_j \right)^{Ind}$ are zero. For verification that the sign patterns of the principal minors of $H^{Dir}$, $H^{Ind}$ and $H^{Tot}$ are equal to the sign patterns of the principal minors of $\bar{H}^{Dir}$, $\bar{H}^{Ind}$ and $\bar{H}^{Tot}$, and that the form of $H^{Dir}$, $H^{Ind}$ and $H^{Tot}$ is as in Eq. 9 see the Appendix which is online supplementary material. Concavity of the direct, indirect and total translog production functions requires that $H^{Dir}$, $H^{Ind}$ and $H^{Tot}$ are negative semi-definite (i.e. all the odd-numbered principal minors are non-positive and all the even-numbered principal minors are non-negative). Outside the sample mean we calculate the proportion of the $H^{Dir}$, $H^{Ind}$ and $H^{Tot}$ matrices which are negative semi-definite over the sample.

The internal, external and total monotonicity propositions for the $k$th input hold at the sample mean and outside the sample mean if the direct, indirect and total elasticities for the $k$th input are $\geq 0$. Outside the sample mean we calculate the proportion of the direct, indirect and total elasticities for the $k$th input which satisfy Propositions 4 – 6.

4 Application to Aggregate Production of European Countries

4.1 Data and the Spatial Weights Matrices

We estimate Eqs. 1 and 2 using 16 inverse distance specifications of $W$ and balanced panel data for 41 European countries for the period 1990 – 2011. All the data was extracted from version 8.0 of the Penn World Table (Feenstra et al., 2015), PWT8.0, which is the first version of the Penn World Table to include data on capital stock. Output is output-side real GDP, $y$ (in 2005 million U.S. dollars at 2005 PPPs, $\text{rgdpo}$), where PWT8.0 notation for the variable is in parentheses. As recommended in the documentation which accompanies PWT8.0 we use $\text{rgdpo}$ to analyse productivity across countries rather than expenditure-side real GDP ($\text{rgdpe}$) or GDP at 2005 national prices ($\text{rgdpm}$) (see page 31 in Feenstra et al., 2013). $x$ is a $(1 \times 2)$ vector of input levels. The first input is the labour input and is the number of people engaged, $x_1 (\text{emp})$. Real capital stock at current PPPs is the second input, $x_2$ (in 2005 million U.S. dollars, $\text{ck}$).\(^{14}\) $z$ is a $(1 \times 3)$ vector of variables. $z_1$ is net exports of merchandise as a share of GDP (where $z_1 = csh\_x + csh\_m$ because all the observations for $csh\_m$ in PWT8.0 are negative to signify that imports are a leakage). $z_2$ is government spending as a share of GDP ($csh\_g$) and $z_3$

\(^{14}\) Following the documentation which accompanies PWT8.0 (see page 13 in Inklaar and Timmer, 2013) we use $\text{ck}$ as our measure of real capital stock rather than real capital stock at 2005 national prices (in 2005 million U.S. dollars, $\text{rkna}$).
is a dummy variable for EU membership. All the continuous variables which are not shares are logged and then mean adjusted so that the first order parameters from the direct, indirect and total translog production functions can be interpreted as elasticities at the sample mean. The descriptive statistics for the continuous variables are presented in Table 1 and are for the raw data.

[Insert Table 1]

All 16 specifications of $W$ are row-normalised inverse distance matrices so the spatial weights are strictly exogenous. Row-normalising adjusts for the absolute scale effects so spillovers are inversely related to the relative great circle distance between countries. The first specification of $W$ is denoted $W_{\text{Dist}}$ and is a dense matrix which is constructed using the inverse distance between each pair of capital cities. The other 15 specifications of $W$ are sparse and are inverse distance matrices with cut-offs. Five specifications of $W$ relate to the nearest $3-7$ capital cities ($W_{3\text{Near}}, ..., W_{7\text{Near}}$). The remaining ten specifications of $W$ are constructed using inverse distances to the capital cities of the biggest $3-7$ import and export partners ($W_{3\text{Import}}, ..., W_{7\text{Import}}$ and $W_{3\text{Export}}, ..., W_{7\text{Export}}$). The biggest $3-7$ import and export partners are based on the average import and average export flows in 2000 U.S. dollars over the period 2000–2011, where the data is from the IMF Direction of Trade Statistics. We do not use the biggest $3-7$ import and export flows, on average, over the period 2000 – 2011 to construct the $W_{3\text{Import}}, ..., W_{7\text{Import}}$ and $W_{3\text{Export}}, ..., W_{7\text{Export}}$ matrices because the spatial weights would be endogenous.

4.2 Overview of the Fitted Models, Curvature and Monotonicity

In this subsection we justify our preference for the $W_{5\text{Import}}$ partial spatial Durbin model by providing an overview of the salient features of the: fitted spatial models, diagnostic test results, marginal effects and returns to scale. We then discuss the concavity and monotonicity results for our preferred partial spatial Durbin model. In the next subsection we discuss in more detail the marginal effects from our preferred partial spatial Durbin model, which is followed by a subsection on the spatial returns to scale estimates. In Table 2 we present the Within estimate of our preferred partial spatial Durbin model. To enable comparisons we also present the Within estimates of the non-spatial model and the $W_{5\text{Import}}$ SAR model. With regard to the results for the time period dummy variables in Table 2, as we noted in footnote 8, when we estimate the non-spatial model using STATA the time period dummy variables for 2000 and 2011 are dropped due to perfect collinearity with the time trend, and the coefficients on the time period dummy variables are relative to the omitted category which is the time period dummy variable for 1990. To remain consistent when estimating the partial spatial Durbin and SAR models using
MATLAB we drop the same time period dummy variables. As is evident from Table 2, a number of time period dummy variables have a significant effect in the three reported models. In the next subsection, for our preferred partial spatial Durbin model we use the direct coefficient on the time trend and the direct coefficients on the time period dummy variables to check if the time period dummy variables represent significant deviations from the time trend.

[Insert Table 2]

All the partial spatial Durbin and SAR models yield non-negligible positive estimates of $\lambda$ ranging from $0.313 - 0.724$, all of which are significant at the 0.1% level. Moreover, to test the null hypothesis that the fixed effects in the partial spatial Durbin and SAR models are not jointly significant (i.e. $\alpha_i = \ldots = \alpha_N = \kappa$) we perform a likelihood ratio (LR) test on each of the fitted models against the corresponding pooled model. The test statistic is chi-squared distributed with degrees of freedom equal to the number of restrictions which must be imposed on the unrestricted model to obtain the restricted model, which in this case is $N - 1$. For all the partial spatial Durbin and SAR models we reject the null at the 0.1% level, thereby justifying the inclusion of fixed effects in Eqs. 1 and 2.

We use the approach which Pfaffermayr (2009) uses to choose between different spatial weights matrices and base our model selection on the Akaike Information Criteria (AIC). To check the robustness of the model selection using the AIC we also use the Bayesian Information Criteria (BIC). We have a strong preference for the $W_{5\text{Import}}$ partial spatial Durbin model as it yields the lowest AIC and BIC values. Furthermore, we use a test for a balanced panel fixed effects estimator to test the residuals from the non-spatial model and the residuals from the $W_{5\text{Import}}$ partial spatial Durbin model for spatial autocorrelation. The test is an extension of Cliff and Ord’s (1972) test for spatial autocorrelation in the residuals from a cross-sectional model. For a formal presentation of the test which we employ see Mutl and Pfaffermayr (2010), where the $I^2$ test statistic for this test follows a chi-squared distribution with one degree of freedom. For a formal presentation of the corresponding test of the residuals from an unbalanced panel fixed effects estimator see Pfaffermayr (2013). For the $W_{5\text{Import}}$ partial spatial Durbin model the $I^2$ test statistic is 0.697 so we cannot reject the null hypothesis of no spatial autocorrelation at the 5% level.\(^{15}\) In contrast, when we perform the same test on the residuals from the non-spatial model, for all 16 specifications of $W$ the $I^2$ test statistic is such that we reject the null at the 5% level.

The $W_{5\text{Import}}$ partial spatial Durbin model yields direct labour and direct capital elasticities at the sample mean of 0.608 and 0.298, respectively, both of which are significant.

\(^{15}\)When we perform the same test on, firstly, the residuals from the other 15 partial spatial Durbin models and, secondly, the residuals from the 16 SAR models, in each case the $I^2$ test statistic is such that we cannot reject the null of no spatial autocorrelation at the 5% level.
at the 0.1% level. Finding that the direct labour elasticity is larger than the direct capital elasticity is in line with the empirical findings of key macroeconomic studies (e.g. Ireland, 2004, and Smets and Wouters, 2003). The direct labour elasticity from the $W_{5\text{Import}}$ partial spatial Durbin model is also within the range reported by the EU Commission for the labour income share for the EU 15 member states ($0.54 - 0.68$ from Table 1 in Arpaia et al., 2009). Furthermore, for the $W_{5\text{Import}}$ partial spatial Durbin model we cannot reject constant internal returns to scale. This is consistent with the assumption of constant returns to scale in the classic non-spatial macroeconomic theories in Ireland (2004) and Smets and Wouters (2003), and is also in line with key empirical evidence (e.g. Burnside et al., 1995). All things considered we think it is reasonable to have a strong preference for the $W_{5\text{Import}}$ partial spatial Durbin model and so the remainder of the empirical analysis focuses on the results from this model.

Turning now to the spatial concavity and spatial monotonicity results. A production function assumes that the $i$th unit’s output is concave in the effect of the $i$th unit’s own inputs. Here there are added issues such as whether the $i$th unit’s output is concave/convex in the effect of the sum of the input vectors across the other $N - 1$ units and concave/convex in the effect of the sum of the input vectors across all $N$ units in the sample. Applying the empirical check of concavity which we set out above to the non-spatial translog production function and the direct, indirect and total translog production functions from the $W_{5\text{Import}}$ partial spatial Durbin model indicates at the sample mean that: (i) the non-spatial function is concave and (ii) the $W_{5\text{Import}}$ partial spatial Durbin model yields concave direct, concave indirect and concave total functions, even though external and total concavity, unlike internal concavity, are not theoretical properties of the spatial model.

Applying the above empirical check of concavity outside the sample mean to the non-spatial translog production function and the direct, indirect and total translog production functions from the $W_{5\text{Import}}$ partial spatial Durbin model is very revealing. We find that the non-spatial translog production function is concave for 91% of the sample. By controlling for spatial dependence using our preferred model specification we observe more evidence of internal concavity outside the sample mean. This is evident because the direct translog production function from the $W_{5\text{Import}}$ partial spatial Durbin model is concave for 100% of the sample. Interestingly, although the direct, indirect and total translog production functions from the $W_{5\text{Import}}$ partial spatial Durbin model are all concave at the sample mean, outside the sample mean we observe far less evidence of external and total concavity than we do internal concavity. We find that the indirect and total translog production functions from the $W_{5\text{Import}}$ partial spatial Durbin model are concave for 66% and 67% of the sample, respectively. This suggests for our preferred spatial model that external concavity is the principal driver of total concavity. This is evident over the entire study period for EU and non-EU countries from Figure 1, which
summarises the external and total concavity results from the $W_{5Import}$ partial spatial Durbin model. Specifically, Figure 1 presents the proportions of the sample for which the indirect and total functions from our preferred spatial model are concave for countries where we do not observe internal, external and total concavity for 100% of the sample.\textsuperscript{16}

[Insert Figure 1]

A production function assumes that the $i$th unit’s output is monotonically increasing in each of the $i$th unit’s inputs. New lines of enquiry which follow from a production function which contains the SAR variable include whether the $i$th unit’s output is monotonically increasing/decreasing in each of the summed inputs across the other $N - 1$ units, and monotonically increasing/decreasing in each of the summed inputs across all $N$ units in the sample. At the sample mean, the non-spatial input elasticities and the direct, indirect and total input elasticities from the $W_{5Import}$ partial spatial Durbin model are all positive. Therefore at the sample mean, the non-spatial model indicates that a country’s output is monotonically increasing in its own labour and own capital, and our preferred spatial model satisfies the internal, external and total monotonicity propositions. The non-spatial labour and capital elasticities outside the sample mean indicate that, on average, a country’s output is monotonically increasing in its labour and capital for 88% and 61% of the sample, respectively. On average, the direct, indirect and total labour (capital) elasticities from the $W_{5Import}$ partial spatial Durbin model satisfy the internal, external and total monotonicity propositions for 93% (79%), 87% (60%) and 91% (64%) of the sample, respectively. We can therefore conclude that, on average, there is much more evidence of internal capital monotonicity from our preferred spatial model than there is capital monotonicity from the non-spatial model. Also, even though external monotonicity and total monotonicity are not properties of spatial production functions, from our preferred spatial model we observe a substantial amount of evidence of external and total labour monotonicity and quite a lot of evidence of external and total capital monotonicity.

4.3 Further Discussion of the Preferred Model

The estimates of $\lambda$ from the fitted spatial models are not spillover elasticities. The spillover elasticities are the indirect marginal effects which are a function of, among other things, $\lambda$ (see Eq. 5b). Direct, indirect and total marginal effects for our preferred $W_{5Import}$ partial spatial Durbin model are reported in Table 3. To enable comparisons we also present in Table 3 the direct, indirect and total marginal effects for the $W_{5Import}$ SAR model. We can see from Table 3 that the indirect input elasticities from the $W_{5Import}$ partial spatial Durbin and SAR models are positive and significant at the 0.1% level. For

\textsuperscript{16}Figure 1 relates to 30 EU and non-EU countries.
both models this suggests for the sample average country there are positive labour and positive capital spillovers coming to the country from the other countries in the sample. In contrast to the corresponding direct labour and direct capital elasticities from the $W_{5Import}$ partial spatial Durbin and SAR models which are similar in magnitude (see Table 3), the corresponding indirect labour and indirect capital elasticities are very different. Specifically, the indirect labour and indirect capital elasticities from the $W_{5Import}$ partial spatial Durbin model are 0.971 and 0.425, respectively, which are much smaller than the indirect labour and indirect capital elasticities from the $W_{5Import}$ SAR model (1.442 and 0.680, respectively). Since the direct and indirect input elasticities from the $W_{5Import}$ partial spatial Durbin and SAR models are positive and significant at the 0.1% level, it follows that the total input elasticities are also positive and significant at the 0.1% level (see Table 3).

From the results reported in Table 3 we can see that a number of the direct coefficients on the time period dummy variables are significant. We can also see for both models in Table 3 that the indirect coefficients on the time period dummy variables (i.e. the spillover effect of a common shock in a particular time period) are in each case larger in absolute magnitude than the corresponding direct coefficient. For both the $W_{5Import}$ partial spatial Durbin model and the $W_{5Import}$ SAR model, $t$-tests of the difference between the direct coefficient on a time period dummy variable and the direct coefficient on the time trend, $t$, indicate that a number of time period dummy variables capture significant deviations from the trend at the sample mean. For example, for the preferred $W_{5Import}$ partial spatial Durbin model, we find that the direct coefficients on the 1992, 2002, 2003, 2007 and 2008 time period dummy variables are significantly different from the direct coefficient on $t$ at the 5% level. Furthermore, it is evident from Table 3 that the direct elasticities for the three $z$ variables from the $W_{5Import}$ partial spatial Durbin model are not significant at the 5% level. This is an important finding because the non-spatial model in Table 2 suggests that government size ($z_2$) has a significant negative effect and EU membership ($z_3$) has a significant positive effect. We can therefore conclude that spatial models can challenge some widely accepted relationships from standard non-spatial models such as the significant negative relationship which Fölster and Henrekson (2001) observe between government size and economic growth.

4.4 Estimates of Internal, External and Total Returns to Scale

Using the labour and capital elasticities at the sample mean from the non-spatial model we compute returns to scale in the usual way which we denote $RTS^{NSp}$. Using the direct, indirect and total labour and capital elasticities at the sample mean from the
spatial models we calculate $RTS^{Int}$, $RTS^{Ext}$ and $RTS^{Tot}$ using Eq. 6. In Table 4 we present estimates of $RTS^{Int}$, $RTS^{Ext}$ and $RTS^{Tot}$ at the sample mean from the partial spatial Durbin models, where the $t$–statistics for one-sided tests of the nulls of constant internal, constant external and constant total returns to scale are in parentheses. For the sample average country, the estimate of $RTS^{NSp}$ is 0.898 and is not significantly less than 1 at the 5% level ($t$–statistic of −1.42). Similarly, it is evident from Table 4 that we cannot reject constant internal returns to scale for the preferred $W_{5Import}$ partial spatial Durbin model at the 5% level.

From Table 4 we can see that we cannot reject constant external returns to scale at the 5% level for the preferred $W_{5Import}$ partial spatial Durbin model. It is also apparent from Table 4 that we reject constant total returns to scale at the 0.1% level for the $W_{5Import}$ partial spatial Durbin model in favour of increasing total returns. This highlights how the combined effect of internal and external returns to scale can result in a classification of total returns to scale which differs from the classification of its two constituent parts. Our findings provide some empirical support for the endogenous growth theories which are based on the assumption of increasing total returns to scale (e.g. Romer, 1986; 1987). It should, however, be noted that in the above endogenous growth theories, increasing total returns to scale are due to knowledge or composite capital spillovers. In contrast, in our empirical analysis increasing total returns are due labour and physical capital spillovers. Furthermore, we observe robust evidence of increasing total returns to scale because for all 16 partial spatial Durbin models and all but one of the 16 SAR models we reject constant total returns to scale in favour of increasing total returns at nominal significance levels (at least the 5% level).

For EU and non-EU countries, in Figure 2 we present annual average estimates of $RTS^{NSp}$ and $RTS^{Int}$ and $RTS^{Tot}$ from the $W_{5Import}$ partial spatial Durbin model. We note that in Figure 2 the difference between $RTS^{Int}$ and $RTS^{Tot}$ is $RTS^{Ext}$. The most striking feature of Figure 2 is the marked fall (rise) in $RTS^{Tot}$ for EU (non-EU) countries as a result of the EU enlargement in 2004. This suggests that, on average, $RTS^{Tot}$ for the 2004 EU accession countries is appreciably lower than $RTS^{Tot}$ for the other countries in the sample. Furthermore, Figure 2 post 2004 indicates that the EU enlargement had, on average, a persistent impact on $RTS^{Tot}$ for EU and non-EU countries.

5 Concluding Remarks

Having estimated spatial translog production functions using panel data for European countries over the period 1990 – 2011 we followed LeSage and Pace (2009) and calculated
the direct, indirect and total marginal effects. Using this approach the effect of geography is related to the factor inputs, which enabled us to extend classic characteristics of production, namely returns to scale and diminishing marginal productivity of factor inputs, to the spatial case. Firstly, we proposed internal, external and total returns to scale. Secondly, since the direct, indirect and total marginal effects can be used to construct three translog production functions, we performed empirical checks to ascertain if the fitted direct, indirect and total functions were concave (i.e. internal, external and total concavity, respectively). These two contributions are particularly appealing because they are not limited to a production function and the translog specification. There is thus scope for wider application of the spatial returns to scale and the empirical check for spatial curvature which we have proposed to other technologies (cost, revenue, standard and alternative profit, and input and output distance functions) and other functional forms (e.g. Cobb-Douglas, Fourier flexible, Generalized McFadden and Generalized Leontief).

**Supplementary material**

Supplementary material (the Appendix) is available online at the OUP website.

**Acknowledgments**

The authors would like to express their thanks to the Associate Editor, Anindya Banerjee, and to Tom Weyman-Jones and two anonymous referees for constructive comments on an earlier version of this paper.

**References**


Figure 1: Concavity results for the preferred partial spatial Durbin model (W5 Import)

Figure 2: Selected average returns to scale for EU and non-EU countries
### Table 1: Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real GDP (2005 million U.S. dollars at 2005 PPPs)</td>
<td>$y$</td>
<td>366,380</td>
<td>589,103</td>
<td>4,049</td>
</tr>
<tr>
<td>Number of people engaged (millions)</td>
<td>$x_1$</td>
<td>8.41</td>
<td>13.04</td>
<td>0.13</td>
</tr>
<tr>
<td>Real capital stock (2005 million U.S. dollars at current PPPs)</td>
<td>$x_2$</td>
<td>1,243,469</td>
<td>2,128,382</td>
<td>9,206</td>
</tr>
<tr>
<td>Exports of merchandise minus imports of merchandise as a share of GDP i.e. net trade openness</td>
<td>$z_1$</td>
<td>-0.05</td>
<td>0.13</td>
<td>-0.59</td>
</tr>
<tr>
<td>Government spending as a share of GDP</td>
<td>$z_2$</td>
<td>0.22</td>
<td>0.08</td>
<td>0.07</td>
</tr>
</tbody>
</table>

### Table 2: Non-spatial model and the partial spatial Durbin and spatial autoregressive W5 import models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real GDP (2005 million U.S. dollars at 2005 PPPs)</td>
<td>$y$</td>
<td>366,380</td>
<td>589,103</td>
<td>4,049</td>
</tr>
<tr>
<td>Number of people engaged (millions)</td>
<td>$x_1$</td>
<td>8.41</td>
<td>13.04</td>
<td>0.13</td>
</tr>
<tr>
<td>Real capital stock (2005 million U.S. dollars at current PPPs)</td>
<td>$x_2$</td>
<td>1,243,469</td>
<td>2,128,382</td>
<td>9,206</td>
</tr>
<tr>
<td>Exports of merchandise minus imports of merchandise as a share of GDP i.e. net trade openness</td>
<td>$z_1$</td>
<td>-0.05</td>
<td>0.13</td>
<td>-0.59</td>
</tr>
<tr>
<td>Government spending as a share of GDP</td>
<td>$z_2$</td>
<td>0.22</td>
<td>0.08</td>
<td>0.07</td>
</tr>
</tbody>
</table>

PSDM denotes the partial spatial Durbin model and SAR denotes the spatial autoregressive model.

* * * denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

LL denotes log-likelihood and the t-statistics are in parentheses.
Table 3: Marginal effects from the partial spatial Durbin and spatial autoregressive W5 import models

<table>
<thead>
<tr>
<th></th>
<th>PSDM W5Impact</th>
<th>SAR W5Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marginal Eff</td>
<td>t-stat</td>
</tr>
<tr>
<td>ln z1 Direct</td>
<td>0.608***</td>
<td>8.79</td>
</tr>
<tr>
<td>Indirect</td>
<td>0.971***</td>
<td>4.79</td>
</tr>
<tr>
<td>Total</td>
<td>1.579***</td>
<td>6.18</td>
</tr>
<tr>
<td>ln z2 Direct</td>
<td>0.298***</td>
<td>7.56</td>
</tr>
<tr>
<td>Indirect</td>
<td>0.425***</td>
<td>3.36</td>
</tr>
<tr>
<td>Total</td>
<td>0.723***</td>
<td>4.67</td>
</tr>
<tr>
<td>(ln z1)² Direct</td>
<td>-0.253**</td>
<td>-9.72</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.606***</td>
<td>-6.04</td>
</tr>
<tr>
<td>Total</td>
<td>-0.861***</td>
<td>-7.36</td>
</tr>
<tr>
<td>ln z1 ln z2</td>
<td>0.415**</td>
<td>13.96</td>
</tr>
<tr>
<td>Direct</td>
<td>0.752***</td>
<td>4.99</td>
</tr>
<tr>
<td>Total</td>
<td>1.226***</td>
<td>7.18</td>
</tr>
<tr>
<td>1</td>
<td>-0.012***</td>
<td>-4.39</td>
</tr>
<tr>
<td>Direct</td>
<td>-0.042***</td>
<td>-3.75</td>
</tr>
<tr>
<td>Total</td>
<td>-0.037***</td>
<td>-4.02</td>
</tr>
<tr>
<td>1²</td>
<td>0.000</td>
<td>-1.33</td>
</tr>
<tr>
<td>Direct</td>
<td>-0.001</td>
<td>-1.30</td>
</tr>
<tr>
<td>Total</td>
<td>-0.001</td>
<td>-1.32</td>
</tr>
<tr>
<td>ln z1 t Direct</td>
<td>-</td>
<td>-0.012***</td>
</tr>
<tr>
<td>Indirect</td>
<td>-</td>
<td>-0.029***</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>-0.040***</td>
</tr>
<tr>
<td>ln z2 t Direct</td>
<td>-</td>
<td>-0.012***</td>
</tr>
<tr>
<td>Indirect</td>
<td>-</td>
<td>-0.029***</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>-0.041***</td>
</tr>
<tr>
<td>z1</td>
<td>0.014</td>
<td>0.16</td>
</tr>
<tr>
<td>Direct</td>
<td>0.003</td>
<td>0.02</td>
</tr>
<tr>
<td>Indirect</td>
<td>0.016</td>
<td>0.07</td>
</tr>
<tr>
<td>Total</td>
<td>0.016</td>
<td>0.07</td>
</tr>
<tr>
<td>z2</td>
<td>-0.194</td>
<td>-1.82</td>
</tr>
<tr>
<td>Direct</td>
<td>-0.320</td>
<td>-1.44</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.514</td>
<td>-1.59</td>
</tr>
<tr>
<td>Total</td>
<td>-0.514</td>
<td>-1.59</td>
</tr>
<tr>
<td>ψ₁₁ Direct</td>
<td>-0.055</td>
<td>-1.94</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.042</td>
<td>1.09</td>
</tr>
<tr>
<td>Total</td>
<td>0.064</td>
<td>1.00</td>
</tr>
<tr>
<td>ψ₁₂ Direct</td>
<td>-0.028</td>
<td>-1.08</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.012</td>
<td>-1.07</td>
</tr>
<tr>
<td>Total</td>
<td>-0.030</td>
<td>-1.15</td>
</tr>
<tr>
<td>ψ₁₃ Direct</td>
<td>-0.069**</td>
<td>-2.64</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.130**</td>
<td>-2.60</td>
</tr>
<tr>
<td>Total</td>
<td>-0.200**</td>
<td>-2.64</td>
</tr>
<tr>
<td>ψ₂₁ Direct</td>
<td>-0.102</td>
<td>-1.95</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.156</td>
<td>-1.96</td>
</tr>
<tr>
<td>Total</td>
<td>-0.258</td>
<td>-2.84</td>
</tr>
<tr>
<td>ψ₂₂ Direct</td>
<td>-0.047</td>
<td>-1.75</td>
</tr>
<tr>
<td>Indirect</td>
<td>-0.088</td>
<td>-1.75</td>
</tr>
<tr>
<td>Total</td>
<td>-0.135</td>
<td>-1.76</td>
</tr>
</tbody>
</table>

PSDM denotes the partial spatial Durbin model and SAR denotes the spatial autoregressive model. *, **, *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.
<table>
<thead>
<tr>
<th>Model</th>
<th>PSDM RTS\textsuperscript{int}</th>
<th>PSDM RTS\textsuperscript{ext}</th>
<th>PSDM RTS\textsuperscript{tot}</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{\text{Dist}}$</td>
<td>1.102</td>
<td>0.910</td>
<td>2.012***</td>
</tr>
<tr>
<td></td>
<td>(1.58)</td>
<td>(-0.30)</td>
<td>(3.15)</td>
</tr>
<tr>
<td>$W_{3\text{Near}}$</td>
<td>0.991</td>
<td>0.365***</td>
<td>1.356**</td>
</tr>
<tr>
<td></td>
<td>(-0.13)</td>
<td>(-6.04)</td>
<td>(2.33)</td>
</tr>
<tr>
<td>$W_{4\text{Near}}$</td>
<td>1.011</td>
<td>0.465***</td>
<td>1.476**</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(-4.51)</td>
<td>(2.92)</td>
</tr>
<tr>
<td>$W_{5\text{Near}}$</td>
<td>1.064</td>
<td>0.617**</td>
<td>1.682***</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(-2.74)</td>
<td>(3.74)</td>
</tr>
<tr>
<td>$W_{6\text{Near}}$</td>
<td>1.063</td>
<td>0.767</td>
<td>1.829***</td>
</tr>
<tr>
<td></td>
<td>(0.93)</td>
<td>(-1.51)</td>
<td>(4.29)</td>
</tr>
<tr>
<td>$W_{7\text{Near}}$</td>
<td>1.069</td>
<td>0.862</td>
<td>1.930***</td>
</tr>
<tr>
<td></td>
<td>(1.01)</td>
<td>(-0.85)</td>
<td>(4.65)</td>
</tr>
<tr>
<td>$W_{3\text{Import}}$</td>
<td>0.978</td>
<td>0.401***</td>
<td>1.378*</td>
</tr>
<tr>
<td></td>
<td>(-0.34)</td>
<td>(-4.48)</td>
<td>(2.23)</td>
</tr>
<tr>
<td>$W_{4\text{Import}}$</td>
<td>0.961</td>
<td>0.910</td>
<td>1.871***</td>
</tr>
<tr>
<td></td>
<td>(-0.57)</td>
<td>(-0.41)</td>
<td>(3.35)</td>
</tr>
<tr>
<td>$W_{5\text{Import}}$</td>
<td>0.906</td>
<td>1.397</td>
<td>2.302***</td>
</tr>
<tr>
<td></td>
<td>(-1.40)</td>
<td>(1.45)</td>
<td>(4.13)</td>
</tr>
<tr>
<td>$W_{6\text{Import}}$</td>
<td>0.914</td>
<td>1.399</td>
<td>2.312***</td>
</tr>
<tr>
<td></td>
<td>(-1.25)</td>
<td>(1.45)</td>
<td>(4.17)</td>
</tr>
<tr>
<td>$W_{7\text{Import}}$</td>
<td>0.949</td>
<td>1.155</td>
<td>2.103***</td>
</tr>
<tr>
<td></td>
<td>(-0.73)</td>
<td>(0.64)</td>
<td>(3.91)</td>
</tr>
<tr>
<td>$W_{3\text{Export}}$</td>
<td>0.991</td>
<td>0.441***</td>
<td>1.432**</td>
</tr>
<tr>
<td></td>
<td>(-0.14)</td>
<td>(-1.25)</td>
<td>(2.61)</td>
</tr>
<tr>
<td>$W_{4\text{Export}}$</td>
<td>0.986</td>
<td>0.880</td>
<td>1.867***</td>
</tr>
<tr>
<td></td>
<td>(-0.20)</td>
<td>(-0.61)</td>
<td>(3.65)</td>
</tr>
<tr>
<td>$W_{5\text{Export}}$</td>
<td>1.057</td>
<td>0.717</td>
<td>1.774***</td>
</tr>
<tr>
<td></td>
<td>(0.86)</td>
<td>(-1.54)</td>
<td>(3.61)</td>
</tr>
<tr>
<td>$W_{6\text{Export}}$</td>
<td>1.075</td>
<td>0.925</td>
<td>2.000***</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(-0.35)</td>
<td>(3.94)</td>
</tr>
<tr>
<td>$W_{7\text{Export}}$</td>
<td>1.063</td>
<td>0.798</td>
<td>1.861***</td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(-1.03)</td>
<td>(3.73)</td>
</tr>
</tbody>
</table>

*, **, *** denote that we reject the null hypotheses of constant internal, constant external and constant total returns to scale at the 5%, 1% and 0.1% levels, respectively, where the t-statistics for the tests are in parentheses.
Returns to scale and curvature in the presence of spillovers: evidence from European countries

Anthony J. Glass\textsuperscript{1}, Karligash Kenjegalieva\textsuperscript{2} and Robin C. Sickles\textsuperscript{3}

Appendix: Derivation of the $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$ matrices

To check that the direct, indirect and total translog production functions are concave at and outside the sample mean we check that the sign patterns of the direct, indirect and total Hessian matrices ($H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$) are negative semi-definite. This involves recognising that the sign patterns of the principal minors of $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$ are equal to the sign patterns of the principal minors of three related matrices, $H_{\text{Dir}}^{\text{Ind}}$, $H_{\text{Ind}}^{\text{Ind}}$ and $H_{\text{Tot}}^{\text{Ind}}$, which we obtain from $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$. We must: (i) derive $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$; (ii) show that $H_{\text{Dir}}^{\text{Ind}}$, $H_{\text{Ind}}^{\text{Ind}}$ and $H_{\text{Tot}}^{\text{Ind}}$ can be obtained from $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$; (iii) show that the semi-definiteness of $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$ equals the semi-definiteness of $H_{\text{Dir}}^{\text{Ind}}$, $H_{\text{Ind}}^{\text{Ind}}$ and $H_{\text{Tot}}^{\text{Ind}}$; and (iv) show that $H_{\text{Dir}}$, $H_{\text{Ind}}$ and $H_{\text{Tot}}$ take the form given in Eq. 9. The approach for (i)-(iv) is the same for the direct, indirect and total cases. To illustrate the approach we consider the indirect translog production function.

The indirect translog production function in scalar form is:

$$\text{TL} \left( \sum_{k=1}^{K} \sum_{j=1}^{N} x_{k,j} \right)^{\text{Ind}} = \sum_{k=1}^{K} \left( \gamma^{\text{Ind}}_k \sum_{j=1}^{N} \ln x_{k,j} \right) + \frac{1}{2} \sum_{k=1}^{K} \sum_{l=1}^{K} \left( \theta^{\text{Ind}}_{kl} \sum_{j=1}^{N} \ln x_{k,j}^{\prime} \sum_{j=1}^{N} \ln x_{l,j}^{\prime} \right)$$

and in matrix form is:

$$\text{TL} \left( \sum_{j=1}^{N} x_{j} \right)^{\text{Ind}} = \gamma^{\text{Ind}} \sum_{j=1}^{N} \ln x_{j} + \frac{1}{2} \sum_{j=1}^{N} \ln x_{j}^{\prime} \Theta^{\text{Ind}} \sum_{j=1}^{N} \ln x_{j},$$

where $\gamma^{\text{Ind}} = (\gamma^{\text{Ind}}_1, ..., \gamma^{\text{Ind}}_K)$ is a vector of indirect parameters; $\ln x_{j}^{\prime} = (\ln x_{1,j}, ..., \ln x_{K,j})$ is a vector of logged inputs indexed $k$ or $l = 1, ..., K$ for the $j$th neighbouring unit in the sample; and

\textsuperscript{1}School of Business and Economics, Loughborough University, Leicestershire, UK, LE11 3TU; e-mail: A.J.Glass@lboro.ac.uk.
\textsuperscript{2}School of Business and Economics, Loughborough University; e-mail: K.A.Kenjegalieva@lboro.ac.uk.
\textsuperscript{3}Department of Economics, Rice University, Houston, TX, and School of Business and Economics, Loughborough University; e-mail: rsickles@rice.edu
\[ \Theta^{\text{Ind}} = \begin{pmatrix} \theta^{\text{Ind}}_{11} & \cdots & \theta^{\text{Ind}}_{1K} \\ \vdots & \ddots & \vdots \\ \theta^{\text{Ind}}_{K1} & \cdots & \theta^{\text{Ind}}_{KK} \end{pmatrix} \]

is a matrix of indirect parameters.

\( H^{\text{Ind}} \) is the following symmetric matrix:

\[
H^{\text{Ind}} = \begin{pmatrix}
\frac{\partial^2 y_i}{\partial \left( \sum_{j=1}^N x_{1,j} \right)^2} & \cdots & \frac{\partial^2 y_i}{\partial \sum_{j=1}^N x_{1,j} \partial \sum_{j=1}^N x_{K,j}} \\
\vdots & \ddots & \vdots \\
\frac{\partial^2 y_i}{\partial \sum_{j=1}^N x_{K,j} \partial \sum_{j=1}^N x_{1,j}} & \cdots & \frac{\partial^2 y_i}{\partial \left( \sum_{j=1}^N x_{K,j} \right)^2}
\end{pmatrix}.
\]

To derive \( H^{\text{Ind}} \) we rely on the following definitions and translog properties, (a)-(d):

(a) \[
\frac{\partial y_i}{\partial \sum_{j=1}^N x_{k,j}} = \frac{y_i}{\sum_{j=1}^N x_{k,j}} \frac{\partial \ln y_i}{\partial \sum_{j=1}^N \ln x_{k,j}} = \frac{y_i}{\sum_{j=1}^N x_{k,j}} \left[ \gamma^{\text{Ind}}_k \sum_{j=1}^N \left( \theta^{\text{Ind}}_{kl} \sum_{j=1}^N \ln x_{l,j} \right) \right] = \frac{y_i}{\sum_{j=1}^N x_{k,j}} \epsilon^{\text{Ind}}_k,
\]

where \( \epsilon^{\text{Ind}}_k = \gamma^{\text{Ind}}_k + \sum_{l=1}^K \left( \theta^{\text{Ind}}_{kl} \sum_{j=1}^N \ln x_{l,j} \right) \) and \( \epsilon^{\text{Ind}}_k \) is the indirect elasticity of \( y_i \) with respect to \( \sum_{j=1}^N x_{k,j} \).

(b) The underlying indirect production function (not its logarithmic transformation):

\[
y_i = \exp \left[ \sum_{k=1}^K \left( \gamma^{\text{Ind}}_k \sum_{j=1}^N \ln x_{k,j} \right) + \frac{1}{2} \sum_{k=1}^K \sum_{l=1}^K \left( \theta^{\text{Ind}}_{kl} \sum_{j=1}^N \ln x_{k,j} \sum_{j=1}^N \ln x_{l,j} \right) \right]
\]

is used to evaluate the second order partial derivatives. The second order partial derivatives are given in (c) and (d), where \( x_{k,j} \) and \( x_{m,j} \) denote two distinct inputs of the \( j \)th neighbouring unit:

(c) 

\[
\frac{\partial^2 y_i}{\partial \left( \sum_{j=1}^N x_{k,j} \right)^2} = \frac{\partial}{\partial \sum_{j=1}^N x_{k,j}} \frac{\partial y_i}{\partial \sum_{j=1}^N x_{k,j}} \epsilon^{\text{Ind}}_k = \frac{\sum_{j=1}^N x_{k,j}}{\left( \sum_{j=1}^N x_{k,j} \right)^2} \left( \sum_{j=1}^N x_{k,j} \right) \left( \frac{\partial y_i}{\partial \sum_{j=1}^N \ln x_{k,j}} - \frac{\partial y_i}{\partial \sum_{j=1}^N \ln x_{k,j}} \right) + \epsilon^{\text{Ind}}_k \sum_{j=1}^N x_{k,j} \frac{\partial \ln y_i}{\partial \sum_{j=1}^N \ln x_{k,j}} - (y_i \epsilon^{\text{Ind}}_k \times 1) = \frac{y_i}{\left( \sum_{j=1}^N x_{k,j} \right)^2} \left[ \theta^{\text{Ind}}_{kk} + \left( \epsilon^{\text{Ind}}_k \right)^2 - \epsilon^{\text{Ind}}_k \right].
\]
\[
\frac{\partial^2 y_i}{\partial \sum_{j=1}^{N} x_{k,j} \partial \sum_{j=1}^{N} x_{m,j}} = \frac{\partial}{\partial \sum_{j=1}^{N} x_{m,j}} \frac{y_i}{\sum_{j=1}^{N} x_{k,j}} e^{\text{Ind}}_k = \frac{\sum_{j=1}^{N} x_{k,j} \partial^2 y_i}{\sum_{j=1}^{N} x_{m,j}} - y_i e^{\text{Ind}}_k \frac{\sum_{j=1}^{N} x_{k,j}}{\left(\sum_{j=1}^{N} x_{m,j}\right)^2} - \frac{\partial^2 y_i (y_i e^{\text{Ind}}_k)}{\sum_{j=1}^{N} x_{k,j} \sum_{j=1}^{N} x_{m,j}} - \left(\sum_{j=1}^{N} x_{k,j}\right)^2 \frac{\partial e^{\text{Ind}}_k}{\sum_{j=1}^{N} x_{m,j}} \left(y_i e^{\text{Ind}}_k \times 0 \right)
\]

Collecting the second order partial derivatives:

\[
H^{\text{Ind}} = \begin{pmatrix}
\frac{y_i}{\sum_{j=1}^{N} x_{k,j}} \left(\theta^{\text{Ind}}_{11} + e^{\text{Ind}}_1 \left(e^{\text{Ind}}_1 - 1\right)\right) & \cdots & \frac{y_i}{\sum_{j=1}^{N} x_{k,j}} \left(\theta^{\text{Ind}}_{1K} + e^{\text{Ind}}_1 e^{\text{Ind}}_K\right) \\
\vdots & \ddots & \vdots \\
\frac{y_i}{\sum_{j=1}^{N} x_{k,j}} \left(\theta^{\text{Ind}}_{K1} + e^{\text{Ind}}_K e^{\text{Ind}}_1\right) & \cdots & \frac{y_i}{\sum_{j=1}^{N} x_{k,j}} \left[\theta^{\text{Ind}}_{KK} + e^{\text{Ind}}_K \left(e^{\text{Ind}}_K - 1\right)\right]
\end{pmatrix}
\]

and defining \(\widehat{H}^{\text{Ind}}\) as follows:

\[
\widehat{H}^{\text{Ind}} = \begin{pmatrix}
\theta^{\text{Ind}}_{11} + e^{\text{Ind}}_1 \left(e^{\text{Ind}}_1 - 1\right) & \cdots & \theta^{\text{Ind}}_{1K} + e^{\text{Ind}}_1 e^{\text{Ind}}_K \\
\vdots & \ddots & \vdots \\
\theta^{\text{Ind}}_{K1} + e^{\text{Ind}}_K e^{\text{Ind}}_1 & \cdots & \theta^{\text{Ind}}_{KK} + e^{\text{Ind}}_K \left(e^{\text{Ind}}_K - 1\right)
\end{pmatrix}.
\]

Since in the above \(H^{\text{Ind}}\) matrix \(y_i / \left(\sum_{j=1}^{N} x_{k,j}\right)^2\) and \(y_i / \sum_{j=1}^{N} x_{k,j} \sum_{j=1}^{N} x_{m,j}\) are positive we have the result: **sign definiteness of \(H^{\text{Ind}} = \text{sign definiteness of } \widehat{H}^{\text{Ind}}.**

Rewriting \(\widehat{H}^{\text{Ind}}\) as:

\[
\widehat{H}^{\text{Ind}} = \begin{pmatrix}
\theta^{\text{Ind}}_{11} & \cdots & \theta^{\text{Ind}}_{1K} \\
\vdots & \ddots & \vdots \\
\theta^{\text{Ind}}_{K1} & \cdots & \theta^{\text{Ind}}_{KK}
\end{pmatrix} - \begin{pmatrix}
e^{\text{Ind}}_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & e^{\text{Ind}}_K
\end{pmatrix} - \begin{pmatrix}
\left(e^{\text{Ind}}_1\right)^2 & \cdots & e^{\text{Ind}}_1 e^{\text{Ind}}_K \\
\vdots & \ddots & \vdots \\
e^{\text{Ind}}_K e^{\text{Ind}}_1 & \cdots & \left(e^{\text{Ind}}_K\right)^2
\end{pmatrix}.
\]

Also, using \(\Upsilon^{\text{Ind}}\) to denote the diagonal matrix in the above expression for \(\widehat{H}^{\text{Ind}}\) and \(e^{\text{Ind}}\) to denote the vector of indirect input elasticities, \(\left(e^{\text{Ind}}_1, \ldots, e^{\text{Ind}}_K\right)\), then as in Eq. 9, \(\widehat{H}^{\text{Ind}} = \Theta^{\text{Ind}} - \Upsilon^{\text{Ind}} + e^{\text{Ind}} e^{\text{Ind}}\).