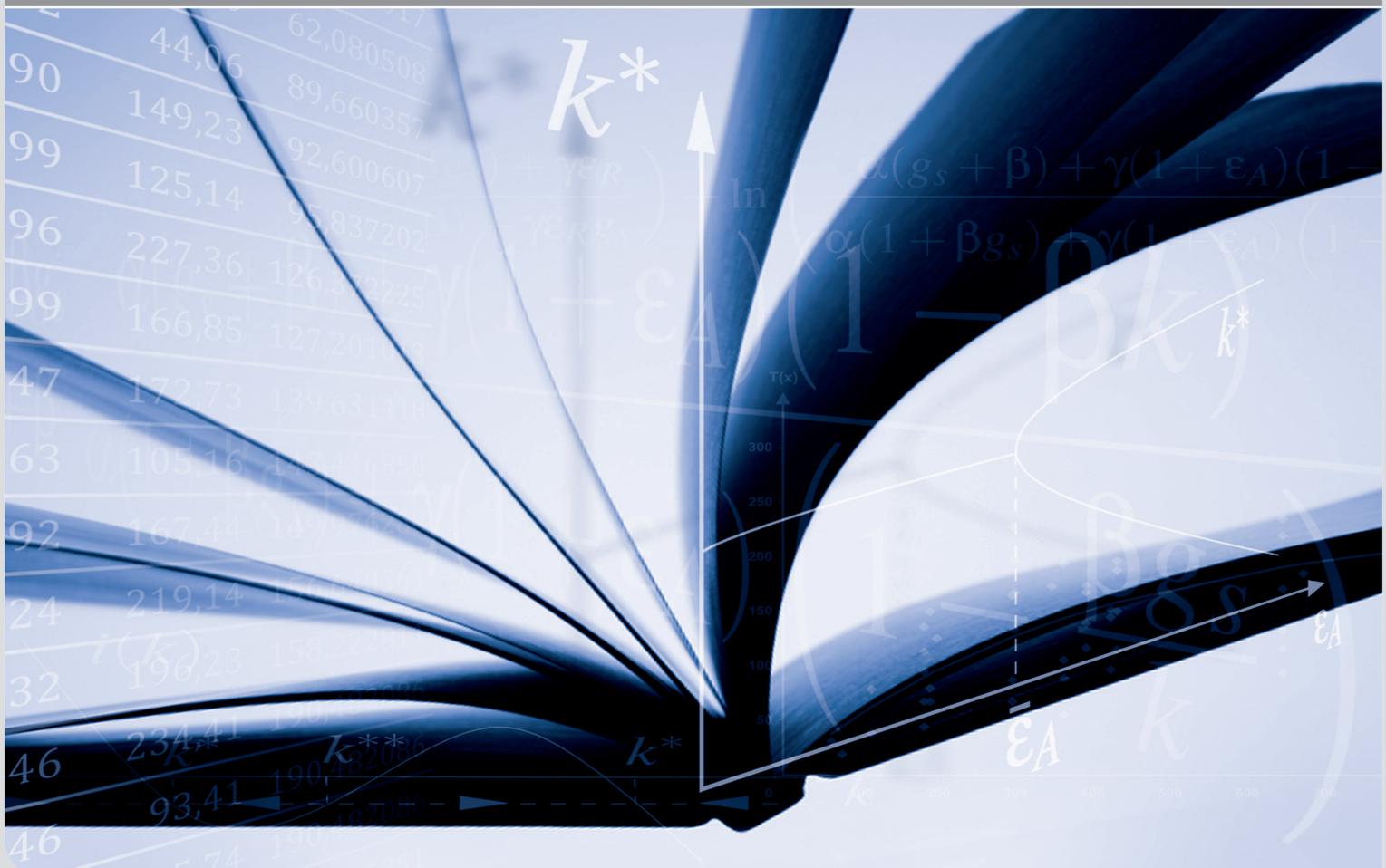


Knowledge spillovers: On the impact of genetic distance and data revisions

by Tim Deeken

No. 74 | NOVEMBER 2015

WORKING PAPER SERIES IN ECONOMICS



Impressum

Karlsruher Institut für Technologie (KIT)
Fakultät für Wirtschaftswissenschaften
Institut für Volkswirtschaftslehre (ECON)

Schlossbezirk 12
76131 Karlsruhe

KIT – Universität des Landes Baden-Württemberg und
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

Working Paper Series in Economics
No. 74, November 2015

ISSN 2190-9806

econpapers.wiwi.kit.edu

Knowledge Spillovers: On the Impact of Genetic Distance and Data Revisions

Tim Deeken*

November 9, 2015

Abstract

This paper assesses the robustness of the empirical results in Ertur and Koch (2007), who develop a model, which accounts for technological interdependence among countries through spatial externalities. The original version models interdependence via an interaction matrix based on geographic distance. In contrast, in this paper, data on genetic distance, defined as the time since two populations have shared a common ancestor, from Spolaore and Wacziarg (2009) is used to construct the interaction matrix. It is found that, whereas in the original model indirect spillovers from capital investment were insignificant, using genetic distance, these spillovers now have a significant effect on steady-state income per worker. However, the version of the model with an interaction matrix based on genetic distance implies an implausibly large capital share of income. In addition, the model is subjected to a further series of robustness checks. The original version relies on data from Penn World Table (PWT) Version 6.1. More recent versions are currently available, and the data has been extensively revised (Johnson et al., 2013). It is shown that results are in general not robust across different versions of the PWT. Furthermore, the estimation results are highly sensitive both to the measure used to model interaction between countries (genetic or geographic distance) and to the specific functional form on which the weights in the interaction matrix are based.

Keywords: Knowledge spillovers; genetic distance; spatial externalities; technological interdependence

JEL: O47, C31, C52

*Karlsruhe Institute of Technology, Institute for Economics, Schlossbezirk 14, 76 131 Karlsruhe, Germany. e-mail: tim.deeken@kit.edu, phone: +49 (0) 721 608-43073, fax: +49 (0) 721 608-45255

An abridged version of this paper is available on the website of the XVIII Applied Economics Meeting in Alicante (<http://www.alde.es/encuentros/trabajos/d/pdf/156.pdf> (accessed: 19 August, 2015)).

1 Introduction

Countries do not develop in isolation from each other, but are connected and interact in many different ways. A key aspect of this interdependence concerns technology, in particular technological knowledge spillovers. Accounting for this technological interdependence both on an empirical and theoretical level requires a notion of how to model the interaction between countries. Empirical evidence suggests that knowledge spillovers decline with the distance in geographic terms between countries (Keller, 2002, 136). This insight has, for instance, been picked up by Ertur and Koch (2007), who develop a theoretical model of economic growth that incorporates technological knowledge spillovers between countries. In the empirical part of their paper, they employ a specification, which qualitatively replicates the effect identified by Keller (2002).

However, geographic distance is only one possible measure to model interaction between countries. The concept is more general and encompasses “any kind of network structure” (Ertur and Koch, 2011, 236). For example, data on genetic distance, which is defined as the time, since two populations have shared a common ancestor (Spolaore and Wacziarg, 2009, 470), can be used to build this structure.

The general possibility of implementing this concept in this way is noted briefly by Ertur and Koch (2011, 236-237, 249), and it follows as Spolaore and Wacziarg (2009) demonstrate that genetic distance has an effect on cross-country income differences. They propose the following mechanism for this result and also provide empirical evidence consistent with it:¹ Within populations, characteristics like habits, implicit beliefs or conventions are transmitted across generations biologically and culturally, and genetic distance can be viewed as a summary statistic that picks up a divergence across populations in characteristics that are slowly changing over time (see, also Spolaore and Wacziarg (2015)). The next step in their argument is the assumption that these differences in characteristics between populations introduce barriers to communication and understanding which then hinder the diffusion of technology. Hence, by using genetic distance, this paper contributes to the literature by providing an important robustness check for the empirical results in the influential model by Ertur and Koch (2007) which relies on geographic distance to model interaction.

A further motivation for employing data on genetic distance is that this approach captures interactions between economies that geographic distance is missing. For instance, Lindner and Strulik (2014, 18) note (without any reference to genetic distance) that it might be

¹It needs to be pointed out though that their empirical results are not uncontroversial, and have, for instance, been challenged by Campbell and Pyun (2015).

the case that knowledge exchange between the United States and the United Kingdom is higher than between the United States and Guatemala even though geographic distance would suggest otherwise. By modeling interaction through genetic distance instead of geographic distance however, stronger knowledge spillovers between the United States and the United Kingdom compared to between the United States and Guatemala would be in line with the data on genetic distance, as the United States and United Kingdom populations are genetically closer to each other than the ones in the United States and Guatemala.

The second contribution of this paper is the assessment of the robustness of the results by Ertur and Koch (2007) to data revisions. In their econometric analysis, they rely on data from Penn World Table (PWT) Version 6.1 (Heston et al., 2002). Since the publication of their article, newer versions of the PWT have become available, and in each update the data has been revised. Ideally, empirical results should be robust to different versions of the PWT. However, this is not a foregone conclusion, and Ponomareva and Katayama (2010) find that conclusions from cross-country growth studies might change even for the same period and units of observation, depending on the version of the PWT. More recently, Johnson et al. (2013) have also investigated this issue. They find that some data revisions have been relatively minor. For instance, the average growth rate of GDP over the period 1975-1999 for Morocco was 1.6% when calculating it using PWT 6.1 and 1.7% when basing the calculations on PWT 6.2 (Johnson et al., 2013, Table 1). Other revisions were drastic, showing high variability in the estimates, as exemplified by the case of Equatorial Guinea. Taking the data from PWT 6.1, its average GDP growth rate in the period 1975-1999 was -2.7% , making it the worst performing of 40 African countries that are covered in both PWT 6.1 and 6.2. On the other hand, for the data from PWT 6.2 its average GDP growth rate over the same period was 4%, thereby becoming the second-best performer in the list of 40 African countries after Botswana (Johnson et al., 2013, 255-256). Hence, the fact that robustness to different versions is an issue for some studies is not too surprising. However, they also argue, based on the results of a series of replication exercises for prominent articles investigating economic growth that results from cross-sectional estimations tend to be robust to changing the version of the PWT (Johnson et al., 2013, 273). This paper investigates whether this is also the case for the results in Ertur and Koch (2007) by estimating the model for the same set of countries and the same time period (1960-1995), but with data taken from PWT Versions 6.2 and 7.1. The importance of checking the robustness of a study's results to data revisions has also been highlighted, for example, in the debate on the relationship between public debt levels and economic growth (see Reinhart and Rogoff (2010) and Herndon et al. (2014)). In this regard, the implications for providing policy advice based

on results that are, for instance, sensitive to the specific version of the data set that is used, cannot be neglected. Policy makers' awareness of this issue needs to be raised.

The third contribution of this paper lies in the quantification of the strength of the indirect (spillover) effects from, for instance, physical capital investment on steady-state per capita income in the model by Ertur and Koch (2007). In the original study, only the magnitude of the direct effects is presented. New methods have been developed by LeSage and Pace (2009) that are applied here which allows for providing important results concerning knowledge spillovers which are not highlighted in Ertur and Koch (2007).

The paper is organized as follows: Section 2 introduces the concept of genetic distance. The following section briefly motivates the need to incorporate knowledge spillovers in theoretical models, introduces the concept of spatial dependence and provides indicative evidence for its existence before presenting the model by Ertur and Koch (2007) in detail. In Section 4, the empirical specification and estimation strategy are discussed, and Section 5 presents and discusses the estimation results. Section 6 concludes.

2 Genetic Distance

Genetic data is increasingly used in economic studies.² Nonetheless, a brief summary of relevant concepts might be helpful in order to better understand the measure of genetic distance employed in the empirical part of this paper. A gene, i.e. a string of DNA encoding a protein, can exist in numerous forms, and a particular form of this gene is called an allele (Giuliano et al., 2014, 182). Individuals with different alleles may have different observable (phenotypic) traits, for instance, eye color; although different alleles between individuals need not result in different observable characteristics (ibid.). It is important to note that the frequency of alleles is not constant across populations, as this is the information used to calculate measures of the genetic distance between populations (Spolaore and Wacziarg, 2009, 480). In principle, on which particular genes' allelic frequency³ this computation is based would not matter. In practice, however, the measure is based on neutral genes. These are genes that do not endow an individual with a selective advantage (Giuliano et al., 2014, 182). This implies that the measure of genetic distance provides no information about specific genes that have a direct impact on fitness and survival or income and productivity (Spolaore and Wacziarg, 2009, 470).

²See, for instance, Spolaore and Wacziarg (2009), Giuliano et al. (2014), Desmet et al. (2011) or Ashraf and Galor (2013).

³A database on allele frequencies is available under: <http://alfred.med.yale.edu> (accessed: 11 August, 2015).

The particular index of genetic distance mainly considered in this paper, F_{ST} distance, measures the probability that the alleles for a gene selected at random from two populations will be different (Spolaore and Wacziarg, 2009, 481).⁴ For identical allele distributions this index equals zero, and it increases with differences in the distributions.⁵ As Spolaore and Wacziarg (2009) argue, these allele differences increase due to the presence of random (or genetic) drift. This concept may be illustrated through an example by Masel (2011, R837): Imagine a population of 5,000 people in which, due to the general diploid nature of human somatic cells (the gametes, ovum and sperm, in contrast, are haploid), 10,000 copies of each gene exist.⁶ If now, for instance, 3,000 of those copies are of a particular form or allele, then in the next generation there might be more or fewer than 3,000 copies, as out of all possible gametes, only some are randomly picked out. When populations become separated, and for constant drift rates (see Kimura (1968) for evidence on this), genetic distance can then be used to measure the time that has passed, since populations have become separated (or, in other words, their degree of genealogical relatedness). It is in this sense that genetic distance can be understood as the time that has elapsed, since populations have shared a common ancestor. Spolaore and Wacziarg (2009, 470-471) furthermore hypothesize that populations that are genetically more distant, have diverged more strongly in characteristics that are variably transmitted across generations, like habits, norms, or implicit beliefs, and that this divergence hinders, for instance, communication and understanding and thereby creates barriers to the diffusion of development or technology. Applying this line of thought to the example mentioned in the introduction: The United States are genetically closer to the United Kingdom than to Guatemala (the pairwise genetic distances are 0.033 and 0.091, respectively) so that with regard to this concept fewer barriers to knowledge diffusion should exist between the United States and the United Kingdom than between the United States and Guatemala.⁷ Note that the stated genetic distances in this example are weighted F_{ST} genetic distances, which take into account that some countries, like the United States or Australia, consist of genetically distant subpopulations (see Spolaore and Wacziarg, 2009, 484-485).⁸

⁴Data from Spolaore and Wacziarg (2009) on an index with different theoretical properties, Nei's distance (see Nei (1972) and Cavalli-Sforza et al. (1994)), which however is highly correlated with F_{ST} distance, will be used to assess the robustness of the empirical results as well.

⁵This index from Cavalli-Sforza et al. (1994) uses the frequency of 128 alleles that are related to 45 genes, which fulfill the conditions that they are both selectively neutral and easy to collect (Giuliano et al., 2014, 183).

⁶A human cell is called haploid if its nucleus has a single set of 23 chromosomes and it is diploid if its nucleus has a double set of 23 chromosomes.

⁷Considering geographic distances between the country capitals suggests that Washington, D.C. is closer to Guatemala City (distance = 3,007km) than to London (distance = 5,909km). See Equation (B.2) in Appendix B for the general formula to calculate these distances.

⁸See Appendix A for a formal definition of F_{ST} genetic distance. The formula for the weighted version is provided in Equation (A.3).

3 Spatial Dependence and Model Setup

This section motivates the need for including technological interdependence across geographical regions in theoretical models and introduces the spatial Solow model as it was developed by Ertur and Koch (2007). Section 3.1 develops the concept of spatial dependence and illustrates the concept with a brief example. Section 3.2 describes in detail the specification of technological progress and technological interdependence in this model before Section 3.3 investigates the transition dynamics and derives an equation for the steady-state income per worker.

3.1 Spatial Dependence

Knowledge spillovers have been discussed by economists for quite a long time, going back to Marshall (1890). His description of these effects was completely verbal however, and the first attempts to incorporate these effects within a theoretical model are due to Arrow (1962) and Romer (1986). These authors made the assumption that knowledge generated in a single firm is not confined to this particular firm, but might spill over to other firms in a given geographical region as knowledge is considered a non-rival input. While this is an improvement on earlier models like the one by Solow (1956), it remains unclear why knowledge diffusion should stop at a given border. Learning-by-doing, for instance, can result as a by-product of mergers and acquisitions, be a result of interfirm cooperation or the meeting of different people at conferences and seminars (Fischer, 2011, 420). None of these activities is necessarily confined within an arbitrary geographical unit. With respect to physical capital externalities, for example, López-Bazo et al. (2004, 44), note that “there is no a priori reason to constrain spillovers within the barriers of the economy where the agent making the investment is located”. Diffusion of these knowledge spillovers across boundaries can then be viewed as a spatial externality, implying that, for instance, the economic development of neighboring countries is related. Before presenting how this effect is picked up in a theoretical model, the concept of spatial dependence will be introduced to provide indicative evidence for the relevance of these spatial externalities.

As countries interact with each other in numerous ways, it is straightforward to assume that the development of one country may be influenced by the development of nearby countries. This latter idea is captured in the (spatial) econometric literature by the concept of spatial dependence. More precisely, spatial dependence captures situations in which the values observed in e.g. country i depend on the values observed in neighboring countries (see LeSage and Pace, 2009, 2).

As an example consider Figure 1, which depicts total factor productivity (TFP) levels relative to those of the United States for 110 countries for the year 2007 in shades of green. Visual inspection of this figure suggests that the TFP levels are not distributed

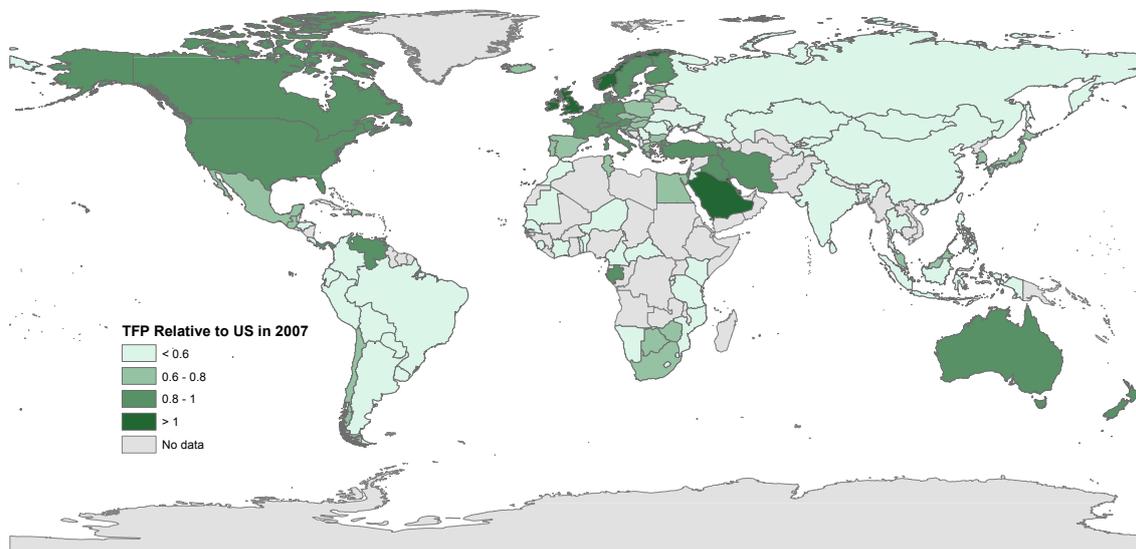


Figure 1: Total Factor Productivity Relative to the United States for 110 Countries in 2007 (Data from Feenstra et al., 2013b).

randomly. Countries with comparatively low TFP levels (less than 80% of the value for the United States) are, for instance, concentrated in South America, southeastern Europe or east and southeast Asia, whereas regions with higher TFP levels (above 80% of the level in the United States) can be found in northwestern Europe.

An alternative visualization of these data is provided by a Moran scatterplot in Figure 2. When interpreting this figure, it is important to note that the variables are in deviations-from-the-mean form. The meaning of the variable “Spatial Lag of TFP” on the ordinate might not be immediately clear. In general, a spatial lag is a weighted average of the values for a variable from countries that are neighbors to country i (see LeSage and Pace, 2009, 8).⁹

The Spatial Lag of TFP on the ordinate in Figure 2 thus has the interpretation that for a given observation i , this variable shows the deviation of the TFP for country i from

⁹For expository reasons the term “neighbor” will be slightly abused in this section. In fact, in the calculation of the spatial lag of TFP for country i , all countries for which data is available are included and not only neighboring countries. However, countries that are geographically closer to country i receive a higher weight in the calculation of the spatial lag. The precise formal specification of this idea is provided in Section 5.1. Note that countries whose TFP levels exceed those of the US are mainly oil-rich countries, like Saudi Arabia, Qatar (QAT) or Kuwait (KWT) for which TFP will be overstated as data is lacking to include also “subsoil assets” in the underlying methodology (see Feenstra et al. (2013a, 35-36) and Inklaar and Timmer (2013)) as well as Singapore or the Special Administrative Regions of Hong Kong and Macao.

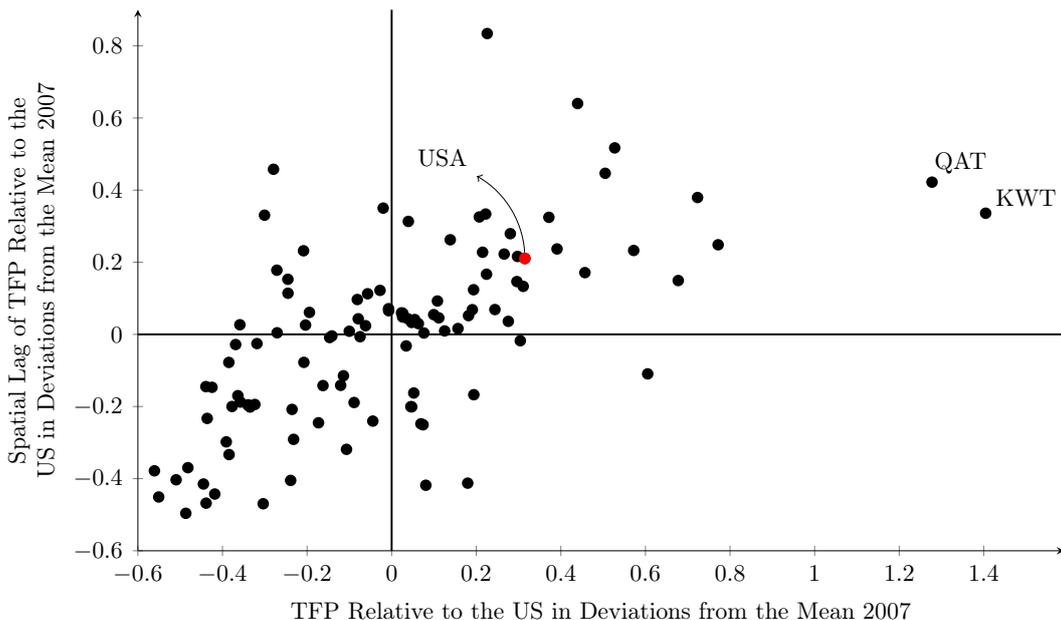


Figure 2: Moran Scatterplot of TFP Relative to the United States for 110 Countries in 2007 (Data from Feenstra et al., 2013b).

Note: The variables are in the form of deviations from the mean so that the value 0 on the abscissa is equivalent to the mean value of 0.685.

the mean of TFP of its neighbors. Hence, in the lower left quadrant of the figure, one finds countries for which not only their own TFP is below the mean, but also the TFP of its neighbors is below the mean. Whereas, in the upper right quadrant countries cluster whose own TFP as well as the one of its neighbors is above mean.¹⁰

Figures 1 and 2 have provided indicative evidence of spatial dependence (or spatial autocorrelation) in country-level data. Theoretical models should therefore not disregard this characteristic of the data, but instead try to represent it. The following section shows one possible way to achieve this.

3.2 Specification of Technological Progress

The aggregate production for each country $i = 1, \dots, N$ at time t in the model developed by Ertur and Koch (2007) is described by the Cobb-Douglas production function

$$Y_i(t) = A_i(t)K_i(t)^\alpha L_i(t)^{1-\alpha} \quad \text{with} \quad 0 < \alpha < 1 \quad (1)$$

¹⁰Note that this result of spatial dependence is not particular to country-level data. Looking, for example, at the distribution of the TFP levels of European NUTS 2 regions gives a similar result (Derbyshire et al., 2011).

where output, $Y_i(t)$, is produced with the three input factors labor, $L_i(t)$, physical capital, $K_i(t)$ and technology, $A_i(t)$. This function is linearly homogenous in the two input factors capital and labor and thus has constant returns to scale with respect to these two factors. The aggregate level of technology in country i is described by

$$A_i(t) = \Omega(t)k_i(t)^\phi \prod_{j \neq i}^N A_j(t)^{\gamma w_{ij}}. \quad (2)$$

Basically, overall technological progress is assumed to be due to three different factors in Equation (2) which are (imperfect) substitutes. The first factor, $\Omega(t)$, reflects exogenous (Harrod-neutral) technological progress as modeled in the original contributions by Solow (1956, 85) and Swan (1956). In formal terms, this is captured by the equation

$$\Omega(t) = \Omega(0)e^{\mu t},$$

with μ as the constant rate of technological progress and $\Omega(0)$ the initial level.

The second term models the influence of the physical capital per worker, $k_i(t) = \frac{K_i(t)}{L_i(t)}$, on aggregate technology in country i . The level of technology increases with the level of capital per worker $k_i(t)$, modeling the assumption that physical capital externalities exist in general. Their strength is governed by the parameter ϕ for which $0 \leq \phi < 1$ holds so that perfect knowledge spillovers from a capital investment in a given firm in country i to the remaining firms in this country are ruled out, as diffusion is not frictionless and some knowledge is “lost in transmission”. The assumption that all firms in a country gain a higher level of technology, if one firm increases its physical capital per worker is due to Arrow (1962) and Romer (1986). As has been mentioned above, the assumption that these knowledge spillovers should be constrained within a single region or country is tenuous. Why should knowledge diffuse only within a country but not across countries? The strength of the spillovers might be dampened (and there is indeed empirical evidence to that extent¹¹), but they should be present nonetheless.

The third factor in Equation (2) picks this up. From a formal perspective, this factor is a weighted geometric mean of the level of technology in all countries $j = 1, \dots, N$ connected to country i . The strength of these cross-border spillovers or spatial externalities is governed by two factors. The parameter γ , for which $0 \leq \gamma < 1$ holds, gauges which fraction of knowledge generated in, for example, country j' spills over into country i . This value is the same for all units of observation. The second factor concerns the weights w_{ij} . In general, these are allowed to differ across countries, and they specify the way in

¹¹See, for example, Keller (2002), who estimates that at a distance of about 1,200 kilometers from the country in which the knowledge originates, 50% is still available.

which countries are connected to each other. It is important to note that how strong country i benefits from knowledge spillovers depends on the way it is connected to all other countries under consideration. This implies that the net effect on a country's level of technology due to spatial spillovers will differ across countries. For a given degree of spillovers, relatively isolated countries will benefit less than more integrated countries. With respect to the spatial weights, it is assumed that these are non-negative, which leaves open the possibility that countries might not be connected to each other at all so that spatial externalities are absent between particular pairs of countries, non-stochastic, implying that the weights are fixed over time, and finite. In addition, the weights w_{ij} lie in the interval $[0, 1]$ and for $i = j$ $w_{ij} = 0$ holds, excluding the case of self-influence. Finally, the weights sum to one.¹² Summarized, the spatial weight matrix or more generally interaction matrix, \mathbf{W} , is thus row-stochastic (LeSage and Pace, 2009, 9-10).¹³

Applying the natural logarithm to Equation (2), it can be rewritten as

$$\ln A_i(t) = \ln \Omega(t) + \phi \ln k_i(t) + \gamma \sum_{j \neq i}^N w_{ij} \ln A_j(t). \quad (3)$$

Stacking the equations for all countries $i = 1, \dots, N$ at time t , the level of technology can be expressed as

$$\underbrace{\begin{pmatrix} \ln A_1(t) \\ \vdots \\ \ln A_N(t) \end{pmatrix}}_{\substack{=\mathbf{A} \\ (N \times 1)}} = \underbrace{\begin{pmatrix} \ln \Omega(t) \\ \vdots \\ \ln \Omega(t) \end{pmatrix}}_{\substack{=\mathbf{\Omega} \\ (N \times 1)}} + \phi \underbrace{\begin{pmatrix} \ln k_1(t) \\ \vdots \\ \ln k_N(t) \end{pmatrix}}_{\substack{=\mathbf{k} \\ (N \times 1)}} + \gamma \underbrace{\begin{pmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{pmatrix}}_{\substack{=\mathbf{W} \\ (N \times N)}} \cdot \underbrace{\begin{pmatrix} \ln A_1(t) \\ \vdots \\ \ln A_N(t) \end{pmatrix}}_{\substack{=\mathbf{A} \\ (N \times 1)}} \quad (4)$$

$$\iff \mathbf{A} = \mathbf{\Omega} + \phi \mathbf{k} + \gamma \mathbf{W} \mathbf{A}.$$

Given that spatial dependence is positive, $\gamma \neq 0$, and that the inverse $(\mathbf{I} - \gamma \mathbf{W})^{-1}$ exists¹⁴ the previous equation is equivalent to

$$\mathbf{A} = (\mathbf{I} - \gamma \mathbf{W})^{-1} \mathbf{\Omega} + \phi (\mathbf{I} - \gamma \mathbf{W})^{-1} \mathbf{k}. \quad (5)$$

¹²On the assumptions for the spatial weights see Ertur and Koch (2007, 1036, Footnote 2) and Fischer and Wang (2011, 20).

¹³An illustration of two spatial weight matrices is given in Appendix B, which also describes the calculation of the spatial weights based on great circle distances between country capitals in detail.

¹⁴This inverse exists if $\frac{1}{\gamma}$ is not an eigenvalue of the spatial weight matrix. However, it is not necessarily guaranteed that the inverse exists for the parameter space for γ assumed here. See Appendix C for a proof that the inverse exists in this case as well.

From this expression, it follows that the level of technology for a given country i can be written as

$$A_i(t) = \Omega(t)^{\frac{1}{1-\gamma}} k_i(t)^\phi \prod_{j=1}^N k_j(t)^\phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ij} \quad (6)$$

where $(\mathbf{W}^r)_{ij}$ are the individual entries in row i and column j of the matrix \mathbf{W} taken to the power of r . Since the derivation of Equation (6) is not immediately obvious, important intermediate results are provided in Appendix D.1. In particular, it is proved that the inverse matrix $(\mathbf{I} - \gamma \mathbf{W})^{-1}$, which is also called the inverse spatial transformation (Le Gallo, 2014, 1515), can be written as an infinite series, i.e.

$$(\mathbf{I} - \gamma \mathbf{W})^{-1} = \sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r.$$

With respect to Equation (5), it follows then that the level of technology in every country is correlated with the level of technology in every other country and closer countries are more closely related.¹⁵ The effect of the inverse spatial transformation is often referred to as the spatial multiplier effect (see, for example, Ertur and Koch (2007, 1044) or Le Gallo (2014, 1515)).

The results derived with respect to the level of technology are helpful in rewriting the production function. This function exhibits constant returns to scale in capital and labor, which implies that Equation (1) can be written in per capita terms

$$y_i(t) = A_i(t) k_i(t)^\alpha \quad (7)$$

where $y_i(t) = \frac{Y_i(t)}{L_i(t)}$. Inserting the expression for the level of technology in Equation (6) into the per worker production function leads to

$$y_i(t) = \Omega(t)^{\frac{1}{1-\gamma}} \cdot k_i(t)^{\alpha + \phi(1 + \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ii})} \cdot \prod_{j \neq i}^N k_j(t)^\phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ij}.$$

Now define

$$u_{ii} \equiv \alpha + \phi \left(1 + \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ii} \right) \quad \text{and} \quad u_{ij} \equiv \phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ij} \quad (8)$$

and substitute for the exponents of physical capital per worker so that the per worker

¹⁵As Anselin (2003, 155) mentions in a slightly different context, which is nonetheless applicable here, this is in effect a reformulation of the first law of geography by Tobler, which states that “everything is related to everything else, but near things are more related than distant things” (1970, 236).

production function can be written more compactly as

$$y_i(t) = \Omega(t)^{\frac{1}{1-\gamma}} \cdot k_i(t)^{u_{ii}} \cdot \prod_{j \neq i}^N k_j(t)^{u_{ij}}. \quad (9)$$

From this function it can be seen that in contrast to the standard Solow model, the model presented here implies heterogeneity in the social elasticities of income per worker with respect to capital per worker. If, for instance, country i increases its own stock of physical capital per worker, the social return (or elasticity) is¹⁶

$$\frac{\partial y_i(t)}{\partial k_i(t)} \frac{k_i(t)}{y_i(t)} = u_{ii}.$$

In case all countries except country i simultaneously increase their stocks of physical capital per worker, then the corresponding elasticity is

$$\sum_{j \neq i}^N \frac{\partial y_i(t)}{\partial k_j(t)} \frac{k_j(t)}{y_i(t)} = \sum_{j \neq i}^N u_{ij}.$$

Hence, if all countries $i = 1, \dots, N$ together increase their stocks of physical capital per worker, then

$$\frac{\partial y_i(t)}{\partial k_i(t)} \frac{k_i(t)}{y_i(t)} + \sum_{j \neq i}^N \frac{\partial y_i(t)}{\partial k_j(t)} \frac{k_j(t)}{y_i(t)} = u_{ii} + \sum_{j \neq i}^N u_{ij} = \alpha + \frac{\phi}{1-\gamma} < 1 \quad (10)$$

is the social output elasticity per worker in the situation in which all countries simultaneously increase their capital stock per worker.¹⁷ The inequality $\alpha + \frac{\phi}{1-\gamma} < 1$ is an

¹⁶The term social in contrast to private is warranted in this case, as the elasticity calculated here includes the physical capital externalities, ϕ , within a country (see the definitions in Equation (8)).

¹⁷The result before the inequality follows since

$$\begin{aligned} u_{ii} + \sum_{j \neq i}^N u_{ij} &= \alpha + \phi + \phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ii} + \sum_{j \neq i}^N \phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ij} \\ &= \alpha + \phi \left(1 + \sum_{j=i}^N \cdot \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ij} \right). \end{aligned}$$

The matrices \mathbf{W}^r are Markov matrices, and in this case it is the rows that sum to one, meaning that $\sum_{j=i}^N \mathbf{W}^r_{ij} = 1 \forall r$ so that $u_{ii} + \sum_{j \neq i}^N u_{ij} = \alpha + \phi (1 + \sum_{r=1}^{\infty} \gamma^r)$, and the term in parentheses can be rewritten as $1 + \sum_{r=1}^{\infty} \gamma^r = 1 + \sum_{r=1}^{\infty} \gamma^r + \gamma^0 - \gamma^0 = 1 + \sum_{r=0}^{\infty} \gamma^r - \gamma^0 = \frac{1}{1-\gamma}$. With this result, the social returns are

$$u_{ii} + \sum_{j \neq i}^N u_{ij} = \alpha + \frac{\phi}{1-\gamma}.$$

assumption made by Ertur and Koch (2007, 1037), since otherwise the per worker production function in Equation (9) would not have decreasing returns to (all) physical capital, and the model would exhibit endogenous growth.

3.3 Transition Dynamics and Steady State

Capital accumulation is described by the fundamental dynamic equation of the Solow model, i.e.

$$\dot{k}_i(t) = s_i y_i(t) - (n_i + \delta) k_i(t) \quad (11)$$

where $\dot{k}_i(t) = dk_i(t)/dt$ denotes a time derivative, s_i is the country specific constant saving rate (the fraction of output invested in physical capital), n_i is the constant growth rate of labor for country i , and δ is the depreciation rate, which is assumed to be identical for all countries.

Due to the decreasing returns to capital per worker (it holds that $0 < \alpha < 1$, see Equation (1)), $k_i(t)$ converges monotonically to its steady-state value or value on the balanced growth path, $k_i^*(t)$.¹⁸ When this value is reached, capital (and by implication output) per worker grow at the balanced growth rate $g = \mu [(1 - \alpha)(1 - \gamma) - \phi]^{-1}$.¹⁹ This rate increases if, for instance, ϕ , the parameter indicating the strength of knowledge spillovers within a country, increases or if γ increases so that knowledge spillovers between countries are stronger.²⁰

The steady-state value $k_i^*(t)$ can be calculated by noting that from Equation (11) on the balanced growth path

$$g = s_i \frac{y_i^*(t)}{k_i^*(t)} - (\delta + n_i) \iff k_i^*(t) = \frac{s_i}{n_i + \delta + g} y_i^*(t) \quad (12)$$

holds. Inserting Equation (9) into the right-most expression above and solving for $k_i^*(t)$ yields

$$k_i^*(t) = \Omega(t)^{\frac{1}{1-\gamma(1-u_{ii})}} \left(\frac{s_i}{n_i + \delta + g} \right)^{\frac{1}{1-u_{ii}}} \prod_{j \neq i}^N (k_j^*(t))^{\frac{u_{ij}}{1-u_{ii}}}.$$

The steady-state value of real income per capita in country i , $y_i^*(t)$, can be derived by

¹⁸Similar to Fischer (2011, 425), the balanced growth path is defined as a situation in which the physical capital per worker grows at rate g , the investment rate for physical capital, the employment growth rate and the growth rate of the exogenous part of technology are constant.

¹⁹The rate can be calculated by taking the derivative of Equation (13) with respect to time, then using $\ln k_i^*(t)/dt = g = \ln y_i^*(t)/dt$, and solving the derivative for g .

²⁰These results hold due to the inequality $\alpha + \frac{\phi}{1-\gamma} < 1$.

first taking the logarithm of the production function in Equation (7), writing it in matrix form (compare Equation (4)) to obtain

$$\mathbf{y}^* = \mathbf{A}^* + \alpha \mathbf{k}^*,$$

where the asterisks denote steady-state values, and then inserting the expression for \mathbf{A} from Equation (5) evaluated at steady state therein and finally solving for \mathbf{y}^* , which yields

$$\mathbf{y}^* = \mathbf{\Omega} + (\alpha + \phi) \mathbf{k}^* - \alpha \gamma \mathbf{W} \mathbf{k}^* + \gamma \mathbf{W} \mathbf{y}^*.$$

Writing this equation for a single country i at time t results in

$$\ln y_i^*(t) = \ln \Omega(t) + (\alpha + \phi) \ln k_i^*(t) - \alpha \gamma \sum_{j \neq i}^N w_{ij} \ln k_j^*(t) + \gamma \sum_{j \neq i}^N w_{ij} \ln y_j^*(t). \quad (13)$$

Inserting now the expression for the capital-output ratio on the balanced growth path from Equation (12) into this expression and solving for $y_i^*(t)$ leads to the final result

$$\begin{aligned} \ln y_i^*(t) &= \frac{1}{1 - \alpha - \phi} \ln \Omega(t) + \frac{\alpha + \phi}{1 - \alpha - \phi} \ln s_i \\ &\quad - \frac{\alpha + \phi}{1 - \alpha - \phi} \ln(n_i + g + \delta) - \frac{\alpha \gamma}{1 - \alpha - \phi} \sum_{j \neq i}^N w_{ij} \ln s_j \\ &\quad + \frac{\alpha \gamma}{1 - \alpha - \phi} \sum_{j \neq i}^N w_{ij} \ln(n_j + g + \delta) + \frac{\gamma(1 - \alpha)}{1 - \alpha - \phi} \sum_{j \neq i}^N w_{ij} \ln y_j^*(t). \end{aligned} \quad (14)$$

In line with the standard Solow model, this equation states that the per worker income in steady state in country i is positively influenced by an increase in its own saving rate, s_i , since an increase in savings leads to more investment and a higher capital stock per worker, which in turn leads to a higher per worker income in steady state. Increases in the labor force (note that g and δ are constant) reduce steady-state income, since for a given saving rate the capital stock must now be spread over more workers so that $k_i^*(t)$ falls, implying a decrease of $y_i^*(t)$. In addition to these standard effects, Equation (14) suggests that the steady-state value also depends negatively on increases in the saving rates of the other countries and positively on the increases in the population growth rate and steady-state levels of the remaining countries. Why this should be the case is not immediately obvious. However, at this point it needs to be taken into account that the steady-state values in the neighboring countries of i depend, for instance, positively on their own saving rates. Higher capital stocks in neighboring countries lead to a higher level of technology in these countries (see Equation (2)). A fraction of this knowledge

spills over into country i , which therefore benefits via these spatial externalities. The elasticity $\eta_{s_j}^i$ of income per worker in steady state in country i with respect to the saving rate in the neighboring countries is given by²¹

$$\eta_{s_j}^i = \frac{\phi}{(1-\alpha)(1-\alpha-\phi)} \sum_{r=1}^{\infty} (\mathbf{W}^r)_{ij} \left[\frac{\gamma(1-\alpha)}{1-\alpha-\phi} \right]^r. \quad (15)$$

This expression is clearly positive (compare Equation (10)). The corresponding elasticity with respect to population growth, $\eta_{n_j}^i$, equals the expression above with a negative sign. A further point to note is that the effect on a country's per capita income from increasing its own saving rate (or decreasing its own population growth rate) is higher in this model than in the standard Solow model. This elasticity is given by

$$\eta_{s_i}^i = \frac{\alpha + \phi}{1-\alpha-\phi} + \frac{\phi}{(1-\alpha)(1-\alpha-\phi)} \sum_{r=1}^{\infty} (\mathbf{W}^r)_{ii} \left[\frac{\gamma(1-\alpha)}{1-\alpha-\phi} \right]^r. \quad (16)$$

The reason is that the knowledge generated by the increase in the capital stock per worker diffuses to the neighboring countries, leading to a higher income per worker in these countries, which again has a positive impact on the income per worker in country i . This feedback effect follows from the model's setup, since even though the diagonal entries of the spatial weight matrix \mathbf{W} are zero, this is not the case for higher orders of the matrix, as, for instance, each country is a second-order neighbor to itself or in other words a neighbor to its first-order neighbor (see LeSage and Pace, 2009, 9).

4 Empirical Specification, Estimation Strategy, and Model Interpretation

This section presents details on the empirical specification of the model from Section 3, develops the spatial econometric estimation strategy, and addresses the interpretation of parameters from the estimation. It will first be shown that ordinary least squares (OLS) estimators of the model's parameters are biased and inconsistent. Thereafter, maximum likelihood estimators (ML) will be presented as an alternative to OLS.

²¹See Appendix D.2 for the derivation.

4.1 Econometric Specification of the Model

Equation (14) from Section 3 has the empirical counterpart at $t = 0$ (both the time index and the star to indicate the steady-state value of per worker income are now dropped to enhance readability)

$$\begin{aligned} \ln y_i = & \beta_0 + \beta_1 \ln s_i + \beta_2 \ln(n_i + g + \delta) + \theta_1 \sum_{j \neq i}^N w_{ij} \ln s_j \\ & + \theta_2 \sum_{j \neq i}^N w_{ij} \ln(n_j + g + \delta) + \rho \sum_{j \neq i}^N w_{ij} \ln y_j + \varepsilon_i \end{aligned} \quad (17)$$

where $\frac{1}{1-\alpha-\phi} \ln \Omega(0) = \beta_0 + \varepsilon_i$ for $i = 1, \dots, N$ and β_0 is a constant and ε_i is a country-specific shock. From the development of the theoretical model, the empirical specification above implies the following constraints on the coefficients $\beta_1 + \beta_2 = 0$ and $\theta_1 + \theta_2 = 0$ (see Equation (14)).

In matrix form, Equation (17) is equivalent to²²

$$\mathbf{y} = \boldsymbol{\iota}_N \beta_0 + \mathbf{X} \boldsymbol{\beta} + \mathbf{W} \mathbf{X} \boldsymbol{\theta} + \rho \mathbf{W} \mathbf{y} + \boldsymbol{\varepsilon}. \quad (18)$$

and the definitions of the respective variables are in order of appearance in the equation above provided in the list below

\mathbf{y} is an $N \times 1$ vector of real income per worker in logarithms,

$\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones,

β_0 is a scalar (constant parameter),

\mathbf{X} is an $N \times 2$ matrix of the exogenous explanatory variables (investment rate and population growth rate) in logarithms for the N observations,

$\boldsymbol{\beta}$ is a 2×1 vector [$\boldsymbol{\beta} = (\beta_1, \beta_2)'$] of the regression parameters for the investment rate and population growth rate,

\mathbf{W} is the $N \times N$ spatial weight matrix in row-standardized form,

$\mathbf{W} \mathbf{X}$ is the $N \times 2$ matrix of the spatially lagged explanatory variables,

²²The notation here and in the list below follows Fischer (2011) and thus differs slightly from the one in Ertur and Koch (2007). The reason for this is to be precise and clear in the notation. In particular, by using the notation in Fischer (2011), having \mathbf{X} denote two different matrices depending on context, is avoided.

$\boldsymbol{\theta}$ is a 2×1 vector [$\boldsymbol{\theta} = (\theta_1, \theta_2)'$] of the regression parameters for the spatially lagged explanatory variables,

ρ is the spatial autoregressive coefficient, $\rho = \frac{\gamma(1-\alpha)}{1-\alpha-\phi}$,

$\mathbf{W}\mathbf{y}$ is an $N \times 1$ vector representing the spatially lagged endogenous variable,

$\boldsymbol{\varepsilon}$ is an $N \times 1$ vector of errors, for which the assumption of normal and identical distribution with mean zero and variance $\sigma^2\mathbf{I}$ holds, i.e. $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$.

Equation (18) includes spatial lags of both the endogenous variable and the explanatory variables on the right-hand side. This specification is called a Spatial Durbin Model (SDM) (see e.g. Anselin, 1988, 111). By redefining $\mathbf{Z} = [\iota_N \mathbf{X} \mathbf{W} \mathbf{X}]$ and $\boldsymbol{\delta} = [\beta_0, \boldsymbol{\beta}, \boldsymbol{\theta}]'$, this model can be rewritten as (see, for instance, LeSage and Pace, 2009, 46)

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \quad (19)$$

which is a spatial autoregressive (SAR) model. This specification will be used to demonstrate that the OLS estimates are biased and to derive the ML estimates for this model.²³

In reduced form (i.e. solved for the endogenous variable), the specification in Equation (19) can be expressed as²⁴

$$\mathbf{y} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{Z}\boldsymbol{\delta} + (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}.$$

This specification implies that the spatial lag of the endogenous variable and the error term are correlated with each other, as

$$\text{Cov}[(\mathbf{W}\mathbf{y}), \boldsymbol{\varepsilon}] = E[(\mathbf{W}\mathbf{y})\boldsymbol{\varepsilon}'] - E[\mathbf{W}\mathbf{y}] = \mathbf{W}(\mathbf{I} - \rho \mathbf{W})^{-1} \sigma^2$$

so that the OLS parameter estimators are biased and inconsistent (Davidson and MacKinnon, 2004) and an alternative estimation strategy is thus necessary.

²³The SAR model is nested in the SDM model and so with the above rewriting their likelihood functions coincide (LeSage and Pace, 2009, 46). Using the SAR model here is simply done to save on notation.

²⁴Note that this only holds if $(\mathbf{I} - \rho \mathbf{W})$ is non-singular. See Appendix C for the proof.

4.2 Estimation Strategy

Given these problems, LeSage and Pace (2009, 45) note with reference to Lee (2004) that maximum likelihood is a viable alternative to OLS.²⁵ Assuming that the errors are normally distributed, the specification in Equation (19) has the following log-likelihood function.

$$\begin{aligned} \ln L(\mathbf{y}; \boldsymbol{\delta}, \rho, \sigma^2) = & -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln |\mathbf{I} - \rho \mathbf{W}| \\ & - \frac{1}{2\sigma^2} [(\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{Z}\boldsymbol{\delta}]' [(\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{Z}\boldsymbol{\delta}]. \end{aligned}$$

Finding the maximum for this function requires calculating the partial derivatives with respect to all parameters, setting these necessary conditions equal to zero, and solving the system for the parameters. Instead, yielding identical results, this multivariate optimization problem can be reduced to a univariate optimization problem by concentrating the log-likelihood function with respect to the parameters $\boldsymbol{\delta}$ and σ^2 (LeSage and Pace, 2009, 47). This concentrated log-likelihood function depends, in addition to the sample data, only on the single parameter ρ and is given by

$$\ln L(\mathbf{y}; \rho) = -\frac{N}{2} [\ln(2\pi) + 1] + \ln |\mathbf{I} - \rho \mathbf{W}| - \frac{N}{2} \ln \left[\frac{(\hat{\mathbf{e}}_O - \rho \hat{\mathbf{e}}_L)'(\hat{\mathbf{e}}_O - \rho \hat{\mathbf{e}}_L)}{N} \right] \quad (20)$$

where $\hat{\mathbf{e}}_O$ are the estimated residuals from a regression of \mathbf{y} on \mathbf{Z} and $\hat{\mathbf{e}}_L$ those from a regression of $\mathbf{W}\mathbf{y}$ on \mathbf{Z} (see Fischer, 2011, 427). Maximizing Equation (20) yields a ML estimate $\hat{\rho}$, which can then be used to compute the ML estimates $\hat{\boldsymbol{\delta}}$ and $\hat{\sigma}^2$.

4.3 Model Interpretation

Due to the presence of the spatial lags $\mathbf{W}\mathbf{X}$ and $\mathbf{W}\mathbf{y}$ in Equation (18), the interpretation of the parameters is a bit more complicated than in standard linear regression models, since the feedback effects mentioned in Section 3.3 need to be taken into account. The partial derivatives of Equation (18) with respect to, for example, the investment rate, are given by

$$\frac{\partial \mathbf{y}}{\partial \mathbf{X}'_1} = (\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{I}\beta_1 + \mathbf{W}\theta_1). \quad (21)$$

This expression is an $N \times N$ matrix, which represents the non-linear impacts on all countries that result from a change in the investment rate in any country (Fischer, 2011). As

²⁵Other approaches like instrumental variables (IV), generalized methods of moments (GMM) or Bayesian Markov Chain Monte Carlo (MCMC) might be alternatives (see Elhorst, 2010, 15).

LeSage and Pace (2009, 36) point out, in general the impact of a change in an explanatory variable in this type of model will not be identical across all observations. Therefore, they suggest a summary measure of these differing impacts. The row sums in the matrix in Equation (21) represent the total impact to an observation, i.e. the impact of a change in the investment rate in all countries on steady-state income in country $i = 1, \dots, N$. The average of these row sums is then labeled the average total impact to an observation by LeSage and Pace (2009). On the main diagonal of the matrix are the own partial derivatives or direct impacts from a change in the explanatory variable. These derivatives capture the effect of a change in, for example, the investment rate in country i on steady-state income in country i , and these impacts are summarized via averaging the entries on the diagonal of the matrix. LeSage and Pace (2009, 37) note that this corresponds, at least to a certain extent, to the typical interpretation of regression coefficients. Finally, the off-diagonal elements in the matrix are the cross-partial derivatives and represent the indirect (or spillover) impacts, which are again summarized by averaging the row sums of the respective matrix elements. In other words, this measure records the effect on the steady-state level income in country i resulting from a change in the investment rate in all countries except country i . Hence, the average indirect impact is given by the difference between the average total impact and the average direct impact.

5 Data, Estimation Results, and Robustness

This section starts by providing information on the data sources used to assemble the data set for the empirical analyses and on how the variables were constructed from the source data. Thereafter, robustness checks on the results in Ertur and Koch (2007) are conducted and discussed. The first set of robustness checks in Section 5.2 considers the sensitivity of the results to changing the version of the PWT from 6.1 to 6.2 and 7.1, respectively. Next, in Section 5.3, estimation results are reported and discussed for the specification in which technological interdependence is modeled via genetic distance. Again, sensitivity of the results is assessed by estimating the model with data from the three different versions of the PWT.

5.1 Data

The main data source for the replication exercise is PWT 6.1 (Heston et al., 2002), while for the robustness checks PWT 6.2 (Heston et al., 2006) and 7.1 (Heston et al., 2012) are used. Additional versions of the PWT exist as well, for example, PWT 6.3 and

7.0. However, Breton (2012) has noted substantial issues with Version 7.0. Moreover, as Johnson et al. (2013, 257) point out, the exposition would soon become intractable if one aimed at a comparison between every single version.²⁶ As in Ertur and Koch (2007, 1042), the initial sample covers the 91 countries of the non-oil sample in Mankiw et al. (1992), for which data is available over the period 1960-1995.²⁷ In contrast to the theoretical model, GDP per capita and GDP per worker are not in fact identical, as not the whole population in a country is employed. Hence, for the empirical exercise, the dependent variable, y , is real GDP (evaluated via the chain method) per worker (variable *rgdpwok* in PWT). The investment rate, s , is the real share of investment in real GDP (variable *ki* in PWT) averaged over the respective years. For the average growth rate of workers, n , no directly corresponding variable is available in PWT. A number for the size of the working-age population can be recovered however by noting that the series for real GDP per capita and population are available so that the number of workers can be calculated by multiplying real GDP per capita (*rgdch* in PWT) by the size of the population (*pop* in PWT) and dividing the result by the value of real GDP per worker (Ertur and Koch (2007, 1042) refer to Caselli (2005, 685) for this method). The average growth rate of the working-age population is then calculated as an approximation (though this is not stated explicitly in Ertur and Koch (2007)) by taking the natural logarithm of the number of workers in 1995, subtracting the natural logarithm of the number of workers in 1960, and dividing the result by the number of years, i.e. 35.

For the construction of the interaction matrices the general assumptions made in Section 3.2 are valid. An additional important point to note is that the weights in these matrices should be exogenous with respect to model (Ertur and Koch, 2007, 1042). This makes geographic and genetic distance ideal candidates.²⁸ The matrices that are based on spatial distances use as weights the great circle distances, d_{ij} , between country capitals i and j . There is however some scope in pinning down the latitude and longitude of a capital, and Ertur and Koch provide no information for their source of this data. In this paper, in all

²⁶See Table 2 in Johnson et al. (2013) for an overview of the evolution of the PWT up to Version 7. More recent versions of the Penn World Table (8.0 and 8.1, respectively) are also available (Feenstra et al., 2015). These data sets will however not be used in this analysis, as these versions lack data on the real share of investment in real GDP. See Table A3 in the document “variable correspondence” available under <http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0> (accessed: 11 August, 2015), which states that not only is this variable not reported in PWT 8.0, but neither is it possible to construct it from the source data. This continues to be valid for PWT 8.1 (see <http://www.rug.nl/research/ggdc/data/pwt/pwt-8.1.xml> (accessed: 11 August, 2015)).

²⁷Appendix F lists these countries.

²⁸Another interesting variable on which to base the weights would be, for example, a measure of technological proximity between countries. However, this measure could not be considered exogenous to the model for the sample period considered in this paper and it would be problematic in the case of technology to disentangle the effects on income per worker due to changes in \mathbf{X} from those effects due to changes in \mathbf{W} .

calculations that rely on latitude and longitude, the coordinates are taken from the CIA’s World Factbook (Central Intelligence Agency, 2013), and the distances are calculated as described in Appendix B.²⁹ As a final step, the weights for the interaction matrices are given by $w_{ij}(1) = w_{ij}^*(1)/\sum_j w_{ij}^*(1)$ and $w_{ij}(2) = w_{ij}^*(2)/\sum_j w_{ij}^*(2)$, and the weights are based on the following functional forms

$$w_{ij}^*(1) = \begin{cases} 0 & \text{if } i = j \\ d_{ij}^{-2} & \text{otherwise} \end{cases} \quad (22a)$$

$$w_{ij}^*(2) = \begin{cases} 0 & \text{if } i = j \\ e^{-2d_{ij}} & \text{otherwise.} \end{cases} \quad (22b)$$

Applying the transformations $w_{ij}(1) = w_{ij}^*(1)/\sum_j w_{ij}^*(1)$ and $w_{ij}(2) = w_{ij}^*(2)/\sum_j w_{ij}^*(2)$, ensures that the row entries in the interaction matrix indeed sum to one. Adopting the inverse of the squared distance as a functional form in Equation (22a) reflects a gravity function (Ertur and Koch, 2007, 1042) and captures that the effect of the spatial externalities weakens more than proportionally with distance; a result that has received support in the empirical literature (see e.g. Keller, 2002). The spatial weight matrix based on the weights in Equation (22a) is called \mathbf{W}_1 and the one in Equation (22b), which Ertur and Koch (2007) employ as a robustness check, is \mathbf{W}_2 .³⁰

The data on genetic distance is taken from the data set of Spolaore and Wacziarg (2009), who rely on data assembled by Cavalli-Sforza et al. (1994). Following the construction of the original weight matrices based on geographic distance, the functional form in Equation (22a) has been chosen for the interaction matrices based on genetic distances as well so that a straightforward robustness check is possible. For the interaction matrix \mathbf{W}_3 , the distances d_{ij} are based on the concept of weighted F_{ST} genetic distance (see Section 2 and Appendix A for this measure) and for matrix \mathbf{W}_4 on weighted Nei’s genetic distance (Nei, 1972, 1973).

²⁹For some countries the capital has moved to a different city over the period from 1960 to 1995. The capital for Côte d’Ivoire, for example, has moved from Abidjan to Yamoussoukro in 1983. In these cases, the coordinates of the city, which was the capital over the longer period with respect to the sample horizon was used. Recent versions of the World Factbook, however, lack geographic coordinates for former capitals so that for these capitals the coordinates have been gathered via Google Maps. This approach has also been employed for cities like Hong Kong that are not listed in the World Factbook.

³⁰This latter matrix is also the one that has been used to calculate the spatial lag of TFP in the Moran scatterplot in Figure 2.

5.2 Results – Interaction Matrix Based on Geographic Distance

Estimation results are presented in Table 1.³¹ The first two columns replicate the results from Table 1 in Ertur and Koch (2007) and serve as a benchmark compared to which all subsequent robustness analyses will be assessed.³²

In this analysis, the interaction matrix based on the weights in Equation (22a) has been used. Column 1 shows in the upper half the results for the standard Solow model estimated by ordinary least squares (OLS). The estimated coefficients on the investment rate and on the population growth rate have the signs expected from the theoretic model and in addition are highly significant. In the lower half, this model is estimated with the restriction $\beta_1 = -\beta_2$ imposed. This restriction is tested with a Wald test and rejected (p -value = 0.038). Furthermore, the implied value for the capital share, $\alpha = 0.58$, is too high compared to empirical estimates. Gollin (2002, 458), for example, estimates that the capital shares for most countries lie in the range of 20% to 35%. Also, Moran's I test indicates spatial autocorrelation in the error term. Based on these results, Ertur and Koch (2007, 1046) thus conclude that the standard model is misspecified as it does not account for physical capital externalities and technological interdependence between countries.

Column 2 shows that the estimation results support the implications of the spatially augmented model. All coefficients have the signs predicted from theory (compare Equation (17)), even though, for instance, the estimated coefficient associated with the spatial lag of the population growth rate is insignificant (p -value = 0.479). The likelihood ratio test does not reject the joint theoretical restriction $\beta_1 + \beta_2 = 0$ and $\theta_1 + \theta_2 = 0$, as the p -value is 0.419, which supports the validity of the spatially augmented model. In addition, the (significant) implied value for the capital share of income is $\alpha = 0.284$ and thus falls approximately right in middle of the range of estimates by Gollin (2002). Furthermore, the parameter ϕ reflecting physical capital externalities is positive and significant at the 10%-level. Also, the implied value for γ , which gauges the degree of technological interdependence among the countries is positive and highly significant, implying that this

³¹All estimations have been carried out in Matlab using the Spatial Econometrics Toolbox by LeSage, which is publicly available under: <http://www.spatial-econometrics.com/> (accessed: 11 August, 2015).

³²Note that since the analysis here is based on the geographic coordinates from the CIA's World Factbook, and these coordinates differ in some cases slightly from the ones in Ertur and Koch (2007), the values for Moran's I test in the unrestricted and restricted versions of the standard Solow model in Column 1 as well as the values for the spatially augmented Solow model in Column 2 are somewhat different. Qualitatively, the results are not affected though. Also, there is a small mistake in Ertur and Koch's Table I, as the values for Moran's I test in the unrestricted Solow model belongs to the restricted Solow model and vice versa.

characteristic indeed needs to be taken into account in growth models, as economies cannot be considered as independent observations (Fischer, 2011, 432). Finally, the value of $\alpha + \phi/(1 - \gamma)$ is below 1, implying that the externalities in the model are not strong enough to lead to endogenous growth (Ertur and Koch, 2007, 1048). In sum, the estimation results therefore provide rather strong support for the model developed by Ertur and Koch.

The next columns in Table 1 assess the sensitivity of these results when changing the underlying data source to more recent versions of the PWT. Due to missing data, for instance, values for the variable capital investment are not available for some countries in PWT 6.2, the sample size needs to be reduced to 83 countries in the estimations based on this data source.³³ In order to obtain estimation results for a balanced sample across all three versions of the PWT considered in this paper, Columns 3 and 4 show estimation results for the 83-country sample with data from PWT 6.1. For the standard model, the results are virtually identical (Column 3) to those from the full sample. However, dropping these 8 observations from the sample affects the results in the spatial model. The implied values for α and ϕ are comparable in size to the full sample with 91 countries, but they are insignificant in the smaller sample. Hence, dropping these 8 countries from the sample already puts a small dent in the robustness of the results obtained by Ertur and Koch (2007).

Columns 5 and 6 change the data source to PWT 6.2. In Column 5 of Table 1 the estimation results are in line with those from Columns 1 and 3. The only exception is that for this data source the restriction $\beta_1 = -\beta_2$ is not rejected (p -value = 0.476), suggesting a good fit between the model and the data, except that the implied value for the capital share is still too high with $\alpha = 0.576$. For the unconstrained estimation of the spatial model, Column 6 shows that compared to Columns 2 and 4, the coefficient for the population growth rate still has the sign implied by the theoretical model, but is now insignificant (p -value = 0.347). The results from the estimation with the joint parameter restriction applied, show that, as for the results for the 83-country sample with data from PWT 6.1, the implied share of capital income and the parameter for the physical capital externalities are insignificant (p -values of 0.403 and 0.213, respectively). Hence, changing the data source from PWT 6.1 to 6.2, suggests that while many results (e.g. concerning the implied value of γ or the test of the joint restriction) are not sensitive to this change, the original results by Ertur and Koch (2007) concerning the implied capital share of income and the parameter ϕ are not robust.

More drastic changes to the original results are visible when moving to PWT 7.1 in

³³See Appendix F for the eight countries with missing data.

Table 1: Estimation Results for the Standard and Spatially Augmented Solow Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_1 (Geographic Distance).

Data set	PWT 6.1		PWT 6.2		PWT 7.1			
	Stand.	Spatial	Stand.	Spatial	Stand.	Spatial		
Model								
Number of observations	91	91	83	83	83	83		
<i>Unconstrained estimation:</i>								
Constant	4.651 (0.010)	0.886 (0.635)	4.609 (0.017)	0.518 (0.796)	7.130 (0.000)	2.780 (0.181)	2.976 (0.189)	1.828 (0.399)
$\ln s_i$	1.276 (0.000)	0.836 (0.000)	1.234 (0.000)	0.789 (0.000)	1.319 (0.000)	0.876 (0.000)	1.697 (0.000)	0.944 (0.000)
$\ln(n_i + 0.05)$	-2.709 (0.000)	-1.538 (0.006)	-2.701 (0.000)	-1.449 (0.021)	-1.835 (0.008)	-0.689 (0.347)	-3.428 (0.000)	-1.441 (0.081)
$\mathbf{W} \ln s_j$	—	-0.347 (0.057)	—	-0.314 (0.137)	—	-0.160 (0.514)	—	0.710 (0.110)
$\mathbf{W} \ln(n_j + 0.05)$	—	0.591 (0.479)	—	0.343 (0.705)	—	-0.191 (0.843)	—	-0.298 (0.793)
$\mathbf{W} \ln y_j$	—	0.742 (0.000)	—	0.732 (0.000)	—	0.608 (0.000)	—	0.595 (0.000)
Moran's I test	0.432 (0.000)	—	0.397 (0.000)	—	0.346 (0.000)	—	0.389 (0.000)	—
<i>Constrained estimation:</i>								
Constant	8.375 (0.000)	2.118 (0.000)	8.407 (0.000)	2.220 (0.000)	8.465 (0.000)	3.158 (0.000)	7.321 (0.000)	1.939 (0.004)
$\ln s_i - \ln(n_i + 0.05)$	1.379 (0.000)	0.855 (0.000)	1.354 (0.000)	0.813 (0.000)	1.356 (0.000)	0.871 (0.000)	1.904 (0.000)	0.958 (0.000)
$\mathbf{W} [\ln s_j - \ln(n_j + 0.05)]$	—	-0.292 (0.098)	—	-0.230 (0.270)	—	-0.149 (0.527)	—	0.692 (0.109)
$\mathbf{W} \ln y_j$	—	0.735 (0.000)	—	0.721 (0.000)	—	0.613 (0.000)	—	0.608 (0.000)
Moran's I test	0.415 (0.000)	—	0.4397 (0.000)	—	0.342 (0.000)	—	0.377 (0.000)	—
Test of restriction	4.427 (0.038)	1.738 (0.419)	4.066 (0.047)	1.474 (0.479)	0.514 (0.476)	0.127 (0.938)	3.805 (0.055)	0.358 (0.836)
Implied α	0.580 (0.000)	0.284 (0.012)	0.575 (0.000)	0.242 (0.120)	0.576 (0.000)	0.196 (0.403)	0.656 (0.000)	8.261 (0.852)
Implied ϕ	—	0.177 (0.082)	—	0.206 (0.139)	—	0.270 (0.213)	—	-7.772 (0.861)
Implied γ	—	0.554 (0.000)	—	0.525 (0.000)	—	0.408 (0.009)	—	-0.043 (0.868)
$\alpha + \frac{\phi}{1-\gamma}$	—	0.680 (0.000)	—	0.676 (0.000)	—	0.651 (0.000)	—	0.808 (0.000)

Note: p -values are given in parentheses. For the standard Solow model the restriction is tested with the Wald test and for the spatially augmented model the restriction is tested with the likelihood ratio (LR) test.

Columns 7 and 8. For the standard model in Column 7, the signs of the coefficient estimates have the expected signs, and Moran's I test indicates misspecification with

respect to spatial correlation in the error term. In accordance with the results for the PWT 6.1 sample, the parameter restriction $\beta_1 = -\beta_2$ is rejected (p -value = 0.055), though this time at the 10% level instead of the 5% level. However, in the spatially augmented model in Column 8, the constrained estimation implies an implausibly large share of capital income. The estimated value for α is 8.261 (although this value is not significant with a p -value of 0.852). Moreover, the value for the physical capital externalities is now negative, but also not significant ($p = 0.861$). The same holds qualitatively for the parameter measuring technological interdependence. These estimates imply that using a more recent data source, leads to drastic changes in the empirical results compared to the benchmark results.³⁴

It needs to be kept in mind though that in addition to the results in Table 1 the model's interpretation relies on the calculation of the direct and indirect effects from changes in the exogenous variables via the approach presented in Section 4.3. The results for these impacts are presented in Table 2 for all four samples considered in this paper. In the paper by Ertur and Koch only the direct effects are reported (though without any reference to the significance of these estimates). Here, a richer analysis is presented by also reporting estimates for the indirect and total impacts on steady-state per worker income due to changes in the exogenous variables and by providing information about the significance of all three impacts as well.

Concerning the direct impacts, the results show that across all four samples an increase in the investment rate in physical capital is approximately comparable in size and significance. The estimated coefficients are highly significant and imply, due to the logarithmic specification of the model, that a 10% increase in the investment rate would result in an increase in per capita income between 8.6% and 11.6%. The results for the indirect impacts of changes in the investment rate, resulting from spatial spillovers, differ however across the samples. Whereas these impacts are comparable in size for the first three samples, the impact is only significant for the PWT 6.2 sample at the 10%-level. For the PWT 7.1 sample, this effect has tripled in size compared to the other estimates and is significant at the 1%-level. These findings indicate again that the results in Ertur and Koch are not robust with respect to changing to more recent versions of the Penn World Table. It is however interesting to note that at least for the first three samples the direct

³⁴That changing the data source from e.g. PWT Version 6.1 or 6.2 to 7.1 can lead to different results in models similar to the one considered here has also been pointed out by Johnson et al. (2013, 270). They find that in the Solow model augmented with human capital, developed by Mankiw et al. (1992), the coefficient on the investment share is reduced in size close to zero, when the estimation is based on a more recent version of the PWT (7.0 in their case). This finding is attributed to the investment series being more variable in this version of the PWT. However, they also state that the reason for this higher variability is unclear (Johnson et al., 2013, 270).

Table 2: Estimation Results for the Direct, Indirect and Total Impacts in the Spatial Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_1 (Geographic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Direct impacts:</i>				
$\ln s_i$	0.916 (0.000)	0.859 (0.000)	0.941 (0.000)	1.158 (0.000)
$\ln(n_i + 0.05)$	-1.693 (0.005)	-1.636 (0.013)	-0.793 (0.269)	-1.635 (0.043)
<i>Indirect impacts:</i>				
$\mathbf{W} \ln s_j$	1.030 (0.118)	0.960 (0.198)	0.915 (0.057)	3.012 (0.004)
$\mathbf{W} \ln(n_j + 0.05)$	-2.008 (0.484)	-2.559 (0.423)	-1.458 (0.476)	-2.709 (0.218)
<i>Total impacts:</i>				
$\ln s_i + \mathbf{W} \ln s_j$	1.945 (0.007)	1.820 (0.023)	1.856 (0.000)	4.170 (0.000)
$\ln(n_i + 0.05) + \mathbf{W} \ln(n_j + 0.05)$	-3.701 (0.230)	-4.196 (0.220)	-2.251 (0.294)	-4.343 (0.054)

Note: p -values are given in parentheses. These were constructed using a set of 500,000 random draws from the estimation.

and indirect impacts from the investment rate contribute both approximately 50% to the total impact of this variable. Table 2 also shows that the results concerning the impacts of the population growth rate are not robust across samples.

Finally, the estimates in Table 2 illustrate that basing the interpretation of the model on the estimates in Table 1 would lead to incorrect conclusions. For instance, interpreting the coefficient associated with the spatial lag of the investment rate, $\mathbf{W} \ln s_j$ in Column 2 of Table 1, as an indicator for the indirect impact would lead to the inference that this effect is negative (-0.347), implying that an increase in the investment rate in neighboring countries would result in a decrease of per capita income in the country under consideration. The true impact estimate in Column 1 of Table 2, however, suggests that the spillover effect is positive, though marginally insignificant at the 10%-level.

Before turning to the estimation results for the interaction matrix based on genetic distance, it should be remembered that Ertur and Koch have also employed an interaction matrix based on the specification in Equation (22b) to assess the sensitivity of their results using their initial choice of weight matrix. The detailed results of the robustness analysis for this interaction matrix across all four samples are delegated to Appendix E.

Tables E.1 and E.2 in this appendix demonstrate again that some of the original results fail to hold when estimating the model across the different samples.

Concerning interaction matrix \mathbf{W}_2 , an important comment needs to be made. This matrix does not seem to correspond exactly to the specification Ertur and Koch (2007) actually use in their empirical analysis. From the Matlab code on the article’s website,³⁵ it is clear that their estimation results are obtained by dividing the geographic distances d_{ij} by 1,000. A reason for this transformation is not given however, and it turns out that the estimation results are highly sensitive to this alternative specification (see the results in Tables E.3 and E.4 in Appendix E). For instance, without dividing the distances by 1,000, the estimation results imply highly significant negative values for the parameters ϕ and γ , and the implied capital share of income increases to an unreasonably, but highly significant value of 90%. With respect to the impact estimates, the values for the direct and total impacts are approximately comparable across both specifications, the indirect effects however turn from being not significant in the specification as implemented by Ertur and Koch to being strongly significant in the specification as claimed in the article (i.e. without the division by 1,000).

5.3 Results – Interaction Matrix Based on Genetic Distance

This section presents the estimation results for the model in which the interaction matrix is based on genetic distance. The general specification of the weights is given by the one in Equation (22a), and the analyses use weighted F_{ST} genetic distance.³⁶ As the results for the standard model do not depend on the interaction matrix, Table 3 shows only the results from the estimation of the spatial Durbin model. The results for the standard model are suppressed in order to avoid duplication.³⁷

Column 1 provides the results for the full sample of 91 countries for data taken from PWT 6.1. In contrast to the benchmark, i.e. the original results in Ertur and Koch, the estimates based on genetic distance show, for instance, that the coefficient associated with the spatial lag of the investment rate is now positive and highly significant (p -value = 0.000). The results for the constrained estimation also differ from the ones

³⁵See <http://qed.econ.queensu.ca/jae/datasets/ertur001/> (accessed: 23 July, 2014).

³⁶For the results based on (weighted) Nei’s distance see Tables E.5 and E.6 in Appendix E. The estimation results are comparable to the ones shown in Tables 3 and 4, which might be explained by the fact that the correlation between the two measures of genetic distance is 93.9% (Spolaore and Wacziarg, 2009, 482) and thus very high.

³⁷It is worth pointing out however, that in all samples the standard model continues to be misspecified, as the values for Moran’s I test suggest spatial autocorrelation in the error term (p -values are 0.000 in all four tests) also when using interaction matrix \mathbf{W}_3 in these tests.

Table 3: Estimation Results for the Spatial Durbin Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_3 (Weighted F_{ST} Genetic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Unconstrained estimation:</i>				
Constant	8.654 (0.001)	8.246 (0.000)	5.941 (0.011)	-1.932 (0.423)
$\ln s_i$	0.820 (0.000)	0.945 (0.000)	0.888 (0.000)	0.972 (0.000)
$\ln(n_i + 0.05)$	-1.034 (0.054)	-0.871 (0.099)	-0.148 (0.790)	-0.930 (0.153)
$\mathbf{W} \ln s_j$	0.901 (0.000)	0.665 (0.001)	0.725 (0.002)	-0.009 (0.983)
$\mathbf{W} \ln(n_j + 0.05)$	0.651 (0.500)	0.431 (0.632)	-1.625 (0.096)	-1.912 (0.078)
$\mathbf{W} \ln y_j$	0.322 (0.006)	0.327 (0.002)	0.198 (0.128)	0.556 (0.000)
<i>Constrained estimation:</i>				
Constant	5.520 (0.000)	5.452 (0.000)	6.013 (0.000)	2.452 (0.000)
$\ln s_i - \ln(n_i + 0.05)$	0.785 (0.000)	0.856 (0.000)	0.870 (0.000)	0.996 (0.000)
$\mathbf{W}[\ln s_j - \ln(n_j + 0.05)]$	0.850 (0.000)	0.653 (0.001)	0.655 (0.003)	0.130 (0.743)
$\mathbf{W} \ln y_j$	0.280 (0.019)	0.296 (0.005)	0.245 (0.045)	0.605 (0.000)
Test of restriction	2.450 (0.294)	1.949 (0.377)	1.862 (0.394)	3.708 (0.157)
Implied α	1.491 (0.003)	1.830 (0.052)	1.598 (0.038)	-0.273 (0.803)
Implied ϕ	-1.052 (0.039)	-1.360 (0.155)	-1.133 (0.147)	0.772 (0.474)
Implied γ	-0.319 (0.098)	-0.189 (0.196)	-0.219 (0.212)	0.238 (0.288)
$\alpha + \frac{\phi}{1-\gamma}$	0.694 (0.000)	0.686 (0.000)	0.669 (0.000)	0.740 (0.000)

Note: p -values are given in parentheses. The restriction for the spatially augmented model is tested with the likelihood ratio (LR) test.

with an interaction matrix using geographic distance, as the implied value for α is now implausibly large and highly significant. Moreover, the estimate for γ , measuring the degree of technological interdependence is now negative and marginally significant at the 10%-level, which seems implausible.³⁸ Similar results are also obtained for the other samples. When using data from PWT 7.1 for instance, the implied value for the capital

³⁸The implied values for α , ϕ , and γ are of approximately similar size in the estimation based on Nei's distance for this sample, though neither value is significant at the 10%-level. This is the exception from the claim about comparable estimation results for both measures of genetic distance made in Footnote 36.

share of income in Column 4 actually turns negative (although the p -value is 0.803). In neither sample, based on a likelihood ratio test, the joint parameter restriction $\beta_1 + \beta_2 = 0$ and $\theta_1 + \theta_2 = 0$ is rejected though.

Table 4: Estimation Results for the Direct, Indirect and Total Impacts in the Spatial Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_3 (Weighted F_{ST} Genetic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Direct impacts:</i>				
$\ln s_i$	0.887 (0.000)	1.004 (0.000)	0.918 (0.011)	1.035 (0.000)
$\ln(n_i + 0.05)$	-1.009 (0.058)	-0.856 (0.101)	-0.198 (0.722)	-1.229 (0.060)
<i>Indirect impacts:</i>				
$\mathbf{W} \ln s_j$	1.673 (0.000)	1.401 (0.000)	1.098 (0.000)	1.116 (0.156)
$\mathbf{W} \ln(n_j + 0.05)$	0.506 (0.701)	0.260 (0.830)	-1.996 (0.062)	-5.246 (0.001)
<i>Total impacts:</i>				
$\ln s_i + \mathbf{W} \ln s_j$	2.560 (0.000)	2.405 (0.000)	2.015 (0.000)	2.151 (0.012)
$\ln(n_i + 0.05) + \mathbf{W} \ln(n_j + 0.05)$	-0.503 (0.707)	-0.560 (0.627)	-2.193 (0.035)	-6.475 (0.003)

Note: p -values are given in parentheses. These were constructed using a set of 500,000 random draws from the estimation.

Despite these results from the estimation of the constrained model across the four samples, the impact estimates in Table 4, which are calculated from the unconstrained estimation results will be briefly discussed. Across all samples, the estimates for the direct impact of a change in the investment rate on steady-state per capita income is comparable to the results for the model with interaction matrix \mathbf{W}_1 . One important difference to the results in Table 2 concerns the spillovers from a change in the investment rate for the full sample of 91 countries. The estimated effects are now highly significant and imply that a change of 1% in the investment rate in all countries except country i would result in an increase of approximately 1.7% in per capita income in country i . Another interesting result is that these spillovers are not significant in the sample for PWT 7.1 in Column 4 of Table 4, whereas the reverse holds for this sample in the estimation with geographic distance. There (see Column 4 in Table 2), this estimate is not only highly significant, but also large in size. Table 4 furthermore clearly shows that the impacts with respect to

the population growth rate are highly sensitive to the particular version of the PWT used in the estimation. Hence, it can be concluded that the original results by Ertur and Koch are challenged strongly both by changing the measure in the interaction matrix from geographic to genetic distance and also by substituting the data set from older versions of the Penn World Table for more recent versions.

6 Conclusion

This paper has presented the growth model with technological interdependence among countries developed by Ertur and Koch (2007) and subjected their empirical results to a series of robustness checks. In contrast to the original specification, which uses an interaction matrix based on geographic distance, here data on measures of genetic distance from Spolaore and Wacziarg (2009) has been used to construct an alternative interaction matrix. Furthermore, additional robustness checks have been conducted to assess the sensitivity of the original results across different versions of the Penn World Table for the same period and the same set of countries. The analyses show that the original results by Ertur and Koch are only robust to a certain extent. While the hypothesis that countries need to be analyzed in an interdependent system is supported, other results are highly sensitive to the version of the Penn World Table that is used in the empirical estimation. Ertur and Koch (2007) estimate, for instance, an implied capital share of income slightly below 30%, which is significant at the 5% level. However, this result is not robust when estimating the model with data from PWT 6.2 or 7.1 instead.

Furthermore, whereas Ertur and Koch only provide estimates of the direct impacts on per worker income associated with changes in the exogenous variables, in this paper values for the indirect and total impacts have been calculated as well. The results again indicate non-robustness across different versions of the PWT, as, for example, the indirect impacts (or spillovers) associated with changes in the investment rate of physical capital on per worker income in steady state are not significant in the PWT 6.1 sample, but significant in the ones based on PWT 6.2 and 7.1, respectively. Results have also been shown to be highly sensitive to the precise specification of the weights in the interaction matrix based on geographical distance.

Based on theoretical and technical considerations, genetic distance has been introduced as an alternative measure to geographic distance, on which to base an interaction matrix. Concerning the empirical results for this alternative matrix, it is found that, whereas in the original model indirect spillovers from capital investment were insignificant in the PWT 6.1 sample, using a measure of genetic distance, these spillovers now have a

significant effect on steady-state income per worker. However, the version of the model with an interaction matrix based on genetic distance implies an implausibly large capital share of income. In addition, also in the case of genetic distance, estimation results are sensitive to the version of the data set. It can thus be stated that the empirical results in Ertur and Koch are highly sensitive to the measure on which the weights in the interaction matrix are based (geographic or genetic distance) as well as to the concrete specification of the weights in the interaction matrix. This conclusion vividly demonstrates the importance of sound data when it comes to, for instance, giving policy advice.

In this paper, as, for example, in Fischer (2011), only level regressions have been addressed. Future work will also investigate the sensitivity of the estimates for the growth regressions in Ertur and Koch (2007), as well as the impact of introducing human capital into this model. Results from Ertur and Koch (2006) suggest that this factor is not related to growth within this framework. However, as the results in this paper clearly demonstrate, that this holds across different versions of the PWT need not necessarily be the case. It should also be pointed out that an endogenous version of the model framework exists (Ertur and Koch, 2011), which for a smaller set of countries and a shorter time period provides empirical support, based in part on data from PWT 6.2, in favor of the endogenous version. But again, this is no guarantee that this necessarily needs to hold across different versions of the PWT. Robustness should be assessed for this finding as well. As this paper has also clearly demonstrated the sensitivity of the empirical results to the precise choice of the interaction matrix, further research will be devoted to this issue. In particular, the method of Bayesian Model Averaging will be used to address the uncertainty concerning the specification of the interaction matrix in this model.

Before concluding this paper, a brief remark on policy concerning the role of geographic and genetic distance in determining the strength of knowledge spillovers: As humans have demonstrated numerous times over the course of history in often horrible ways, neither the geographic distances (through expansionary wars or state collapses) nor the genetic distances (via genocide or the slave trade) between countries are fixed in the long run.³⁹ Abstracting from these, policy can, however, still have an impact by e.g. fostering openness and thereby removing barriers to the diffusion of knowledge (Spolaore and Wacziarg, 2009, 524).

³⁹Also another, maybe at first glance more innocuous possibility, comes to mind in light of the dramatically falling costs for gene sequencing. These are currently below \$4,500 for the whole human genome compared to nearly \$100,000,000 in 2001 (Wetterstrand, 2015). I do not think that it is completely unrealistic that personal genetic data will be used by states to influence, for instance, visa decisions.

Appendix

A Definition of F_{ST} Genetic Distance

This appendix provides a formal definition of the concept of F_{ST} genetic distance which was developed by Wright (1951).⁴⁰ Consider as an example the case of two populations, A and B , that are of equal size and a single gene which can either have the form of allele 1 or of allele 2.⁴¹ Denote the gene frequency of allele 1 in population A by p_A and the one for allele 2 by q_A . The probability that two randomly selected alleles at the locus under consideration will be identical (i.e. homozygosity occurs) is given by $p_A^2 + q_A^2$. The case of heterozygosity (i.e. two randomly selected alleles will differ) then is $het_A = 1 - (p_A^2 + q_A^2) = 2p_Aq_A$. This follows as $p_A + q_A = 1$ and hence $(p_A + q_A)^2 = p_A^2 + q_A^2 + 2p_Aq_A = 1$ holds. For population B , the equivalent expressions for homozygosity and heterozygosity are given by $p_B^2 + q_B^2$ and $het_B = 1 - (p_B^2 + q_B^2) = 2p_Bq_B$, respectively.

Denoting the average gene frequencies of the two alleles in the two populations as $\bar{p} = \frac{1}{2}(p_A + p_B)$ and $\bar{q} = \frac{1}{2}(q_A + q_B)$, it follows that in the sum of the two populations heterozygosity is given by $het_{AB} = 1 - (\bar{p}^2 + \bar{q}^2) = 2\bar{p}\bar{q}$. The average heterozygosity in the two populations is $het_{mean} = \frac{1}{2}(het_A + het_B)$. By comparing het_{mean} to het_{AB} , F_{ST} is a measure for the “variation in gene frequencies of populations” (Spolaore and Wacziarg, 2009, 525)

$$F_{ST} = \frac{het_{AB} - het_{mean}}{het_{AB}} = 1 - \frac{p_Aq_A + p_Bq_B}{2\bar{p}\bar{q}} = \frac{1}{4} \frac{(p_A - p_B)^2}{\bar{p}(1 - \bar{p})}. \quad (\text{A.1})$$

It now follows that the genetic distance between two populations is zero, if their allele frequencies at the given locus are identical (i.e. $p_A = p_B$) and that F_{ST} equals one if the respective frequencies are completely different (i.e. $p_A = 1$ and $p_B = 0$ or vice versa).⁴²

In the construction of the interaction matrix based on genetic distance in Section 5, weighted F_{ST} distances are used to account for the fact that populations in, for instance,

⁴⁰For a compact review of F_{ST} that covers additional details, see Holsinger and Weir (2009).

⁴¹Compare Spolaore and Wacziarg (2009, 524-525) for this approach and Cavalli-Sforza et al. (1994, 26-27) for extensions to more than two alleles and two populations.

⁴²Note that Cavalli-Sforza et al. (1994, 29), for instance, provide the following formula for F_{ST} genetic distance

$$F_{ST} = \frac{V_p}{\bar{p}(1 - \bar{p})} \quad (\text{A.2})$$

where \bar{p} are the average gene frequencies across the populations under consideration, and V_p indicates the variance between gene frequencies across these populations. If now $p_A \equiv \bar{p} + \sigma$ and $p_B \equiv \bar{p} - \sigma$ with $\sigma \geq 0$ and the variance is denoted by σ^2 , then the formula in Equation (A.1) is equivalent to the one provided in Equation (A.2) (see Spolaore and Wacziarg (2015, 6-7) for this derivation).

the United States or the United Kingdom consist of many subpopulations. If now the United States contains the populations $i = 1, \dots, I$ and the United Kingdom the populations $j = 1, \dots, J$ and s_{1i} is the share of population i in the United States and s_{2j} the share of population j in the United Kingdom, then the weighted F_{ST} genetic distance between these two countries is given by (see, for example, Spolaore and Wacziarg (2009, 484-485))

$$F_{ST}^W = \sum_{i=1}^I \sum_{j=1}^J (s_{1i} \times s_{2j} \times d_{ij}) \quad (\text{A.3})$$

where d_{ij} denotes the F_{ST} genetic distance between populations i and j .

B Spatial Weight Matrices and Great Circle Distances

Numerous possibilities exist to model spatial connectivity via a spatial weight matrix. This appendix illustrates two possibilities and provides details on the calculation of great circle distances that are commonly used in empirical work. As an example, for how to model the spatial relationship between geographic regions, consider the four NUTS 2 regions⁴³ Schleswig-Holstein (SH), Hamburg (HH), Lüneburg (LÜ), and Mecklenburg-Vorpommern (MP) depicted in the map in Figure B.1.



Figure B.1: NUTS 2 Regions in Northern Germany.

Defining now those regions that share a common border as neighboring regions, the

⁴³NUTS is an acronym of the French Nomenclature des Unités territoriales statistiques, i.e. the Nomenclature of territorial units for statistics of the EU, and the NUTS 2 level comprises government regions.

following spatial weight matrix of first-order neighbors can be constructed

$$\mathbf{W}_A = \begin{pmatrix} & \text{SH} & \text{HH} & \text{LÜ} & \text{MV} \\ \text{SH} & 0 & 1 & 1 & 1 \\ \text{HH} & 1 & 0 & 1 & 0 \\ \text{LÜ} & 1 & 1 & 0 & 1 \\ \text{MV} & 1 & 0 & 1 & 0 \end{pmatrix}.$$

In row-standardized form, the result is

$$\mathbf{W}_B = \begin{pmatrix} & \text{SH} & \text{HH} & \text{LÜ} & \text{MV} \\ \text{SH} & 0 & 1/3 & 1/3 & 1/3 \\ \text{HH} & 1/2 & 0 & 1/2 & 0 \\ \text{LÜ} & 1/3 & 1/3 & 0 & 1/3 \\ \text{MV} & 1/2 & 0 & 1/2 & 0 \end{pmatrix}. \quad (\text{B.1})$$

Here, queen contiguity is used as a concept to determine the spatial weights.⁴⁴ However, a spatial weight matrix can also be constructed based on geographic distance between units of observations. One relatively straightforward alternative in this vein is the geographic distance between country capitals. Consider, for example, the four capitals Berlin, Buenos Aires, Moscow and Seoul and take as spatial weights the shortest distance between them. Since the earth is (approximately) a sphere, this distance is not a straight line, but the shortest path between the cities along the surface of the earth, i.e. an arc of a great circle. Figure B.2 shows the great circle distances between Berlin and Buenos Aires and between Moscow and Seoul on a Plate Carée projection of the earth, and Panels (a) and (b) in Figure B.3 show the same distances on the surface of a spherical earth.⁴⁵

In general, the great circle distance between two points i and j can be calculated by using the spherical law of cosines (Shekhar and Xiong, 2008, 639-642)

$$d_{ij} = R_{\oplus} \times \arccos[\cos lat_i \cos lat_j \cos(long_i - long_j) + \sin lat_i \sin lat_j] \quad (\text{B.2})$$

where $R_{\oplus} = 6,378.1\text{km}$ is the (rounded value of the) earth's equatorial radius (see Ahrens (1995, 36)), and the values for latitude and longitude are in decimal degrees.

⁴⁴The name follows from the movement of the queen on the chessboard. Other criteria for determining spatial weights, whose names have the same origin are bishop and rook contiguity (see, for instance, (Anselin, 1988, 18)).

⁴⁵These figures were drawn with ArcGIS.



Figure B.2: Great Circle Distances between Berlin and Buenos Aires and between Moscow and Seoul on a Plate Carée Projection of the Earth.



(a) Great Circle Distance between Berlin and Buenos Aires on the Surface of a Spherical Earth.



(b) Great Circle Distance between Moscow and Seoul on the Surface of a Spherical Earth.

Figure B.3: Great Circle Distances on the Surface of a Spherical Earth.

The geographic coordinates for the four capitals under consideration are as follows:

Berlin: N52°31', E13°24'

Buenos Aires: S34°35', W58°40'

Moscow: N55°45', E37°36'

Seoul: N37°33', E126°59'

These coordinates are taken from the CIA's World Factbook (Central Intelligence Agency, 2013) and before inputting them into the formula above they need to be converted from degree-minute format to decimal degrees by dividing the entry for minutes by 60 and adding it to the value for the degree (see Peterson and Smith (2012, 458)). In decimal-degree form south and west are denoted with negative values (Shekhar and Xiong, 2008, 639) so that the coordinates from above now read⁴⁶

Berlin: 52.52° , 13.4°

Buenos Aires: -34.58° , -58.67°

Moscow: 55.75° , 37.6°

Seoul: 37.55° , 126.98°

The distance between Berlin and Buenos Aires, for instance, can then be calculated as

$$\begin{aligned} d_{B,BA} &= R_{\oplus} \times \arccos [\cos(52.52^\circ) \cos(-34.58^\circ) \cos(13.4^\circ - (-58.68^\circ)) \\ &\quad + \sin(52.52^\circ) \sin(-34.58^\circ)] \\ \iff d_{B,BA} &= R_{\oplus} \times 107.2356^\circ. \end{aligned}$$

Converting now from degrees to radians by multiplying the angle with $\pi/180$ gives the distance between Berlin and Buenos Aires as

$$d_{B,BA} = 6,378.1\text{km} \cdot 107.2356^\circ \cdot \frac{\pi}{180} = 6,378.1\text{km} \cdot 1.8716 = 11,937.25\text{km}. \quad (\text{B.3})$$

The complete spatial weight matrix for the four capitals thus reads

$$\mathbf{W}_C = \begin{pmatrix} & \text{Berlin} & \text{Buenos Aires} & \text{Moscow} & \text{Seoul} \\ \text{Berlin} & 0 & 11,936 & 1,610 & 8,138 \\ \text{Buenos Aires} & 11,936 & 0 & 13,505 & 19,431 \\ \text{Moscow} & 1,610 & 13,505 & 0 & 6,616 \\ \text{Seoul} & 8,138 & 19,431 & 6,616 & 0 \end{pmatrix}. \quad (\text{B.4})$$

The difference between the value in Equation (B.3) and the corresponding value in Equation (B.4) stems from rounding the results of the trigonometric functions in the calculation above. In the matrix, the values have been calculated by implementing the formula in

⁴⁶These conversions can easily be done in Mathematica or Matlab using the functions FromDMS and dm2degrees, respectively. The values here are rounded to two decimal points.

Matlab directly. After row-standardizing the matrix, it could be used in a straightforward manner in the econometric exercise in the main text.

Note that as an alternative to the equatorial radius the mean radius of the earth could have been used in the calculation. In its Earth Fact Sheet⁴⁷ the National Aeronautic and Space Administration (NASA) gives a value of 6,371km for the mean radius. Substituting this value in Equation (B.2) would not change the relative distances between capitals, however.

It should as well be kept in mind that the distances have been calculated by assuming the earth is a sphere, although it is better described by an oblate spheroid and hence, for instance, Vincenty's formulae would be more accurate (Vincenty, 1975). For the distances considered here, the gain in accuracy is negligible though and using a spatial weights matrix based on Vincenty's formulae would not change the qualitative results in the main text.

C Proof that $(\mathbf{I} - \gamma\mathbf{W})^{-1}$ exists

This appendix demonstrates that the inverse of $(\mathbf{I} - \gamma\mathbf{W})$ exists for the assumed parameter space of γ , given that $\gamma \neq 0$ and that $\frac{1}{\gamma}$ is not an eigenvalue of \mathbf{W} . The first condition is obvious as it simply posits the existence of spatial externalities.⁴⁸ In general, $(\mathbf{I} - \gamma\mathbf{W})$ will have an inverse, if it is non-singular, implying that $|\mathbf{I} - \gamma\mathbf{W}| \neq 0$. The matrix will thus be singular and have no inverse if $|\mathbf{I} - \gamma\mathbf{W}| = 0$. Applying the rules for determinants, (see, for example, Sydsæter et al., 2008, 5) this expression is equivalent to

$$\left| \frac{1}{\gamma}\mathbf{I} - \mathbf{W} \right| = 0 \quad \iff \quad \left| \mathbf{W} - \frac{1}{\gamma}\mathbf{I} \right| = 0.$$

The second equation is the characteristic or eigenvalue equation of \mathbf{W} , demonstrating that $(\mathbf{I} - \gamma\mathbf{W})$ will not have an inverse if $\frac{1}{\gamma}$ is an eigenvalue of \mathbf{W} .

Having established the general conditions under which the inverse exists, it will now be shown that it exists for $0 \leq \gamma < 1$.⁴⁹

The last step in the proof will use the result that the eigenvalues of \mathbf{W} will be less than or

⁴⁷This is available under: <http://nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html> (accessed: 11 August, 2015).

⁴⁸For $\gamma = 0$, the model would reduce to a standard Solow model with physical externalities.

⁴⁹It will actually be shown that the inverse exists for $|\gamma| < 1$. This naturally includes the parameter space described by the inequality in the main text, where one could exclude $\gamma = 0$.

equal to 1 in absolute value. This result is now proved via Gerschgorin's Circle Theorem.⁵⁰

This theorem states that the eigenvalues of a matrix $\mathbf{B} \in \mathbb{C}^{n \times n}$ lie in the complex plane in the area that is given by the intersection of the union of all Gerschgorin circles associated with the rows of \mathbf{B} and the union of all Gerschgorin circles associated with the columns of \mathbf{B} . Formally, the Gerschgorin circles (or discs), \mathcal{G}_i^r , associated with the rows are given by

$$\mathcal{G}_i^r = \{z \in \mathbb{C} : |z - b_{ii}| \leq r_i\}, \quad \text{where} \quad r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |b_{ij}| \quad \text{for} \quad i = 1, 2, \dots, n$$

which means that the circles have the entry b_{ii} of the matrix \mathbf{B} as their center and the sum of the absolute values of the off-diagonal entries of the respective row as their radius. The eigenvalues of the matrix are then contained in the union of these n Gerschgorin circles associated with the rows of \mathbf{B} , i.e. in $\bigcup_{i=1}^N \mathcal{G}_i^r$.

The Gerschgorin circles associated with the columns, \mathcal{G}_i^c , are given by⁵¹

$$\mathcal{G}_j^c = \{z \in \mathbb{C} : |z - b_{jj}| \leq c_j\}, \quad \text{where} \quad c_j = \sum_{\substack{i=1 \\ i \neq j}}^n |b_{ij}| \quad \text{for} \quad j = 1, 2, \dots, n$$

and the union of these n Gerschgorin circles is denoted by $\bigcup_{i=1}^N \mathcal{G}_i^c$. Hence, the eigenvalues of \mathbf{B} will be contained in the following intersection

$$\left(\bigcup_{i=1}^N \mathcal{G}_i^r \right) \cap \left(\bigcup_{i=1}^N \mathcal{G}_i^c \right).$$

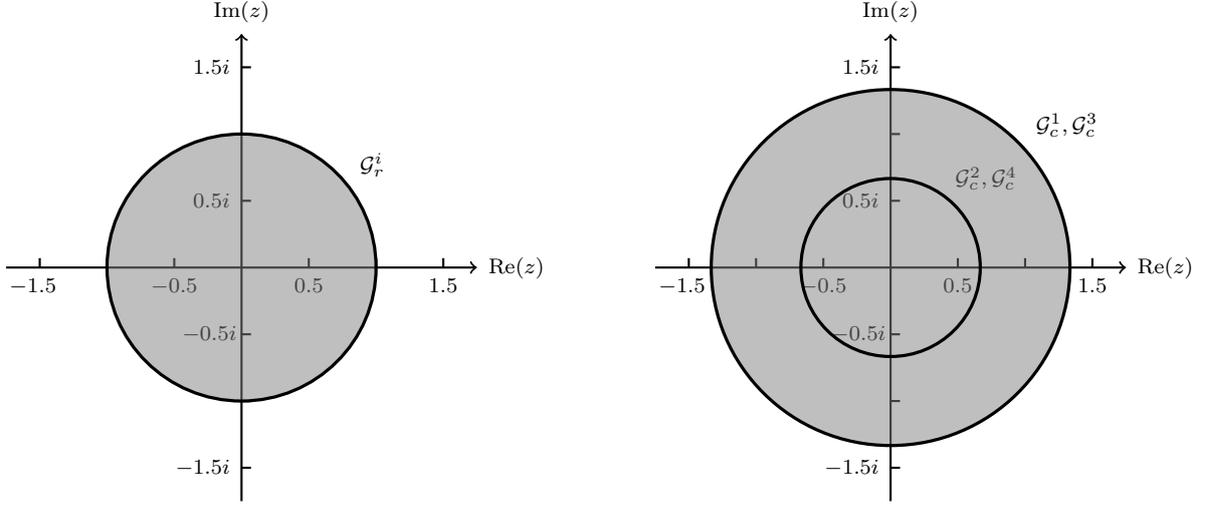
Before applying this theorem to the spatial weight matrix \mathbf{W} from the main text, a brief graphical illustration for the matrix in (B.1) from Appendix B will be provided to deepen the understanding of the theorem.

For the matrix \mathbf{W}_B , all Gerschgorin circles are centered around the point $(0, 0)$ in the complex plane, and since the matrix is row standardized all circles associated with the rows have a radius of 1. The union of these circles is shown in Panel (a) of Figure C.4. Naturally, the circles associated with the columns are also centered around $(0, 0)$ and, since $c_1 = c_3 = 4/3$ and $c_2 = c_4 = 2/3$, there are in effect only two circles for the columns, which have radii of $4/3$ and $2/3$, respectively. These circles and their union are depicted

⁵⁰The original statement is due to Gerschgorin (1931). Here, I rely on the presentations in Meyer (2000, 498) and Cheney and Kincaid (2008, 347-349).

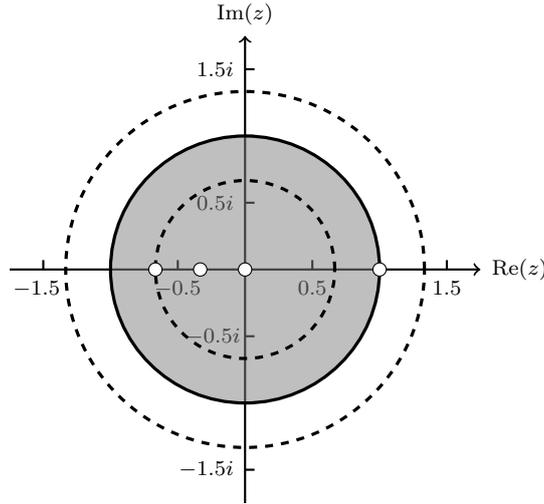
⁵¹That the eigenvalues of \mathbf{B} also are contained in the circles associated with the columns follows, since the calculation of the eigenvalues involves the determinant, which is identical for a matrix and its transpose (Meyer, 2000, 463).

in Panel (b) of Figure C.4. Finally, Panel (c) of Figure C.4 overlays the two results, showing that all eigenvalues⁵² (which are depicted with a white circle in the figure) will be contained within the unit circle.



(a) Union of Gerschgorin Circles Associated with the Rows of \mathbf{W}_B , i.e. $\bigcup_{i=1}^N \mathcal{G}_i^r$.

(b) Union of Gerschgorin Circles Associated with the Columns of \mathbf{W}_B , i.e. $\bigcup_{i=1}^N \mathcal{G}_i^c$.



(c) Area in which the Eigenvalues of the Matrix \mathbf{W}_B are Contained, i.e. $\left(\bigcup_{i=1}^N \mathcal{G}_i^r\right) \cap \left(\bigcup_{i=1}^N \mathcal{G}_i^c\right)$.

Figure C.4: Illustration of Gerschgorin's Circle Theorem for the Matrix \mathbf{W}_B .

This brief illustration provides an insight into why the eigenvalues λ_i of the matrix \mathbf{W} from the main text will be equal to or less than 1 in absolute value. The result hinges on the assumption that the spatial weight matrix is row standardized so that $r_i = 1$ for $i = 1, \dots, N$, implying that $|\lambda_i| \leq 1$ will hold for the eigenvalues.

⁵²The eigenvalues of \mathbf{W}_B are $\lambda_1 = 1, \lambda_2 = -1/3, \lambda_3 = -2/3$, and $\lambda_4 = 0$.

In order to finally show that $(\mathbf{I} - \gamma\mathbf{W})^{-1}$ exists a last intermediate result is helpful. From Schur's Triangularization Theorem it follows that via a similarity transformation⁵³ every square matrix can be made upper triangular (Meyer, 2000, 508). Hence, an invertible $N \times N$ matrix \mathbf{P} exists so that

$$\mathbf{P}^{-1}\mathbf{W}\mathbf{P} = \mathbf{T}, \quad \text{with} \quad \mathbf{T} = \begin{pmatrix} t_{11} & \cdots & t_{1N} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t_{NN} \end{pmatrix}$$

where as an implication of Schur's Triangularization Theorem the eigenvalues of \mathbf{W} are the diagonal entries of the matrix \mathbf{T} . This matrix will now be used to prove that $\mathbf{I} - \gamma\mathbf{W}$ is non-singular.

Proof.

$$\begin{aligned} |\mathbf{I} - \gamma\mathbf{W}| &= |\mathbf{P}\mathbf{P}^{-1}(\mathbf{I} - \gamma\mathbf{W})| = |\mathbf{P}(\mathbf{I} - \gamma\mathbf{W})\mathbf{P}^{-1}| \\ \iff |\mathbf{I} - \gamma\mathbf{W}| &= |\mathbf{P}\mathbf{I}\mathbf{P}^{-1} - \gamma\mathbf{P}\mathbf{W}\mathbf{P}^{-1}| = |\mathbf{I} - \gamma\mathbf{T}| \\ \iff |\mathbf{I} - \gamma\mathbf{W}| &= (1 - \gamma t_{11}) \cdots (1 - \gamma t_{NN}) \\ \iff |\mathbf{I} - \gamma\mathbf{W}| &\neq 0 \end{aligned} \tag{C.1}$$

The last line follows if $|\gamma\lambda_{ii}| \neq 1$, which holds since $t_{ii} = \lambda_i$ and $|\lambda_{ii}| \leq 1$ from Gerschgorin's Theorem, and also $|\gamma| < 1$ holds. \square

Here, the product rule on determinants as well as the fact that $|\mathbf{P}\mathbf{P}^{-1}| = 1$ (see Meyer, 2000, 508) has been used, and to obtain Equation (C.1) the rule for the determinant of a triangular matrix has been employed (see Meyer, 2000, 462).

An implication of applying Gerschgorin's Theorem in this case is that it rules out that $\frac{1}{\gamma}$ is an eigenvalue of \mathbf{W} . This follows since it has been established that the eigenvalues of \mathbf{W} are in the interval $[-1, 1]$ and $|\gamma| < 1$ so that the hypothetical eigenvalue $\frac{1}{\gamma}$ would be larger than 1.

⁵³Two square matrices \mathbf{A} and \mathbf{B} are similar "whenever there exists a nonsingular matrix \mathbf{P} such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$. The product $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$ is called a **similarity transformation** on \mathbf{A} " (Meyer, 2000, 506, emphasis in the original).

D Derivations and Proofs of Selected Model Results

In this appendix the expressions for the spatial multiplier and the elasticities are derived.

D.1 Derivation of Equation (6)

This appendix proves that $(\mathbf{I} - \gamma\mathbf{W})^{-1}$ equals $\sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r$ and uses the result to derive the expression for the level of technology in country i in Equation (6) in Section 3.2. The crucial step in this proof will be to show that $\lim_{n \rightarrow \infty} (\gamma\mathbf{W})^n = \mathbf{0}$. If this holds, then $\mathbf{I} - \gamma\mathbf{W}$ has an inverse, thereby providing an alternative proof for its existence, and it follows that this inverse can be written as $\sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r$.⁵⁴

Proof. As a first step, the following result is helpful

$$(\mathbf{I} - \gamma\mathbf{W}) (\mathbf{I} + \gamma\mathbf{W} + (\gamma\mathbf{W})^2 + \dots + (\gamma\mathbf{W})^n) = \mathbf{I} - (\gamma\mathbf{W})^{n+1}$$

where, if $\lim_{n \rightarrow \infty} (\gamma\mathbf{W})^n = \mathbf{0}$ the right-hand side tends to \mathbf{I} as $n \rightarrow \infty$. Left multiplying this equation by $(\mathbf{I} - \gamma\mathbf{W})^{-1}$ then leads to the Neumann series

$$\mathbf{I} + \gamma\mathbf{W} + (\gamma\mathbf{W})^2 + \dots + (\gamma\mathbf{W})^n = \sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r = (\mathbf{I} - \gamma\mathbf{W})^{-1}.$$

It remains to be shown that $\lim_{n \rightarrow \infty} (\gamma\mathbf{W})^n = \mathbf{0}$ indeed holds. This is equivalent to the statement that the spectral radius of the matrix $\gamma\mathbf{W}$ is strictly smaller than 1 (see Meyer, 2000, 618). Since the spectral radius of a matrix is given by its largest eigenvalue in absolute value (Meyer, 2000, 497), a straightforward application of Gerschgorin's Circle Theorem to the matrix $\gamma\mathbf{W}$ shows that its spectral radius is smaller than 1. This follows, since for the matrix \mathbf{W} the largest eigenvalue is 1, and multiplying each matrix entry by $|\gamma| < 1$ would reduce the radii of the Gerschgorin circles. \square

The expression $(\mathbf{I} - \gamma\mathbf{W})^{-1} = \sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r$ is also referred to as the spatial multiplier (Ertur and Koch, 2007, 1044) and using this result in Equation (5) leads to

$$\mathbf{A} = \sum_{r=0}^{\infty} \lambda^r \mathbf{W}^r \mathbf{\Omega} + \phi \sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r \mathbf{k}.$$

⁵⁴This is called a Neumann Series and can be used to approximate the inverse (Meyer, 2000, 126).

This equation can be simplified by repeatedly substituting the result $\mathbf{W}\Omega = \Omega$.⁵⁵ The equation then reads

$$\mathbf{A} = \frac{1}{1-\gamma} \cdot \Omega + \phi \sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r \mathbf{k}.$$

The last term on the right-hand side can be shown to equal

$$\phi \sum_{r=0}^{\infty} \gamma^r \mathbf{W}^r \mathbf{k} = \phi \begin{pmatrix} \ln k_1(t) \\ \vdots \\ \ln k_N(t) \end{pmatrix} + \begin{pmatrix} \prod_{j=1}^N \ln k_j(t)^{\phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{1j}} \\ \vdots \\ \prod_{j=1}^N \ln k_j(t)^{\phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{Nj}} \end{pmatrix}$$

so that after first collecting the terms in logarithms and then applying the exponential transformation, the level of technology for a given country i is given by

$$A_i(t) = \Omega(t)^{\frac{1}{1-\gamma}} \cdot k_i(t)^{\phi} \cdot \prod_{j=1}^N k_j(t)^{\phi \sum_{r=1}^{\infty} \gamma^r (\mathbf{W}^r)_{ij}}$$

which is Equation (6) in Section 3.2.

D.2 Derivation of the Elasticities

Define \mathbf{S} as the $N \times 1$ vector of investment rates, s_i , in logarithms and \mathbf{N} as the $N \times 1$ vector of the effective depreciation rates, $n_i + g + \delta$, also in logarithms, then Equation (14) can be rewritten in matrix form as

$$\mathbf{y} = \frac{1}{1-\alpha-\phi} \Omega + \frac{\alpha+\phi}{1-\alpha-\phi} \mathbf{S} - \frac{\alpha+\phi}{1-\alpha-\phi} \mathbf{N} - \frac{\alpha\gamma}{1-\alpha-\phi} \mathbf{W}\mathbf{S} - \frac{\alpha\gamma}{1-\alpha-\phi} \mathbf{W}\mathbf{N} + \frac{\gamma(1-\alpha)}{1-\alpha-\phi} \mathbf{W}\mathbf{y}.$$

Solving this equation for \mathbf{y} , yields

$$\mathbf{y} = \frac{1}{1-\alpha-\phi} \left[\mathbf{I} - \frac{\gamma(1-\alpha)}{1-\alpha-\phi} \mathbf{W} \right]^{-1} \Omega + \left[\frac{\gamma(1-\alpha)}{1-\alpha-\phi} \mathbf{W} \right]^{-1} \left(\frac{\alpha+\phi}{1-\alpha-\phi} \mathbf{I} - \frac{\alpha\gamma}{1-\alpha-\phi} \mathbf{W} \right) \mathbf{S} + \left[\frac{\gamma(1-\alpha)}{1-\alpha-\phi} \mathbf{W} \right]^{-1} \left(\frac{\alpha+\phi}{1-\alpha-\phi} \mathbf{I} - \frac{\alpha\gamma}{1-\alpha-\phi} \mathbf{W} \right) \mathbf{N}$$

⁵⁵That this holds can be seen by writing out the details of the matrix multiplication and then using the assumption that \mathbf{W} is row standardized so that $\sum_{j=1}^N w_{1j} = 1$.

and taking now the derivative with respect to \mathbf{S} leads to a matrix for the elasticities of steady-state income with respect to the investment rate

$$\boldsymbol{\eta}_s = \frac{\alpha + \phi}{1 - \alpha - \phi} \mathbf{I} + \left[\frac{\phi}{(1 - \alpha)(1 - \alpha - \phi)} \right] \sum_{r=1}^{\infty} \mathbf{W}^r \left(\frac{\gamma(1 - \alpha)}{1 - \alpha - \phi} \right)^r = -\boldsymbol{\eta}_n$$

where $\boldsymbol{\eta}_n$ denotes the corresponding matrix of elasticities with respect to the vector of effective depreciation rates \mathbf{N} . The elasticities given in Section 3.3 in Equations (15) and (16) for a country i then follow directly from the equation above.

E Further Robustness Checks

This appendix gathers detailed estimation results for a series of specifications mentioned in the main text. The results in Tables E.1 and E.2 demonstrate that the original results by Ertur and Koch are not robust across different versions of the Penn World Table based on the specification for the interaction matrix \mathbf{W}_2 as actually implemented in their estimation. Tables E.3 and E.4 show that the estimation results are highly sensitive to division of the geographic distances between country capitals by 1,000 in the weights of interaction matrix \mathbf{W}_2 . Finally, Tables E.5 and E.6 depict the results when the weights in the interaction matrix using genetic distances between countries are based on weighted Nei's genetic distance. In this case, the estimation results are not robust across the different samples, but comparable to the ones based on weighted F_{ST} distance in the main text with the exception mentioned in Footnote 38.

Table E.1: Estimation Results for the Standard and Spatially Augmented Solow Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_2 (Geographic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Unconstrained estimation:</i>				
Constant	0.546 (0.771)	0.214 (0.911)	3.042 (0.135)	1.139 (0.586)
$\ln s_i$	0.804 (0.000)	0.757 (0.000)	0.836 (0.000)	0.936 (0.000)
$\ln(n_i + 0.05)$	-1.471 (0.08)	-1.263 (0.030)	-0.496 (0.479)	-1.094 (0.146)
$\mathbf{W} \ln s_j$	-0.381 (0.021)	-0.370 (0.031)	-0.132 (0.530)	0.152 (0.669)
$\mathbf{W} \ln(n_j + 0.05)$	0.158 (0.840)	-0.145 (0.856)	-0.595 (0.500)	-0.623 (0.520)
$\mathbf{W} \ln y_j$	0.657 (0.000)	0.659 (0.000)	0.516 (0.000)	0.577 (0.000)
<i>Constrained estimation:</i>				
Constant	2.769 (0.000)	3.022 (0.000)	4.086 (0.000)	2.649 (0.000)
$\ln s_i - \ln n_i$	0.826 (0.000)	0.791 (0.000)	0.839 (0.000)	0.970 (0.000)
$\mathbf{W}[\ln s_j - \ln(n_j + 0.05)]$	-0.318 (0.045)	-0.232 (0.174)	-0.085 (0.679)	0.252 (0.457)
$\mathbf{W} \ln y_j$	0.665 (0.000)	0.635 (0.000)	0.510 (0.000)	0.578 (0.000)
Test of restriction	2.378 (0.305)	2.090 (0.352)	0.652 (0.722)	0.543 (0.762)
Implied α	0.323 (0.001)	0.268 (0.042)	0.143 (0.612)	-0.773 (0.695)
Implied ϕ	0.129 (0.126)	0.174 (0.138)	0.313 (0.235)	1.265 (0.519)
Implied γ	0.538 (0.000)	0.484 (0.000)	0.324 (0.026)	0.166 (0.406)
$\alpha + \frac{\phi}{1-\gamma}$	0.603 (0.000)	0.605 (0.000)	0.606 (0.000)	0.743 (0.000)

Note: p -values are given in parentheses. The restriction for the spatially augmented model is tested with the likelihood ratio (LR) test.

Table E.2: Estimation Results for the Direct, Indirect and Total Impacts in the Spatial Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_2 (Geographic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Direct impacts:</i>				
$\ln s_i$	0.847 (0.000)	0.799 (0.000)	0.897 (0.000)	1.091 (0.000)
$\ln(n_i + 0.05)$	-1.709 (0.004)	-1.583 (0.012)	-0.667 (0.332)	-1.378 (0.064)
<i>Indirect impacts:</i>				
$\mathbf{W} \ln s_j$	0.388 (0.286)	0.337 (0.386)	0.562 (0.049)	1.487 (0.037)
$\mathbf{W} \ln(n_j + 0.05)$	-2.128 (0.244)	-2.561 (0.168)	-1.588 (0.252)	-2.696 (0.110)
<i>Total impacts:</i>				
$\ln s_i + \mathbf{W} \ln s_j$	1.236 (0.004)	1.136 (0.014)	1.459 (0.000)	2.578 (0.003)
$\ln(n_i + 0.05) + \mathbf{W} \ln(n_j + 0.05)$	-3.837 (0.066)	-4.143 (0.055)	-2.255 (0.143)	-4.073 (0.029)

Note: p -values are given in parentheses. These were constructed using a set of 500,000 random draws from the estimation.

Table E.3: Estimation Results for the Spatially Augmented Solow Model According to the Specification of \mathbf{W}_2 as Implemented (Column 1) vs. as Claimed (Column 2) in Ertur and Koch (2007).

Data set	PWT 6.1	
	$e^{-2d_{ij}/1,000}$	$e^{-2d_{ij}}$
Weight in the interaction matrix \mathbf{W}_2		
Number of observations	91	91
<i>Unconstrained estimation:</i>		
Constant	0.546 (0.771)	5.140 (0.005)
$\ln s_i$	0.804 (0.000)	1.238 (0.000)
$\ln(n_i + 0.05)$	-1.471 (0.008)	-2.475 (0.000)
$\mathbf{W} \ln s_j$	-0.381 (0.021)	0.826 (0.004)
$\mathbf{W} \ln(n_j + 0.05)$	0.158 (0.840)	-1.570 (0.012)
$\mathbf{W} \ln y_j$	0.657 (0.000)	-0.236 (0.073)
<i>Constrained estimation:</i>		
Constant	2.769 (0.000)	8.884 (0.000)
$\ln s_i - \ln n_i$	0.826 (0.000)	1.093 (0.000)
$\mathbf{W}[\ln s_j - \ln(n_j + 0.05)]$	-0.318 (0.045)	2.021 (0.000)
$\mathbf{W} \ln y_j$	0.665 (0.000)	-0.236 (0.000)
Implied α	0.323 (0.001)	0.895 (0.000)
Implied ϕ	0.129 (0.126)	-0.373 (0.000)
Implied γ	0.538 (0.000)	-1.078 (0.000)
$\alpha + \frac{\phi}{1-\gamma}$	0.603 (0.000)	0.716 (0.000)

Note: p -values are given in parentheses. The likelihood ratio (LR) could not be performed for this matrix, as no value for the log-likelihood was returned in this model.

Table E.4: Estimation Results for the Direct, Indirect and Total Impacts in the Spatial Model According to the Specification of \mathbf{W}_2 as Implemented (Column 1) vs. as Claimed (Column 2) in Ertur and Koch (2007).

Data set	PWT 6.1	
	$e^{-2d_{ij}/1,000}$	$e^{-2d_{ij}}$
Weight in the interaction matrix \mathbf{W}_2		
Number of observations	91	91
<i>Direct impacts:</i>		
$\ln s_i$	0.847 (0.000)	1.210 (0.000)
$\ln(n_i + 0.05)$	-1.709 (0.004)	-2.412 (0.001)
<i>Indirect impacts:</i>		
$\mathbf{W} \ln s_j$	0.388 (0.286)	0.463 (0.033)
$\mathbf{W} \ln(n_j + 0.05)$	-2.128 (0.244)	-0.850 (0.001)
<i>Total impacts:</i>		
$\ln s_i + \mathbf{W} \ln s_j$	1.236 (0.004)	1.673 (0.000)
$\ln(n_i + 0.05) + \mathbf{W} \ln(n_j + 0.05)$	-3.837 (0.066)	-3.262 (0.000)

Note: p -values are given in parentheses. These were constructed using a set of 500,000 random draws from the estimation.

Table E.5: Estimation Results for the Spatial Durbin Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_4 (Weighted Nei's Genetic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Unconstrained estimation:</i>				
Constant	8.023 (0.003)	8.745 (0.000)	5.864 (0.013)	-2.162 (0.373)
$\ln s_i$	0.847 (0.000)	0.962 (0.000)	0.901 (0.000)	1.005 (0.000)
$\ln(n_i + 0.05)$	-1.091 (0.047)	-0.857 (0.115)	-0.137 (0.806)	-0.892 (0.172)
$\mathbf{W} \ln s_j$	0.725 (0.003)	0.728 (0.001)	0.739 (0.003)	-0.027 (0.949)
$\mathbf{W} \ln(n_j + 0.05)$	0.595 (0.550)	0.396 (0.676)	-1.747 (0.074)	-2.041 (0.065)
$\mathbf{W} \ln y_j$	0.330 (0.006)	0.285 (0.015)	0.181 (0.184)	0.557 (0.000)
<i>Constrained estimation:</i>				
Constant	5.407 (0.000)	5.750 (0.000)	6.119 (0.000)	2.325 (0.001)
$\ln s_i - \ln n_i$	0.812 (0.000)	0.898 (0.000)	0.884 (0.000)	1.031 (0.000)
$\mathbf{W}[\ln s_j - \ln(n_j + 0.05)]$	0.668 (0.005)	0.688 (0.001)	0.671 (0.004)	0.047 (0.905)
$\mathbf{W} \ln y_j$	0.308 (0.019)	0.262 (0.024)	0.232 (0.070)	0.624 (0.000)
Test of restriction	1.710 (0.425)	1.857 (0.395)	2.082 (0.353)	3.945 (0.139)
Implied α	1.855 (0.105)	1.615 (0.023)	1.529 (0.030)	-0.081 (0.913)
Implied ϕ	-1.407 (0.221)	-1.141 (0.115)	-1.060 (0.138)	0.588 (0.417)
Implied γ	-0.199 (0.294)	-0.225 (0.158)	-0.233 (0.207)	0.284 (0.188)
$\alpha + \frac{\phi}{1-\gamma}$	0.681 (0.000)	0.683 (0.000)	0.669 (0.000)	0.741 (0.000)

Note: p -values are given in parentheses. The restriction for the spatially augmented model is tested with the likelihood ratio (LR) test.

Table E.6: Estimation Results for the Direct, Indirect and Total Impacts in the Spatial Model According to Three Different Versions of the Penn World Table Based on Interaction Matrix \mathbf{W}_4 (Weighted Nei's Genetic Distance).

Data set	PWT 6.1		PWT 6.2	PWT 7.1
Number of observations	91	83	83	83
<i>Direct impacts:</i>				
$\ln s_i$	0.909 (0.000)	1.011 (0.000)	0.923 (0.000)	1.074 (0.000)
$\ln(n_i + 0.05)$	-1.071 (0.050)	-0.843 (0.118)	-0.185 (0.741)	-1.227 (0.063)
<i>Indirect impacts:</i>				
$\mathbf{W} \ln s_j$	1.454 (0.000)	1.365 (0.000)	1.080 (0.000)	1.122 (0.152)
$\mathbf{W} \ln(n_j + 0.05)$	0.390 (0.776)	0.254 (0.834)	-2.101 (0.046)	-5.462 (0.008)
<i>Total impacts:</i>				
$\ln s_i + \mathbf{W} \ln s_j$	2.363 (0.000)	2.376 (0.000)	2.001 (0.000)	2.195 (0.011)
$\ln(n_i + 0.05) + \mathbf{W} \ln(n_j + 0.05)$	-0.681 (0.629)	-0.589 (0.630)	-2.286 (0.026)	-6.689 (0.003)

Note: p -values are given in parentheses. These were constructed using a set of 500,000 random draws from the estimation.

F List of Countries Included in the Empirical Analyses

This appendix lists the countries that are included in the empirical analyses. Country codes are given as well. In the analyses with 83 countries, Angola, Bangladesh, Botswana, Central African Republic, Mauritania, Papua New Guinea, Sierra Leone, and Democratic Republic of Congo have been dropped to achieve a balanced sample over PWT Versions 6.1, 6.2, and 7.1.

Table F.7: Alphabetical List of the 91 Countries from PWT 6.1 Included in the Empirical Analyses.

Country	Code	Country	Code
Angola	AGO	Mali	MLI
Argentina	ARG	Mauritania	MRT
Australia	AUS	Mauritius	MUS
Austria	AUT	Mexico	MEX
Bangladesh	BGD	Morocco	MAR
Belgium	BEL	Mozambique	MOZ
Benin	BEN	Nepal	NPL
Bolivia	BOL	Netherlands	NLD
Botswana	BWA	New Zealand	NZL
Brazil	BRA	Nicaragua	NIC
Burkina Faso	BFA	Niger	NER
Burundi	BDI	Nigeria	NGA
Cameroon	CMR	Norway	NOR
Canada	CAN	Pakistan	PAK
Central African Republic	CAF	Panama	PAN
Chad	TCD	Papua New Guinea	PNG
Chile	CHL	Paraguay	PRY
Colombia	COL	Peru	PER
Costa Rica	CRI	Philippines	PHL
Côte d'Ivoire	CIV	Portugal	PRT
Democratic Republic of the Congo	ZAR	Republic of the Congo	COG
Denmark	DNK	Republic of Korea	KOR
Dominican Republic	DOM	Rwanda	RWA
Ecuador	ECU	Senegal	SEN
Egypt	EGY	Sierra Leone	SLE
El Salvador	SLV	Singapore	SGP
Ethiopia	ETH	South Africa	ZAF
Finland	FIN	Spain	ESP
France	FRA	Sri Lanka	LKA
Ghana	GHA	Sweden	SWE
Greece	GRC	Switzerland	CHE
Guatemala	GTM	Syria	SYR
Honduras	HND	Tanzania	TZA

Table F.7: *(Continued)*

Country	Code	Country	Code
Hong Kong	HKG	Thailand	THA
India	IND	Togo	TGO
Indonesia	IDN	Trinidad and Tobago	TTO
Ireland	IRL	Tunisia	TUN
Israel	ISR	Turkey	TUR
Italy	ITA	Uganda	UGA
Jamaica	JAM	United Kingdom	GBR
Japan	JPN	United States	USA
Jordan	JOR	Uruguay	URY
Kenya	KEN	Venezuela	VEN
Madagascar	MDG	Zambia	ZMB
Malawi	MWI	Zimbabwe	ZWE
Malaysia	MYS		

References

- AHRENS, T. J. (editor) (1995). *Global Earth Physics: A Handbook of Physical Constants*, Volume 1 of *AGU Reference Shelf Series*, American Geophysical Union, Washington.
- ANSELIN, L. (1988). *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- (2003). “Spatial Externalities, Spatial Multipliers, and Spatial Econometrics,” *International Regional Science Review*, 26(2), 153–166.
- ARROW, K. J. (1962). “The Economic Implications of Learning by Doing,” *The Review of Economic Studies*, 29, 155–173.
- ASHRAF, Q. AND O. GALOR (2013). “The “Out of Africa” Hypothesis, Human Genetic Diversity, and Comparative Economic Development,” *American Economic Review*, 103(1), 1–46.
- BRETON, T. R. (2012). “Penn World Table 7.0: Are the data flawed?” *Economics Letters*, 117, 208–210.
- CAMPBELL, D. L. AND J. H. PYUN (2015). “The Diffusion of Development: Along Genetic or Geographic Lines?” *Journal of International Development*.
- CASELLI, F. (2005). “Accounting for Cross-Country Income Differences,” in P. Aghion and S. N. Durlauf (editors), *Handbook of Economic Growth*, Volume 1A, Chapter 9, Elsevier B.V., Oxford, 679–741.
- CAVALLI-SFORZA, L. L., P. MENOZZI, AND A. PIAZZA (1994). *The History and Geography of Human Genes*, Princeton University Press, Princeton.
- CENTRAL INTELLIGENCE AGENCY (2013). *The World Factbook, 2013-14*, Central Intelligence Agency, Washington, DC.
URL: <https://www.cia.gov/library/publications/the-world-factbook/index.html>, accessed 22 Januar, 2015.
- CHENEY, W. AND D. KINCAID (2008). *Numerical Mathematics and Computing*, 6th Edition, Thomson Higher Education, Belmont.
- DAVIDSON, R. AND J. G. MACKINNON (2004). *Econometric Theory and Methods*, Oxford University Press, Oxford.
- DERBYSHIRE, J., B. GARDINER, AND S. WAIGHTS (2011). “Estimating the capital stock for the NUTS 2 regions of the EU-27,” *European Union Regional Policy Working Papers*, (1).
- DESMET, K., M. LE BRETON, I. ORTUÑO-ORTÍN, AND S. WEBER (2011). “The stability and breakup of nations: a quantitative analysis,” *Journal of Economic Growth*, 16, 183–213.

- ELHORST, P. J. (2010). “Applied Spatial Econometrics: Raising the Bar,” *Spatial Economic Analysis*, 5(1), 9–28.
- ERTUR, C. AND W. KOCH (2006). “Convergence, Human Capital and International Spillovers,” Working Paper, LEG 2006-03, Université de Bourgogne, France.
- (2007). “Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence,” *Journal of Applied Econometrics*, 22(6), 1033–1062.
- (2011). “A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions,” *Journal of Economic Growth*, 16, 215–255.
- FEENSTRA, R. C., R. INKLAAR, AND T. MARCEL (2013a). “PWT 8.0 - a user guide,” *mimeo*.
URL: <http://www.rug.nl/research/ggdc/data/pwt/pwt-research-papers>,
accessed: 18 August, 2015.
- FEENSTRA, R. C., R. INKLAAR, AND M. P. TIMMER (2013b). “The Next Generation of the Penn World Table,” .
URL: <http://www.rug.nl/research/ggdc/data/pwt/pwt-8.1>, *accessed: 18 August, 2015.*
- (2015). “The Next Generation of the Penn World Table,” *American Economic Review*, (forthcoming).
- FISCHER, M. M. (2011). “A spatial Mankiw-Romer-Weil model: theory and evidence,” *The Annals of Regional Science*, 47, 419–436.
- FISCHER, M. M. AND J. WANG (2011). *Spatial Data Analysis: Models, Methods and Techniques*, SpringerBriefs in Regional Science, Springer, Heidelberg.
- GERSCHGORIN, S. (1931). “Über die Abgrenzung der Eigenwerte einer Matrix,” *Bulletin de L’Académie des Sciences de L’URSS*, 749–754.
- GIULIANO, P., A. SPILIMBERGO, AND G. TONON (2014). “Genetic distance, transportation costs, and trade,” *Journal of Economic Geography*, 14, 179–198.
- GOLLIN, D. (2002). “Getting Income Shares Right,” *Journal of Political Economy*, 110(2), 458–474.
- HERNDON, T., M. ASH, AND R. POLLIN (2014). “Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff,” *Cambridge Journal of Economics*, 38, 257–279.
- HESTON, A., R. SUMMERS, AND B. ATEN (2002). “Penn World Table Version 6.1,” Center for International Comparison of Production, Income and Prices at the University of Pennsylvania.
- (2006). “Penn World Table Version 6.2,” Center for International Comparison of Production, Income and Prices at the University of Pennsylvania.

- (2012). “Penn World Table Version 7.1,” Center for International Comparison of Production, Income and Prices at the University of Pennsylvania.
- HOLSINGER, K. E. AND B. S. WEIR (2009). “Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} ,” *Nature Reviews Genetics*, 10, 639–650.
- INKLAAR, R. AND M. P. TIMMER (2013). “Capital, labor and TFP in PWT8.0,” *mimeo*.
URL: <http://www.rug.nl/research/ggdc/data/pwt/pwt-research-papers>,
accessed: 18 August, 2015.
- JOHNSON, S., W. LARSON, C. PAPAGEORGIOU, AND A. SUBRAMANIAN (2013). “Is newer better? Penn World Table Revisions and their impact on growth estimates,” *Journal of Monetary Economics*, 60, 255–274.
- KELLER, W. (2002). “Geographic localization of international technology diffusion,” *American Economic Review*, 92(1), 120–142.
- KIMURA, M. (1968). “Evolutionary Rate at the Molecular Level,” *Nature*, 217, 624–626.
- LE GALLO, J. (2014). “Cross-Section Spatial Regression Models,” in M. M. Fischer and P. Nijkamp (editors), *Handbook of Regional Science*, Chapter 76, Springer-Verlag, Berlin, 1511–1533.
- LEE, L. (2004). “Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models,” *Econometrica*, 72(6), 1899–1925.
- LESAGE, J. AND R. K. PACE (2009). *Introduction to Spatial Econometrics*, Chapman and Hall/CRC, Boca Raton.
- LINDNER, I. AND H. STRULIK (2014). “The great divergence: a network approach,” Discussion Papers, Center for European Governance and Economic Development Research, No. 193.
- LÓPEZ-BAZO, E., E. VAYÓ, AND M. ARTÍS (2004). “Regional Externalities and Growth: Evidence from European Regions,” *Journal of Regional Science*, 44(1), 43–73.
- MANKIW, N. G., D. ROMER, AND D. N. WEIL (1992). “A Contribution to the Empirics of Economic Growth,” *The Quarterly Journal of Economics*, 107, 407–437.
- MARSHALL, A. (1890). *Principles of Economics*, Macmillan, London.
- MASEL, J. (2011). “Genetic drift,” *Current Biology*, 21(20), R837–R838.
- MEYER, C. D. (2000). *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia.
- NEI, M. (1972). “Genetic Distance between Populations,” *The American Naturalist*, 106(949), 283–292.
- (1973). “Analysis of Gene Diversity in Subdivided Populations,” *Proceedings of the National Academy of Sciences*, 70(12), 3321–3323.

- PETERSON, J. C. AND R. D. SMITH (2012). *Introductory Technical Mathematics*, 6th Edition, Cengage Learning, Clifton Park.
- PONOMAREVA, N. AND H. KATAYAMA (2010). “Does the version of the Penn World Tables matter? An analysis of the relationship between growth and volatility,” *Canadian Journal of Economics*, 43, 152–179.
- REINHART, C. M. AND K. S. ROGOFF (2010). “Growth in a Time of Debt,” *American Economic Review: Papers & Proceedings*, 100, 573–578.
- ROMER, P. M. (1986). “Increasing Returns and Long-Run Growth,” *Journal of Political Economy*, 94, 1002–1037.
- SHEKHAR, S. AND H. XIONG (editors) (2008). *Encyclopedia of GIS*, Springer, New York.
- SOLOW, R. M. (1956). “A Contribution to the Theory of Economic Growth,” *Quarterly Journal of Economics*, 70(1), 65–94.
- SPOLAORE, E. AND R. WACZIARG (2009). “The Diffusion of Development,” *The Quarterly Journal of Economics*, 124(2), 469–529.
- (2015). “Ancestry, Language and Culture,” Working Paper 21242, National Bureau of Economic Research.
- SWAN, T. W. (1956). “Economic Growth and Capital Accumulation,” *The Economic Record*, 32(2), 334–361.
- SYDSÆTER, K., P. HAMMOND, A. SEIERSTAD, AND A. STRØM (2008). *Further Mathematics for Economic Analysis*, 2nd Edition, Prentice Hall, Harlow.
- TOBLER, W. R. (1970). “A Computer Movie Simulating Urban Growth in the Detroit Region,” *Economic Geography*, 46, 234–240.
- VINCENTY, T. (1975). “Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations,” *Survey Review*, XXII, 88–93.
- WETTERSTRAND, K. (2015). “DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP),” .
URL: <http://www.genome.gov/sequencingcosts/>, accessed: 17 August, 2015.
- WRIGHT, S. (1951). “The Genetical Structure of Populations,” *Annals of Eugenics*, 15, 323–354.

Working Paper Series in Economics

recent issues

- No. 74** *Tim Deeken*: Knowledge spillovers: On the impact of genetic distance and data revisions, November 2015

- No. 73** *Christian Feige and Karl-Martin Ehrhart*: Voting and transfer payments in a threshold public goods game, November 2015

- No. 72** *Steffen Huck, Nora Szech, Lukas M. Wenner*: More effort with less pay: On information avoidance, belief design and performance, September 2015

- No. 71** *Florian Kreuchauff and Vladimir Korzinov*: A patent search strategy based on machine learning for the emerging field of service robotics, August 2015

- No. 70** *Christian Feige*: Success rates in simplified public goods games - a theoretical model, June 2015

- No. 69** *Markus Fels*: Mental accounting, access motives, and overinsurance, May 2015

- No. 68** *Ingrid Ott and Susanne Soretz*: Green attitude and economic growth, May 2015

- No. 67** *Nikolaus Schweizer and Nora Szech*: Revenues and welfare in auctions with information release, April 2015

- No. 66** *Andranik Tangian*: Decision making in politics and economics: 6. Empirically constructing the German political spectrum, April 2015

- No. 65** *Daniel Hoang and Martin Ruckes*: The effects of disclosure policy on risk management incentives and market entry, November 2014

- No. 64** *Sebastian Gatzler, Daniel Hoang, Martin Ruckes*: Internal capital markets and diversified firms: Theory and practice, November 2014