Journal of Development Economics 101 (2013) 297–315

Contents lists available at SciVerse ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Spatial spillovers in the development of institutions☆

Harry H. Kelejian [a], Peter Murrell [a,*], Oleksandr Shepotylo [b]

[a] Department of Economics, University of Maryland, College Park, MD 20742, USA
[b] Kyiv School of Economics and Kyiv Economics Institute, Kyiv, Ukraine

ARTICLE INFO

Article history:
Received 20 August 2009
Received in revised form 11 May 2012
Accepted 19 December 2012
Available online 28 December 2012

JEL classification:
C5
O1
O5
P5

Keywords:
Institutions
Spatial econometrics
Governance
Neighborhood effects
Spatial spillovers

ABSTRACT

We examine spatial spillovers in institutional development. Dependent variables are institutional measures reflecting politics, law, and governmental administration. The explanatory variable of interest is the level of institutions in bordering countries—a spatial lag of the dependent variable. Our spatial model directly leads to the identification strategy for the endogenous spatial lag. We implement new results in spatial econometrics to counter missing-data problems usually rife in spatial empirics. Spatial institutional spillovers are statistically significant and economically important. A counter-factual exercise – the non-existence of the USSR – reveals large direct and indirect spillovers. Numerous robustness exercises bolster conclusions, including yearly cross-section regressions, fixed effects estimates, and adding many extra explanatory variables. Moreover, we provide a new theoretical result showing the robustness of estimates in the presence of omitted variables. We extend the core model, allowing different effects for better and worse neighbors, using inverse distance weights, estimating the spatial-Durbin model, and using Polity's institutional measure.

1. Introduction

This paper weaves together three rapidly developing areas of research by using spatial econometrics to examine neighborhood effects between countries thereby deriving novel empirical results on the determinants of institutional strength. Recent advances within spatial econometrics have led to more powerful tools for the estimation of neighborhood effects.[1] These advances complement the revival of interest in economic geography, which has prompted analyses of direct spillovers between countries (Easterly and Levine, 1998; Fujita et al., 1999; Persson and Tabellini, 2006; Simmons and Elkins, 2004). In turn, economic geography provides new insights into the determinants of institutions, which have been the focus of much recent attention (Acemoglu and Robinson, 2006; Beck and Laeven, 2006).

We analyze how levels of institutions in one country are affected by levels in bordering countries. To this relationship, we add fundamental determinants of institutions that are standard in the literature, such as legal origin, religion, ethnic fractionalization, natural resources, and GDP per capita. Our measures of institutions – the dependent variables – are three of the institutional indicators constructed by Kaufmann et al. (2009), one focusing on politics, one on law, and the third on governmental administration. We use a panel of countries from 1998 to 2008.

Spillovers of institutional development between countries are captured by using a spatially lagged dependent variable as an explanatory variable. In our formulation, the spatial lag is a weighted average of institutional levels in bordering countries. As a consequence, institutional development in each country reflects that of its neighbors: the spatial lag is endogenous. Moreover, there is the possibility that a random shock occurring in one country has effects that spill over to other countries. We allow for this possibility using a spatially correlated error term. Thus, we estimate our model by an instrumental variable procedure that accounts both for the endogeneity of the spatially lagged dependent variable and for the spatially correlated error term. The instruments are the predetermined variables in our model and their spatial lags. In the literature, our procedure has been labeled "generalized spatial two stage least squares" (GS2SLS) and was developed in a series of articles by Kelejian and Prucha (1998, 1999, 2004). It is a particular form of GMM. We also implement a recent result of Kelejian and Prucha (2010) that facilitates the use of larger samples – 20% larger in this paper – in the face of the very demanding data requirements of spatial models.

Our core results show that the level of institutions in a country's immediate neighbors has a quantitatively important impact on the country's institutions. For example, the direct spatial effect of having

☆ We would like to thank Janet Hao, Nona Karalashvili, Martin Schmidt, and Jing Zhang for research assistance, and Donald Clarke, Peter Grajzl, Aart Kraay, Nuno Limao, Ingmar Prucha, two anonymous referees, and the associate editor for helpful comments.
* Corresponding author.
E-mail addresses: kelejian@econ.umd.edu (H.H. Kelejian), murrell@econ.umd.edu (P. Murrell), oshepotylo@kse.org.ua (O. Shepotylo).
[1] See for example Kelejian and Prucha (1998, 1999, 2004), Lee (2003, 2004), and LeSage and Pace (2009), and references therein.

0304-3878/$ – see front matter
http://dx.doi.org/10.1016/j.jdeveco.2012.12.003

neighbors with Romanian-level rule-of-law rather than the Swiss-level is equivalent in magnitude to the effect of the country having had a socialist legal heritage rather than the common law. The direct impact is statistically significant for all three measures of institutions.

A large number of robustness tests all bolster the conclusions derived from the core results. One reason why we conduct extensive robustness tests is that there is a common belief that estimates of spatial effects are especially prone to omitted-variable bias. This belief seems to arise because many variables affecting institutional levels are themselves spatially correlated, raising the suspicion that if one of these variables is unobserved the spatial lag of the dependent variable simply proxies this missing variable. In a new theoretical result, we show that this line of reasoning exaggerates the possibility of bias. For our IV procedure, the consistency of the estimate of the spatial lag coefficient is unaffected by the omission of a wide class of spatially-correlated explanatory variables. Thus, contrary to common intuition, our estimates of spatial spillovers are less prone to omitted-variable bias than are estimates of the coefficients of standard explanatory variables.

Notwithstanding this reassuring result, the paper contains many estimates examining the sensitivity of our core results to changes in specification, explanatory variables, sample, and model estimated. The size and significance of the spatial effect are preserved in cross-sectional yearly regressions, in fixed-effect analyses, when we either treat GDP per capita as endogenous or omit it, and when we add a large number of explanatory variables. Indeed, the size and significance of the spatial effect are usually larger in the robustness exercises than in our core estimates.

We extend the horizons of our core model by estimating alternative versions. One alternative allows better neighboring institutions to have a different effect than worse ones. Better neighbors have the stronger effect, but both types of neighbors exhibit significant spatial effects. Another variation – the spatial-Durbin model – adds spatial lags of explanatory variables as additional explanatory variables. The core results are preserved. We investigate a different form of spatial lag, with the strength of spatial spillovers proportional to the inverse of the distance between countries on the same continent. Similar results are obtained. Finally, we use a different dependent variable – from the Polity IV Project (2011) – which also lengthens the time period under consideration. The spatial spillover effect is stronger than in our core results, both in terms of economic size and statistical significance.

The spatial model implies that fundamental determinants of institutions in one country indirectly affect institutions in all countries. To provide an example, we simulate what would have happened to levels of institutions had the USSR never existed. General equilibrium effects, those taking into account spatial spillovers, are substantially larger than the direct effect which does not include such spillovers. In the case of the rule of law, total effects are 44% higher than direct effects. The general equilibrium effect means that the Soviet system had significant consequences for countries that were never part of the Soviet bloc, or even bordered the Soviet bloc.

Our results also bear on the effects of previously studied determinants of institutions. For example, while we find that a country's legal origin and the resource curse have their expected effects, ethnic fractionalization appears to be unimportant. Human capital and Muslim prevalence have a much stronger effect on political institutions than on legal or governmental ones.

We proceed as follows. Section 2 reviews elements of the literature that strongly suggest that spatial effects are important determinants of institutions. Section 3 provides an overview of the pertinent spatial econometrics and describes the method used to calculate the general equilibrium effects of changes in fundamentals in one country on institutions everywhere. Section 4 describes variables and data. Section 5 presents estimates of our core spatial relationship, documenting the strength of the direct spatial effect. Section 6 shows the effect on institutions had the Soviet system never existed. Section 7 contains standard robustness tests—yearly regressions, fixed effects regressions, and the addition of other explanatory variables. Section 8 expands the horizons of our model, allowing for asymmetric effects of better and worse institutions, introducing the spatial-Durbin model, and using the alternative institutional (Polity IV) variable over an expanded time period. Appendixes A and B provide new theoretical results, establishing the validity of new methods to preserve sample size in the presence of missing data and showing that estimates of the spatial effect are relatively robust in the presence of omitted variables. Conclusions appear in Section 9.

2. Spatial processes in institutional change

The existing literature justifies an empirical investigation of neighborhood effects in institutional development. We show this in two steps, first examining how institutions diffuse between countries and how this diffusion is affected by distance and then reviewing examples of existing empirical and theoretical work that contain the seeds of a spatial approach to institutional development.

2.1. Institutional diffusion and the costs of distance

Institutions spread from country to country by a variety of mechanisms. Governments and citizens learn from each other. Imperialist nations impose their arrangements. Countries compete to provide trade- and investment-friendly institutions. Foreign economic agents spread institutional knowledge and create a demand for the institutions with which they are familiar.

For many reasons, institutional diffusion is subject to a cost of distance, suggesting that it occurs more often and with greater strength between bordering countries.[2] *Ceteris paribus*, it is easier to conquer a neighbor than a far-distant nation. Civil chaos is more likely to spill over a border than to cross continents. Economic agents from neighboring countries are more numerous than other foreigners, having a greater effect on the diffusion of institutional knowledge and more sway on institutional choice. Governments of bordering nations compete to provide an institutional environment that fosters trade and investment, with institutions in neighboring countries providing a salient target to be emulated.

Moreover, knowledge spreads more easily between individuals with similar cultural, historical, and economic backgrounds, and bordering countries are more likely than distant ones to share a common history, culture, and language. Hence transaction costs in the exchange of institutional knowledge increase with distance, meaning that attempts to spread and copy institutions occur more readily between neighbors. In a similar vein, Crafts and Venables (2001, p. 33) argue that proximity is especially important where complexities of information require skilled labor, as in institution building.

2.2. Spatial processes in institutional diffusion

As early as the eleventh century the activities of professional merchants led to the growth of a customary cross-border commercial law, the *lex mercatoria*, which was gradually absorbed into formal law.[3] Similar processes abound in modern times with changes in commercial laws spurred by private initiatives to harmonize institutions between countries (Rosett, 1992, p. 683).[4] In pursuit of regional markets,

[2] See Crafts and Venables (2001) for a broad-ranging argument that the cost of distance is a fundamental force in world development and trade. Our emphasis on the cost of distance follows that approach closely.

[3] North and Thomas (1970) provide a similar example describing the early growth of commerce in Western Europe and the institutional innovation that it precipitated, in which ideas were spread by merchants crossing borders, with countries competing on institutions.

[4] Casella and Feinstein (2002) examine this process theoretically and find that increasing exchange will lead to harmonization of institutions for smaller countries, but not necessarily for large ones.

multinationals participate in policy formation, seeking to make local law conform to common principles (Leebron, 1996). Governments compete for foreign investment by improving institutions (Qian and Roland, 1998). Institutional diffusion also takes place outside government and the legal system through international coordination of standards by private organizations (Casella, 1996). Given the cost of distance in knowledge transfer, all such effects are larger between neighbors.[5]

Institution builders learn from each other. In the data of Berkowitz et al. (2003), institutional transplants between geographically close countries are more likely to be receptive than transplants between distant lands. Mukand and Rodrik (2005) model the institutional learning decision, with countries choosing between experimentation and imitation. Because neighbors share similar characteristics, countries closer to a successful one choose imitation, resulting in better institutions.[6]

The decision to copy the institutions of neighbors arises from a number of sources. Informational cascades (Bikhchandani et al., 1992) lead to imitation, with the most likely models provided by countries with similar cultural and economic characteristics, most often neighbors (Murrell et al., 1996). Simmons and Elkins (2004) argue that politicians reap reputational payoffs by conforming to policies recognized as effective outside their own country.

A more formal route for the copying of institutions is via negotiations and bargains between governments. For example, countries with more stringent environmental or labor standards attempt to change the standards of their trading partners (Bhagwati and Hudec, 1996). EU accession requires acceptance of laws and institutions similar to those in existing member countries. Even before accession, institutional change is a prerequisite for participation in preferential trade agreements, which were particularly important for the EC/EU's neighbors in the 1990s (Grilli, 1997; Winters, 1993).

Imperialism, war, revolution, and civil strife all have a spatial component that affects institutional diffusion. The Napoleonic code was first spread via war and occupation. The imposition of Soviet-style institutions on Eastern Europe was a direct result of the geography of World War II. The revolutions of 1989 against the Soviet system had a clear geographical component. Persson and Tabellini (2006) make a direct case that democracy in one country bolsters democracy in its neighbors.

In sum, the existing literature provides a secure basis for an empirical investigation of spatial diffusion of institutions. Related empirical studies indirectly indicate the same. Easterly and Levine (1998) find that there is a strong relationship between growth in neighboring African countries, suggesting that the copying of policies might be partially responsible for this relationship. Bosker and Garretsen (2009) show that levels of institutions in neighboring countries increase a country's level of development. Simmons and Elkins (2004) find that switches between policy regimes can be explained by policy choices in similar countries.

When specifying the formal model, there is a choice to be made on how to characterize spatial effects. Between which countries do spillovers occur? The above argues that spillovers are strongest between bordering countries. Pragmatically therefore, we adopt a simple characterization of spatial effects, assuming that institutions in a given country are directly affected by institutions only in bordering countries. To examine the robustness of our results to this assumption, Section 8 considers an alternative, with the strength of spillovers inversely proportional to the distance between countries.

3. The spatial econometric model of institutional diffusion: an overview

The purpose of this section is three-fold. First, we specify a generic spatial model and interpret it in terms of the above discussion.[7] In this interpretation, a given country's institutional level (the dependent variable) is directly related to the fundamentals of that country (e.g. religion) and to the institutional levels of bordering countries through the spatial lag of the dependent variable.[8] Since this relationship holds for all countries, it implies via a reduced form that each country's institutional level is related to the fundamentals of all countries.[9] The power of the spatial specification is that it captures these multitudinous interactions in a parsimonious way: the hypothesis that there are no spillovers between countries is easily tested.

Second, we describe our estimation procedure including the specification of the instruments that are used for the endogenous spatially lagged dependent variable. Our procedure is one that is now in widespread use in spatial estimation, namely "generalized spatial two stage least squares." It is a spatial form of a GMM estimator and was developed by Kelejian and Prucha (1998, 1999, 2004). Its underlying basis is intuitive as it is built on standard instrumental variable and Cochrane–Orcutt transformation methods. It does not depend upon distributional assumptions, and its asymptotic properties are well documented. It is relatively simple to implement.[10] Instrument selection follows directly from the spatial structure of the model. Its flexibility is seen in Section 9, where we estimate a variety of versions of the basic spatial model, with the intuition of the estimation procedure for each version exactly matching that of our core model. Moreover, the procedure is easily adapted to implement a recent result of Kelejian and Prucha (2010) that facilitates the use of larger samples, an important property in the face of the demanding data requirements of spatial models (see Appendix A). Finally, in Appendix B we show that consistency of the estimate of the spatial lag coefficient is unaffected by the incorrect omission of a wide class of spatially-correlated explanatory variables.

Third, we illustrate how the effects of a change in one or more of the fundamentals in a given country emanate to other countries and show how these emanating effects are calculated.[11]

3.1. Model specifications

The model that we estimate reflects two different spatial processes, direct spatial spillovers between institutional levels and spillovers between idiosyncratic features of the environment that affect institutions. We are primarily interested in the former (captured in the parameter λ below), which reflects the types of diffusion of institutional decisions between countries that was discussed above. However, inclusion of the latter (captured in the parameter ρ below) is justified on an economic basis, because neighboring countries do share idiosyncratic characteristics, and as a consequence their inclusion in the model is necessary for consistent estimation of the standard errors of other parameters.

Our model (to be applied separately to each of the three institutional measures) is, for time t:

$$\begin{aligned} y_t &= X_t\beta_1 + H\beta_2 + \lambda W y_t + u_t \\ u_t &= \rho W u_t + \varepsilon_t, \quad t = 1, \ldots, T \end{aligned} \tag{1}$$

where y_t is an $n \times 1$ vector of observations on institutional levels for n countries. X_t is a $n \times k_1$ matrix of observations on k_1 fundamentals

[5] On the political side, Eichengreen and Leblang, 2006; Spilimbergo, 2006 provide examples of interchange spurring democratization.

[6] See Grajzl and Dimitrova-Grajzl (2009) for a similar finding in the context of lawmaking.

[7] There is a large literature on spatial models. Classic references are Cliff and Ord (1973, 1981), Anselin (1988) and Cressie (1993). Some important theoretical contributions since 1993 are Conley (1996, 1999), Kelejian and Prucha (1998, 1999, 2001, 2004), Lee (2003, 2004), Pinske et al. (2002), Baltagi et al. (2003) and LeSage and Pace (2009).

[8] This term is widely used in the literature, and was originally introduced by Anselin (1988). Its specification will become clear below.

[9] The term "fundamentals" is used purely for convenience. In fact, spatial variables are as fundamental as others.

[10] Since the first version of this paper was written, it has been implemented in STATA.

[11] The first published reference to these effects was by Kelejian et al. (2006). In that paper the effects were called emanating effects. Since that time, others have calculated these effects and have used other terms, such as "impact measures." Two interesting studies along these lines are LeSage and Pace (2009) and LeSage and Fischer (2008).

whose values vary over time; H is an $n\times k_2$ matrix of observations on k_2 fundamentals whose values do not vary over time; W is an $n\times n$ weighting matrix, described below; u_t is the $n\times 1$ disturbance vector; and ε_t is the corresponding innovation vector, which determines the disturbance vector as indicated by the second equation in Eq. (1). β_1 and β_2 are, respectively, $k_1\times 1$ and $k_2\times 1$ parameter vectors and λ and ρ are scalar parameters. The term Wy_t in Eq. (1) is the spatial lag of the dependent variable. Note that since y_t depends directly on the disturbance vector u_t, Wy_t also depends on u_t and so must be viewed as endogenous in the estimation of Eq. (1).[12]

For simplicity of presentation in this section, we ignore estimation problems associated with missing observations. As is evident from Eq. (1), the calculation of the spatial lag for each country requires observations on the dependent variable for all of its neighbors. Some of these observations are not available. To address this problem, we implement the method suggested by Kelejian and Prucha (2010), described briefly in Appendix A. Use of this methodology increases the sample size by 20%.

Consistent with the argument of Section 2, we define the weighting matrix W so that the value of the dependent variable for each country depends directly on an average of that variable for bordering countries. Suppose the i^{th} country has ϕ_i countries that border it. Then, the i^{th} row of W has zeroes everywhere except in ϕ_i positions corresponding to ϕ_i neighbors. In these positions, the values in the i^{th} row are $1/\phi_i$. Such a weighting matrix is row normalized: the elements of each and every row sum to unity. The diagonal elements of W are zero—no country is viewed as its own neighbor.

Assume that $|\lambda|<1$ and $|\rho|<1$ and let $G=(I-\lambda W)^{-1}$. Then the reduced form solution of Eq. (1) for y_t is:[13]

$$\begin{aligned} y_t &= GX_t\beta_1 + GH\beta_2 + Gu_t \\ u_t &= (I-\rho W)^{-1}\varepsilon_t \end{aligned} \tag{2}$$

Clearly Eq. (2) implies that the value of the dependent variable for each country depends in general upon all of the fundamentals of each and every country, as well as on the innovation shocks in all of the countries. If $\lambda=0$ there are no spillovers in institutional levels between countries: a test for the absence of spatial spillovers is therefore straightforward.

3.2. Disturbance terms

Assume the innovation vector ε_t is independently and identically distributed (i.i.d.) over time and the elements of ε_t are i.i.d. over countries with mean and variance $(0, \sigma^2)$.[14] Given these and previous assumptions, Eq. (2) implies that the variance–covariance matrix of the disturbance vector u_t, say VC_u, is

$$VC_u = \sigma^2(I-\rho W)^{-1}(I-\rho W')^{-1} \tag{3}$$

VC_u is not diagonal: the elements of the disturbance vector u_t are spatially correlated.

3.3. The estimation procedure: An overview

Because the spatial lag in Eq. (1) is endogenous, an IV procedure is natural. In this section, we focus on formalities and show how the specification of Eq. (1) dictates which instruments should be used. The estimating procedure also accounts for the spatially correlated error terms.[15]

There are three steps in estimation. First, the regression parameters in Eq. (1) are estimated by an IV procedure that does not account for the spatial correlation of the disturbances but does account for the endogeneity of the spatial lag. Second, these estimated parameters are used to estimate the disturbances which in turn are used to estimate the autoregressive parameter in the disturbance process (ρ) using a GMM procedure (Kelejian and Prucha, 1999). In the third step, the estimate of ρ is used to transform the model in a spatial version of a Cochrane–Orcutt procedure. This transformed model is then estimated by an IV procedure using the same instruments.

3.4. The instruments

Using the above assumptions,

$$(I-\lambda W)^{-1} = I + \lambda W + \lambda^2 W^2 + \ldots \tag{4}$$

This implies that the mean of the dependent variable vector is

$$\begin{aligned} E[y_t] &= (I-\lambda W)^{-1}[X_t\beta_1 + H\beta_2] \\ &= \left(I + \lambda W + \lambda^2 W^2 + \ldots\right)[X_t\beta_1 + H\beta_2] \end{aligned} \tag{5}$$

and so the mean of the spatial lag, Wy_t, is

$$E[Wy_t] = W\left(I + \lambda W + \lambda^2 W^2 + \ldots\right)[X_t\beta_1 + H\beta_2] \tag{6}$$

The model in Eq. (1), and the expression in Eq. (6), suggest the instruments X_t, H, and products of these matrices with powers of the weighting matrix (Kelejian and Prucha, 1998). It is standard practice in the literature to use the instruments X_t, H, WX_t, and WH to estimate models similar to our Eq. (1). Das et al. (2003) give results which suggest that instruments based on such a linear truncation of the mean of Wy_t are almost as efficient as the maximum likelihood estimator, which is based on the normality assumption. Our IV estimator does not require normality.

3.5. The calculations of emanating effects: A special case

If one of the fundamentals in (X_t,H) changes in a given country, the calculation of its direct effect is straightforward in terms of the first part of Eq. (1). These direct effects are then transmitted to other countries via the spatial lag and ultimately feed back to the given country, leading to indirect effects whose calculation rests on the solution of the model as given in Eq. (2). Clearly the implication is that if $\lambda\neq 0$ the levels of institutions in every country depend upon the fundamentals of all countries.[16]

As an illustration, suppose there is a change in country 1 in the fundamental that corresponds to the first column of X_t, that is, the value of the first regressor at time t for country 1. Simplifying notation, denote the value of that first fundamental at time t for country 1 as $x_{t,1}$. Also, denote the J^{th} element of y_t as $y_t^{(J)}$. Then the change in $y_t^{(J)}$ with respect to $x_{t,1}$ is

$$\frac{\partial E\left[y_t^{(J)}\right]}{\partial x_{t,1}} = b_{1,1}G_{J1}, J = 2, \ldots, N \tag{6}$$

where $b_{1,1}$ is the first element of β_1 and G_{J1} is the (J,1)$^{\text{th}}$ element of G. In the literature, these changes are termed emanating effects because

[12] The model and methods we use follow directly from the processes examined in Section 2. Of course, other approaches are possible. For example, Graham (2008) presents a novel framework for testing for spatial processes.

[13] Gesgorin's Theorem implies that the characteristic roots of a row normalized weighting matrix are less than or equal to 1 in absolute value (Horn and Johnson, 1985, pp. 344–345). Therefore, for $|\lambda|<1$ the roots of λW are less than 1 in absolute value and hence $(I-\lambda W)$ is nonsingular—see Kelejian and Prucha (1998).

[14] A complete set of formal assumptions is given in Kelejian and Prucha (2010).

[15] This procedure was developed by Kelejian and Prucha (1998, 1999), who do not use a panel framework, but extension to a panel framework is straightforward.

[16] These effects are rarely estimated by spatial modelers. For exceptions, see Kelejian et al. (2006); LeSage and Fischer (2008), LeSage and Pace (2009), and Ward and Gleditsch (2008).

they describe how a fundamental in one country affects the dependent variables of all of the other countries (Kelejian et al., 2006).[17] These emanating effects are defined with respect to the mean of the dependent variable conditional on the exogenous variables, which therefore implies that error terms or parameters such as ρ in Eq. (1) are not involved in emanating effects.

3.6. Calculations of Emanating Effects: The General Case

More generally, one can calculate emanating effects due to a change in a set of variables in one or more countries. For example, ignoring the disturbance term, let $(GX_t\beta_1+GH\beta_2)|_\Delta$ be the $n\times1$ vector of new values of the reduced form expression Eq. (2) that corresponds to the hypothetical new values of (X_t,H). Let $(GX_t\beta_1+GH\beta_2)$ be the vector of values of that reduced form corresponding to the initial values of (X_t,H). The resulting vector of emanating effects on all n countries due to the hypothetical change in the values of the set of fundamentals is therefore

$$\left(GX_{t,1}\beta_1+GH_1\beta_2\right)\Big|_\Delta-\left(GX_{t,1}\beta_1+GH_1\beta_2\right) \tag{7}$$

The expression in Eq. (7) includes the feedback effects of spillovers on the countries in which the original changes in the fundamentals took place. We implement this form of emanating effects in Section 6.

4. The variables and the data

As data for institutions, we use the well-known measures produced by Kaufmann et al. (2009). The data cover 1998–2008.[18] We focus on three of the six measures. “Voice and Accountability” captures whether citizens can participate in selecting their government, as well as freedom of expression, association and the media . “Rule of law” measures whether citizens have confidence in and abide by the rules of society, particularly focusing on contract enforcement, the police, and the courts. “Government effectiveness” measures the quality of public services, the civil service, and policy formulation and implementation.

Kaufmann et al. (2009) have data on six indicators in total. We chose those three indicators that best represented political, legal, and administrative institutions, the three distinct areas most emphasized in discussions of economic development. Political stability, regulatory quality, and control of corruption are the three that we omit from this study. The results for these three indicators are broadly similar to those we present here, in terms of qualitative economic effects, but somewhat weaker from the perspective of statistical significance.[19]

In the core results of Section 5, we use for X_t and H variables that have already come into common usage in the literature on institutions, reviewing each briefly in the paragraphs below. But we also undertake extensive robustness tests in Section 7, adding other elements to X_t and H. We describe the variables used in these robustness tests when they are introduced. Details on definitions, years of coverage, and sources appear in Table 1. Summary statistics are in Table 2.

La Porta et al. (1998), plus many subsequent empirical studies, argue that the origin of a country's legal system has a profound effect on institutional performance. Generally, a distinction is made between five legal origins, English, French, German, Scandinavian, and socialist. In our model, legal origin is captured by dummy variables, with the omitted dummy corresponding to English legal origin. Socialist legal origin proxies more generally for a heritage of Soviet-type institutions.

Culture attracts increasing attention in the literature (Guiso et al., 2006). Recent studies often use religious affiliation either as a cultural measure or as an exogenous determinant of more ephemeral aspects of culture. We therefore use variables capturing the proportion of the population belonging to three major religions, Protestantism, Catholicism, and Islam. For obvious reasons, the proportion of the population not subscribing to any of these three is omitted.

Ethnic fractionalization is usually included in regressions aiming to explain institutional levels (Alesina et al., 2003). Ethnic tensions lead to a focus on redistribution, rather than efficiency, and make it more difficult to settle on effective institutional arrangements. We use the probability that two members of the population belong to the same ethnic group, where ethnicity reflects a combination of racial and linguistic attributes (Alesina et al., 2003).

The resource curse appears when the dominance of natural resources in an economic activity focuses politics on the struggle for economic rents. When the gains from capturing those rents are high enough, unproductive institutional arrangements arise (Acemoglu and Robinson, 2006; North, 1990). Our resource-curse variable is the log of proven oil reserves at the end of 1994 divided by 1994 GDP. We comment below on possible problems of endogeneity arising from the use of GDP.

Spatial effects have been labeled geography of the second nature, that is advantages deriving from location (Krugman, 1993). Geography of the first nature can also be important, that is, factors intrinsic in a country, independent of its neighbors. Two variables are popular in the institutional literature in this respect. Absolute latitude is used for a variety of reasons, for example, because colonizers who brought their own institutions had a preference for certain types of climate and topography (Hall and Jones, 1999; Rodrik et al., 2004). Acemoglu et al. (2001) suggest that latitude might function as a proxy for early colonial experience. Hence, our use of latitude compensates for the necessary omission of settler mortality (Acemoglu et al., 2001) when using a sample that includes countries that are not ex-colonies. A second geographical variable is a dummy for whether a country is landlocked (Gallup et al., 1999). Being landlocked reduces the flow of labor that could bring institutional knowledge and makes countries more susceptible to destructive military pressure from aggressive neighbors.

Glaeser et al. (2004) argue that human capital accumulation precedes the generation of productive political institutions. Alesina and Perotti (1996) find that higher levels of human capital are associated with higher levels of political stability while Chong and Zanforlin (2000) suggest a link between institutions and human capital. We measure human capital by the percentage of the population completing tertiary education.

High levels of income provide resources to build institutions (Keefer and Knack, 1996; La Porta et al., 1999).[20] Therefore, we include the log of GDP per capita as an explanatory variable. We assume that longer-term trends are important, rather than year-to-year fluctuations, and therefore measure GDP per capita as the natural log of the mean of annual GDP per capita from 1991 to 1995, the years before the measurement of institutions. Of course, the use of GDP brings concerns about endogeneity, even if lagged. Thus, in Section 7, we probe these concerns, finding that the results relating to our variable of interest are unaffected if we either omit GDP or instrument it.

Lastly, we include the log of population. On the one hand, larger entities might find it more difficult to reach workable institutional arrangements because of the logic of collective action. On the other, institutional public goods become cheaper as population grows.

5. Core estimates of direct spatial effects

Table 3 presents estimates of Eq. (1), using each of the three institutional measures separately as dependent variables. Our identification strategy follows the logic of Section 3: we use as instruments

[17] Ward and Gleditsch (2008) refer to these as equilibrium effects and provide examples of their estimation.

[18] Kaufmann et al. (2009) also include data for 1996, but we exclude that year because country coverage is worse than subsequently.

[19] Results on the three omitted indicators are available from the authors on request.

[20] But see Kaufmann and Kraay (2002) and Acemoglu et al. (2008) for some counter-evidence.

Table 1
Definitions of variables and sources of data.

Variable	Description	Source
Voice and accountability	Measures whether citizens can participate in selecting their government, as well as freedom of expression, association and the media. 1998, 2000, 2002, 2003, 2004, 2005, 2006, 2007, 2008	Kaufmann et al. (2009)
Government effectiveness	Measures the quality of public services, the civil service, and policy formulation and implementation. 1998, 2000, 2002, 2003, 2004, 2005, 2006, 2007, 2008	Kaufmann et al. (2009)
Rule of law	Measures whether citizens have confidence in and abide by the rules of society, particularly focusing on contract enforcement, the police, and the courts 1998, 2000, 2002, 2003, 2004, 2005, 2006, 2007, 2008	Kaufmann et al. (2009)
Polity	Combined Polity Score: The POLITY score is computed by subtracting the AUTOC score from the DEMOC score; the resulting unified polity scale ranges from +10 (strongly democratic) to 10 (strongly autocratic). The index is normalized to take values between -1 and 1.	Polity IV Project (2011)
Common border (used in weighting matrix)	1 if countries have common border, 0 otherwise	CEPII (2006)
Legal origin	Dummy variables identifying the legal origin of the company law or commercial law of each country. There are five possible origins: (1) English Common Law; (2) French Commercial Code; (3) German Commercial Code; (4) Scandinavian Commercial Code; (5) socialist/communist Laws	La Porta et al. (1999)
Religious affiliation	The proportion of the population of each country that belonged to the three most widely spread world religions in 1980. The three religions identified are: (1) Roman Catholic; (2) Protestant; (3) Islam.	La Porta et al. (1999)
Ethnic fractionalization	Index of ethnic fractionalization: the probability that two randomly chosen citizens will be from different ethnic groups. Various years 1981–2001.	Alesina et al. (2003)
Oil reserves to GDP ratio	Natural log of (0.1 + proven reserves of oil)/GDP. Reserves are measured in thousand million barrels at end of 1994. GDP is 1994 PPP GDP measured in 2000 US$.	BP (2006)
Absolute latitude	Absolute value of latitude of a country scaled from 0 to 1.	La Porta et al. (1999)
Landlocked	Dummy variable equals 1 if the country is landlocked	CEPII (2006)
Human capital	Percent of population 25 and over that has completed tertiary education in 1990	Barro and Lee (2001)
Log GDP per capita (Tables 3–14)	Natural log of mean PPP GDP per capita 1991–1995, measured in 2000 US$.	World Bank (2005)
Log GDP per capita (Table 15)	Natural log of mean PPP GDP per capita in 2005 US$, lagged 5 years relative to institutional measure.	Heston et al. (2011)
Log population	Log of population for the same years as the measures of institutions.	United Nations (2007)
Continents	Dummy variables identifying continents. There are five variables: Africa, Americas, Asia, Pacific, and Europe.	CEPII (2006)
Trade openness	Natural log of (1 + (exports + imports)/GDP)	Heston et al. (2009)
Colonizers	Dummy variables identifying the colonial power. There are six variables: Colonizer Britain, Colonizer France, Colonizer Spain, Colonizer Russia, Colonizer Other, and Not Colonized.	CEPII (2006)
Coastal access	Proportion of population within 100 km of an ice-free coast	Gallup et al. (2001)
Tropical land	Percent of land area within geographical tropics	Gallup et al. (2001)
Malaria prevalence	Percent of 1995 population living in areas with malaria in 1982	Gallup et al. (2001)
Log of land area	Natural log of land area	Gallup et al. (2001)
Buddhist proportion	Proportion of population that is Buddhist	Barro (2006)
Confucian proportion	Proportion of population that is Confucian	Barro (2006)

the spatial lags of the predetermined variables and, of course, the predetermined variables themselves.

The coefficients of the spatially lagged dependent variables are significant, statistically and economically. To understand the economic size of the direct spatial effect, focus on the rule-of-law equation. Suppose that countries A and B are alike in all respects except that A is surrounded by countries at Romanian levels of rule-of-law in 2008 while B is surrounded by countries at Swiss 2008 levels, Romania versus Switzerland providing a striking within-Europe contrast. Then the difference between the rule-of-law in A and B would be roughly the same as that between Egypt and Poland or between Cyprus and the UK. The effect of having Romanian-type rather than Swiss-type neighbors is roughly three times the effect of having French legal origin rather than English legal origin. It is nearly as big as the effect of having socialist legal origin rather than English legal origin and it is 50% bigger than the effect of moving from the mean level of oil dependence to the maximum level of oil dependence. Spatial spillovers in the form of diffusion of institutional decisions are indeed profound.

Table 3 presents overidentification tests for the instruments in the form of the values of a Hansen *J*-statistic.[21] Although such tests can never be used definitively to rule out endogeneity, it is notable that none of these statistics raise concerns about biases due to endogeneity.

One particular form of such concerns about biases arises from noting that neighboring countries often share 'common factors' such as cultures, aspects of which are unmeasured. Thus, it is tempting to suppose that the estimated spillover effect between institutions is spurious, reflecting the spatial correlation between the unmeasured variables. In Section 7 we take these 'omitted common factors' concerns very seriously, providing a theoretical result and a battery of robustness tests. The results in Section 7 all serve to bolster the validity of those in Table 3, supporting our conclusion that spatial interactions between institutional levels are statistically significant and economically meaningful.

Although the results on fundamentals are only of secondary interest here, some are noteworthy. Consistency with existing studies appears in a number of ways. Socialist legal origin is statistically significant with a large negative effect. The resource curse is always statistically significant, and economically large. As in many other studies, absolute latitude is significant, perhaps proxying for varieties of colonial experience.

Differences between the results for the three institutional indicators also provide interesting observations. French legal origin has significant effects on law and government, but not on politics. The resource variable has a larger effect on politics and government than law, suggesting that the resource curse works through the destructive politics of rent-seeking. Consistent with previous results (Mobarak, 2005), the strongest result for religion, statistically and economically, is the effect of the prevalence of Muslims on political institutions.

Some variables found to be significant in previous studies are not significant in ours, for example, ethnic fractionalization and the landlocked variable. The proportion of the population subscribing to Protestantism is never significantly different from the proportion of the population not subscribing to the three major religions. One conjecture therefore is that omission of the spatial effect in previous studies might have led to omitted variable bias. Indeed, ethnic fractionalization has a strong negative correlation with the spatial lag of institutions, which might account

[21] Although informative, it should be noted that the Hansen *J*-test was not developed in the context of a spatial model containing spatial lags in both the dependent variable and the error term.

Table 2
Summary statistics.

	N	Mean	Standard deviation	Min	Max
Institutional measures					
Voice and accountability (VA)	1161	0.063	0.966	−2.112	1.826
Government effectiveness (GE)	1161	0.129	1.006	−1.893	2.531
Rule of law (RL)	1161	0.025	1.000	−2.110	2.116
Baseline explanatory variables					
Spatial lag of VA, using contiguity weights	1161	−0.127	0.804	−2.353	1.769
Spatial lag of GE, using contiguity weights	1161	−0.054	0.797	−2.121	2.167
Spatial lag of RL, using contiguity weights	1161	−0.186	0.771	−1.872	1.979
English legal origin	129	0.318	0.467	0	1
French legal origin	129	0.395	0.491	0	1
Socialist legal origin	129	0.209	0.408	0	1
German legal origin	129	0.039	0.194	0	1
Scandinavian legal origin	129	0.039	0.194	0	1
Catholic proportion	129	0.330	0.364	0	0.973
Muslim proportion	129	0.203	0.341	0	0.994
Protestant proportion	129	0.138	0.224	0	0.978
Other religion proportion	129	0.329	0.312	0.003	1
Ethnic fractionalization	129	0.420	0.247	0.002	0.930
Log of oil resources to GDP ratio	129	−5.029	1.980	−10.23	0.768
Absolute latitude	129	0.317	0.195	0	0.722
Landlocked	129	0.217	0.414	0	1
Log GDP per capita	129	8.486	1.065	6.260	10.292
Human capital	129	5.950	4.611	0.100	27.300
Log population	1161	16.098	1.695	10.600	21.004
Additional explanatory variables for robustness tests					
Colonizer Britain	129	0.326	0.470	0	1
Colonizer France	129	0.109	0.312	0	1
Colonizer Spain	129	0.132	0.340	0	1
Colonizer Russia	129	0.109	0.312	0	1
Colonizer other	129	0.147	0.356	0	1
Not colonized	129	0.178	0.384	0	1
Trade openness	1024	0.592	0.248	−0.218	1.717
Coastal access	121	0.423	0.368	0	1
Tropical land	121	0.426	0.473	0	1
Malaria prevalence	123	0.353	0.439	0	1
Log of land area	121	12.292	1.680	8.488	16.623
Buddhist proportion	129	0.025	0.116	0	0.853
Confucian proportion	129	0.016	0.058	0	0.447
Africa	129	0.240	0.429	0	1
Americas	129	0.217	0.414	0	1
Asia	129	0.233	0.424	0	1
Europe	129	0.279	0.450	0	1
Pacific	129	0.031	0.174	0	1
Spatial lag of English legal origin	129	0.190	0.336	0	1
Spatial lag of French legal origin	129	0.385	0.417	0	1
Spatial lag of Socialist legal origin	129	0.210	0.357	0	1
Spatial lag of German legal origin	129	0.039	0.136	0	1
Spatial lag of Scandinavian legal origin	129	0.020	0.121	0	1
Spatial lag of Catholic proportion	129	0.308	0.349	0	0.969
Spatial lag of Muslim proportion	129	0.172	0.273	0	0.991
Spatial lag of Protestant proportion	129	0.093	0.147	0	0.955
Spatial lag of other religion proportion	129	0.272	0.255	0	0.976
Spatial lag of ethnic fractionalization	129	0.378	0.239	0	0.856
Spatial lag of log of oil resources/GDP	129	−4.438	2.337	−9.814	0
Spatial lag of absolute latitude	129	0.273	0.213	0	0.700
Spatial lag of landlocked	129	0.186	0.247	0	1
Spatial lag of log GDP per capita	129	7.111	3.173	0	10.282
Spatial lag of human capital	129	4.896	4.259	0	27.300
Spatial lag of log population	1161	14.214	6.177	0	21.004
Spatial lag of VA, using distance weights	1161	−0.029	0.722	−1.063	1.235
Spatial lag of GE, using distance weights	1161	−0.009	0.619	−0.943	1.366
Spatial lag of RL, using distance weights	1161	−0.078	0.574	−0.943	1.208
Polity regression variables					
Polity	2926	0.215	0.755	−1	1
Spatial lag of polity	2926	0.060	0.494	−0.863	0.963
Log GDP per capita at 2005 PPP prices, 5-year lag	2926	8.341	1.223	5.390	10.654
Log of population	2926	16.258	1.453	12.929	20.968

A note on numbers of observations: We list the number of observations used in our regressions rather than the number available in original sources. The summary information on spatial lags of explanatory variables reflects data after missing observations are ignored in the manner suggested in Appendix A. Using the notation of that Appendix, the summary information on the baseline explanatory variables reflects the data on $X_{t,1}$ and H_1, which are observed. However, complete data on the spatial lags of these variables for all 129 countries would require data on $(W_{11}X_{t,1}+W_{12}X_{t,2})$ and $(W_{11}H_2+W_{12}H_2)$. But $X_{t,2}$ and H_2 are missing. Therefore for the spatial lags of the explantory variables, the table presents data on $W_{11}X_{t,1}$ and $W_{11}H_1$, which is what is used in the paper.

Table 3
Core estimates of the spatial determinants of institutions.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.191[b]	0.165[a]	0.316[c]
	(0.010)	(0.056)	(0.000)
French legal origin	−0.143	−0.234[b]	−0.182[a]
	(0.223)	(0.017)	(0.064)
Socialist legal origin	−0.481[c]	−0.506[c]	−0.596[c]
	(0.002)	(0.001)	(0.000)
German legal origin	−0.075	−0.096	0.174
	(0.672)	(0.640)	(0.275)
Scandinavian legal origin	−0.176	0.294	0.169
	(0.356)	(0.218)	(0.504)
Catholic proportion	0.135	−0.123	−0.310[b]
	(0.418)	(0.389)	(0.035)
Muslim proportion	−0.635[c]	−0.143	−0.101
	(0.001)	(0.355)	(0.519)
Protestant proportion	0.379	0.066	−0.021
	(0.125)	(0.800)	(0.940)
Ethnic fractionalization	0.096	−0.048	−0.081
	(0.608)	(0.790)	(0.675)
Log of oil resources to GDP ratio	−0.095[c]	−0.087[c]	−0.060[b]
	(0.000)	(0.000)	(0.015)
Absolute latitude	0.885[b]	0.711[a]	0.651[a]
	(0.011)	(0.076)	(0.080)
Landlocked	−0.141	−0.005	−0.048
	(0.271)	(0.963)	(0.655)
Log GDP per capita	0.305[c]	0.540[c]	0.420[c]
	(0.000)	(0.000)	(0.000)
Human capital	0.013	0.008	0.008
	(0.124)	(0.348)	(0.425)
Log population	−0.112[c]	−0.038	−0.081[c]
	(0.000)	(0.127)	(0.004)
Constant	−1.347[a]	−4.256[c]	−2.368[c]
	(0.081)	(0.000)	(0.001)
Adjusted R^2	0.667	0.775	0.779
Number of countries	129	129	129
Observations	1161	1161	1161
Spatial autocorrelation coefficient (ρ)	0.114	0.145	0.019
Hansen *J*-statistic	10.877	15.744	13.131
p value of *J*-statistic	0.621	0.263	0.438

Notes: Omitted variables are the English legal origin dummy and other-religion proportion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

for the difference between our studies and others using that variable. But this is conjecture that needs serious further investigation.

6. The magnitude of spillovers in institutional development: general equilibrium effects

We now analyze the general equilibrium effects of spillovers using the methods laid out in Section 3.5 and the estimates of Table 3. Since socialist legal origin is strongly significant for all three institutional indicators, both economically and statistically, one interesting experiment imagines how much institutions would be different if the socialist legal system had never existed.[22]

Table 4 presents the results. We focus on Europe and the former Soviet Union in 1998. We estimate the direct effect on each country of that country never having had a socialist legal origin, as well as the general equilibrium effect on each country of no country ever having had a socialist legal origin. Direct effects are based on Eq. (1), while general equilibrium effects are based on Eq. (2), with indirect effects reflecting the difference between these two.

[22] Since the omitted dummy variable for legal origin is English common law, our exercise calculates what would have happened had the formerly socialist countries instead had a heritage of the common law.

Table 4
The effects on selected countries of the past existence of the socialist system.

	Levels in 1998			Direct effects			Total effects including spillovers		
	VA	GE	RL	ΔVA	ΔGE	ΔRL	ΔVA	ΔGE	ΔRL
Armenia	−0.38	−0.57	−0.46	0.48	0.51	0.6	0.54	0.56	0.74
Austria	1.24	1.55	1.93	0	0	0	0.06	0.06	0.15
Azerbaijan	−1.04	−0.86	−0.93	0.48	0.51	0.6	0.55	0.57	0.77
Belarus	−1.07	−0.9	−1.17	0.48	0.51	0.6	0.59	0.6	0.86
Bosnia–Herz.	−1.2	−0.86	−1.12	0.48	0.51	0.6	0.59	0.61	0.87
Bulgaria	0.43	−1.03	−0.31	0.48	0.51	0.6	0.54	0.56	0.75
Croatia	−0.32	0.23	−0.2	0.48	0.51	0.6	0.59	0.6	0.85
Czech Rep.	1.08	0.69	0.59	0.48	0.51	0.6	0.54	0.56	0.74
Denmark	1.4	2.2	1.92	0	0	0	0.01	0	0.02
Estonia	0.78	0.45	0.44	0.48	0.51	0.6	0.59	0.6	0.86
Finland	1.39	2.12	1.98	0	0	0	0.04	0.03	0.1
Georgia	−0.44	−0.47	−0.84	0.48	0.51	0.6	0.56	0.58	0.79
Germany	1.25	1.78	1.85	0	0	0	0.03	0.03	0.07
Greece	0.82	0.74	0.65	0	0	0	0.07	0.06	0.17
Hungary	1.09	0.76	0.73	0.48	0.51	0.6	0.57	0.59	0.82
Italy	1.17	0.99	1.03	0	0	0	0.03	0.03	0.07
Kazakhstan	−0.8	−0.77	−0.91	0.48	0.51	0.6	0.59	0.6	0.86
Kyrgyz	−0.52	−0.33	−0.76	0.48	0.51	0.6	0.59	0.6	0.86
Latvia	0.73	0.24	−0.04	0.48	0.51	0.6	0.59	0.61	0.87
Lithuania	0.86	0.2	0.07	0.48	0.51	0.6	0.59	0.61	0.86
Macedonia	0.01	−0.42	−0.43	0.48	0.51	0.6	0.54	0.56	0.74
Moldova	−0.13	−0.56	−0.16	0.48	0.51	0.6	0.59	0.61	0.87
Poland	1	0.84	0.49	0.48	0.51	0.6	0.58	0.59	0.82
Romania	0.2	−0.63	−0.35	0.48	0.51	0.6	0.59	0.6	0.86
Russia	−0.26	−0.62	−0.9	0.48	0.51	0.6	0.57	0.59	0.82
Slovakia	0.37	0	0.07	0.48	0.51	0.6	0.57	0.59	0.81
Slovenia	0.86	0.68	0.86	0.48	0.51	0.6	0.54	0.56	0.75
Switzerland	1.44	2.48	2.27	0	0	0	0.01	0	0.02
Tajikistan	−1.47	−1.45	−1.53	0.48	0.51	0.6	0.59	0.6	0.86
Turkey	−0.97	−0.38	−0.01	0	0	0	0.06	0.06	0.15
Turkmenistan	−1.68	−1.46	−1.28	0.48	0.51	0.6	0.56	0.58	0.79
Ukraine	−0.15	−1	−0.88	0.48	0.51	0.6	0.59	0.6	0.86
Uzbekistan	−1.6	−1.33	−1.13	0.48	0.51	0.6	0.59	0.6	0.86

VA = voice and accountability; GE = government effectiveness; RL = rule of law; Δ = changes in these variables had the socialist system never existed

We use the difference between levels of institutions in Romania and Germany in 1998 as a metric to judge economic magnitude, these two countries now being at opposite ends of the European Union in terms of institutional development.[23] For each formerly socialist country, the direct effect of not having had the socialist legal system is 48% of the Romania–Germany difference for voice and accountability, 24% for government effectiveness, and 37% for the rule of law. For each formerly socialist country surrounded by formerly socialist countries, such as Belarus or Bosnia, the total effect is 23% larger than the direct effect for voice and accountability, 20% larger for government effectiveness, and 45% larger for rule of law. For each of those countries, the total effect of having had the socialist legal system is 59% of the Romania–Germany difference on voice and accountability, 29% on government effectiveness, and 54% on rule of law. These are large spillovers, with indirect effects adding considerably to direct ones.[24]

Spillovers affect countries other than those that were socialist. Neighbors of Soviet-bloc countries were affected. For example, there is an appreciable effect on Austria: 6% of the Romania–Germany difference on voice and accountability, 3% on government effectiveness, and 9% on rule of law. There are even discernible effects on neighbors of neighbors, such as Denmark, which did not border any socialist countries. These negative effects should not necessarily be viewed as arising from countries copying bad institutions, but more likely they are a result

[23] The Kaufmann et al. (2009) institutional measures are normalized to have mean zero, making elasticities not meaningful.
[24] Ward and Gleditsch (2008) provide an example of such general equilibrium effects, using a much more parsimonious model than the one used her. For the effect of GDP per capita on institutions, they find relatively modest indirect effects.

of the absence (relative to the counter-factual) of better neighboring institutions that could set a standard and spur competition.

7. Robustness exercises: Examining possible sources of bias

In this section we consider the robustness of the core estimates. The central concern raised by readers of the first version of this paper arose from the plausible reasoning that estimates of the spatial effect (λ) might be biased because of the existence of variables with three characteristics: they are unobserved, they are common to a group of neighboring countries and therefore spatially correlated, and they directly affect institutions. We refer to such variables as omitted common factors.

We argue that these concerns could well be unfounded. First, we provide a new theoretical result showing that the presence of omitted common factors is not sufficient to cause bias in the estimate of λ. Second, we subject the core estimates to many robustness tests, none of which raises any concerns.

7.1. A theoretical result

Appendix B proves that the estimate of λ is consistent if Eq. (1) is used when it is not the correct model due to an omitted common factor with plausible characteristics. We summarize the basic intuition here.

Omitted common factors will likely be correlated with included explanatory variables: for example, many studies show that deep cultural characteristics are related to religion. Omitted common factors also could reflect spatially correlated phenomena independent of the ones that we measure. Thus, our theoretical exercise assumes that the omitted common factors are a function of three elements: a linear combination of the explanatory variables included in the model, X_t and H, a disturbance term orthogonal to X_t and H, and the spatial lag of that disturbance term. Note that such common factors are spatially correlated for two reasons: directly because of the third term and indirectly because some of the variables in X_t and H are spatially correlated.

Suppose that the true model is Eq. (1) with an extra explanatory variable, an omitted common factor conforming to the above. Appendix B shows that under reasonable conditions the estimate of λ will be consistent when estimated using the IV procedure described above applied to Eq. (1) rather than the true model. This result is not intuitive because the omitted common factor is correlated with some of the variables in X_t and H, as well as with the spatial lag of y_t which is also an explanatory variable in Eq. (1). The reasoning for the consistency of λ proceeds in two steps. First, the estimator of λ in Eq. (1) is basically a 2SLS estimator, and so, very intuitively speaking, it eliminates the effect of both X_t and H. Thus, the only part of the omitted common factor that matters for the estimation of λ is its disturbance term and the spatial lag of that term. However, we estimate our model by an instrumental variable procedure using a set of instruments that are orthogonal to that disturbance term, as well as to its spatial lag. Thus, there is no part of the omitted common factor that is of consequence in large samples in the estimation of λ: it is as if the omitted common factor is not in the true model.[25] This is proven rigorously in Appendix B. We now turn to empirical demonstrations of robustness.

[25] We ran Monte Carlo simulations examining whether the insights from the theory, which reflect asymptotics, carry over into estimations with samples sizes similar to those employed in our empirical analyses. The simulations were based on the data used in Table 3 combined with simulated data on a common factor. Different trials were generated by varying separately both the element of the common factor correlated with the explanatory variables and the element orthogonal to the explanatory variables. For the value of λ equal to the estimated coefficient of the spatial lag of voice and accountability in Table 3 (0.191), we found a range of 0.006 to 0.012 for the square root of the mean squared difference between the estimates of λ for the true model and for the model that (incorrectly) omits the spatially correlated common factor. The range did not change systematically with the size of the common factor.

7.2. The possible endogeneity of GDP

The log of GDP per capita appears in Table 3 on the theory that wealth provides resources to create better institutions. This assumption is common in the literature (Chong and Zanforlin, 2000; Keefer and Knack, 1996; La Porta et al., 1999), although Kaufmann and Kraay (2002) and Acemoglu et al. (2008) present counter-evidence. Because better institutions spur increases in GDP (Acemoglu et al., 2001; Hall and Jones, 1999; Rodrik et al., 2004), the consistency of the estimates in Table 3 requires that the GDP variable (the mean from 1991–1995) be predetermined. This could be the case if increased production capacity translates into more effective institutions with a lag. However, it is natural to question this assumption. Therefore, we conduct two robustness exercises in order to examine whether our chosen strategy has any effect on the qualitative conclusions concerning the estimate of the spatial coefficient.

First, we omit the log of GDP per capita and normalize oil reserves by population rather than GDP. The results appear in Table 5. The patterns of estimated coefficients and their levels of statistical significance are broadly similar to those in Table 3. The economic and statistical significance of the spatial variable is greater in Table 5 than in Table 3.

Second, we treat GDP per capita as endogenous. Since the literature lacks agreement on variables that affect GDP and not institutions, we estimate our model with the set of instruments we have already used, dropping GDP and its spatial lag from this set. The results appear in Table 6. The patterns of estimated coefficients and their levels of statistical significance are broadly similar to those previously estimated. In fact,

Table 5
Estimates without GDP per capita in the spatial equation.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.321[c]	0.469[c]	0.537[c]
	(0.000)	(0.000)	(0.000)
French legal origin	−0.208	−0.253[b]	−0.234[b]
	(0.103)	(0.036)	(0.034)
Socialist legal origin	−0.929[c]	−0.861[c]	−0.868[c]
	(0.000)	(0.000)	(0.000)
German legal origin	0.202	0.585[c]	0.557[c]
	(0.203)	(0.007)	(0.001)
Scandinavian legal origin	−0.382	−0.008	0.028
	(0.068)	(0.978)	(0.920)
Catholic proportion	0.055	−0.119	−0.227
	(0.782)	(0.514)	(0.174)
Muslim proportion	−0.817[c]	−0.274	−0.196
	(0.000)	(0.111)	(0.253)
Protestant proportion	0.164	−0.079	−0.175
	(0.528)	(0.799)	(0.551)
Ethnic fractionalization	−0.269	−0.440[b]	−0.500[b]
	(0.196)	(0.050)	(0.020)
Log of oil resources per capita	−0.024	0.001	0.007
	(0.204)	(0.949)	(0.766)
Absolute latitude	1.839[c]	1.648[c]	1.386[c]
	(0.000)	(0.001)	(0.002)
Landlocked	−0.342[b]	−0.308[c]	−0.238[b]
	(0.012)	(0.007)	(0.028)
Human capital	0.030[c]	0.033[c]	0.026[b]
	(0.000)	(0.001)	(0.016)
Log population	−0.113[c]	−0.056[b]	−0.086[c]
	(0.000)	(0.027)	(0.002)
Constant	1.851[c]	0.960[b]	1.529[c]
	(0.000)	(0.037)	(0.001)
Adjusted R^2	0.663	0.739	0.764
Number of countries	129	129	129
Spatial autocorrelation coefficient (ρ)	0.011	−0.052	−0.106
Hansen *J*-statistic	14.464	16.114	9.710
p value of *J*-statistic	0.272	0.186	0.641
Observations	1161	1161	1161

Notes: Omitted variables are the English legal origin dummy and other-religion proportion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 6
Estimates with GDP per capita treated as endogenous.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.276[c]	0.328[c]	0.433[c]
	(0.000)	(0.000)	(0.000)
French legal origin	−0.049	−0.196[a]	−0.225[b]
	(0.704)	(0.059)	(0.026)
Socialist legal origin	−0.366[b]	−0.590[c]	−0.773[c]
	(0.043)	(0.001)	(0.000)
German legal origin	0.002	0.166	0.340[a]
	(0.992)	(0.504)	(0.071)
Scandinavian legal origin	−0.135	0.326	0.130
	(0.518)	(0.209)	(0.627)
Catholic proportion	0.132	−0.054	−0.274[a]
	(0.449)	(0.739)	(0.074)
Muslim proportion	−0.538[b]	−0.207	−0.132
	(0.011)	(0.194)	(0.423)
Protestant proportion	0.382	0.058	−0.098
	(0.118)	(0.839)	(0.725)
Ethnic fractionalization	0.309	−0.045	−0.312
	(0.149)	(0.836)	(0.207)
Log of oil resources to GDP ratio	−0.105[c]	−0.069[c]	−0.043
	(0.000)	(0.002)	(0.108)
Absolute latitude	0.997[c]	1.433[c]	1.026[b]
	(0.004)	(0.001)	(0.043)
Landlocked	−0.136	−0.158	−0.148
	(0.275)	(0.144)	(0.220)
Log GDP per capita	0.259[c]	0.220[c]	0.169
	(0.000)	(0.000)	(0.399)
Human capital	0.015[a]	0.024[c]	0.022
	(0.058)	(0.009)	(0.172)
Log population	−0.103[c]	−0.037	−0.093[c]
	(0.000)	(0.187)	(0.001)
Constant	−1.069[c]	−1.475[c]	0.027
	(0.010)	(0.007)	(0.988)
Adjusted R^2	0.614	0.722	0.776
Number of countries	129.000	129.000	129.000
Spatial autocorrelation coefficient (ρ)	0.228	0.182	−0.039
Hansen *J*-statistic	5.161	10.697	9.884
p value of *J*-statistic	0.923	0.469	0.541
Observations	1161	1161	1161

Notes: Omitted variables are the English legal origin dummy and other-religion proportion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

our core estimates are the weakest, statistically and economically, of all those in Tables 3, 5, and 6. In sum, there are no concerns that the treatment of GDP affects basic conclusions on how the spatial variable affects institutions.

7.3. Seeking important omitted variables

In implementing Eq. (1) to obtain Table 3, we chose the set of variables that we judge to be the main candidates to explain levels of institutions in a country-panel context, based on our reading of the literature. Nevertheless, given the vast number of ways in which countries differ, readers will surely have their own candidates for favorite explanatory variables, thus suspecting omitted-variable bias, notwithstanding the theory of Appendix B. Hence, in this subsection, we add a number of variables individually to the regressions of Table 3, focusing solely on how the estimates of λ change.

The literature on determinants of institutions within a country-panel context is still in its infancy and no existing paper provides a comprehensive set of explanatory variables. Thus, to find candidates for causes of omitted-variable bias we turn to the cross-country growth literature. Since growth and institutional development are intertwined, studies of growth that do not directly include measures of institutions could reveal candidates to add to the explanatory variables of Table 3.

Sala-i-Martin et al. (2004) analyze cross-country growth regressions and identify variables that are significantly and robustly partially correlated with long-term growth. They identify eighteen leading candidates. Five of these are already included in Table 3 and three are not relevant to an analysis of institutions, leaving ten that we examine: three continental dummies, a colonizer dummy, trade openness, coastal access, tropical land, malaria prevalence, Buddhist proportion, and Confucian proportion. We also examine log of land area since this variable and population density have appeared previously in studies of institutions. Combining log of land area with the already included log of population implies the inclusion of density.

The results appear in Tables 7(a), (b), and 8, which, for brevity, include only estimates of the coefficient of the spatial lag, the coefficients of the added variables, the spatial autocorrelation coefficient, and the *p*-value of the *J*-statistic.[26] There are 27 new estimates of λ (7 new variables and 2 new sets of dummies, in equations for each of the 3 institutional indicators). Statistical significance increases (relative to Table 3) in 13 cases and falls in 14. The numerical size of the estimate of λ increases in 12 cases and decreases in 15. In 23 of the 27 cases the estimate of λ is significant at the 90% level. In all 9 instances of added variables at least two of the estimates of λ are significant at the 99% level. The results on the spatial-lag are usually stronger when a significant variable is added and weaker when the added variable is not significant, suggesting the main effect of adding variables is simply a change in precision.

In sum, adding variables to those in Table 3 tends, if anything, to strengthen the conclusion that institutions in neighboring countries have an important effect on institutions in a country.

7.4. Cross-sectional or within-country effects?

Our sample is a panel, with estimates reflecting both variations between countries and within countries. Given the usual supposition that institutions change slowly, it is natural to assume that cross-sectional variations dominate. However, our data span a period of vigorous institutional reform, especially in the transition countries (Murrell, 2008). A significant number of countries do experience large changes in institutions in this short time period. Therefore, we investigate whether spatial effects are separately present both in cross-section and within countries over time.

To focus on within-country variations, we use country fixed-effects for *H*, the time-invariant regressors. The fixed effects then replace variables in Table 3 that have more intuitive interpretations such as oil dependence and legal origin. Estimates appear in Tables 9 and 10. Table 9 uses a sample identical to that in Table 3. Table 10 uses a sample that is not limited by the availability of data on the explanatory variables of Table 3, thereby including all countries with data on the spatial lag.

Consistent estimation of the fixed effects version of our model should account for the endogeneity of the spatial lag. However, given model Eq. (1), the only available instruments are the spatial lags of population since this is the only time varying variable. Not surprisingly, these instruments are weak. Therefore, we estimate the fixed effects version of our spatial model by OLS, using the same three-step procedure as above but without instruments.[27]

Within-country effects are strong. There is only one case (rule of law in the larger sample) for which the estimate of the spatial coefficient is smaller than in Table 3. Significance levels are much higher for all coefficients. Since fixed-effect estimates reflect only within-country variations, these are strong and unexpected results given the prevailing view in the literature that institutions change slowly.

[26] A full set of results is available on request to the authors.
[27] When comparing OLS and IV estimates for the voice and accountability and rule of law regressions reported in Table 3, the IV estimates of the spatial lag coefficient are larger than the corresponding OLS estimates. This is an indication that the estimates of the spatial lag coefficients for voice and accountability and rule of law that appear in Tables 9 and 10 are likely downward biased due to the use of OLS, an observation bolstering our interpretation of the results.

Table 7
Estimates of the spatial lag coefficient when adding variables to the spatial equation.

Added concept	Estimated statistics	Voice and accountability	Government effectiveness	Rule of law
Confucianism	Coefficient on spatial lag (λ)	0.175[b]	0.170[b]	0.365[c]
		(0.021)	(0.033)	(0.000)
	Coefficient on Confucianism dummy	−1.066[b]	2.655[c]	2.165[c]
		(0.045)	(0.000)	(0.000)
	Spatial autocorrelation coefficient (ρ)	0.149	0.062	−0.079
	p value of J-statistic	0.634	0.183	0.432
Buddhism	Coefficient on spatial lag (λ)	0.188[b]	0.201[b]	0.365[c]
		(0.012)	(0.023)	(0.000)
	Coefficient on Buddhism dummy	−0.245	−0.149	0.453
		(0.251)	(0.546)	(0.125)
	Spatial autocorrelation coefficient (ρ)	0.122	0.125	−0.039
	p value of J-statistic	0.642	0. 145	0.329
Log of land area	Coefficient on spatial lag (λ)	0.188[b]	0.137[a]	0.303[c]
		(0.010)	(0.067)	(0.001)
	Coefficient on Log of land area	−0.022	−0.056[a]	−0.027
		(0.456)	(0.083)	(0.504)
	Spatial autocorrelation coefficient (ρ)	0. 084	0. 093	−0. 010
	p value of J-statistic	0. 632	0. 386	0. 657
Ratio of population within 100 km of ice-free coast to total population	Coefficient on spatial lag (λ)	0.227[c]	0.166[b]	0.337[c]
		(0.001)	(0.032)	(0.000)
	Coefficient on ratio of population within 100 km of ice-free coast to total population	0.482[c]	0.385[b]	0.322[a]
	Spatial autocorrelation coefficient (ρ)	0. 079	0.039	−0. 026
	p value of J-statistic	0. 594	0. 385	0. 663
% land area in geographical tropics	Coefficient on spatial lag (λ)	0.176[b]	0.116	0.339[c]
		(0.014)	(0.143)	(0.000)
	Coefficient on % land area in geographical tropics	0.164	0.289	0.143
		(0.553)	(0.214)	(0.563)
	Spatial autocorrelation coefficient (ρ)	0. 072	0. 083	−0. 033
		0. 620	0.222	0.572
	p value of J-statistic			
% of 1995 population living in areas with malaria, 1982	Coefficient on spatial lag (λ)	0.142[b]	0.112	0.279[c]
		(0.050)	(0.165)	(0.001)
	Coefficient on % of 1995 population living in areas with malaria, 1982	−0.185	−0.059	0.093
		(0.185)	(0.656)	(0.543)
	Spatial autocorrelation coefficient (ρ)	0.123	0.146	0.016
	p value of J-statistic	0.050	0.163	0.119
Trade openness	Coefficient on spatial lag (λ)	0.203[c]	0.098	0.299[c]
		(0.006)	(0.270)	(0.001)
	Coefficient on trade openness	−0.229	0.302[a]	0.035
		(0.220)	(0.095)	(0.846)
	Spatial autocorrelation coefficient (ρ)	0. 072	0.185	0.021
	p value of J-statistic	0.575	0.337	0.410

Notes: Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 11 contains the results of regressions for each separate year, reporting only the main estimates of interest. Cross-sectional variation is an important source of spatial effects. Indeed, despite the much smaller sample sizes, the spatial-lag coefficient is statistically significant in 21 of the 27 cases. The estimates of that coefficient are comparable in size to those in Table 3.

8. Expanding the horizons of the core spatial model

In this section, we examine whether results change qualitatively when we move farther from the framework used for Table 3. We examine four separate extensions of the model—the possibility of a different effect between worse and better neighboring institutions; adding spatial lags of explanatory variables to the set of explanatory variables (the spatial Durbin model); using a weighting matrix with distance weights; and employing a different measure of institutions (over a different time period).

8.1. Asymmetric effects

The effect of neighboring institutions could depend on whether those institutions are better or worse than in the country itself. For example, it might be easier to learn by emulating better examples rather than avoiding worse ones. In contrast, setting up populist institutions might be more tempting when conditions are worse. Ultimately, which effect is stronger is an empirical question, which we address by allowing for the possibility of asymmetric effects. The primary objective here is to ascertain whether there is any change in the qualitative conclusions reached so far.

Table 8
Estimates of the spatial lag coefficient when adding continental and colonial dummy variables to the spatial equation.

Added concept	Estimated statistics	Voice and accountability	Government effectiveness	Rule of law
Continents	Coefficient on spatial lag (λ)	0.097	0.213[b]	0.254[c]
		(0.265)	(0.013)	(0.003)
	Coefficient on Africa dummy	−0.199	0.407[b]	0.298
		(0.394)	(0.022)	(0.142)
	Coefficient on America dummy	−0.192	0.036	−0.172
		(0.216)	(0.825)	(0.379)
	Coefficient on Asia dummy	−0.461[b]	0.154	0.046
		(0.024)	(0.308)	(0.794)
	Coefficient on Pacific dummy	−0.361	0.173	0.079
		(0.200)	(0.608)	(0.834)
	Spatial autocorrelation coefficient (ρ)	0.171	0.062	0.073
	p value of J-statistic	0.270	0.221	0.242
Colonies	Coefficient on spatial lag (λ)	0.235[c]	0.257[c]	0.224[c]
		(0.002)	(0.001)	(0.003)
	Coefficient on Colonizer UK dummy	0.233[a]	0.324[c]	0.404[c]
		(0.062)	(0.001)	(0.002)
	Coefficient on Colonizer France dummy	−0.080	0.371[b]	0.167
		(0.661)	(0.013)	(0.281)
	Coefficient on Colonizer Spain dummy	−0.055	0.019	−0.326[b]
		(0.713)	(0.897)	(0.039)
	Coefficient on Colonizer Russia dummy	−0.091	0.034	−0.278
		(0.644)	(0.832)	(0.111)
	Spatial autocorrelation coefficient (ρ)	0.095	0.050	0.122
	p value of J-statistic	0.620	0.483	0.265

Notes: Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

The model to be estimated is now

$$\begin{aligned} y_t &= X_t\beta_1 + H\beta_2 + \lambda_1 W_1 y_t + \lambda_2 W_2 y_t + u_t \\ u_t &= \rho W u_t + \varepsilon_t, \quad t = 1, \dots, T \end{aligned} \tag{10}$$

where W_1 is a contiguity weighting matrix constructed as a row-normalization of the weighting matrix $\tilde{W}_1$, where $\tilde{w}_{1,ij} = 1$ if countries i and j share a common border and $y_{i,1998} \le y_{j,1988}$, and equals 0 otherwise. W_2 is the contiguity weighting matrix constructed as a row-normalization of the weighting matrix $\tilde{W}_2$, where $\tilde{w}_{2,ij} = 1$ if countries i and j share a common border and $y_{i,1998} > y_{j,1988}$, and equals 0 otherwise. That is, for neighboring countries the weights in W_1 (W_2) are equal to the reciprocal of the number of neighboring countries with worse (better) institutions, and zero for non-neighboring countries. The model is estimated by an obvious variant of our above-described procedure, with instruments in this case being X_t, W_1X_t, W_2X_t, H, W_1H, and W_2H. Under reasonable conditions, these estimates are consistent and asymptotically normal (Kelejian and Prucha, 2004).

Table 9
Fixed effects estimates of the spatial lag coefficient using the restricted sample.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.499[c]	0.343[c]	0.421[c]
	(0.000)	(0.000)	(0.000)
Log population	0.091	−0.112	−0.102
	(0.508)	(0.348)	(0.323)
Constant	−1.707	2.343	2.113
	(0.417)	(0.194)	(0.179)
Adjusted R^2	0.972	0.975	0.982
Number of countries	129	129	129
Observations	1161	1161	1161
Spatial autocorrelation coefficient (ρ)	−0.095	−0.030	−0.064

Notes: Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 10
Fixed effects estimates of the spatial lag coefficient using the expanded sample.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.582[c]	0.181[c]	0.200[c]
	(0.000)	(0.000)	(0.000)
Log population	0.025	−0.218[b]	0.016
	(0.790)	(0.031)	(0.866)
Constant	−1.196	2.371	−2.008
	(0.529)	(0.176)	(0.228)
Adjusted R^2	0.975	0.966	0.974
Number of countries	179	179	179
Observations	1611	1611	1611
Spatial autocorrelation coefficient (ρ)	−0.154	−0.002	−0.020

Notes: Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 11
Estimates of the spatial lag coefficient for each year.

Year	Estimated statistics	Voice and accountability	Government effectiveness	Rule of law
1998	Coefficient on spatial lag (λ)	0.133	0.155[a]	0.336[c]
		(0.109)	(0.076)	(0.000)
	Spatial autocorrelation coefficient (ρ)	0.024	0.052	−0.013
	p value of J-statistic	0.224	0.516	0.097
2000	Coefficient on spatial lag (λ)	0.109	0.197[b]	0.262[c]
		(0.175)	(0.018)	(0.002)
	Spatial autocorrelation coefficient (ρ)	0.128	0.014	0.049
	p value of J-statistic	0. 509	0.383	0.155
2002	Coefficient on spatial lag (λ)	0.183[b]	0.213[b]	0.329[c]
		(0.023)	(0.012)	(0.002)
	Spatial autocorrelation coefficient (ρ)	0.073	0.070	−0.005
	p value of J-statistic	0.843	0.430	0.211
2003	Coefficient on spatial lag (λ)	0.176[b]	0.133	0.346[c]
		(0.034)	(0.161)	(0.000)
	Spatial autocorrelation coefficient (ρ)	0.069	0.112	−0.034
	p value of J-statistic	0.745	0.239	0.471
2004	Coefficient on spatial lag (λ)	0.222[c]	0.164[a]	0.343[c]
		(0.004)	(0.097)	(0.000)
	Spatial autocorrelation coefficient (ρ)	0.011	0.099	−0.104
	p value of J-statistic	0.483	0.282	0.518
2005	Coefficient on spatial lag (λ)	0.203[c]	0.135	0.340[c]
		(0.008)	(0.139)	(0.000)
	Spatial autocorrelation coefficient (ρ)	−0.025	0.084	−0.069
	p value of J-statistic	0.382	0.133	0.478
2006	Coefficient on spatial lag (λ)	0.187[b]	0.131	0.269[c]
		(0.020)	(0.134)	(0.002)
	Spatial autocorrelation coefficient (ρ)	0.025	0.020	−0.025
	p value of J-statistic	0.618	0.178	0.618
2007	Coefficient on spatial lag (λ)	0.196[b]	0.132	0.305[c]
		(0.013)	(0.143)	(0.000)
	Spatial autocorrelation coefficient (ρ)	0.033	0.020	−0.054
	p value of J-statistic	0.781	0.173	0.498
2008	Coefficient on spatial lag (λ)	0.190[b]	0.185[b]	0.307[c]
		(0.020)	(0.033)	(0.000)
	Spatial autocorrelation coefficient (ρ)	0.037	−0.051	−0.062
	p value of J-statistic	0.900	0.329	0.566

Notes: Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 12
Estimates of two spatial lag coefficients, assuming the effect of better neighboring institutions is different from the effect of worse neighboring institutions.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag reflecting only neighbors with better institutions (λ_1)	0.312[c] (0.000)	0.337[c] (0.000)	0.411[c] (0.000)
Spatial lag reflecting only neighbors with worse institutions (λ_2)	0.156[c] (0.000)	0.094[b] (0.031)	0.126[b] (0.022)
French legal origin	−0.049 (0.626)	−0.215[b] (0.010)	−0.160[a] (0.083)
Socialist legal origin	−0.502[c] (0.000)	−0.519[c] (0.000)	−0.665[c] (0.000)
German legal origin	−0.137 (0.418)	−0.122 (0.541)	0.075 (0.549)
Scandinavian legal origin	−0.300[a] (0.054)	0.169 (0.438)	0.182 (0.379)
Catholic proportion	0.031 (0.831)	−0.008 (0.950)	−0.068 (0.618)
Muslim proportion	−0.696[c] (0.000)	−0.129 (0.302)	0.008 (0.948)
Protestant proportion	0.428[b] (0.026)	0.207 (0.364)	−0.068 (0.781)
Ethnic fractionalization	0.052 (0.761)	−0.141 (0.389)	−0.165 (0.343)
Log of oil resources to GDP ratio	−0.077[c] (0.000)	−0.072[c] (0.000)	−0.041[b] (0.043)
Absolute latitude	1.195[c] (0.000)	1.104[c] (0.001)	1.089[c] (0.002)
Landlocked	−0.164 (0.100)	0.008 (0.932)	−0.058 (0.532)
Log GDP per capita	0.221[c] (0.001)	0.475[c] (0.000)	0.380[c] (0.000)
Human capital	0.008 (0.195)	0.007 (0.360)	0.013 (0.191)
Log population	−0.095[c] (0.000)	−0.023 (0.252)	−0.076[c] (0.001)
Constant	−0.773 (0.193)	−3.902[c] (0.000)	−2.104[c] (0.001)
Adjusted R^2	0.724	0.816	0.780
Number of countries	129	129	129
Spatial autocorrelation coefficient (ρ)	0.105	0.193	0.160
Hansen J-statistic	35.712	36.597	35.067
p value of J-statistic	0.097	0.081	0.110
χ^2 Statistic for null hypothesis $\lambda_1=\lambda_2$	5.810	10.469	11.710
p value of χ^2statistic	0.003	0.003	0.003
Observations	1161	1161	1161

Notes: Omitted variables: English legal origin and other-religion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 12 contains the results. Notably, all six spatial coefficients are strongly significant. The effect of better institutions is larger than the estimated effect in Table 3, while the effect of worse institutions is smaller. Thus, a marginal improvement in a neighbor's institutions has a larger effect if that neighbor already has better institutions. This difference between the effect of better and worse neighbors is significant.[28] Note that the effect of worse institutions is positive, improvements in worse neighboring institutions still spur institutional development.[29]

8.2. A spatial Durbin model

It is possible that institutions might be directly affected by conditions in a neighboring country rather than indirectly by the effect of those conditions on the neighbor's institutions. For example, a higher level of GDP in country A might stimulate better institutions in country B induced by the desire in B to take advantage of the better trading environment with A. We have not allowed for such direct effects so far. To do so, we add spatial lags of the explanatory variables as additional explanatory variables. The resultant framework is called the spatial Durbin model. Formally, we now estimate:

$$\begin{aligned} y_t &= X_t\beta_1 + H\beta_2 + WX_t\beta_3 + WH\beta_4 + \lambda Wy_t + u_t \\ u_t &= \rho Wu_t + \varepsilon_t, \quad t = 1, \ldots, T \end{aligned} \tag{11}$$

by the standard two-stage procedure, using X, H, WX, WH, WWX, and WWH as instruments. Note that β_3 and β_4 are respectively $k_1 \times 1$ and $k_2 \times 1$ parameter vectors. Again, under reasonable conditions, the estimates are consistent and asymptotically normal (Kelejian and Prucha, 2004).

The results appear in Table 13. The estimates of the coefficient of the spatial lag of institutions in the government effectiveness and the rule-of-law equations are much larger than in Table 3, with increases in statistical significance. The estimate of the coefficient on the spatial lag of institutions in the voice and accountability equation is now small and insignificant. However, for voice and accountability none of the estimates of the coefficients on the spatial lags of the explanatory variables are significant resulting not surprisingly in an insignificant Wald statistic for the test of $(\beta_3, \beta_4) = 0$. This indicates that the spatial Durbin model is not appropriate for voice and accountability: Eq. (1) is preferred to Eq. (11). Because so many coefficients are now estimated and many are insignificant, it is difficult to see patterns in the results on the non-institutional coefficients. Nevertheless, the results in Table 13 clearly underscore the results on spatial spillovers in Table 3.

8.3. Using inverse distance weights

The specification of the weighting matrix (W) in Eq. (1) reflected the assumption that institutions in those countries bordering a given country most directly affect institutions in a given country. But many different weighting matrixes are plausible. We present results for one variant common in the literature, where weights decline with distance. For this version, the weight applied to the effect of institutions in country j on institutions in i is proportional to the inverse of the distance between the two countries if i and j are located on the same continent and 0 otherwise.

Table 14 contains the results. The qualitative features of the estimates of λ are similar to those with border weights. However, the coefficient on the government effectiveness equation is no longer significant, while the voice and accountability coefficient increases in significance and size. The results of Table 14 broadly support our core results, which rely on a contiguity weighting matrix.

8.4. An alternative institutional indicator and an alternative time period

We now examine whether our results depend critically upon the specific institutional indicators used and the particular time period covered. To address this question, we use an institutional dependent variable that reflects a much longer time period. Using the Polity IV Project (2011) database, we construct our "polity" measure by subtracting the autocracy score from the democracy score and normalizing so that the measure takes on values between −1 and 1. We were able to assemble a balanced panel of 77 countries from 1965 to 2002.[30] The weighting matrix is the distance-matrix as in the previous section: inverse distance

[28] See the chi-squared test statistic in the bottom panel of Table 12.
[29] It is tempting to think that the effect should be negative, dragging a country down, but that intuition is not correct. The correct intuition is that a country's institutions are made better if its (bad) neighbors improve.

[30] A balanced panel of polity indicators from 1965 to 2002 is available for 119 countries. After merging with the 5-year lag of GDP per capita variable, we were left with a balanced panel of 90 countries. Starting the analysis earlier than 1965 would involve too few countries since the sample of countries for the lagged GDP per capita variable diminishes considerably before then. After merging the polity and GDP per capita data with the other explanatory variables there remain 77 countries.

Table 13
Estimates of the spatial lag coefficient using the spatial Durbin model.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.024	0.489[c]	0.710[c]
	(0.901)	(0.000)	(0.000)
French legal origin	−0.258[a]	−0.062	0.076
	(0.067)	(0.634)	(0.579)
Socialist legal origin	−0.548[b]	0.160	−0.205
	(0.020)	(0.476)	(0.396)
German legal origin	−0.320	0.246	0.371[a]
	(0.133)	(0.257)	(0.064)
Scandinavian legal origin	−0.308	0.634[a]	0.453
	(0.272)	(0.074)	(0.226)
Catholic proportion	0.253	0.354[b]	0.147
	(0.164)	(0.037)	(0.430)
Muslim proportion	−0.666[c]	−0.134	−0.148
	(0.002)	(0.306)	(0.319)
Protestant proportion	0.485[a]	0.235	0.205
	(0.067)	(0.433)	(0.537)
Ethnic fractionalization	0.123	0.094	0.012
	(0.530)	(0.603)	(0.953)
Log of oil resources to GDP ratio	−0.091[c]	−0.083[c]	−0.052[b]
	(0.001)	(0.000)	(0.034)
Absolute latitude	1.051[a]	0.207	0.347
	(0.060)	(0.737)	(0.610)
Landlocked	−0.075	0.017	0.073
	(0.556)	(0.877)	(0.524)
Log GDP per capita	0.274[c]	0.604[c]	0.467[c]
	(0.001)	(0.000)	(0.000)
Human capital	0.010	−0.003	0.008
	(0.331)	(0.779)	(0.544)
Log population	−0.106[c]	−0.053[a]	−0.096[c]
	(0.002)	(0.066)	(0.003)
Spatial lag French legal origin	−0.064	−0.248	−0.368[a]
	(0.714)	(0.173)	(0.055)
Spatial lag Socialist legal origin	−0.352	−0.556[a]	−0.112
	(0.244)	(0.052)	(0.736)
Spatial lag German legal origin	0.067	−0.160	−0.347
	(0.866)	(0.697)	(0.385)
Spatial lag Scandinavian legal origin	0.269	−0.706	−0.318
	(0.619)	(0.189)	(0.599)
Spatial lag Catholic proportion	0.007	−0.192	−0.112
	(0.978)	(0.420)	(0.638)
Spatial lag Muslim proportion	0.135	0.582[b]	0.448
	(0.688)	(0.021)	(0.113)
Spatial lag Protestant proportion	−0.599	0.068	−0.372
	(0.281)	(0.884)	(0.492)
Spatial lag Ethnic fractionalization	−0.289	0.105	0.458
	(0.419)	(0.712)	(0.186)
Spatial lag log of oil resources to GDP ratio	−0.039	0.096[b]	0.054
	(0.512)	(0.012)	(0.162)
Spatial lag Absolute latitude	0.238	0.424	0.194
	(0.751)	(0.584)	(0.799)
Spatial lag Landlocked	0.162	0.073	−0.126
	(0.514)	(0.720)	(0.522)
Spatial lag log GDP per capita	0.083	−0.154[a]	−0.171[b]
	(0.373)	(0.052)	(0.025)
Spatial lag Human capital	0.007	−0.013	−0.022
	(0.636)	(0.257)	(0.164)
Spatial lag log population	−0.055	0.102[c]	0.095[c]
	(0.225)	(0.001)	(0.007)
Constant	−1.099	−4.653[c]	−2.594[c]
	(0.144)	(0.000)	(0.001)
Adjusted R^2	0.683	0.874	0.875
Number of countries	129	129	129
Spatial autocorrelation coefficient (ρ)	0.106	−0.277	−0.367
Hansen J-statistic	19.898	18.762	13.657
p value of J-statistic	0.098	0.131	0.398
$\chi^2(14)$: $\beta_3=0, \beta_4=0$	11.930	71.093	60.434
p value of χ^2	0.612	0.000	0.000
Observations	1161	1161	1161

Notes: Omitted variables are the English legal origin dummy and other-religion proportion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

Table 14
Estimates of the spatial lag coefficient using inverse distance weights for the spatial lag.

	Voice and accountability	Government effectiveness	Rule of law
Spatial lag (λ)	0.298[c]	0.133	0.322[b]
	(0.003)	(0.315)	(0.043)
French legal origin	−0.232[b]	−0.253[b]	−0.217[a]
	(0.047)	(0.010)	(0.053)
Socialist legal origin	−0.656[c]	−0.593[c]	−0.943[c]
	(0.000)	(0.000)	(0.000)
German legal origin	−0.085	−0.068	0.026
	(0.561)	(0.681)	(0.870)
Scandinavian legal origin	−0.112	0.196	−0.111
	(0.550)	(0.379)	(0.632)
Catholic proportion	0.089	0.084	−0.072
	(0.618)	(0.572)	(0.660)
Muslim proportion	−0.566[c]	−0.045	−0.017
	(0.002)	(0.784)	(0.923)
Protestant proportion	0.106	0.139	0.121
	(0.686)	(0.578)	(0.658)
Ethnic fractionalization	0.291	0.098	−0.030
	(0.147)	(0.592)	(0.879)
Log of oil resources to GDP ratio	−0.091[c]	−0.089[c]	−0.059[b]
	(0.000)	(0.000)	(0.027)
Absolute latitude	0.966[c]	1.161[b]	1.203[b]
	(0.005)	(0.010)	(0.017)
Landlocked	−0.154	0.030	0.005
	(0.228)	(0.780)	(0.966)
Log GDP per capita	0.280[c]	0.551[c]	0.434[c]
	(0.000)	(0.000)	(0.000)
Human capital	0.015[a]	0.010	0.018
	(0.081)	(0.198)	(0.127)
Log population	−0.088[c]	−0.049[b]	−0.085[c]
	(0.001)	(0.037)	(0.006)
Constant	−1.504[b]	−4.505[c]	−2.716[c]
	(0.037)	(0.000)	(0.001)
Adjusted R^2	0.699	0.712	0.689
Number of countries	129	129	129
Spatial autocorrelation coefficient (ρ)	0.071	0.430	0.378
Hansen J-statistic	12.603	22.356	29.057
p value of J-statistic	0.479	0.050	0.006
Observations	1161	1161	1161

Notes: Omitted variables are the English legal origin dummy and other-religion proportion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

weights if country i and country j are located on the same continent and zero weights otherwise, row normalized.[31]

The estimates appear in Table 15. The results on the spatial lag are stronger than in Table 3. A one-standard deviation change in the spatial lag of polity produces 40% of a standard deviation change in the polity score. By contrast, for the strongest effect in Table 3, a one standard deviation increase in the spatial lag of the rule-of-law produces 25% of a standard deviation change in the rule-of-law.

9. Conclusions

We have implemented new methods of estimating the effect of spatial spillovers on institutional development. Our results are based on

[31] This choice of the weighting matrix was dictated by the fact that the contiguity based weighting matrix for 77 countries has far too few positive entries, causing the estimation procedure to fail. We follow many other studies in, essentially, ignoring the problem of missing data when constructing the distance weighting matrix: the weights are constructed assuming the countries with missing institutional data do not exist, an assumption that is rarely justified. This assumption is obviously more tenable when the weights reflect the many countries on a continent. But when the weights reflect a few neighboring countries (as in the analysis for Tables 3–13) it is not tenable, which is why we do not use that assumption for Tables 3–13 and instead employ the missing data econometric approach of Appendix A.

Table 15
Estimates of the spatial lag coefficient using the polity variable.

	POLITY indicator
Spatial lag (λ)	0.610[c]
	(0.000)
French legal origin	−0.421[c]
	(0.000)
Socialist legal origin	−0.867[c]
	(0.000)
German legal origin	−0.147
	(0.478)
Scandinavian legal origin	0.063
	(0.807)
Catholic proportion	−0.101
	(0.487)
Muslim proportion	−0.153
	(0.463)
Protestant proportion	−0.225
	(0.448)
Ethnic fractionalization	−0.084
	(0.609)
Log of oil resources to GDP ratio	−0.038
	(0.136)
Absolute latitude	−0.263
	(0.494)
Landlocked	0.011
	(0.924)
Log GDP per capita, 5 year lag	0.169[a]
	(0.042)
Human capital	0.011
	(0.196)
Log population	0.017
	(0.551)
Constant	−1.279
	(0.067)
Adjusted R^2	0.573
Number of countries	77
Spatial autocorrelation coefficient (ρ)	0.108
Hansen *J*-statistic	16.761
p value of *J*-statistic	0.210
Observations	2926

Notes: Omitted variables are the English legal origin dummy and other-religion proportion. Significance levels are in parentheses.
[a] Significant at the 90% confidence level.
[b] Significant at the 95% confidence level.
[c] Significant at the 99% confidence level.

econometric procedures that properly account for endogenous regressors, for spatial autocorrelation, and for missing values. A new econometric result shows that our estimates are unusually resistant to omitted-variable bias. An extensive set of robustness tests serve to endorse our core results, as well as to expand the horizons of the basic spatial model, for example, by considering asymmetries between the effects of better and worse institutions.

Spatial effects are important.[32] They are statistically significant in our core estimates in Table 3 and in the robustness exercises. They are economically important. For example, the spatial effect of being surrounded by Romania rather than Switzerland would be roughly three times the effect of moving from French legal origin to English legal origin. Because spatial effects spread eventually between all bordering countries, they are found in rather unlikely places: there is a discernible effect on Denmark of the earlier presence of the socialist system far to the East.

Important spatial effects appear in cross-section and over time, when GDP is included and when it is omitted, when many different explanatory variables are added, when we allow for asymmetry between better and worse institutions, when we include spatial lags of explanatory variables as explanatory themselves, and when we look at world experience as far back as the 1960s.

We close in making some suggestions for future research. Our results raise some intriguing issues concerning the design of development policies and aid programs by multilateral organizations. Most obviously, we have uncovered a fundamental externality in world development. The creation of institutions in one country has public good properties, with all of the usual implications for undersupply when the focus is solely on one country's welfare. The spillover effects are large: in the case of the rule-of-law the indirect effects of spurring institutional development are 45% of the direct effects. Thus, our results suggest alternative strategies to spur development in countries that are resistant to institutional change. Perhaps aid to neighbors could break this resistance.

Our results also suggest that research on a system of spatial models might be productive. Our results indicate that the levels of one country's institutions are affected by the institutions of other countries. Moreover, it is generally accepted that institutions affect variables such as GDP and education in the longer run. Therefore, an expansion of our model incorporating equations describing how GDP and educational levels are affected by institutional levels should be of interest.

Appendix A. Spatial estimation methods when there is missing data

The purpose of this section is to describe briefly a method of consistently estimating the standard spatial model, Eq. (1) when there are missing data. For a full presentation, see Kelejian and Prucha (2010). Missing data problems are virtually intrinsic in cross-country economics. In spatial models this leads to extra complications because observations relating to one country contain elements of the data for others through the spatial lag of the dependent and explanatory variables, and the latter are often used as instruments. In the face of such problems, we employ an estimating procedure that explicitly takes into account the structure of missing data rather than simply restricting the sample of countries to those for which full information is available, which could entail disregarding valuable information.

Countries can be grouped into three mutually exclusive and exhaustive sets. In the first set, containing n_1 countries, the dependent variable (the institutional measure) is observed as are all of the explanatory variables appearing on the right-hand-side. The explanatory variables include data on the institutional measure for all countries bordering those n_1 countries. Next, there are some countries that border one or more of the first set of countries but which are not included in the first set because either they are missing data on one or more of the non-spatial regressors or because they border other countries for which institutional measures are not available. Assume there are n_2 countries in this second set. Members of the third set of countries are not in the first set because observations on the regressors of the model corresponding to these countries are not complete, and are not in the second set because they do not border any country in the first set. Countries in this third set border each other and some of the countries in the second set. Let there be n_3 countries in this third set. A more formal description of this scenario is given below.

A.1. Specification

The specifications below account for an incomplete data set. Divide the data vectors and matrices appearing in Eq. (1) into component subvectors and submatrices corresponding to the partition of the observations into the three sets defined above. Thus, $y'_t = (y'_{t,1}, y'_{t,2}, y'_{t,3})$, $X'_t = \left(X'_{t,1}, X'_{t,2}, X'_{t,3}\right)$, $H' = \left(H'_1, H'_2, H'_3\right)$, $u'_t = (u'_{t,1}, u'_{t,2}, u'_{t,3})$, and $\varepsilon'_t = (\varepsilon'_{t,1}, \varepsilon'_{t,2}, \varepsilon'_{t,3})$, where $y_{t,1}$, $y_{t,2}$, and $y_{t,3}$ are $n_1 \times 1$, $n_2 \times 1$, and $n_3 \times 1$ vectors, the $X_{t,i}$, $i = 1,2,3$ are $n_i \times k_1$ matrices, the matrices H_i, $i =$

[32] Seldadyo et al. (2010) reach similar conclusions on the importance of spatial effects in institutional development, using different methods and different data.

1,2,3 are $n_i \times k_2$, and $n = n_1 + n_2 + n_3$. Let W be the $n \times n$ matrix whose i, j^{th} block is W_{ij}, $i,j = 1,2,3$, which is $n_i \times n_j$. Then Eq. (1) becomes:

$$\begin{pmatrix} y_{t,1} \\ y_{t,2} \\ y_{t,3} \end{pmatrix} = \begin{pmatrix} X_{t,1} \\ X_{t,2} \\ X_{t,3} \end{pmatrix} \beta_1 + \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix} \beta_2 + \lambda \begin{pmatrix} W_{11} & W_{12} & 0 \\ W_{21} & W_{22} & W_{23} \\ 0 & W_{32} & W_{33} \end{pmatrix} \begin{pmatrix} y_{t,1} \\ y_{t,2} \\ y_{t,3} \end{pmatrix} + \begin{pmatrix} u_{t,1} \\ u_{t,2} \\ u_{t,3} \end{pmatrix}$$
$$\begin{pmatrix} u_{t,1} \\ u_{t,2} \\ u_{t,3} \end{pmatrix} = \rho \begin{pmatrix} W_{11} & W_{12} & 0 \\ W_{21} & W_{22} & W_{23} \\ 0 & W_{32} & W_{33} \end{pmatrix} \begin{pmatrix} u_{t,1} \\ u_{t,2} \\ u_{t,3} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \\ \varepsilon_{t,3} \end{pmatrix}, t = 1, \ldots, T \tag{A.1}$$

That is, a complete set of observations on $X_{t,1}$ is available, but $X_{t,2}$ or $X_{t,3}$, $t = 1,\ldots,T$ could have missing values. H_1 is observed, but H_2 and H_3 are not necessarily observed. $y_{t,1}$ and $y_{t,2}$ are observed, but $y_{t,3}$ is not necessarily observed. (A.1) follows directly from the categorization of observations into three sets according to the structure of missing data. The economics underlying (A.1) is identical to that underlying Eq. (1). In particular, the 0 submatrices in the partitioned W are a result of the fact that no countries in the third set border those in the first.

A.2. Estimation

In addition to the assumptions made in the text, we assume for the purposes of large sample results that $n_2/n_1 \to 0$. [33] This is reasonable since n_2 corresponds to the number of countries that border the n_1 countries in our set 1, and typically, n_2 will be small relative to n_1.[34] The specification in (A.1) implies, for $t = 1,\ldots,T$

$$y_{t,1} = X_{t,1}\beta_1 + H_1\beta_2 + \lambda\left(W_{11}, W_{22}\right)\begin{pmatrix} y_{t,1} \\ y_{t,2} \end{pmatrix} + u_{t,1} \tag{A.2}$$

$$y_{t,2} = X_{t,2}\beta_1 + H_2\beta_2 + \lambda(W_{21}, W_{22}, W_{23})y_t + u_{t,2} \tag{A.3}$$

$$u_{t,1} = \rho(W_{11}, W_{12})\begin{pmatrix} u_{t,1} \\ u_{t,2} \end{pmatrix} + \varepsilon_{t,1} \tag{A.4}$$

$$u_{t,2} = \rho(W_{21}, W_{22}, W_{23})\begin{pmatrix} u_{t,1} \\ u_{t,2} \\ u_{t,3} \end{pmatrix} + \varepsilon_{t,2} \tag{A.5}$$

Since the data set is complete for (A.2) but not for (A.3), $u_{t,1}$ can be estimated but $u_{t,2}$ cannot.

Following Kelejian and Prucha (2010), the estimation procedure focuses on (A.2) and (A.4). First estimate the parameters of (A.2) by two-stage least squares by stacking the data over $t = 1,\ldots,T$ and using the instruments $P_t = (X_{t,1}, H_1, W_{11}X_{t,1}, W_{11}H_1)$. This method provides consistent parameter estimates. Using these estimates, $u_{t,1}$ $t = 1,\ldots,T$ is estimated from (A.2). The parameter ρ is then estimated by applying the GMM procedure in Kelejian and Prucha (1999) to a variation of the model in (A.4) obtained by treating $q_{t,1}$ as if it were $\varepsilon_{t,1}$:

$$\begin{aligned} u_{t,1} &= \rho W_{11}u_{t,1} + q_{t,1} \\ q_{t,1} &= \varepsilon_{t,1} + \rho W_{12}u_{t,2}, \quad t = 1, \ldots, T \end{aligned}$$

The reason for doing this is that our empirical inferences are based on a large sample theory in which $n_1 \to \infty$ and $n_2/n_1 \to 0$. In this case, given our specification of the weighting matrix,[35] the term $W_{12}u_{t,2}$ is asymptotically negligible.[36]

[33] A complete set of formal assumptions is given in Kelejian and Prucha (2010).

[34] Consider an $n \times n$ checker board of unit squares, with one layer of border units all around. Assuming data on all of the units in the "checker board," in this case the number of border units would be $n_2 = 4n + 4$ and the number of units within these borders would be $n_1 = n^2$ and so $(n/n^2) \to 0$. The values of n_1 and n_2 in our sample depend upon the equation being estimated. The average of these values over the equations being estimated are $\overline{n}_1 = 128$ and $\overline{n}_2 = 64$.

[35] Our specification of the weighting matrix is such that its row and column sums are uniformly bounded in absolute value.

[36] For example, if $n_1 \to \infty$ and $n_2/n_1 \to 0$ it is not difficult to show that $n_1^{-1} q_{t,1}' q_{t,1} \to^P \sigma^2$

Let $\hat{\rho}$ be the estimate of ρ. Then $\hat{\rho}$ is used to transform the model via a spatial variant of the Cochrane–Orcutt procedure and the resulting model is estimated by two-stage least squares. That is, denote the spatial lag appearing in (A.2) as $y_{t,1,2}$:

$$y_{t,1,2} = (W_{11}, W_{12}) \begin{pmatrix} y_{t,1} \\ y_{t,2} \end{pmatrix}$$

and the spatial Cochrane–Orcutt transformations of the variables in (A.2) as

$$\begin{aligned} y_{t,1}^* &= (I - \widehat{\rho}W_{11})y_{t,1} \\ X_{t,1}^* &= (I - \widehat{\rho}W_{11})X_{t,1} \\ H_1^* &= (I - \widehat{\rho}W_{11})H_1 \\ y_{t,1,2}^* &= (I - \widehat{\rho}W_{11})y_{t,1,2} \end{aligned}$$

Then we estimate the parameters of (A.2) by two-stage least squares by regressing $y_{t,1}^*$ on $X_{t,1}^*$ H_1^*, and $y_{t,1,2}^*$ using the same instruments as above, $P_t = (X_{t,1}, H_1, W_{11}X_{t,1}, W_{11}H_1)$. Under reasonable conditions, the resultant parameter estimators are consistent and asymptotically normal and the large sample variance–covariance matrix has the usual form—see, for example, Kelejian and Prucha (2010).

Appendix B. The effect of an omitted common factor

In this appendix we demonstrate that under reasonable conditions, the omission of a "common" factor from a model such as Eq. (1) in the text will not affect the consistency of the estimator of the coefficient of the spatially lagged dependent variable. In a different context, similar conclusions were arrived at by Pace and LeSage (2009).

B.1. Model specifications

Consider the model

$$\begin{aligned} y &= Z\beta + \lambda Wy + u \\ u &= \rho Wu + \varepsilon \end{aligned} \tag{B.1}$$

where y is an $n \times 1$ vector of observations on the dependent variable at time t; Z is an $n \times k$ matrix of observations on exogenous variables some, but not all of which may vary over time; β is a $k \times 1$ parameter vector; W is an $n \times n$ weighting matrix; ε is the innovation to the disturbance term u and λ and ρ are scalar parameters. The model in (B.1) relates to either a single cross section, or a panel—e.g., n can be the product of the number of cross sectional observations and time periods.

Suppose the model in (B.1) is mis-specified because it does not account for an omitted factor. In particular, suppose the true model determining y is

$$\begin{aligned} y &= Z\beta + \lambda Wy + \alpha F + u \\ u &= \rho Wu + \varepsilon \end{aligned} \tag{B.2}$$

where α is a scalar parameter, and F is the $n \times 1$ vector of values on the omitted factor. Below we show that if (B.1) is estimated by our IV procedure, but the true model is (B.2) the estimator of λ will be consistent under reasonable conditions. It will also become evident that the estimator of β will not be consistent.

B.2. A specification of F

In a spatial framework, one would expect the values of the common factor to be spatially correlated, as well as related to the values of the model regressors Z. Assuming that Z contains the constant term, a reasonable specification for F is

$$F = Z\delta + v + bWv \tag{B.3}$$

where δ is a $k \times 1$ parameter vector, and ν is that $n_2 = 4n + 4$ and the number of units within these borders would be part of F which is not linearly related to Z. Note that (B.3) suggests that F will be spatially correlated for two reasons: in general some of the variables in Z as well as $W\nu$ will be spatially correlated.

Assumptions. As in the text, suppose the model in (B.1) is estimated by 2SLS after it is first transformed to correct for spatial correlation. Let Ψ be the matrix of instruments

$$\Psi = \left(Z, WZ, W^2Z\right)_{LI} \tag{B.4}$$

where the subscript LI denotes the linearly independent columns of the indicated matrix.

Our assumptions are

Assumption 1. The elements of Ψ are uniformly bound in absolute value, and

$$Lim_{n\to\infty} n^{-1}\Psi'\Psi = G_{\Psi'\Psi}$$

where $G_{\Psi\Psi}$ is a finite nonsingular matrix.

Assumption 2. (a) $|\lambda|<1$ and $|\rho|<1$. (b) $(I_n - aW)$ is nonsingular for all $|a|<1$. (c) The row and column sums of W and $(I_n - aW)^{-1}$ are uniformly bound in absolute value for all $|a|<1$.

Assumption 3. The elements of ε and ν are, respectively, *i.i.d.* as $N(0,\sigma_\varepsilon^2)$ and $N(0,\sigma_\nu^2)$.[37] Let $A'=(\beta',\lambda)$, $y^*=Wy$, and $P_\Psi=\Psi(\Psi'\Psi)^{-1}\Psi'$. For future reference note that P_Ψ is symmetric idempotent: $P_\Psi P_\Psi=P_\Psi$. Denote the regressor matrix in (B.1) as $Q=(Z,y^*)$, and the regressor matrix which would be used in the 2SLS procedure after the model is first transformed to correct for spatial correlation as $\tilde{Q}$ where:

$$\begin{aligned}\tilde{Q} &= P_\Psi(I_n-\hat\rho W)Q\\ &\equiv\left(\tilde{Z},\tilde{y}^*\right)\\ \tilde{Z} &= P_\Psi[I_n-\hat\rho W]Z\\ \tilde{y}^* &= P_\Psi[I_n-\hat\rho W]y^*\end{aligned} \tag{B.5}$$

Note that since Ψ contains Z and WZ, $\tilde{Z} = P_\Psi[I_n-\hat\rho W]Z = [I_n-\hat\rho W]Z$ so that the exogenous regressor matrix Z is not being viewed as endogenous.

The estimator of $A'=(\beta',\lambda)$ is:

$$\hat{A} = \left(\tilde{Q}'\tilde{Q}\right)^{-1}\tilde{Q}'[I_n-\hat\rho W]y \tag{B.6}$$

Since the true value of y is determined in (B.2), the product of the last two terms in (B.6) is

$$[I_n-\hat\rho W]y = [I_n-\hat\rho W]\left[Z\beta+\lambda y^*+\alpha F+u\right] \tag{B.7}$$

Using matrix partition inversion (Greene, 2003, pp. 126–27), it is straightforward, but tedious, to show that:

$$\begin{aligned}\hat\lambda &= \left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]y\\ &= \left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]\left[Z\beta+\lambda y^*+\alpha F+u\right]\end{aligned} \tag{B.8}$$

where $\tilde{y}._* = (I_n-R_{\tilde{Z}})\tilde{y}^*$, and $R_{\tilde{Z}} = \tilde{Z}\left(\tilde{Z}'\tilde{Z}\right)^{-1}\tilde{Z}'$ so that $(I_n-R_{\tilde{Z}})\tilde{Z}=0$. For future reference note that $R_{\tilde{Z}}$ is symmetric idempotent.

Express the second line of (B.8) as

$$\begin{aligned}\hat\lambda &= \left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]Z\beta + \left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]\lambda y^*\\ &\quad+\left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]\alpha F + \left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]u\end{aligned} \tag{B.9}$$

[37] As will become clear, normality is not needed for our results. It is made only to simplify the presentation.

B.3. The probability limit of $\hat\lambda$

B.3.1. The first term of (B.9)

Consider $\tilde{y}._*'[I_n-\hat\rho W]Z$ in the first line of (B.9). Since $\tilde{y}._*' = \tilde{y}^{*'}(I_n-R_{\tilde{Z}}), \tilde{y}^* = P_\Psi[I_n-\hat\rho W]y^*, \tilde{Z}' = Z'\left[I_n-\hat\rho W'\right]P_\Psi$ and $P_\Psi{}^2 = P_\Psi$

$$\begin{aligned}\tilde{y}._*'[I_n-\hat\rho W]Z &= \tilde{y}^{*'}(I_n-R_{\tilde{Z}})[I_n-\hat\rho W]Z\\ &= y^{*'}\left[I_n-\hat\rho W'\right]P_\Psi(I_n-R_{\tilde{Z}})[I_n-\hat\rho W]Z\end{aligned} \tag{B.10}$$

Note that the product of the middle terms in (B.10) is

$$\begin{aligned}P_\Psi(I_n-R_{\tilde{Z}}) &= P_\Psi\left[I_n-\tilde{Z}\left(\tilde{Z}'\tilde{Z}\right)^{-1}\tilde{Z}'\right] = \left[P_\Psi{}^2-P_\Psi\tilde{Z}\left(\tilde{Z}'\tilde{Z}\right)^{-1}\tilde{Z}'\right]\\ &= \left[P_\Psi{}^2-P_\Psi\tilde{Z}\left(\tilde{Z}'\tilde{Z}\right)^{-1}Z'\left[I_n-\hat\rho W'\right]P_\Psi P_\Psi\right]\\ &= P_\Psi\left[I_n-\tilde{Z}\left(\tilde{Z}'\tilde{Z}\right)^{-1}\tilde{Z}'\right]P_\Psi = P_\Psi(I_n-R_{\tilde{Z}})P_\Psi\end{aligned} \tag{B.11}$$

so that $P_\Psi(I_n-R_{\tilde{Z}}) = P_\Psi(I_n-R_{\tilde{Z}})P_\Psi$. Substituting this expression into the right hand side of (B.10), and recalling that $(I_n-R_{\tilde{Z}})\tilde{Z}=0$

$$y^{*'}\left[I_n-\hat\rho W'\right]P_\Psi(I_n-R_{\tilde{Z}})P_\Psi[I_n-\hat\rho W]Z = y^{*'}\left[I_n-\hat\rho W'\right]P_\Psi(I_n-R_{\tilde{Z}})\tilde{Z} = 0 \tag{B.12}$$

Therefore the first term of (B.9) is zero.

B.3.2. The second term of (B.9)

Going through the same set of manipulations as in (B.10) and (B.11) the second term of (B.9) reduces to

$$\lambda\left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]y^* = \lambda\left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'\tilde{y}._* = \lambda \tag{B.13}$$

B.3.3. The third term of (B.9)

Since $F=Z\phi+\nu+bW\nu$ and by (B.10), $\tilde{y}^{*'}[I_n-\hat\rho W]Z=0$, the third line in (B.9) reduces to

$$\begin{aligned}&\left(\tilde{y}._*'\tilde{y}._*\right)^{-1}\tilde{y}._*'[I_n-\hat\rho W]\alpha F\\ &= \alpha\left(n^{-1}\tilde{y}._*'\tilde{y}._*\right)^{-1}\left(n^{-1}\tilde{y}._*'[I_n-\hat\rho W][\nu+bW\nu]\right)\\ &= \alpha\left(n^{-1}\tilde{y}._*'\tilde{y}._*\right)^{-1}\left(n^{-1}\tilde{y}._*'[I_n-\hat\rho W][I_n+bW]\nu\right)\end{aligned} \tag{B.14}$$

Under further conditions, $\hat\rho \xrightarrow{P} \rho$ (Kelejian and Prucha, 1998, 1999; Mutl and Pfaffermayr, 2011). The results in these studies should also make it clear that $n^{-1}\tilde{y}._*'\tilde{y}._* \xrightarrow{P} G_{\tilde{y}._*'\tilde{y}._*}$ where $G_{\tilde{y}._*'\tilde{y}._*}$ is a finite non-singular matrix. Therefore the third term in (B.9) limits to zero because the last term in parentheses in (B.14) limits to zero, i.e.,

$$n^{-1}\tilde{y}._*'[I_n-\hat\rho W][I_n+bW]\nu \xrightarrow{P} 0 \tag{B.15}$$

To see the limiting result in (B.15), let $\hat\phi = n^{-1}\tilde{y}._*'[I_n-\hat\rho W][I_n+bW]\nu$ and note that

$$\hat\phi = n^{-1}\tilde{y}._*'\nu + bn^{-1}\tilde{y}._*'W\nu - \hat\rho n^{-1}\tilde{y}._*'W\nu - \hat\rho bn^{-1}\tilde{y}._*'W^2\nu \tag{B.16}$$

Since $\hat\rho \to^P \rho$ the limit condition in (B.15) holds if $\phi \to^P 0$ where

$$\phi = n^{-1}\tilde{y}._*'\nu + bn^{-1}\tilde{y}._*'W\nu - \rho n^{-1}\tilde{y}._*'W\nu - \rho bn^{-1}\tilde{y}._*'W^2\nu \tag{B.17}$$

Consider the first term in (B.17), namely $\phi_1 = n^{-1}\tilde{y}._*'\nu$. Since $\tilde{y}._*' = y^{*'}\left(I_n-\hat\rho W'\right)P_\Psi(I_n-R_{\tilde{Z}});\ R_{\tilde{Z}} = \tilde{Z}\left(\tilde{Z}'\tilde{Z}\right)^{-1}\tilde{Z}'$

and, via the first and fifth term in (B.11)

$$P_{\Psi}[I_n - R_{\tilde{Z}}] = P_{\Psi}[I_n - R_{\tilde{Z}}]P_{\Psi}$$

it follows that

$$\begin{aligned}\phi_1 &= n^{-1}y^{*'}\left(I_n - \hat{\rho}W'\right)P_{\Psi}(I_n - R_{\tilde{Z}})P_{\Psi}\nu = n^{-1}\tilde{y}_{.}^{*'}P_{\Psi}\nu \\ &= \left[n^{-1}\tilde{y}_{.}^{*'}\Psi\right]\left[n^{-1}\Psi'\Psi\right]^{-1}\left[n^{-1}\Psi'\nu\right]\end{aligned} \tag{B.18}$$

Consider the last term on the second line of (B.18): $\chi = [n^{-1}\Psi'\nu]$. Note that

$$E(\chi) = 0 \tag{B.19}$$

$$\begin{aligned}E\left(\chi\chi'\right) &= n^{-1}\left[n^{-1}\Psi' E\left(\nu\nu'\right)\Psi\right] \\ &= \sigma_{\nu}^2 n^{-1}\left[n^{-1}\Psi'\Psi\right] \rightarrow 0\end{aligned}$$

since by Assumption 1, $n^{-1}\Psi'\Psi \rightarrow G_{\Psi'\Psi}$, and $n^{-1} \rightarrow 0$. Therefore, by Chebyshev's inequality

$$\chi \xrightarrow{P} 0$$

and so

$$\phi_1 \xrightarrow{P} 0 \tag{B.20}$$

since $n^{-1}\tilde{y}_{.}^{*'}\Psi \rightarrow G_{\tilde{y}_{.}^{*'}\Psi}$, which is a finite matrix under reasonable conditions (Kelejian and Prucha, 1998, 1999).

Using similar manipulations, it is straightforward, but quite tedious to show that the probability limit of the remaining three terms in (B.17) limit to zero. Therefore, the third term in (B.9) limits to zero:

$$\left(\tilde{y}_{.}^{*'}\tilde{y}_{.}^{*}\right)^{-1}\tilde{y}_{.}^{*'}[I_n - \hat{\rho}W]\alpha F \xrightarrow{P} 0 \tag{B.21}$$

B.3.4. The fourth term of (B.9)

Using virtually the same manipulations as those above relating to the third line of (B.9), it is straightforward to show that the fourth line of (B.9) limits to zero. Therefore, under reasonable conditions, the results in (B.9)–(B.18) imply

$$\hat{\lambda} \xrightarrow{P} \lambda \tag{B.22}$$

References

Acemoglu, Daron, Robinson, James A., 2006. Economic Origins of Dictatorship and Democracy. Cambridge, New York.

Acemoglu, Daron, Johnson, Simon, Robinson, James, 2001. The colonial origins of comparative development: an empirical investigation. American Economic Review 91, 1369–1401 (December).

Acemoglu, Daron, Johnson, Simon, Robinson, James A., Yared, Pierre, 2008. Income and democracy. American Economic Review 98, 808–842 (June).

Alesina, Alberto, Perotti, Roberto, 1996. Income distribution political instability, and investment. European Economic Review 40, 1203–1228.

Alesina, Alberto, Devleeschauwer, Arnaud, Easterly, William, Kurlat, Sergio, Wacziarg, Romain, 2003. Fractionalization. Journal of Economic Growth 8, 155–194.

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Boston.

Baltagi, B.H., Song, S.H., Koh, W., 2003. Testing panel data regression models with spatial error correlation. Journal of Econometrics 117, 123–150.

Barro, Robert J., 2006. Religion Adherence Data. http://www.economics.harvard.edu/faculty/barro/data_sets_barro.

Barro, Robert J., Lee, Jong-Wha, 2001. International data on educational attainment: updates and implications. Oxford Economic Papers 53, 541–563.

Beck, Thorsten, Laeven, Luc, 2006. Institutions and growth in transition economies. Journal of Economic Growth 11 (2).

Berkowitz, Daniel, Pistor, Katharina, Richard, Jean-Francois, 2003. Economic development, legality and the transplant effect. European Economic Review 47 (1), 165–195 (February).

Bhagwati, Jagdish, Hudec, Robert (Eds.), 1996. Fair Trade and Harmomization. Prerequisites for Free Trade? Vol. 2 Legal Analysis. MIT Press, Cambridge Mass.

Bikhchandani, Sushil, Hirshleifer, David, Welch, Ivo, 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100 (5), 992–1026 (October).

Bosker, Maarten, Garretsen, Harry, 2009. Economic development and the geography of institutions. Journal of Economic Geography 9 (3), 295–328.

BP, 2006. Statistical Review of World Energy 2006. Available at http://www.bp.com/productlanding.do?categoryId=6842&contentId=7021390.

Casella, Alessandra, 1996. Free trade and evolving standards. In: Bhagwati, Jagdish, Hudec, Robert (Eds.), Fair Trade and Harmonization. : Prerequisites for Free Trade? Vol. 1 Economic Analysis. MIT Press, Cambridge Mass, pp. 119–156.

Casella, Alessandra, Feinstein, Jonathan S., 2002. Public goods in trade on the formation of markets and jurisdictions. International Economic Review 43 (2), 437–462.

CEPII, 2006. CEPII's distance measures and geographical data. Centre d'Etudes Prospectives et d'Informations Internationales (Available at http://www.cepii.fr/anglaisgraph/bdd/distances.htm).

Chong, Alberto, Zanforlin, Luisa, 2000. Law tradition and institutional quality: some empirical evidence. Journal of International Development 12, 1057–1068.

Cliff, A., Ord, J., 1973. Spatial Autocorrelation. Pion, London.

Cliff, A., Ord, J., 1981. Spatial Processes, Models and Applications. Pion, London.

Conley, T., 1996. Econometric Modeling of Cross Sectional Dependence, Ph.D. Thesis. Chicago: University of Chicago.

Conley, T., 1999. GMM estimation with cross sectional dependence. Journal of Econometrics 92, 1–45.

Crafts, N., Venables, A.J., 2001. Globalization in history: a geographical perspective. CEPR Discussion Paper, 3079. Centre for Economic Policy Research (London).

Cressie, N., 1993. Statistics of Spatial Data. Wiley, New York.

Das, D., Kelejian, H., Prucha, I., 2003. Finite sample properties of estimators of spatial autoregressive models with autoregressive disturbances. Papers in Regional Science 82, 1–26.

Easterly, W., Levine, R., 1998. Troubles with the neighbours: Africa's problem, Africa's opportunity. Journal of African Economics 7 (1), 120–142.

Eichengreen, Barry, Leblang, David, 2006. Democracy and globalization. Economics and Politics 20 (3), 289–334.

Fujita, M., Krugman, P., Venables, A., 1999. The Spatial Economy: Cities, Regions and International Trade. MIT Press.

Gallup, John Luke, Sachs, Jeffrey D., Mellinger, Andrew, 1999. Geography and economic development. International Regional Science Review 22 (2), 179–232 (August).

Gallup, John Luke, Sachs, Jeffrey D., Mellinger, Andrew, 2001. Geography Datasets, Center for International Development, Harvard University. http://www.cid.harvard.edu/ciddata/geographydata.htm.

Glaeser, Edward, LaPorta, Rafael, Lopes-de-Silanes, Florencio, Shleifer, Andrei, 2004. Do institutions cause growth? Journal of Economic Growth 9 (3), 271–303.

Graham, Bryan S., 2008. Identifying social interactions through conditional variance restrictions. Econometrica 76 (3), 643–660 (May).

Grajzl, Peter, Dimitrova-Grajzl, Valentina, 2009. The choice in the lawmaking process: legal transplants vs. indigenous law. Review of Law and Economics 5 (1), 615–660.

Greene, William, 2003. Econometric Analysis, 5th Edition. Prentice Hall.

Grilli, E., 1997. EU trade and aid policies towards the LDCs and CEECs. In: El-Agra, Ali M. (Ed.), Economic Integration Worldwide. MacMillan Press, London, pp. 135–174.

Guiso, L., Sapienza, P., Zingales, L., 2006. Does culture affect economic outcomes? Journal of Economic Perspectives 20 (2), 23–48.

Hall, Robert, Jones, Charles, 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114 (1), 83–116.

Heston, Alan, Summers, Robert, Aten, Bettina, 2009. Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Heston, Alan, Summers, Robert, Aten, Bettina, 2011. Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Horn, R., Johnson, C., 1985. Matrix Analysis. Cambridge University Press, Cambridge.

Kaufmann, Daniel, Kraay, Aart, 2002. Growth without governance. Economia 3 (1), 169–215 (Fall).

Kaufmann, Daniel, Kraay, Aart, Mastruzzi, Massimo, 2009. Governance matters VIII: aggregate and individual governance indicators for 1996–2008. The World Bank, June. World Bank Policy Research Working Paper No. 4978 (Available at SSRN: http://ssrn.com/abstract=1424591).

Keefer, P., Knack, S., 1996. Political stability and economic stagnation. In: Borner, S., Paldam, M. (Eds.), The Political Dimension of Economic Growth. MacMillan, Basingstoke.

Kelejian, H.H., Prucha, I.R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. Journal of Real Estate Finance and Economics 99–121.

Kelejian, H.H., Prucha, I.R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review 40, 509–533.

Kelejian, H.H., Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104, 219–257.

Kelejian, H.H., Prucha, I.R., 2004. Estimation of simultaneous systems of spatially interrelated cross sectional equations. Journal of Econometrics 118, 27–50.

Kelejian, H., Prucha, I., 2010. Spatial models with spatially lagged dependent variables and incomplete data. Journal of Geographical Systems 50 (2), 592–614 (May).

Kelejian, H., Tavlas, G., Hondroyiannis, 2006. A spatial modelling approach to contagion among emerging economies. Open Economies Review 17 (4/5), 423–443.
Krugman, Paul, 1993. First nature, second nature, and metropolitan location. Journal of Regional Science 33.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1998. Law and finance. Journal of Political Economy 106, 1113–1155.
La Porta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, Vishny, Robert, 1999. The quality of government. Journal of Law, Economics and Organization, vol. 15(1). Oxford University Press, pp. 222–279 (April).
Lee, L.F., 2003. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. Econometric Reviews 22, 307–335.
Lee, L., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica 72, 1899–1926.
Leebron, David, 1996. Lying down with procrustes: an analysis of harmomization. In: Bhagwati, Jagdish, Hudec, Robert (Eds.), Fair Trade and Harmomization. : Prerequisites for Free Trade? Volume 1 Economic Analysis. MIT Press, Cambridge Mass, pp. 41–118.
LeSage, James P., Fischer, Manfred M., 2008. Spatial growth regressions: model specification, estimation and interpretation. Spatial Economic Analysis 3 (3), 275–304.
LeSage, James, Pace, R. Kelley, 2009. Introduction to Spatial Econometrics. CRC Press, Taylor and Frances Group, New York.
Mobarak, Ahmed Mushfiq, 2005. Democracy, volatility, and economic development. The Review of Economics and Statistics 87 (2), 348–361.
Mukand, S.W., Rodrik, D., 2005. In search of the holy grail: policy convergence, experimentation, and economic performance. American Economic Review 95 (1), 374–383.
Murrell, Peter, 2008. Institutions and transition, In: Durlauf, Steven N., Blume, Lawrence E. (Eds.), The New Palgrave Dictionary of Economics, 2nd Edition.
Murrell, Peter, Dunn, Karen Turner, Korsun, Georges, 1996. The culture of policy-making in the transition from socialism: price policy in Mongolia. Economic Development and Cultural Change 45 (1), 175–194.
Mutl, J., Pfaffermayr, M., 2011. The Hausman test in a Cliff and Ord panel model. The Econometrics Journal 14, 48–76.
North, Douglass C., 1990. Institutions, Institutional Change and Economic Performance. Cambridge University Press, New York.
North, Douglass, Thomas, Robert, 1970. An economic theory of the growth of the Western world. The Economic History Review 23 (1), 1–17.
Pace, R. Kelley, LeSage, James P., 2009. Omitted variable bias of OLS and Spatial Lag Models. In: Pez, A., LeGallo, J., Buliung, R., DallErba, S. (Eds.), Spatial Analysis: Theory and Computation, and Thematic Applications. Springer, Berlin.
Persson, Torsten, Tabellini, Guido, 2006. Democratic capital: the nexus of political and economic change. CEPR Discussion Paper No. 5654.
Pinske, J., Slade, M.E., Brett, C., 2002. Spatial price competition: a semiparametric approach. Econometrica 70, 1111–1153.
Polity IV Project, 2011. Political Regime Characteristics and Transitions, 1800–2009. http://www.systemicpeace.org/inscr/inscr.htm.
Qian, Y., Roland, G., 1998. Federalism and the soft budget constraint. American Economic Review 88, 1143–1162.
Rodrik, Dani, Subramanian, Arvind, Trebbi, Francesco, 2004. Institutions rule: the primacy of institutions over geography and integration in economic development. Journal of Economic Growth 9 (2).
Rosett, Arthur, 1992. Unification harmonization, restatement, codification, and reform in international commercial law. American Journal of Comparative Law 40, 683–697.
Sala-i-Martin, Xavier, Doppelhofer, Gernot, Miller, Ronald I., 2004. Determinants of long-term growth: A Bayesian Averaging of Classical Estimates (BACE) approach. American Economic Review 94 (4), 813–835 (September).
Seldadyo, Harry, Paul Elhorst, J., de Haan, Jakob, 2010. Geography and governance: does space matter? Papers in Regional Science 89 (3), 625–640 (August).
Simmons, B.A., Elkins, Z., 2004. The globalization of liberalization: policy diffusion in the international political economy. American Political Science Review 98 (1), 171–189.
Spilimbergo, Antonio, 2006. Democracy and Foreign Education. CEPR Discussion Paper No. 5934 (Available at SSRN: http://ssrn.com/abstract=954588).
United Nations, 2007. National Accounts Main Aggregates Database. Available at http://unstats.un.org/unsd/snaama/Introduction.asp.
Ward, Michael D., Gleditsch, Kristian Skrede, 2008. Spatial Regression Models. Sage, Los Angeles.
Winters, A., 1993. Expanding EC membership and association accords. In: Anderson, K., Blackhurst, R. (Eds.), Regional Integration and the World Trading System. Harvester Wheatsheaf, pp. 104–125.
World Bank, 2005. World Development Indicators 2005 . (Washington D.C.).