‘Dual’ gravity: Using spatial econometrics to control for multilateral resistance

Kristian Behrens* Cem Ertur† Wilfried Koch‡

February 13, 2009

Abstract

We derive a quantity-based gravity equation system that yields an estimating equation with both cross-sectional interdependence and spatially lagged error terms, and which can be estimated using spatial econometric techniques. To illustrate our methodology, we apply it to the Canada-U.S. data set used previously, among others, by Anderson and van Wincoop (2003) and Feenstra (2002, 2004). Our key result is to show that controlling directly for spatial interdependence across trade flows, as suggested by theory, reduces border effects by capturing ‘multilateral resistance’. Using a spatial autoregressive moving average specification, we find that border effects between the U.S. and Canada are smaller than suggested by previous studies: about 7 for Canadian provinces and about 1.3 for U.S. states. Heterogeneous coefficient estimations further reveal that border effects and distance elasticities of trade flows vary widely across provinces and states.

Keywords: gravity equation system; border effects; spatial econometrics

JEL Classification: C31; F12; R12

*Department of Economics, Université du Québec à Montréal, Canada; CIRPÉE, Canada; and CORE, Université catholique de Louvain, Belgium. E-mail address: `behrens.kristian@uqam.ca`

†*Corresponding author:* LEO, Faculté de Droit d’Economie et de Gestion, Université d’Orléans, Rue de Blois – BP 6739, 45067 Orléans Cedex2, France. E-mail address: `cem.ertur@univ-orleans.fr`

‡LEG, Université de Bourgogne, France. E-mail address: `kochwilfriedfr@aol.com`

"Improved econometric techniques based on careful consideration of the error structure are likely to pay off. Recent literature on spatial econometrics [...] may be helpful." (Anderson and van Wincoop, 2004, p.713)

1 Introduction

The gravity equation offers a remarkably good framework for analyzing bilateral trade flows. Having been derived from various formal trade models under a wide range of modeling assumptions, it is nowadays firmly rooted in mainstream economic theory and has, as such, become an essential part of every applied trade theorist's toolbox.[1] Despite its wide applicability and excellent fit, the gravity equation suffers from several well-know and from several less well-known shortcomings. The former category comprises mainly empirical issues, such as the treatment of zero trade flows, the construction of own absorption, the measurement of internal distances, and concerns about the theoretical plausibility of various parameter estimates. These problems have been extensively discussed elsewhere in the literature (see Anderson and van Wincoop, 2004, pp.729-733, for a recent overview). We instead focus on one of the less well-known problems that plague the gravity equation: how to take into account the interdependence between trade flows and estimate, as consistently as possible, the general equilibrium system?

Anderson and van Wincoop (2003) have recently argued that dealing with the interaction structure across regions is important when estimating gravity equation systems. They show that the proper inclusion of *multilateral resistance* terms, i.e., terms which capture the fact that bilateral trade flows do not only depend on bilateral trade barriers but also on trade barriers across all trading partners, is crucial for the results one obtains. In other words, bilateral predictions do not readily extend to a multilateral world because of complex interactions linking all the trading partners. Although such a finding is hardly surprising in a general equilibrium setting, it has been largely neglected until now in applied work. Interdependence has, however, to be somehow controlled for in the gravity equation to obtain consistent estimates. Some previous studies aim at doing so by including ad hoc remoteness indices, even if there is no theoretical foundation to such an approach. Other studies try to capture interdependence among trade flows by including origin- and destination-specific fixed effects. The somewhat worriesome feature of both of these approaches

[1] For various instances of the gravity equation see, e.g., Anderson (1979), Helpman and Krugman (1985), Feenstra (2002, 2004), Anderson and van Wincoop (2003), Helpman *et al.* (2008), Melitz and Ottaviano (2008), and Chaney (2008).

lies in the implicit assumption that trade flows between two trading partners are *independent from what happens to the rest of the trading world*. This is clearly a very strong assumption that is not likely to hold and, therefore, may lead to biased estimates of the gravity equation.

This paper offers a contribution to the rapidly expanding literature on the theory-based estimation of gravity equation systems. Building upon the observation that the structural estimation of a CES-based gravity equation system crucially hinges on the correct treatment of the *unobservable* price indices, our modeling strategy consists in taking a 'dual' approach that relies on *observable* trade flows only. More concretely, we derive a gravity equation from the quantity-based version of the CES model by exploiting the property that the price indices are themselves implicit functions of trade flows. Using a linearization of the resulting equilibrium system allows us to recover, quite naturally, an econometric specification in which bilateral trade flows between two regions depend on trade flows involving all the other trading partners. Put differently, the model has a *spatial autoregressive structure in trade flows*. Since varieties are substitutes in the CES model, the sales from any region to a market negatively depend on the sales from the other regions to that same market, which themselves depend on the whole distribution of bilateral trade barriers. Controlling for interdependence with the help of spatial econometric techniques then amounts to control for multilateral resistance and yields improved estimates of the gravity equation. Although the idea of applying spatial econometrics to gravity equation systems has been in the air recently we provide, to the best of our knowledge, the first attempt at doing so starting from a theory-based trade model.[2] On top of controlling directly for cross-sectional interdependence across trade flows, our approach has several additional advantages. First, it does not enforce a strict pattern of spatial interdependence as the model by Anderson and van Wincoop (2003) does. Our approach is, therefore, more robust to potential misspecifications concerning the form of the spatial correlations. Second, it reveals that all coefficients, including the spatial autoregressive ones for both the spatially lagged endogenous variable and the error terms, are generally region-specific. Hence, the full

[2]See Anselin and Bera (1998) for an overview of spatial econometrics. The asymptotic properties of some spatial estimators have been derived by Kelejian and Prucha (1998) and Lee (2004). Spatial econometric techniques have been applied to a wide range of topics including growth and convergence (Moreno and Trehan, 1997; Ertur and Koch, 2007), spatial patterns of foreign direct investment (Bloningen *et al.*, 2004, 2005), retail price competition (Pinkse *et al.*, 2002), and interactions between local governments (Case *et al.*, 1993; Brueckner, 1998). To the best of our knowledge there are, until now, no applications to trade and the gravity equation, which may be due to the fact that origin-destination interdependencies have not yet been much developed in the spatial econometrics literature (yet, see LeSage and Pace, 2006, for an extension of the standard theory to origin-destination interactions in a migration-based context).

estimation of the model requires the use of local techniques that can deal with parameter heterogeneity across observations. Our approach allows us to do so and provides statistical inference on region-specific border effects and distance elasticities. Last, our procedure allows us to model more carefully the error structure, thereby controlling for cross-sectional correlations that may arise in the error terms. This last point has not received much attention until now.

We illustrate our methodology by applying it to the well-known Canada-U.S. dataset used by, among others, Anderson and van Wincoop (2003) and Feenstra (2002, 2004). We first estimate a baseline specification where all coefficients are constrained to be identical across regions ('homogeneous coefficient case'). Doing so simplifies the econometric implementation and yields results that are comparable to those in the literature. We then provide estimates for a model with region-specific coefficients, except for the spatial autoregressive ones which we assume to be country specific ('heterogeneous coefficient case').

Our key results may be summarized as follows. First, we show that there remains a significant amount of spatial autocorrelation in the OLS residuals of the gravity equation, even when including origin- and destination-specific fixed effects. Put differently, OLS estimates are at best inefficient and at worst inefficient and biased, because the fixed effects fail to capture the spatial interdependence among trade flows. This finding vindicates the use of spatial econometric techniques and a more careful modeling of the error structure. Second, we estimate the homogeneous coefficient specification of the model and show that, as predicted by theory, there is significant negative spatial autocorrelation between trade flows.[3] Once this autocorrelation is controlled for, the border effects between the U.S. and Canada are shown to be smaller than in previous studies: about 7 for Canadian provinces and about 1.3 for U.S. states. Our approach reveals how spatial econometrics can deal with the 'border effect puzzle' by controlling for multilateral resistance in a novel way. Last, we provide results for the heterogeneous coefficient specification of the model under the restriction of country-specific autoregressive parameters. Our estimates reveal significant variations in both distance elasticities and border effects across provinces and states. Interestingly, the theoretical prediction that distance elasticities are negatively linked to own market size is supported by our

[3]It is worth pointing out that, although the possible existence of such a negative spatial autocorrelation is acknowledged in the literature, it is usually considered more a 'textbook case' than of empirical relevance. The reason is that, to the best of our knowledge, negative spatial autocorrelation is never derived from a structural model and, therefore, hard to interpret convincingly. This contrasts starkly with theory where the existence of negative interdependence in, for example, exchange networks is fairly well known and has been established experimentally (see, e.g., Bonacich, 1987).

empirical results. Furthermore, border effects are the largest for provinces and states close to the border. While border effects for most U.S. states are small and statistically insignificant, those for Canadian provinces are generally larger and statistically significant.

The remainder of the paper is organized as follows. Section 2 presents the model and derives our gravity equation system. We then propose, in Section 3, a spatial econometric estimating equation derived from the linearized version of the theoretical model. We also show how we can theoretically decompose and retrieve the border effects. In Section 4, we first present our data and briefly review previous estimation methods. We then discuss our empirical results for both the homogenous and the heterogeneous coefficient cases. Section 5 concludes.

2 A 'dual' gravity model

We begin by presenting a novel way of deriving a system of gravity equations that does not depend on unobservable price indices, yet encapsulates the general equilibrium interdependencies of the full trading system. The basic idea is to get rid of prices and price indices by using the inverse demand functions and by exploiting the fact that price indices depend on trade flows. Doing so allows us to obtain an implicit equation system *that depends on observables only* and that can be estimated with spatial econometric techniques. In a nutshell, whereas Anderson and van Wincoop (2003) derive a gravity equation system with nonlinear constraints in the unobservable price indices, we derive an 'unconstrained' gravity equation in which the observable trade flows are spatially autocorrelated. To do so, we build upon a CES trade model à la Dixit and Stiglitz (1977) and Krugman (1980), with an arbitrary number n of regions. Each region i is endowed with L_i consumers/workers, who each supply inelastically one unit of labor. Labor is the only production factor so that L_i stands for both the size of, and the aggregate labor supply in, region i.[4]

[4]Our structural model features neither intermediate goods nor multiple factors of production (however, intermediate goods could be incorporated in the CES model following Ethier, 1982). Hence, we cannot account for several phenomena like multi-stage and multi-state processing, the co-location of up- and down-stream firms, vertical production chains, and hub-and-spoke entrepôt trade. One may thus wonder why we need a structural model when we could use instead a reduced form gravity model without structural interpretation. The key reason we opt for the structural approach is that it is to date unclear how to model the spatial interdependence across trade flows in a reduced-form model (except by specifying a priori an ad hoc structure). Using a structural model clarifies how spatial interdependence may appear in equilibrium and how trade flows are likely to be linked.

2.1 Preferences

All consumers have identical preferences over a continuum of horizontally differentiated product varieties. A representative consumer in region j solves the following problem:[5]

$$\max \quad U_j \equiv \sum_i \int_{\Omega_i} q_{ij}(v)^{\frac{\sigma-1}{\sigma}} \mathrm{d}v \qquad \text{subject to} \qquad \sum_i \int_{\Omega_i} q_{ij}(v) p_{ij}(v) \mathrm{d}v = y_j,$$

where $\sigma > 1$ denotes the constant elasticity of substitution between any two varieties; y_j stands for individual income in region j; $p_{ij}(v)$ and $q_{ij}(v)$ denote the consumer (i.e., the delivered) price and per capita consumption of variety v produced in region i; and where Ω_i denotes the set of varieties produced in region i. Since varieties produced in the same region are assumed to be symmetric, in what follows we alleviate notation by dropping the variety index v. Let m_k stand for the measure of Ω_k (i.e., the mass of varieties produced in region k). It is readily verified that the aggregate inverse demand functions for each variety are then given by

$$p_{ij} = \frac{Q_{ij}^{-1/\sigma}}{\sum_k m_k Q_{kj}^{1-1/\sigma}} Y_j, \tag{1}$$

where $Q_{ij} \equiv L_j q_{ij}$ denotes the aggregate demand in region j for a variety produced in region i; and where $Y_j \equiv L_j y_j$ stands for the aggregate income in region j.

2.2 Technology

Each firm produces only a single product variety. Thus, there is a one-to-one correspondence between varieties and firms, i.e., m_k also stands for the mass of firms operating in region k. To produce q units of output requires $cq + F$ units of labor, where c is the marginal and F is the fixed input requirement. Shipping varieties *both within and across* regions is costly. More precisely, shipping one unit of any variety between regions j and k requires to dispatch $\tau_{jk} > 1$ units from the region of origin, while the rest 'melts away' in transportation (the so-called 'iceberg' cost). It is worth pointing out at this stage that we need not make a priori any assumption on either the value of intraregional trade costs τ_{ii}, nor on symmetry of trade costs across regions.

[5]Following previous work by Anderson (1979), Anderson and van Wincoop (2003) derive a gravity equation from a CES expenditure system with goods that are differentiated by region of origin and the supply of which is fixed. We instead rely on the monopolistic competition specification with free entry, since it allows us to control for factor price differences in the empirical part. Note also that including home bias parameters in the utility function is irrelevant for the empirical analysis as they cannot be isolated from population size and trade costs.

A firm located in country j maximizes its profit, given by

$$\pi_j = \sum_k (p_{jk} - cw_j\tau_{jk}) Q_{jk} - Fw_j,$$

with respect to the quantities Q_{jk} and subject to the inverse demand schedule (1). Because price and quantity competition are equivalent when there is a continuum of firms, the profit maximizing prices display a constant markup over marginal cost: $p_{jk} = \tau_{jk}p_j$, where $p_j \equiv cw_j\sigma/(\sigma-1)$ stands for the producer (i.e., the mill) price in region j. Free entry and exit drive profits to zero, which implies that each firm must produce the break-even quantity

$$\sum_k \tau_{jk}Q_{jk} = \frac{F(\sigma-1)}{c} \equiv \overline{Q}, \tag{2}$$

irrespective of the region j it is located in.[6]

2.3 Equilibrium

To derive the gravity equation system requires to determine the value of trade flows from i to j. The latter is given by $X_{ij} \equiv m_i p_{ij} Q_{ij}$ which, using (1) can be expressed as follows:

$$X_{ij} = m_i \frac{Q_{ij}^{1-1/\sigma}}{\sum_k m_k Q_{kj}^{1-1/\sigma}} Y_j. \tag{3}$$

Aggregate income constraints, the equilibrium prices, and the zero profit condition (2) then imply that

$$Y_i = \sum_k m_i p_{ik} Q_{ik} = m_i p_i \overline{Q}.$$

Solving for $m_i = Y_i/(p_i\overline{Q})$ and substituting into (3), we can eliminate the unobservable mass of firms to obtain

$$X_{ij} = Y_i Y_j \frac{Q_{ij}^{1-1/\sigma}}{\sum_k \frac{p_i}{p_k} Y_k Q_{kj}^{1-1/\sigma}}. \tag{4}$$

By definition of the trade flows X_{ij} and the mass of firms m_i, it must be that

$$Q_{ij} = \frac{X_{ij}}{m_i p_{ij}} = \frac{X_{ij}\overline{Q}}{Y_i \tau_{ij}}. \tag{5}$$

[6] Strictly speaking, this equilibrium condition only holds for interior equilibria. In what follows, we focus exclusively on such equilibria as they are the empirically relevant ones for our subsequent analysis.

Plugging (5) into (4) and simplifying then yields

$$X_{ij} = Y_i Y_j \frac{\left(\frac{X_{ij}\overline{Q}}{Y_i \tau_{ij}}\right)^{1-1/\sigma}}{\sum_k \frac{p_i}{p_k} Y_k \left(\frac{X_{kj}\overline{Q}}{Y_k \tau_{kj}}\right)^{1-1/\sigma}} = Y_j \frac{\tau_{ij}^{1/\sigma-1}\left(\frac{X_{ij}}{Y_i}\right)^{1-1/\sigma}}{\sum_k \frac{L_k}{L_i} \tau_{kj}^{1/\sigma-1}\left(\frac{X_{kj}}{Y_k}\right)^{1-1/\sigma}}, \tag{6}$$

where we have used the equilibrium relationship $p_i/p_k = w_i/w_k$ and the aggregate income constraint $w_i = Y_i/L_i$. Expression (6) can be rewritten as follows:

$$X_{ij} = Y_j^{\sigma} \left[\sum_k \frac{L_k}{L_i} \left(\frac{\tau_{kj}}{\tau_{ij}} \frac{Y_k}{Y_i} \right)^{1/\sigma-1} X_{kj}^{1-1/\sigma} \right]^{-\sigma} \quad \forall i, j \tag{7}$$

which defines a system of equations capturing the interdependence of all trade flows going towards region j. To close the general equilibrium system finally requires to impose the aggregate income constraints

$$Y_i - \sum_k X_{ik} = 0, \quad \forall i. \tag{8}$$

As can be seen from expressions (7) and (8), the GDP $Y_i \equiv f_i(\mathbf{L}, \sigma, \mathbf{T})$ of each region can generally be expressed as a function of technology f_i, the vector of endowments $\mathbf{L} = (L_i)$, preferences σ, and the matrix of trade frictions $\mathbf{T} = (\tau_{ij})$. As can be further seen from (7) and (8), all trade flows X_{ij} (including own absorption X_{ii}) are linked in equilibrium, both directly (since varieties are substitutes) and indirectly (via the aggregate income constraints). Formally, such a system can be represented as a directed graph, where the X_{ij} are the flows between regions (the 'nodes') along trading routes (the 'edges'), and where the Y_i play the role of flow conservation constraints. Figure 1 illustrates the equilibrium relationships in a simple three-region world.

Insert Figure 1 about here.

As should be clear from Figure 1, any comparative static exercise must take into account the equilibrium interdependence of the trade flows and GDPs. This seems especially relevant for gravity equations, since the estimated coefficients are usually interpreted as providing precisely these comparative static results. Yet, taking into account all the equilibrium relationships unfortunately yields a system that does not allow for any tractable empirical specification. In what follows, we therefore only control for a part of the interdependence, namely that arising between the different trade flows X_{ij}. We thus stick closely to the existing literature which traditionally considers that regional GDPs are exogenous to the analysis.[7]

[7]It has been realized since a long time that the full interdependence of the system should be somehow taken

3 Econometric specification

We now propose an econometric method for estimating the gravity equation system (7). This method builds on the foregoing observation that trade flows are spatially interdependent and that this interdependence needs to be somehow taken into account. Our approach draws quite naturally on spatial econometric techniques, which are precisely designed to deal with cross-sectional interdependence in both variables and error terms. When compared to other estimation methods, we believe that ours offers essentially four advantages:

1. it directly accounts for cross-sectional interdependence among trade flows, as implied by our theoretical model;

2. it uses a more careful modeling of the error structure, thereby controlling for possible cross-sectional correlations in the error terms;

3. it reveals that all coefficients, including the distance elasticities and border effects, are generally region-specific (see Anderson and Smith, 1999; Helpman *et al.*, 2008) and allows for statistical inference on estimated border effects and distance elasticities;

4. it allows us to sidestep the thorny issue of how to chose σ by working with a linearized version of the equilibrium system.[8]

into consideration. Bergstrand (1985, p.474) argues that "the gravity equation is a reduced form from a partial equilibrium subsystem of a general equilibrium model". Anderson and Smith (1999, p.29) claim that "SUR is an appropriate econometric technique", yet they do not estimate the gravity equation using that technique because of problems with handling own absorption X_{ii}. As argued by Bergstrand (1985), exogeneous GDPs amount to assuming that regions are small enough so that they cannot affect GDPs by any one trade flow. Though this may be the case for the X_{ij}, the same does not hold true for the X_{ii}. These constitute indeed a quite large share of GDP for most regions, so that changes in them are bound to affect regional GDPs. A few recent papers deal with gravity equations and endogenous GDPs (see, e.g., Balistreri and Hillberry, 2007; Behrens *et al.*, 2008).

[8]The usual problem in CES-based models is that "without knowing σ we cannot infer the size of the trade barrier, and without knowing the size of the barrier we cannot infer σ" (Hummels, 2001, p.9). Estimation results for σ depend both on the level of aggregation and the estimation method, and vary widely. For example, Hanson (2005), using aggregate U.S. data, obtains about 7 with non-linear least squares and about 2 with GMM. Estimates in Hummels (2001) vary from 2 to 5.26. Using extremely disaggregated data, Broda and Weinstein (2006) estimate several thousand elasticities of substitution, which range, depending on the industry and the level of aggregation, from 1.3 (telecommunication equipments) to 22.1 (crude oil).

We now linearize the model of Section 2, derive a spatial econometric specification and discuss in more detail the error structure.

3.1 Linearization and matrix form

We start with the theory-based specification of the model. Taking equation (7) in logarithmic form, we readily obtain:[9]

$$\ln X_{ij} = \sigma \ln Y_j - \sigma \ln \left[\sum_k \frac{L_k}{L_i} \left(\frac{\tau_{kj} Y_k}{\tau_{ij} Y_i} \right)^{\frac{1}{\sigma}-1} X_{kj}^{1-\frac{1}{\sigma}} \right] \equiv f(\sigma). \tag{9}$$

Clearly, there is spatial interdependence as X_{ij} depends negatively on the nominal sales of the other regions in market j. To obtain a specification that can be estimated with the help of spatial econometric techniques, we linearize f around $\sigma = 1$.[10] As shown in Appendix A, doing so yields the following equation:

$$\begin{aligned} \ln X_{ij} \;=\; & \sigma \sum_k \frac{L_k}{L} \ln \frac{L_k}{L} + \sigma \ln Y_j - (\sigma - 1) \left[\ln \tau_{ij} - \sum_k \frac{L_k}{L} \ln \tau_{kj} \right] \\ & -\sigma \left[\ln w_i - \sum_k \frac{L_k}{L} \ln w_k \right] + \left[\ln Y_i - \sum_k \frac{L_k}{L} \ln Y_k \right] - (\sigma - 1) \sum_k \frac{L_k}{L} \ln X_{kj}, \end{aligned} \tag{10}$$

where $L \equiv \sum_k L_k$ denotes the total population. Expression (10) reveals the essence of spatial interdepence in the gravity equation system: *the trade flow X_{ij} from region i to region j also depends on all the trade flows from the other regions k to region j.*

Several comments are in order. First, as expected, trade flows from i to j increase with destination GDP Y_j. Yet, by contrast to more standard gravity equations, the coefficient on partner GDP exceeds unity. Second, trade flows from i to j are affected by relative trade barriers, as measured by the deviation of bilateral trade barriers τ_{ij} from the population weighted average

[9] As recently argued by Santos Silva and Tenreyro (2006), the log-linearization may bias some estimates in the presence of heterogeneity. Yet, the Poisson pseudo maximum likelihood estimator that these authors suggest cannot be readily implemented in our specification with lagged endogenous variables. Whether the omission of spatial interdependence is preferable to the log-linearization of the estimating equation is unclear and beyond the scope of this paper. Yet, it is worth emphasizing that our weight matrix gives more weight to larger regions (as measured by its population share $L_i / \sum_k L_k$) which, as argued by Santos Silva and Tenreyro (2006) is desirable because trade data for larger regions is usually more accurate.

[10] When compared to Anderson and van Wincoop (2003), our approach has the potential drawback to require a linearization in order to obtain an exploitable econometric specification. Similar linearizations are commonly used in empirical growth and convergence, as well as in estimating CES production functions (see, e.g., Kmenta, 1967).

(third term in brackets). Put differently, *relative accessibility matters.* Third, trade flows from i to j are negatively affected by wages w_i in the origin region, measured again by the deviation from the population weighted average (fourth term in brackets). Above-average wages raise production costs and make region i's firms less competitive in market j. Fourth, trade flows from i to j increase in own GDP Y_i, yet again only as measured by the deviation from the population weighted average (fifth term in brackets). The intuition is that a larger region hosts more firms, because of the 'home market effect', yet that the presence of other large regions reduces that mass by providing equally attractive export bases. Last, trade flows from i to j decrease with the value of sales X_{kj} from any third region k into the destination market, because varieties are substitutes. This effect is stronger the closer substitutes the varieties are (i.e., the larger the value of σ). Since spatial interdependence is captured by the *spatial autoregressive coefficient* in our estimating equation, this coefficient may be interpreted as a measure of 'spatial competition' encapsulating both aspects of market power and consumer preference for diversity.[11] It is worth pointing out that when $\sigma \to 1$, the linear approximation of the model improves but that the spatial autoregressive term disappears, whereas the approximation gets worse when σ is large but there is more spatial interdependence. When σ gets very large, trade flows fall to zero, which corresponds to an extreme form of spatial interdependence where trade frictions almost completely inhibit interregional exchange.

To make notation more compact, we recast (10) into matrix form as follows:

$$\mathbf{X} = \sigma\zeta\mathbb{1} + \sigma\mathbf{Y_d} + \underbrace{(\mathbf{I}-\mathbf{W})\mathbf{Y_o}}_{\equiv\, \widetilde{\mathbf{Y}}_\mathbf{o}} - (\sigma-1)\underbrace{(\mathbf{I}-\mathbf{W})\tau}_{\equiv\, \widetilde{\tau}} - \sigma\underbrace{(\mathbf{I}-\mathbf{W})\mathbf{w}}_{\equiv\, \widetilde{w}} - (\sigma-1)\mathbf{W}\mathbf{X}. \tag{11}$$

In expression (11), we have defined:

$\mathbf{X} \equiv (\ln X_{ij})$ as the $n^2 \times 1$ vector of the logarithms of trade flows;

$\zeta \equiv \sum_k \frac{L_k}{L} \ln \frac{L_k}{L}$, which is the entropy of the population distribution;

$\mathbb{1}$ as the $n^2 \times 1$ vector whose components are all equal to 1;

$\mathbf{Y_d} \equiv (Y_j)$ as the $n^2 \times 1$ vector of the logarithms of destination GDPs;

$\mathbf{I}$ as the $n^2 \times n^2$ identity matrix;

[11] We do not seek to disentangle market power from preference for diversity in our model as both are observationally equivalent by reducing trade flows between i and j. See Benassy (1996) for further discussion on how to disentangle market power from preference for diversity.

$\mathbf{W}$ as the $n^2 \times n^2$ *spatial weight matrix*, whose contents will be made precise below;

$\mathbf{Y_o} \equiv (Y_i)$ as the $n^2 \times 1$ vector of the logarithms of origin GDPs;

$\tau \equiv (\ln \tau_{ij})$ as the $n^2 \times 1$ vector of the logarithms of trade costs;

$\mathbf{w} \equiv (\ln w_i)$ as the $n^2 \times 1$ vector of the logarithms of origin wages.

Note from expressions (10) and (11) that all variables superscripted with a tilde are measured as deviations from their population weighted averages. We stick to this notation in the remainder of the paper to ease the exposition. Some simple algebraic manipulations show that the structure of the *theory-based spatial weight matrix* is given by: $\mathbf{W} = [\mathbf{S}\,\mathrm{diag}\,(\mathbf{L})] \otimes \mathrm{I_n}$, where $\mathbf{S}$ is the $n \times n$ matrix whose elements are all equal to 1; where $\otimes$ denotes the Kronecker (tensor) product; and where diag($\mathbf{L}$) is defined as the $n \times n$ diagonal matrix of the L_k/L terms. It is worth pointing out that, by construction, $\mathbf{W}$ is row-standardized (i.e., the rows sum to one).

Turning to the functional form of trade costs, we follow standard practice by assuming that τ_{ij} is a log-linear function of distance and border effects as follows:[12]

$$\tau_{ij} \equiv d_{ij}^{\gamma} \mathrm{e}^{\xi b_{ij}} \tag{12}$$

where d_{ij} denotes the distance between regions i and j, and where b_{ij} is a dummy variable taking the value 1 if the flow X_{ij} crosses the Canada-U.S. border, and 0 otherwise. Taking logarithms of (12), we can rewrite this expression in matrix form as follows:

$$\tau = \gamma \mathbf{d} + \xi \mathbf{b}, \tag{13}$$

where $\mathbf{d} \equiv (\ln d_{ij})$ is the $n^2 \times 1$ vector of the logarithms of distance; and where $\mathbf{b}$ is the $n^2 \times 1$ vector of dummy variables indicating cross-border flows. Substituting (13) into (11) then yields the following estimating equation:

$$\mathbf{X} = \beta_0 \mathbb{1} + \beta_1 \mathbf{Y_d} + \beta_2 \widetilde{\mathbf{Y}}_\mathbf{o} + \beta_3 \widetilde{\mathbf{d}} + \beta_4 \widetilde{\mathbf{w}} + \theta \widetilde{\mathbf{b}} + \rho \mathbf{W}\mathbf{X}, \tag{14}$$

where $\beta_0 \equiv \sigma\zeta < 0$ is the constant term; $\beta_1 \equiv \sigma > 1$ is the coefficient for destination GDP; $\beta_2 \equiv 1$ is the coefficient for origin GDP; $\beta_3 \equiv -(\sigma - 1)\gamma < 0$ is the distance coefficient (which, because of the implicit structure of the model, differs from the true distance elasticity); and where $\beta_4 \equiv -\sigma < 1$ is the coefficient for wage in the origin region. Note that β_2, β_3 and β_4 all capture

[12] Henderson and Millimet (2006) show that this linearity assumption cannot be rejected.

deviations from population weighted averages, as explained in the foregoing. Turning to the border effects, their coefficient is given by $\theta \equiv -(\sigma-1)\xi < 0$. How to precisely compute and decompose the border effects into intra- and international components is explained more fully in Section 5. Finally, the spatial autoregressive coefficient $\rho \equiv -(\sigma-1) < 0$ is the smaller the closer substitutes the varieties are. Hence, ρ provides an intuitive measure of 'spatial competition'.

3.2 Spatial econometric specification

To obtain a specification that can be estimated by spatial econometric techniques requires to rewrite (14) in explicit form, i.e., to move all of the $\ln X_{ij}$ terms to the left-hand side. Let $\mathbf{W_{diag}} \equiv \mathrm{diag}\,(\mathbf{L}) \otimes \mathbf{I_n}$ denote the matrix containing only the diagonal elements of $\mathbf{W}$, each repeated n times by block. Recalling that $\rho \equiv -(\sigma-1)$, equation (14) can then be rewritten as follows:

$$(\mathbf{I} - \rho\mathbf{W_{diag}})\,\mathbf{X} = \beta_0 \mathbb{1} + \beta_1 \mathbf{Y_d} + \beta_2 \widetilde{\mathbf{Y}}_{\mathbf{o}} + \beta_3 \widetilde{\mathbf{d}} + \beta_4 \widetilde{\mathbf{w}} + \theta \widetilde{\mathbf{b}} + \rho\,(\mathbf{W} - \mathbf{W_{diag}})\,\mathbf{X}.$$

Because $\mathbf{I} - \rho\mathbf{W_{diag}}$ is, by construction, an invertible diagonal matrix, we can premultiply by its inverse to obtain the following expression:

$$\mathbf{X} = \overline{\beta}_{\mathbf{0}} \mathbb{1} + \overline{\beta}_{\mathbf{1}} \mathbf{Y_d} + \overline{\beta}_{\mathbf{2}} \widetilde{\mathbf{Y}}_{\mathbf{o}} + \overline{\beta}_{\mathbf{3}} \widetilde{\mathbf{d}} + \overline{\beta}_{\mathbf{4}} \widetilde{\mathbf{w}} + \overline{\theta} \widetilde{\mathbf{b}} + \overline{\rho}\,(\mathbf{W} - \mathbf{W_{diag}})\,\mathbf{X}. \tag{15}$$

The n elements between positions $i \times n+1$ and $(i+1)\times n$ of $(\mathbf{I} - \rho\mathbf{W_{diag}})^{-1}$, given by $\left[1 + (\sigma-1)\frac{L_i}{L}\right]^{-1}$, depend on the origin index i only which is fixed and identical for all destinations. In expression (15), the components of the transformed (overlined) vectors of coefficients are thus given by:

$$\begin{aligned}
\overline{\beta}_{1i} &\equiv \sigma\,[1 - \rho(L_i/L)]^{-1} > 0, & \overline{\beta}_{2i} &\equiv [1 - \rho(L_i/L)]^{-1} > 0, \\
\overline{\beta}_{3i} &\equiv \rho\,[1 - \rho(L_i/L)]^{-1}\,\gamma < 0, & \overline{\beta}_{4i} &\equiv -\overline{\beta}_{1i} < 0, \\
\overline{\theta}_{i} &\equiv \rho\,[1 - \rho(L_i/L)]^{-1}\,\xi < 0, & \overline{\rho}_{i} &\equiv \rho\,[1 - \rho(L_i/L)]^{-1} < 0.
\end{aligned}$$

We clearly obtain a specification with a distinct set of parameters for each region. The full model, therefore, has a 'club' structure since all parameters (including the spatial autoregressive ones) must be estimated locally for each region. Quite naturally, we refer to this model as the *heterogeneous coefficients model*. Since it is econometrically complex to handle, we first estimate a simpler benchmark in which we constrain all coefficients to be identical across regions, which we refer to as the *homogeneous coefficients model*. Formally, constraining the coefficients to be identical amounts to assuming that the diagonal elements of $\mathbf{W}$ are equal to zero in equation (14).

In that case, the model simplifies substantially and can readily be estimated using standard spatial econometric techniques.

Before turning to the estimation proper, we need to spell out the error structure underlying the model. Though fundamental to the analysis, this modeling aspect has received only little attention until now. This is quite surprising because when the error terms are introduced into the econometric specification via the trade costs τ_{ij} or the trade flows X_{ij}, as usually done in the literature, one must take into account the fact that "the multilateral resistance variables also depend on these error terms" (Anderson and van Wincoop, 2004, p.713). The same holds true for the border effects, since these effects in any region depend in a complex way on a spatially weighted average of the effects in all the other regions. Consequently, *the error terms will exhibit some form of cross-sectional correlation that has to be dealt with.* To the best of our knowledge, this point has largely gone unnoticed until now in the gravity literature.[13] Although "errors can enter the model in many [...] ways of course, about which the theory has little to say" (Anderson and van Wincoop, 2003, p.180), it is likely that the exact way the error terms are introduced into the model is crucial to the estimates one obtains.

In what follows, we introduce the error terms via the trade flows X_{ij}. Doing so can be justified on the basis that regional trade flows are imperfectly observed. Let $X_{ij}^{\text{real}} \equiv X_{ij}^{\text{obs}} \mathrm{e}^{\varepsilon_{ij}}$ stand for the unobserved 'real' trade flow, where X_{ij}^{obs} denotes the observed trade flow and ε_{ij} is an i.i.d. normal error term. Introducing this error specification into (15) yields:

$$\mathbf{X} = \overline{\beta}_{\mathbf{0}} \mathbb{1} + \overline{\beta}_{\mathbf{1}} \mathbf{Y_d} + \overline{\beta}_{\mathbf{2}} \widetilde{\mathbf{Y}}_{\mathbf{o}} + \overline{\beta}_{\mathbf{3}} \widetilde{\mathbf{d}} + \overline{\beta}_{\mathbf{4}} \widetilde{\mathbf{w}} + \overline{\theta} \widetilde{\mathbf{b}} + \overline{\rho} \left(\mathbf{W} - \mathbf{W_{diag}} \right) \mathbf{X} + \mathbf{u} \tag{16}$$

where $\mathbf{X}$ now stands for the vector of observed trade flows and where

$$\mathbf{u} = -\varepsilon + \overline{\rho} \big(\mathbf{W} - \mathbf{W_{diag}} \big) \varepsilon \tag{17}$$

stands for the error term.[14] Note from (17) that the error terms ε_{ij} are spatially correlated under the form of a *first-order moving average* with region-specific correlation coefficients $\overline{\rho}_i$. We hence

[13] As pointed out by Anderson and van Wincoop (2004, p.713), such interdependencies give rise to complex problems since "[structural estimation techniques] would have to be modified since the multilateral resistance variables also depend on these error terms." While introducing origin–destination fixed effects circumvents this problem by using standard estimation techniques (Feenstra, 2002, 2004) we propose, in the remainder of this paper, to explicitly take into account the richer error structure.

[14] Note that since we use different autoregressive coefficients for the spatially lagged variable and the error terms when estimating the model, all of the subsequent developments hold true even when the error terms are introduced via the trade costs ($\tau_{ij} \equiv d_{ij}^{\gamma} \mathrm{e}^{\xi b_{ij}} \mathrm{e}^{\varepsilon_{ij}}$).

have, quite naturally, cross-sectional correlations in the error terms. Although moving averages are quite common in structural econometric models, especially in time series, they are less so in spatial econometrics. A first explanation of this fact is the clear lack of structural models. A second explanation is that the combined estimation of a moving average error structure with an autoregressive part (the so-called SARMA model; Huang, 1984) is complex and thus scarcely used.

4 Empirical implementation

In what follows, we apply our methodology to the well-known Canada-U.S. dataset used, among others, by Anderson and van Wincoop (2003) and Feenstra (2002, 2004). We begin by briefly reviewing the data, previous estimation methods and results. We first estimate the OLS benchmark, both with and without importer-exporter fixed effects and show that the residuals are always spatially autocorrelated. This finding vindicates the use of spatial econometric techniques for estimating such equations because OLS estimators are at best inefficient and at worst inefficient and biased. We then estimate our preferred theory-based specification, namely the SARMA model, under the assumption of homogeneous coefficients. We also run a series of robustness checks by estimating the model using alternative error structures. As will become clear, the empirical results back the theoretical specification. Last, we estimate the SARMA model with heterogeneous coefficients. Since estimating the fully heterogeneous model requires estimating as many spatial autoregressive coefficients for the endogenous lagged variable and the error terms as there are regions, we restrict ourselves to the case where we estimate only country-specific autoregressive coefficients, whereas all other coefficients are allowed to vary.[15]

4.1 Data and controls

The publicly available dataset features bilateral trade flows X_{ij}, regional GDPs Y_i, internal absorption X_{ii} (all measured in million US$), and distances d_{ij} in km between regional and provincial capitals for 30 U.S. states and 10 Canadian provinces. Unlike most gravity equations, which disregard own absorption X_{ii}, we require a measure of internal trade costs because we have to take into account the full spatial interdependence structure. Following Redding and Venables (2004),

[15] The technical details of the procedure are relegated to a separate technical appendix available from the authors upon request. To the best of our knowledge, ours are the first estimates of SARMA models with heterogeneous autoregressive coefficients.

we measure internal trade costs as $\tau_{ii} \equiv \kappa\sqrt{\text{surface}_i/\pi}$. As estimation results are known to be somewhat sensitive to the measure of internal distance (see, e.g., Head and Mayer, 2002) we use the values of 1/3 and 2/3 for the parameter κ as robustness checks.[16] We also use Anderson and van Wincoop's (2003) measure of internal distance, which is given by $d_{ii} \equiv (1/4) \min_{j \neq i} d_{ij}$, as an additional robustness check. Hourly wages across sectors in 1993 by province and state are obtained from Statistics Canada and from the Bureau of Labor Statistics, and the Canadian values are converted to US$ using the average 1993 exchange rate. Finally, we obtain population sizes from the U.S. Census Bureau and from Statistics Canada. In the U.S. case, we approximate the 1993 populations by linearly interpolating the 1990 and 2000 censuses; whereas in the Canadian case, we interpolate the 1991 and 1996 censuses.

Since our estimation method requires the whole information contained in the sample to account for spatial interdependence, we further have to deal with the well-known problem of zero trade flows. Indeed, there are 49 zero observations out of 1600, which requires an appropriate treatment. Since there is no generally agreed-upon method for doing so (Anderson and van Wincoop, 2004; Disdier and Head, 2008), we control for the potential zero flow outliers by including a dummy variable in all regressions. Although this is an admittedly crude way of controlling for zero trade flows, alternative methods like truncating the sample are not known to perform better or to be theoretically more sound.[17]

4.2 Previous estimation methods

We now briefly review previous estimation methods. The first one is based upon the strong assumption that *trade flows are independent*: estimating the determinants of X_{ij} can then be

[16] In unreported results, we also use $\kappa = 1$. Results are little sensitive to this choice. When $\kappa = 2/3$, our measure of internal distance gives the average distance between the center and any point in a disc-shaped country of the specified surface. Lower values of κ correspond to a more concentrated demand pattern, whereas larger values correspond to a more dispersed pattern. Note that we did not try more complex measures of interregional distance as suggested in, e.g., Helliwell and Verdier (2001) or Head and Mayer (2002). We conjecture that our main results are relatively robust to the use of such more complex measures.

[17] Note that our zeros are unlikely to be 'true zeros', as this would imply no aggregate manufacturing trade between several U.S. states. In the case of 'true zeros', such as those present in more disaggregated international trade patterns, a Tobit estimator would perform better (Helpman *et al.*, 2007). See Felbermayr and Kohler (2006) for a recent discussion on the various treatments of zero trade flows. They argue that the neglect of zero trade flows is at the heart of the "puzzling persistance of distance" and they estimate a gravity equation with corner solutions using a Tobit estimator.

done without taking into account any information contained in X_{kl}. McCallum (1995), among others, makes this assumption to estimate by OLS the following empirical gravity equation for Canada-U.S. interregional trade:

$$\ln X_{ij} = \alpha_1 + \alpha_2 \ln Y_i + \alpha_3 \ln Y_j + \alpha_4 \ln \mathrm{d}_{ij} + \alpha_5 b_{ij} + \varepsilon_{ij}. \tag{18}$$

Quite surprisingly, McCallum obtains paradoxically large values for the border coefficient α_5, ranging from 3.07 to 3.30. Consequently, Canadian provinces seem to trade 21.5 to 27 times more with themselves than with U.S. states of equal size and distance, an unrealistically large value for two well-integrated and culturally similar countries like Canada and the U.S.

In columns 1 to 3 of Table 1, we replicate McCallum-type OLS regressions of the form (18) as our benchmark.[18] As can be seen, all coefficients have the correct sign, reasonable magnitudes, and are precisely estimated. Results for the distance elasticity are somewhat sensitive to the definition of internal distance, which is a well-known result in the literature. As can be further seen from Table 1, the magnitude of the border effects for Canadian provinces ranges from about 14.5 to 16, depending on the definition of internal distance. These estimates are in line with the McCallum-type regressions of Anderson and van Wincoop (2003, Table 1, p.173), which obtain border effects of about 15.7.[19]

Insert Table 1 about here.

However, as can also be seen from the last line of Table 1, *not a single OLS specification à la McCallum passes Moran's I test* for the absence of spatial autocorrelation of the residuals (Cliff and Ord, 1981). Stated differently, there remains a significant amount of spatial autocorrelation in the OLS residuals, which leads at best to inefficient and at worst to both inefficient and biased estimates (with omitted variable bias because of the missing spatially lagged variable). The presence of spatial autocorrelation suggests that the use of appropriate econometric techniques dealing with this problem is required.

[18]To stay as closely as possible to the original analysis, we define the border effects as in Anderson and van Wincoop (2003). Hence, we introduce two sets of dummy variables, bordCA_{ij} and bordUS_{ij}, for Canada-U.S. and U.S.-Canada flows, respectively. The implied border effects can, as always, be retrieved as the exponential of minus the coefficient of bordCA_{ij} and bordUS_{ij}.

[19]It is worth pointing out that the slight differences between our OLS estimates and those of Anderson and van Wincoop (2003), despite using the same dataset, are due to: (i) inclusion of own trade flows X_{ii}; (ii) accounting for intra-regional distances; and (iii) controlling for zero trade flows.

To deal with interdependence and with McCallum's 'border effect puzzle', Anderson and van Wincoop (2003) build on the 'price version' of the CES model presented in Section 2. Assuming equal wages and symmetric trade costs that are a log-linear function of bilateral distance and the existence of an international border between i and j, they derive the following instance of a gravity equation system:

$$\ln\left(\frac{X_{ij}}{Y_i Y_j}\right) = k + a_1 \ln \mathrm{d}_{ij} + a_2(1 - b_{ij}) - \ln \widetilde{\mathbb{P}}_i^{1-\sigma} - \ln \widetilde{\mathbb{P}}_j^{1-\sigma} + \varepsilon_{ij}, \tag{19}$$

where $k \equiv -Y_W$ is a constant, with Y_W the 'world' GDP; and where $\widetilde{\mathbb{P}}_i^{1-\sigma}$ and $\widetilde{\mathbb{P}}_j^{1-\sigma}$ are the multilateral resistance terms of regions i and j, which, apart from unitary income elasticities, represent the key difference with equation (18) as estimated by McCallum. The multilateral resistance terms are linked by a system of non-linear equations involving all regions' expenditure shares and the whole trade cost distribution:

$$\widetilde{\mathbb{P}}_i^{1-\sigma} = \sum_k \frac{Y_k}{Y_W} \widetilde{\mathbb{P}}_k^{\sigma-1} \mathrm{e}^{a_1 \ln \mathrm{d}_{ik} + a_2(1-b_{ik})} \quad \forall i. \tag{20}$$

Equations (19) and (20) reveal that the determinants of X_{ij} cannot be consistently estimated without taking into account the conditions prevailing in the origin and destination markets i and j, as captured here by a simple transformation of the CES price indices. These price indices depend themselves on the inverse demands and, therefore, on the different trade flows.[20] Hence, the independence assumption underlying the McCallum-type estimates is clearly invalid and likely to bias the results.

Anderson and van Wincoop (2003) estimate equation (19) using nonlinear least squares, where the multilateral resistance terms are solved for in a first step using (20). While this procedure accounts for interdependence, it has at least three drawbacks. First, it makes strong structural assumptions on the form of the spatial interdependence between the multilateral resistance terms that directly stem from the CES specification. The spatial econometric approach does not impose a priori such a structure on the spatial correlations and is, therefore, more robust to misspecifications. Second, Anderson and van Wincoop's method does not allow for parameter heterogeneity across

[20] Despite their central theoretical role, price indices have been largely neglected in the gravity equation. The main reason for this is that they are unobservable, so that most studies have tried to somehow eliminate them. Notable exceptions are given by Bergstrand (1985) and Baier and Bergstrand (2001), who retain the price indices as explanatory variables using published price data, namely GDP deflators. This method suffers from severe data constraints, especially at subnational levels for which regional GDP deflators are not available. Furthermore, the theoretical link between published price indices and the CES price aggregators is unclear.

regions, whereas ours does. As will be shown later, parameter estimates widely vary across regions, and imposing common coefficients may thus be unwarranted. Last, as the multilateral resistance terms are solved for numerically, their statistical significance is usually not tested. We will provide such tests and show that measured U.S. border effects are mostly insignificant, whereas measured Canadian border effects are not.

A third estimation method has been suggested by Anderson and van Wincoop (2003) and Feenstra (2002) and used, among others, by Rose and van Wincoop (2001). This method replaces the multilateral resistance terms with region-specific importer-exporter fixed effects. In this case, (19) can be written as:

$$\ln\left(\frac{X_{ij}}{Y_i Y_j}\right) = k + a_1 \ln \mathrm{d}_{ij} + a_2(1 - b_{ij}) + \beta_1^i \delta_1^i + \beta_2^j \delta_2^j + \varepsilon_{ij}, \tag{21}$$

where δ_1^i denotes an indicator variable that equals one if region i is the exporter, and zero otherwise; and where δ_2^j denotes an indicator variable that equals one if region j is the importer, and zero otherwise. The coefficients $\beta_1^i = (\sigma - 1)\ln \widetilde{\mathbb{P}}_i$ and $\beta_2^j = (\sigma - 1)\ln \widetilde{\mathbb{P}}_j$ then provide estimates of the multilateral resistance terms.[21]

Although the fixed effects procedure yields theoretically consistent estimates of the average border effect (see Feenstra, 2002), it disregards some part of the spatial interdependence. Hence, while the fixed effects method has the advantage of being simple to implement, as OLS can be used under the traditional assumptions on the error terms ε_{ij}, its main drawback is that *it does not fully capture the spatial interactions* of the model. Columns 4 to 6 of Table 1 show results for OLS fixed effects regression. As can be seen from the last line in Table 1, even after controlling for multilateral resistance by using region-specific importer-exporter fixed effects, there remains a significant amount of spatial autocorrelation in the OLS residuals. This finding suffices to show that fixed effects capture at best some heterogeneity but do not capture spatial interdependence. In other words, although fixed effects allow to partly control for interdependence, they are by no means sufficient from both a theoretical and from an econometric point of view.[22]

[21]Note that equation (19) features only a single multilateral resistance term per region, whereas there are two fixed effects in equation (21). The reason is that Anderson and van Wincoop (2003, p.175) make a symmetry assumption on trade costs (in the general case, there are two terms per region given by expressions (10) and (11) on p.175 of their paper).

[22]Note that Anderson and van Wincoop (2003, p.180) also emphasize that the fixed-effects estimator could be less efficient than the non-linear least squares estimator, which uses information on the full structure of the model. We furthermore show in this paper that it could also be biased. One should keep in mind that fixed effects allow

4.3 Spatial econometric estimations with homogeneous coefficients

We now revisit the estimation of gravity equation systems using spatial econometric techniques. To do so, we build upon our preferred theory-based specification, namely the SARMA model (16) and (17) which takes into account the spatial interdependence among both trade flows and error terms. Columns 1 to 3 in Table 2 summarize the estimation results obtained under the assumption of homogeneous coefficients.

Insert Table 2 about here.

As can be seen from columns 1 to 3 in Table 2, all coefficients, including the spatial autoregressive ones, have the correct signs, plausible magnitudes, and are precisely estimated. To begin with, note that, as predicted by our model, the coefficient for origin GDP, as measure by the deviation from the population weighted average $(\mathbf{I}-\mathbf{W})\ln Y_i$, remains close to unity. As further predicted by our model, the coefficient on destination GDP $\ln Y_j$ clearly exceeds unity. The distance coefficient, measured again as deviation from the population weighted average $(\mathbf{I}-\mathbf{W})\ln d_{ij}$, slightly decreases in absolute value when compared to the OLS specification but remains overall fairly stable. Turning to the wage terms, it is worth noting that they are negative, highly significant, and exceed the distance coefficient in absolute value. Put differently, *higher origin wages reduce trade flows because of increased production costs*. Although one might a priori suspect that interregional wage differentials should not significantly affect interregional trade flows in an integrated economic environment like North America, where interregional wages differentials are relatively small, our results show that this is not the case: interregional wage differentials are significant enough across North American regions to affect trade flows. One of the key empirical results in the SARMA specification is that *there is a significant amount of negative spatial autocorrelation among trade flows* ($\widehat{\rho} < 0$), as predicted by our model.[23] There is also negative spatial autocorrelation among error terms ($\widehat{\lambda} < 0$), which suggests that controlling for cross-sectional correlations is important.

Finally, as can be seen from Table 2, capturing the spatial interdependence of the equilibrium system in the SARMA model reduces the border effects with respect to the OLS estimates, but

to control for heterogeneity, but not for interdependence. This fact is often overlooked in the literature, even when more complex fixed effects specifications are used (e.g., Baltagi *et al.*, 2003).

[23] Note that we cannot identify σ in our estimations. However, our results "suggest" that σ may lie somewhere in between 1.5 and 2.5, which is the lower end of the spectrum of existing estimates (e.g., Hummels, 2001; Hanson, 2005; Broda and Weinstein, 2006). Given the high aggregation level of the data, this result seems plausible and suggests that the linearization around $\sigma = 1$ may be sufficiently accurate.

also with respect to Anderson and van Wincoop (2003). Indeed, in our preferred theory-based specification, the border effects for Canadian provinces range from about 7.3 to 7.8, whereas the ones for U.S. states range from 1.29 to about 1.3 (see Appendix B.1. for a more detailed explanation of how to compute and to decompose the border effects). These values must be compared with Anderson and van Wincoop's measured border effects of 10.5 for Canada and 2.6 for the U.S.

As can be seen from equation (14), the most exact theoretical specification imposes some additional restrictions on the coefficients of the model. In particular, own GDP should have a unit coefficient, whereas the coefficients on relative distance and importer GDP should be identical (up to their sign). In Appendix C, we derive a constrained model that encapsulates these additional restrictions. Columns 4 to 6 in Table 3 give results for the constrained specification. Note that all coefficients have again the correct sign and are precisely estimated. The estimation results of the constrained specification largely confirm those of the unconstrained one, thus suggesting that our results are quite robust.

As stated in the foregoing, there are many ways of modeling the error structure about which theory has little to say. To see how sensitive the results are to the precise nature of the error structure, we now run two robustness checks. First, we approximate the moving average by a more general autoregressive error structure, which leads to the so-called general spatial model (henceforth, GSM; Anselin, 1988). Consider a vector of error terms $\mathbf{u}$ that is spatially correlated according to the autoregressive structure

$$\mathbf{u} = \overline{\lambda}\big(\mathbf{W} - \mathbf{W_{diag}}\big)\mathbf{u} + \varepsilon, \tag{22}$$

where ε is i.i.d. and normally distributed with zero mean and variance $\sigma^2\mathbf{I}$. Provided that $|\overline{\lambda}_i| < 1$ for all i, we then can write

$$\mathbf{u} = \left[\mathbf{I_{n^2}} - \overline{\lambda}(\mathbf{W} - \mathbf{W_{diag}})\right]^{-1}\varepsilon = \sum_{j=1}^{\infty}\left[\overline{\lambda}(\mathbf{W} - \mathbf{W_{diag}})\right]^{j}\varepsilon + \varepsilon.$$

When the successive powers of $\left[\overline{\lambda}(\mathbf{W} - \mathbf{W_{diag}})\right]^{j}$ converge to 0 sufficiently quickly, the spatial autoregressive structure approximates appropriately the first-order moving average, i.e., $\mathbf{u} \approx \varepsilon + \overline{\lambda}(\mathbf{W} - \mathbf{W_{diag}})\varepsilon$ as in (17).[24]

We estimate the GSM specification in the homogeneous case and the results are summarized in columns 1 to 3 of Table 3. Observe that, as in the SARMA model, the spatial autoregressive

[24]Note that this approximation is reasonably accurate provided that: (i) all λ_i are small enough; and (ii) the elements of the successive powers of $\overline{\lambda}(\mathbf{W} - \mathbf{W_{diag}})$ converge to zero sufficiently quickly.

coefficient ρ is negative and highly significant in all estimations, which is in accord with the underlying theory stipulating that varieties are substitutes. The magnitude of ρ is smaller than in the unconstrained theory-based SARMA model and closer to the one obtained in the constrained theory-based SARMA model.

Insert Table 3 about here.

All remaining coefficients are precisely estimated and the signs are identical to the ones obtained under the SARMA specification. Note that the magnitude of both origin and destination GDPs change, with the former now exceeding unity whereas the latter falls short of unity. This result is at odds with the underlying model and suggests that the approximation of the error structure may not be very accurate. The distance coefficients and the border effects remain fairly similar, and the wage coefficient is again negative and highly significant.

As a second robustness check, we re-estimate the model by introducing the error terms in an ad hoc way. The simplest way of doing so is to rewrite (15) as follows:

$$\mathbf{X} = \overline{\beta}_{\mathbf{0}}\mathbb{1} + \overline{\beta}_{\mathbf{1}}\mathbf{Y_d} + \overline{\beta}_{\mathbf{2}}\widetilde{\mathbf{Y}}_{\mathbf{o}} + \overline{\beta}_{\mathbf{3}}\widetilde{\mathbf{d}} + \overline{\beta}_{\mathbf{4}}\widetilde{\mathbf{w}} + \overline{\theta}\widetilde{\mathbf{b}} + \overline{\rho}\left(\mathbf{W} - \mathbf{W_{diag}}\right)\mathbf{X} + \varepsilon, \tag{23}$$

which simply amounts to adding the i.i.d. normal error term ε to the estimating equation. The resulting specification (23) is a standard *spatial autoregressive model* (for short, SAR; Lee, 2004). Table 3 (columns 4–6) summarize estimation results for the SAR with homogeneous coefficients. Observe that, although the other coefficients remain fairly stable, the spatial autoregressive coefficient ρ is not significantly different from zero (with even positive point estimates). This result runs plainly against the underlying theory which predicts a negative spatial autocorrelation across trade flows. Hence, the ad hoc introduction of the error term is not backed by the data in the sense that it is incompatible with the qualitative predictions of the theory presented in Section 2.

Insert Table 4 about here.

Last, we provide an additional robustness check with respect to the magnitude of U.S. flows. As argued by Balistreri and Hillberry (2007), the construction of the state-state trade flows by Anderson and van Wincoop (2003), using shipping data from the Commodity Flow Survey (CFS), relies on a scalar numerical scaling to make the data "comparable" to the Canada-U.S. trade data. However, it is likely that Anderson and van Wincoop's pre-treatment of the data has overcorrected the CFS shipping data and, therefore, understates U.S. state-state trade flows. Consequently, measured Canadian border effects would appear artificially lower since the intra-Canadian trade intensity seems significantly larger when compared to the U.S. one. Table 4 replicates our estimates when all U.S. state-state trade flows are inflated by a factor of 1.3 (i.e., instead of adjusting the CFS data by a factor of 0.53, we adjust it by a factor of only 0.69). As one can see, measured Canadian border effects increase indeed to reflect the larger U.S. trade flows, which is consistent with the findings by Balistreri and Hillberry (2007). Measured border effects hence crucially depend on data pre-treatment, which prompts us to be careful when measuring them.[25]

4.4 Spatial econometric estimations with heterogeneous coefficients

All previous estimates are based upon the strong assumption of homogeneous coefficients. Although, as pointed out by Henderson and Millimet (2006), this assumption does not directly flow from the theory, it has become a staple in estimating gravity equations.[26] Yet, as one can see from equation (15), the theory predicts that coefficients are region specific. This is in accord with recent findings by Helpman *et al.* (2008, p.474), who note that "the elasticities vary widely across different country pairs". We therefore now estimate the model by allowing every region to have different coefficients, as implied by the underlying theoretical specification. In so doing, we use the preferred SARMA specification because OLS have no theoretical foundation, because the GSM approximation is not very accurate, and because SAR runs plainly against the underlying theory.

As already mentioned in the foregoing, we estimate a simpler heterogeneous coefficients model in which only the non-autoregressive parameters $\overline{\beta}_i$ are allowed to vary across regions, whereas the autoregressive coefficients $\overline{\rho}$ and $\overline{\lambda}$ vary by country only ($\overline{\rho}_{CA}$ and $\overline{\lambda}_{CA}$ for Canada; and $\overline{\rho}_{US}$ and $\overline{\lambda}_{US}$ for the U.S.). For estimation purposes, we rewrite the model in a more compact way, as presented in Appendix C.[27] The estimation results for this *β-heterogeneous SARMA model* are summarized in Table 4, where we give the estimated coefficients for the borders and the distances, the true regional distance elasticities $\varepsilon_{d_{ij}}$, the border effects (and their decomposition; see Appendix B.2.), as well as the country-specific autoregressive coefficients. It is worth noting that, in the presence

[25]Balistreri and Hillberry (2007) re-estimate the border effects excluding simply the U.S.-U.S. flows. They find that multilateral resistance reduces measured border effects by only 8%.

[26]Anderson and van Wincoop (2004, p.711) note that: "Implausibly strong regularity (common coefficients) conditions are often implicitly imposed on the trade cost function." As shown in the foregoing, heterogeneity is also likely to affect the other coefficients.

[27]All the technical details for estimating this model, including the derivation of the likelihood function and the information matrix, are relegated to a separate technical appendix available from the authors upon request.

of region-specific coefficients, we can no longer identify the impacts of origin GDP Y_i separately, as it is subsumed by the regional fixed effect.

Insert Table 5 about here.

Several comments are in order. First, it is worth noting that there is *a substantial amount of heterogeneity in the estimated coefficients*, both for distance elasticities, border effects, and autoregressive coefficients. As can be seen from Table 5, the autoregressive coefficient for the U.S. is smaller than that for Canada, thus suggesting that the U.S. market is 'more competitive' than the Canadian market. The estimated distance coefficients (which differ from the true elasticities $\varepsilon_{d_{ij}}$ because of the cross-sectional interdependence and the implicit form of the estimating equation) range from -0.7 for California to -3.3 for Newfoundland.[28] The real elasticities are very similar, with values in the same range.

Insert Figure 2 about here.

Figure 2 depicts the relationship between the true distance elasticities and the size of the local market (L_i/L). There is a clear pattern relating regional sizes to distance elasticities: *larger regions face systematically lower distance elasticities than smaller regions.* As shown in Section 4.2, the model predicts the existence of such a positive and concave relationship between a region's relative size L_i/L and its distance coefficient. The latter is indeed given by $\overline{\beta}_{3i} \equiv \rho \frac{\gamma}{1-\rho\frac{L_i}{L}}$, which is concave and increasing in L_i/L.[29] One possible explanation for this finding is that firms in larger regions predominantly serve the local market, so that export flows are relatively smaller and less sensitive to distance (when measured in percentage changes). Another possible interpretation of this finding, which is in accord with recent developments in the literature on firm heterogeneity, is that smaller markets are less competitive, so that less productive firms are selected into those markets (Melitz,

[28]Because the estimating equation is given in implicit form, the true distance elasticities differ from the estimated coefficients. However, starting from (15), they can be computed (in matrix form) as follows:

$$\varepsilon_d \equiv \left[\mathbf{I} - \overline{\rho}(\mathbf{W} - \mathbf{W_{diag}})\right]^{-1} \overline{\beta}_3,$$

which is derived from the explicit solution to the estimating equation. Note that $\varepsilon_d \equiv (\varepsilon_d)_{ij}$ is the $n^2 \times n^2$ matrix of distance elasticities. Since all distance elasticities with the same origin index are identical, there are only n distinct distance elasticities, i.e., one for each region.

[29]Table 8 in Helpman *et al.* (2007), though very aggregated since countries are just clustered into three broad categories, also exhibits such a positive and concave relationship when size is measured by GDP. We conjecture that further disaggregation of their results would yield a graph similar to the one in Figure 4.

2003; Melitz and Ottaviano, 2008). These firms then have a greater handicap in serving foreign markets, thus facing higher distance elasticities than more productive firms in larger and more competitive markets.

Insert Figure 3 about here.

The regional structure of distance elasticities is depicted in Figure 3. The Canadian core regions (Ontario and Québec), as well as the north-eastern U.S. states (Maryland, New York, Pennsylvania) form a cluster of regions facing small distance elasticities (in absolute value) of trade flows. The same holds true for the western states and provinces (Alberta, British Columbia, Washington, California), whereas the Great Plains and the remote Canadian provinces, face relatively high distance elasticities (in absolute value).

Turning next to the border effects, Table 5 and Figure 4 reveal that, as expected, the Canadian provinces face larger border effects than the U.S. states. The reason for this is as in Anderson and van Wincoop (2003) and explained in detail in Appendix B. As shown by Table 4, the intranational trade-boosting effect of the border is much larger for Canadian provinces than for U.S. states. Put differently, "trade barriers raise size adjusted trade within small countries more than within large countries" (Anderson and van Wincoop, 2003, p.176). Furthermore, the trade-reducing international effect of the border for Canadian exports is larger than that for U.S. exports, which illustrates again that the border has a stronger effect on Canadian firms than on U.S. firms. The reason is that the U.S. internal market is much larger, so that the border affects only a much smaller part of sales from U.S. firms than from Canadian firms.

Insert Figure 4 about here.

Finally, as can be seen from Table 5, most border coefficients for U.S. states are not significant at the 5% level, except for a few regions like Maine and Virginia (and almost North Dakota). On the contrary, the border effects for Canadian provinces are almost all highly significant and there is a lot of variation.[30] Magnitudes for the border effects range from about 0.8 in Newfoundland to 23.2

[30]As in Anderson and Smith (1999), there is a huge amount of variation in border effects. Ontario and Quebec may be viewed as "import platforms" (low value of the international border component), whereas British Columbia appears to be an "export platform" (large value of the international border component). Contrary to Anderson and Smith, our estimates include information on the full sample since we account for interdependencies. Large border effects for Québec and Ontario mirror the economic sizes of these regions and the fact that they trade a lot with the U.S. and, therefore, stand to gain the most from removing the border.

in Ontario, yet most values are clustered between 6 and 12 (with the exception of a few provinces with larger border effects). On the contrary, border effects for the U.S. states are uniformly small, ranging from a low of about 0.68 to a high of about 5.4. Note that although we obtain a large number of positive coefficients for U.S. states, which runs against theory, these are not precisely enough estimated to be significantly different from zero. To sum up, border effects are generally small for U.S. exports to Canada, whereas they do exist for Canadia exports to the U.S. Their magnitude, however, appears to be much smaller than previously thought once 'multilateral resistance' is captured via the spatial interdependence of trade flows.

5 Conclusions

Building on a 'dual' version of the gravity equation, we have shown how spatial econometric techniques provide a natural tool for controlling for cross-sectional interdependence among trade flows. Handling directly such interdependence is a major issue for consistent estimation but has been rather elusive until now. Our results suggest that, as in Anderson and van Wincoop (2003), consistent theory-based estimates of the gravity equation lead to significantly smaller border effects than those obtained with ad hoc specifications or fixed effect methods. Put differently, there is much less of a border effect puzzle once the cross-sectional correlations have been controlled for.

Besides partially solving the 'border effect puzzle', our methodology offers a number of additional advantages when compared to previous approaches: (i) it accounts for cross-sectional interdependence among trade flows, as implied by our model, and thus directly controls for multilateral resistance; (ii) it uses a more careful modeling of the error structure, thereby controlling for possible cross-sectional interdependence in the error terms; (iii) it does not impose a rigid pattern on the spatial interaction structure and is, therefore, more robust to potential misspecifications; and (iv) it reveals that all coefficients are generally region-specific, and allows for statistical inference on estimated regional border effects and distance elasticities.

While we have tried to stay as closely as possible to the structural model, we have not yet estimated the fully heterogeneous model with region-specific spatial autoregressive coefficients. Doing so is a daunting task that we will explore in the future.

Acknowledgements. We thank Peter Egger, Giordano Mion, Yasusada Murata, Giovanni Peri, Michael Pfaffermayr, seminar participants in Louvain-la-Neuve, Innsbruck, and conference participants in Kiel, Toronto, Louvain-la-Neuve, Cambridge, Tokyo, Paris, and the 62nd European Meetings of the Econometric Society for very helpful comments and suggestions. Kristian Behrens gratefully acknowledges financial support from the European Commission under the Marie Curie Fellowship MEIF-CT-2005-024266, and from UQAM (PAFARC 2007-08). Kristian Behrens and Wilfried Koch gratefully acknowledge financial support from CNRS-GIP-ANR (Grant JC-05-439904). Part of the paper was written while Wilfried Koch was visiting the Center for Operations Research and Econometrics (CORE, Université catholique de Louvain), the hospitality of which is greatly appreciated. The usual disclaimer applies.

References

[1] Anderson, J.E. (1979) A theoretical foundation for the gravity equation, *American Economic Review* 69, 106-116.

[2] Anderson, J.E. and E. van Wincoop (2003) Gravity with gravitas: A solution to the border puzzle, *American Economic Review* 93, 170-192.

[3] Anderson, J.E. and E. van Wincoop (2004) Trade costs, *Journal of Economic Literature* 42, 691-751.

[4] Anderson, M.A. and S.L.S. Smith (1999) Canadian provinces in world trade: engagement and detachment, *Canadian Journal of Economics* 32, 22-38.

[5] Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Dordrecht, Netherlands: Kluwer.

[6] Anselin, L. and A. Bera (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A. and D. Giles (eds.), *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, pp. 237-289.

[7] Baier, S.L. and J.H. Bergstrand (2001) The growth of world trade: tariffs, transport costs, and income similarity, *Journal of International Economics* 53, 1-27.

[8] Balistreri, E.J. and R. Hillberry (2007) Structural estimation and the border puzzle, *Journal of International Economics* 72, 451-463.

[9] Baltagi, B.H., P. Egger and M. Pfaffermayr (2003) A generalized design for bilateral trade flow models, *Economics Letters* 80, 391-397.

[10] Behrens, K., G. Mion, Y. Murata and J. Südekum (2008) Trade, wages, and productivity. IZA Discussion Paper #3682, Bonn, Germany.

[11] Benassy, J.-P. (1996) Taste for variety and optimum production patterns in monopolistic competition, *Economics Letters* 52, 41-47.

[12] Bergstrand, J.H. (1985) The gravity equation in international trade: Some microeconomic foundations and empirical evidence, *Review of Economics and Statistics* 67, 474-481.

[13] Blonigen, B.A., R.B. Davies, H.T. Naughton and G.R. Waddell (2004) FDI in space: Spatial autoregressive relationships in foreign direct investment. NBER Discussion Paper #10939.

[14] Blonigen, B.A., R.B. Davies, H.T. Naughton and G.R. Waddell (2005) Spacey parents: Spatial autoregressive patterns in inbound FDI. NBER Discussion Paper #11466.

[15] Bonacich, Ph. (1987) Power and centrality: A family of measures, *American Journal of Sociology* 92, 1170-1182.

[16] Broda, C. and D. Weinstein (2006) Globalization and the gains from variety, *Quarterly Journal of Economics* 121, 541-585.

[17] Brueckner, J.K. (1998) Testing for strategic interaction among local governments: The case of growth controls, *Journal of Urban Economics* 44, 438-467.

[18] Case, A.C., H.S. Rosen and J.R. Hines (1993) Budget spillovers and fiscal policy interdependence : Evidence from the states, *Journal of Public Economics* 52, 285-307.

[19] Chaney, T. (2008) Distorted gravity: The intensive and extensive margins of international trade, *American Economic Review* 98, 1707-1721.

[20] Cliff, A.D. and G.K. Ord (1981) *Spatial Processes: Models and Applications*. London, UK: Pion.

[21] Disdier, A.-C. and K. Head (2008) The puzzling persistence of the distance effect on bilateral trade, *Review of Economics and Statistics*, forthcoming.

[22] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297-308.

[23] Ertur, C. and W. Koch (2007) Growth, technological interdependence and spatial externalities: Theory and evidence, *Journal of Applied Econometrics* 22, 1033-1062.

[24] Feenstra, R.C. (2002) Border effects and the gravity equation: consistent methods for estimation, *Scottish Journal of Political Economy* 49, 491-506.

[25] Feenstra, R.C. (2004) *Advanced International Trade*. Princeton, NJ: Princeton Univ. Press.

[26] Felbermayr, G.J. and W. Kohler (2006) Exploring the intensive and extensive margins of world trade, *Review of World Economics* 142, 642-674.

[27] Hanson, G.H. (2005) Market potential, increasing returns and geographic concentration, *Journal of International Economics* 67, 1-24.

[28] Head, K. and T. Mayer (2002) Illusory border effects: distance mismeasurement inflates estimates of home bias in trade. CEPII Discussion Paper #2002/1.

[29] Helliwell, J.F. and G. Verdier (2001) Measuring internal trade distances: A new method applied to estimate provincial border effects in Canada, *Canadian Journal of Economics* 34, 1024-1041.

[30] Helpman, E. and P.R. Krugman (1985) *Market Structure and Foreign Trade*. Cambridge, MA: Cambridge Univ. Press.

[31] Helpman, E., M.J. Melitz and Y. Rubinstein (2008) Estimating trade flows: Trading partners and trading volumes, *Quarterly Journal of Economics* 123, 441-487.

[32] Henderson, D.J and D. Millimet (2007) Is gravity linear? Forthcoming in *Review of Economics and Statistics*.

[33] Huang, S.J. (1984) The autoregressive moving average model for spatial analysis, *Australian Journal of Statistics* 26, 169-178.

[34] Hummels, D. (2001) Towards a geography of trade costs. Mimeographed, Krannert School of Management, Purdue University.

[35] Kelejian, H.H. and I.R. Prucha (1998) A generalized spatial two-stage least square procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics* 17, 99-121.

[36] Kmenta, J. (1967) On estimation of the CES production function, *International Economic Review* 8, 180-189.

[37] Krugman, P.R. (1980) Scale economies, product differentiation and the pattern of trade, *American Economic Review* 70, 950-959.

[38] Lee, L.-F. (2004) Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models, *Econometrica* 72, 1899-1925.

[39] LeSage, J.P. and R.K. Pace (2006) Spatial econometric modeling of origin-destination flows. Mimeographed.

[40] McCallum, J.M. (1995) National borders matter: Canada-U.S. regional trade patterns, *American Economic Review* 85, 615-623.

[41] Melitz, M.J. (2003) The impacts of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica* 71, 1695-1725.

[42] Melitz, M.J. and G.I.P. Ottaviano (2008) Market size, trade, and productivity, *Review of Economic Studies* 75, 295-316.

[43] Moreno, R. and B. Trehan (1997) Location and the growth of nations, *Journal of Economic Growth* 2, 399-418.

[44] Pinkse, J., M. Slade and C. Brett (2002) Spatial price competition: a semiparametric approach, *Econometrica* 70, 1111-1153.

[45] Redding, S. and A.J. Venables (2004) Economic geography and international inequality, *Journal of International Economics* 62, 53-82.

[46] Rose, A. and E. van Wincoop (2001) National money as a barrier to international trade: The real case for currency union, *American Economic Review Papers and Proceedings* 91, 386-390.

[47] Santos Silva, J.M.C. and S. Tenreyro (2006) The log of gravity. Forthcoming in *Review of Economics and Statistics*.

Appendix A: Linearization of the model

In this appendix, we linearize (9) to obtain the estimable specification (10). The linearization of f around $\sigma = 1$ is given by $\ln X_{ij} = f(1) + (\sigma - 1)f'(1)$. Some straightforward calculation yields

$$f(1) = \ln Y_j - \ln\left[\sum_k \frac{L_k}{L_i}\right] = \ln Y_j - \ln L + \ln L_i, \tag{A.1}$$

where $L \equiv \sum_k L_k$. Turning to the derivative, some longer calculation shows that

$$\begin{aligned} f'(\sigma) &= \ln Y_j - \ln\left[\sum_k \frac{L_k}{L_i}\left(\frac{\tau_{kj}Y_k}{\tau_{ij}Y_i}\right)^{\frac{1}{\sigma}-1} X_{kj}^{1-\frac{1}{\sigma}}\right] \\ &\quad -\sigma\frac{\sum_k \frac{L_k}{L_i}\left(\frac{1}{\sigma^2}\right)\left\{-\left(\frac{\tau_{kj}Y_k}{\tau_{ij}Y_i}\right)^{\frac{1}{\sigma}-1} X_{kj}^{1-\frac{1}{\sigma}}\ln\left(\frac{\tau_{kj}Y_k}{\tau_{ij}Y_i}\right) + \left(\frac{\tau_{kj}Y_k}{\tau_{ij}Y_i}\right)^{\frac{1}{\sigma}-1} X_{kj}^{1-\frac{1}{\sigma}}\ln X_{kj}\right\}}{\sum_l \frac{L_l}{L_i}\frac{\tau_{lj}Y_l}{\tau_{ij}Y_i}^{\frac{1}{\sigma}-1} X_{lj}^{1-\frac{1}{\sigma}}}, \end{aligned}$$

which implies that

$$f'(1) = \ln Y_j - \ln L + \ln L_i + \sum_k \frac{L_k}{L}\ln\frac{\tau_{kj}}{\tau_{ij}} + \sum_k \frac{L_k}{L}\ln\frac{Y_k}{Y_i} - \sum_k \frac{L_k}{L}\ln X_{kj}. \tag{A.2}$$

Using (A.1) and (A.2), linearized equation can then be expressed as follows:

$$\begin{aligned} \ln X_{ij} &= \sigma\ln L_i - \sigma\ln L + \sigma\ln Y_j - (\sigma-1)\ln Y_i - (\sigma-1)\ln\tau_{ij} \\ &\quad +(\sigma-1)\sum_k \frac{L_k}{L}\ln\tau_{kj} + (\sigma-1)\sum_k\frac{L_k}{L}\ln Y_k - (\sigma-1)\sum_k\frac{L_k}{L}\ln X_{kj}, \end{aligned}$$

which, using the aggregate income constraint $Y_i = w_i L_i$, yields:

$$\begin{aligned} \ln X_{ij} &= -\sigma\ln w_i - \sigma\ln L + \sigma\ln Y_j + \ln Y_i - (\sigma-1)\ln\tau_{ij} + (\sigma-1)\sum_k\frac{L_k}{L}\ln\tau_{kj} \\ &\quad +\sigma\sum_k\frac{L_k}{L}\ln w_k + \sigma\sum_k\frac{L_k}{L}\ln L_k - \sum_k\frac{L_k}{L}\ln Y_k - (\sigma-1)\sum_k\frac{L_k}{L}\ln X_{kj}. \end{aligned}$$

Rearranging terms we then readily obtain equation (10).

Appendix B: Border effects

B.1. Homogeneous coefficients. Following Anderson and van Wincoop (2003) we decompose the border effects into two components: the trade-boosting intranational effect and the trade-reducing international effect of the border. To disentangle the two components and to retrieve the full implied border effect (both intranational and international), we proceed as follows. First, we define the border effects as the ratio of trade flows in a world with borders to that which would prevail in a borderless world. Let X_{ij} denote the former and $\overline{X}_{ij}$ the latter. Using (10) and (12), we then have

$$B_{ij} \equiv \frac{X_{ij}}{\overline{X}_{ij}} = \mathrm{e}^{\theta\left[b_{ij} - \sum_k \frac{L_k}{L} b_{kj}\right]} \prod_k \left(\frac{X_{kj}}{\overline{X}_{kj}}\right)^{\rho \frac{L_k}{L}}, \tag{B.1}$$

where the term $\mathrm{e}^{\theta\left[b_{ij} - \sum_k \frac{L_k}{L} b_{kj}\right]}$ subsumes the border frictions as a deviation from their population-weighted average. Note that (B.1) defines a log-linear system of all the relative trade flows, which depend on all border effects. Let $\mathbf{B}$ stand for the $n^2 \times 1$ vector of the $\ln(X_{ij}/\overline{X}_{ij})$ and let $\mathbf{b}$ stand for the $N^2 \times 1$ vector of the $\left[b_{ij} - \sum_k \frac{L_k}{L} b_{kj}\right]$. The log-linearized version of the system has the following solution, $\mathbf{B} = \theta(\mathrm{I} - \rho\mathbf{W})^{-1}\mathbf{b}$, which allows us to retrieve the border effect as the exponential of the foregoing expression.

Note that (B.1) quite naturally depends upon where regions i and j are located. Four cases may therefore arise with respect to Canada-U.S. trade. Let $\text{popCA} \equiv \sum_{k \in \text{CA}} \frac{L_k}{L}$ (resp., $\text{popUS} \equiv \sum_{k \in \text{US}} \frac{L_k}{L}$) stand for the Canadian (resp., the U.S.) population share. It is readily verified that

$$\theta\left[b_{ij} - \sum_k \frac{L_k}{L} b_{kj}\right] = \begin{cases} -\theta\,\text{popUS} & \text{if} \quad i \in \text{CA}, j \in \text{CA} \\ \theta\,\text{popUS} & \text{if} \quad i \in \text{CA}, j \in \text{US} \\ \theta\,\text{popCA} & \text{if} \quad i \in \text{US}, j \in \text{CA} \\ -\theta\,\text{popCA} & \text{if} \quad i \in \text{US}, j \in \text{US} \end{cases} \tag{B.2}$$

The explicit solution for $\ln B_{ij}$ is then given by

$$\ln B_{ij} = \theta\left[(\mathbf{I} - \rho\mathbf{W})^{-1}\right]_i \mathbf{b}, \tag{B.3}$$

where $[(\mathbf{I} - \rho\mathbf{W})^{-1})]_i$ denotes the i-th line of the matrix. Using (B.2) and (B.3), and the fact that $\mathbf{W}$ is row-standardized and has a special structure which implies that $\mathbf{W}\mathbf{b} = 0$, the border effects are finally given as follows:

$$\ln B_{ij} = \begin{cases} -\theta\,\text{popUS} & \text{if} \quad i \in \text{CA}, j \in \text{CA} \\ \theta\,\text{popUS} & \text{if} \quad i \in \text{CA}, j \in \text{US} \\ \theta\,\text{popCA} & \text{if} \quad i \in \text{US}, j \in \text{CA} \\ -\theta\,\text{popCA} & \text{if} \quad i \in \text{US}, j \in \text{US} \end{cases}$$

These expressions for the border effects reveal several interesting points. First, the expressions for CA-CA and U.S.-U.S. can be interpreted as the *trade-boosting* effect of the international border on flows within each country. Indeed, when ξ is positive and ρ is negative (as implied by our model), the trade flows within each country will be larger in a world with border than in a borderless world. The reason is that the border protects domestic firms from import competition and gives them an advantage in the home market. Second, the expressions for CA-U.S. and U.S.-CA can be interpreted as the *trade-reducing* effect of the international border on flows across countries. When ξ is positive and ρ is negative, the trade flows across countries will be smaller in a world with borders than in a borderless world. Third, as in Anderson and van Wincoop (2003), smaller countries will have larger implied border effects than large countries since their magnitude depends positively on the size of the trading partner, as measured by its population share. The reason is that the border affects the small country more than the large country, as it creates trade frictions for a larger share of the total demand served by its firms. Finally, the full border effect (combining the trade-boosting and trade-reducing effects), is given by $\mathrm{e}^{-2\xi\rho\,\mathrm{popUS}}$ for Canadian provinces and by $\mathrm{e}^{-2\xi\rho\,\mathrm{popCA}}$ for U.S. states.

B.2. Heterogeneous coefficients. In the heterogeneous coefficients model, we can retrieve the region-specific border effects in an analogous way to that presented in the foregoing Appendix B.1. Starting from (B.1), taking logarithms and rearranging, we readily obtain:

$$\ln X_{ij} - \ln \overline{X}_{ij} = \underbrace{\frac{\theta}{1-\rho\frac{L_i}{L}}}_{\overline{\theta}_i} \left[b_{ij} - \sum_k \frac{L_k}{L} b_{kj} \right] - \underbrace{\frac{\rho}{1-\rho\frac{L_i}{L}}}_{\overline{\rho}_i} \sum_{k \neq i} \frac{L_k}{L} \left(\ln X_{kj} - \ln \overline{X}_{kj} \right). \tag{B.4}$$

Using the expressions established in Appendix B.1. (which remain unchanged in the heterogeneous coefficient case), as well as the same matrix notation, we then obtain:

$$\ln B_{ij} = \overline{\theta}_i \left[\mathbf{I} - \overline{\rho} \otimes (\mathbf{W} - \mathbf{W_d}) \right]_i^{-1} \mathbf{b}.$$

The only change with respect to the homogeneous coefficient case is that the coefficient $\overline{\theta}_i$ captures the *local border frictions*, whereas $\overline{\rho}$ is a vector of elements accounting for the varying 'thoughness of competition' in the different regional markets.

Appendix C: Constrained specification

In this appendix, we derive a constrained version of equation (10) that integrates all the theoretical restrictions on the coefficients. This specification will be useful for estimation in the presence of heterogeneous coefficients. Starting from (10), we get

$$\begin{aligned}\ln\left(\frac{X_{ij}}{Y_iY_j}\right) &= \sigma\sum_k \frac{L_k}{L}\ln\frac{L_k}{L} + (\sigma-1)\ln Y_j - (\sigma-1)\left[\ln\tau_{ij} - \sum_k \frac{L_k}{L}\ln\tau_{kj}\right] \\ &- \sigma\left[\ln w_i - \sum_k \frac{L_k}{L}\ln w_k\right] - \sum_k \frac{L_k}{L}\ln Y_k - (\sigma-1)\sum_k \frac{L_k}{L}\ln X_{kj}.\end{aligned}$$

Using the aggregate income constraint $Y_i = L_i w_i$, and since $\sum_k (L_k/L) = 1$, we then have

$$\begin{aligned}\ln\left(\frac{X_{ij}}{Y_iY_j}\right) &= \sigma\sum_k \frac{L_k}{L}\ln\frac{L_k}{L} + (\sigma-1)\sum_k \frac{L_k}{L}\ln Y_j - (\sigma-1)\left[\ln\tau_{ij} - \sum_k \frac{L_k}{L}\ln\tau_{kj}\right] \\ &- \sigma\left[\ln w_i - \sum_k \frac{L_k}{L}\ln Y_k + \sum_k \frac{L_k}{L}\ln L_k\right] - \sum_k \frac{L_k}{L}\ln Y_k - (\sigma-1)\sum_k \frac{L_k}{L}\ln X_{kj} \\ &= -\sigma L - (\sigma-1)\left[\ln\tau_{ij} - \sum_k \frac{L_k}{L}\ln\tau_{kj}\right] - \sigma\ln w_i - (\sigma-1)\sum_k \frac{L_k}{L}\ln\left(\frac{X_{kj}}{Y_kY_j}\right).\end{aligned}$$

Some simple rearrangements the yield

$$\ln\left(\frac{X_{ij}L}{Y_iY_j}\right) = -(\sigma-1)\left[\ln\tau_{ij} - \sum_k \frac{L_k}{L}\ln\tau_{kj}\right] - \sigma\ln w_i - (\sigma-1)\sum_k \frac{L_k}{L}\ln\left(\frac{X_{kj}L}{Y_kY_j}\right).$$

The previous expression is a spatial autoregressive model with respect to the transformed explained variable $Z_{ij} \equiv (X_{ij}L)/(Y_iY_j)$:

$$\ln Z_{ij} = -(\sigma-1)\left[\ln\tau_{ij} - \sum_k \frac{L_k}{L}\ln\tau_{kj}\right] - \sigma\ln w_i - (\sigma-1)\sum_k \frac{L_k}{L}\ln Z_{kj}. \quad \text{(C.1)}$$

Note that (C.1) is structurally close to the estimating equations of both Feenstra (2002, 2004) and Anderson and van Wincoop (2003). In the case of local estimates with region-specific coefficients, $\ln w_i$ may be viewed as origin fixed effect, whereas $\sum_k \frac{L_k}{L}\ln Z_{kj}$ is a destination 'fixed effect' that incorporates the spatial equilibrium interdependence.

Table 1 — OLS regressions.

Model	OLS(1)	OLS(2)	OLS(3)	OLS(4)	OLS(5)	OLS(6)
Dependent variable	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln \frac{X_{ij}}{Y_i Y_j}$	$\ln \frac{X_{ij}}{Y_i Y_j}$	$\ln \frac{X_{ij}}{Y_i Y_j}$
Obs.	1600	1600	1600	1600	1600	1600
$\ln Y_i$	1.045	1.044	1.038	—	—	—
	(0.000)	(0.000)	(0.000)	—	—	—
$\ln Y_j$	0.920	0.919	0.913	—	—	—
	(0.000)	(0.000)	(0.000)	—	—	—
$\ln d_{ij}$	−1.172	−1.228	−1.144	−1.289	−1.385	−1.248
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bordCA_{ij}	2.674	2.734	2.655	—	—	—
	(0.000)	(0.000)	(0.000)			
bordUS_{ij}	0.393	0.397	0.412	—	—	—
	(0.000)	(0.000)	(0.000)			
bord_{ij}	—	—	—	−1.482	−1.504	−1.487
				(0.000)	(0.000)	(0.000)
fixed effects†	no	no	no	yes	yes	yes
internal distance	$\frac{1}{3}\sqrt{\frac{\text{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\text{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$	$\frac{1}{3}\sqrt{\frac{\text{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\text{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$
adjusted R^2	0.937	0.935	0.936	0.921	0.919	0.920
Border effect						
Canada	14.496	15.401	14.219	—	—	—
	(0.000)	(0.000)	(0.000)			
U.S.	1.482	1.487	1.510	—	—	—
	(0.000)	(0.000)	(0.000)			
"Average border"	4.634	4.786	4.633	4.404	4.501	4.422
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Moran's I stat.	0.038	0.043	0.035	−0.015	−0.015	−0.015
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are given in parentheses, those for border effect coefficients are computed using the Delta method. OLS(4), OLS(5) and OLS(6) include importer-exporter fixed effects. Following Feenstra (2002, 2004), average border effects are computed as the geometric mean of the individual border effects.

Table 2 — Homogeneous coefficients SARMA regressions.

Model	SARMA(1)	SARMA(2)	SARMA(3)	SARMA(4)	SARMA(5)	SARMA(6)
Dependent variable	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln \frac{X_{ij}}{Y_i Y_j}$	$\ln \frac{X_{ij}}{Y_i Y_j}$	$\ln \frac{X_{ij}}{Y_i Y_j}$
Obs.	1600	1600	1600	1600	1600	1600
Weight matrix	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$
constant	−10.844	−11.002	−10.357	−17.871	−17.515	−17.826
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\ln Y_i$	0.885	0.881	0.888	—	—	—
	(0.000)	(0.000)	(0.000)			
$\ln Y_j$	1.935	1.976	1.853	—	—	—
	(0.000)	(0.000)	(0.000)			
$(\mathrm{I}-\mathbf{W})\ln w_i$ or $\ln w_i$	−1.182	−1.236	−1.159	−0.923	−1.003	−0.953
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\ln d_{ij}$	−1.151	−1.214	−1.124	−1.224	−1.297	−1.190
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\mathrm{b}_{ij}$	−1.120	−1.137	−1.123	−1.062	−1.077	−1.069
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ρ	−0.831	−0.886	−0.755	−0.163	−0.154	−0.166
	(0.000)	(0.000)	(0.000)	(0.002)	(0.004)	(0.002)
λ	−4.404	−4.509	−3.998	−1.450	−1.336	−1.413
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
internal distance	$\frac{1}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$	$\frac{1}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$
AIC	−2.239	−2.267	−2.257	−2.273	−2.304	−2.283
BIC	−2.216	−2.243	−2.233	−2.258	−2.287	−2.266
Border effect[†] (total) (intra) / (inter)						
CA	7.285	7.511	7.331	6.581	6.745	6.967
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CA–CA / CA–US	2.699 / 0.371	2.741 / 0.365	2.708/0.369	2.565 / 0.390	2.598 / 0.385	2.640/0.379
	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)
US	1.289	1.294	1.290	1.272	1.276	1.281
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
US–US / US–CA	1.135 / 0.881	1.137 / 0.879	1.136 / 0.881	1.128 / 0.887	1.130 / 0.885	1.132 / 0.883
	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)
"Average border"	3.064	3.117	3.075	2.893	2.935	2.988
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are given in parentheses, those for border effect coefficients are computed using the Delta method. See Appendix B for an explanation of how to compute the border effects. Following Feenstra (2002, 2004), average border effects are computed as the geometric mean of the individual border effects. AIC and BIC stand for the Akaike and the Schwarz information criteria, respectively.

Table 3 — Homogeneous coefficients GSM and SAR regressions.

Model	GSM(1)	GSM(2)	GSM(3)	SAR(1)	SAR(2)	SAR(3)
Dependent variable	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$
Obs.	1600	1600	1600	1600	1600	1600
Weight matrix	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$
constant	−6.903	−6.952	−6.833	−5.807	−5.757	−5.839
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\ln Y_i$	1.256	1.273	1.244	1.062	1.058	1.066
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln Y_j$	0.959	0.956	0.955	0.980	0.979	0.975
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\ln w_i$ or $\ln w_i$	−1.135	−1.188	−1.129	−1.127	−1.180	−1.124
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
$(\mathrm{I}-\mathbf{W})\ln d_{ij}$	−1.249	−1.320	−1.217	−1.259	−1.331	−1.229
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\mathrm{b}_{ij}$	−1.063	−1.081	−1.071	−1.028	−1.093	−1.037
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ρ	−0.187	−0.214	−0.175	0.009	0.007	0.007
	(0.008)	(0.005)	(0.012)	(0.793)	(0.843)	(0.839)
λ	0.484	0.511	0.476	—	—	—
	(0.000)	(0.000)	(0.000)			
internal distance	$\frac{1}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$	$\frac{1}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$
AIC	−8.508	−8.538	−8.518	−2.290	−2.321	−2.299
BIC	−8.484	−8.514	−8.494	−2.267	−2.297	−2.275
Border effect† (total) (intra) / (inter)						
CA	6.593	6.796	6.680	6.190	6.361	6.287
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CA–CA / CA–US	2.568 / 0.389	2.601 / 0.384	2.585 / 0.387	2.488 / 0.402	2.522 / 0.397	2.507 / 0.399
	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)
US	1.272	1.277	1.275	1.262	1.267	1.265
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
US–US / US–CA	1.128 / 0.887	1.130 / 0.885	1.129 /0.886	1.124 / 0.890	1.125 / 0.889	1.125 / 0.889
	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)
"Average border"	2.896	2.946	2.918	2.795	2.984	2.820
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are given in parentheses, those for border effect coefficients are computed using the Delta method. See Appendix B for an explanation of how to compute the border effects. Following Feenstra (2002, 2004), average border effects are computed as the geometric mean of the individual border effects. AIC and BIC stand for the Akaike and the Schwarz information criteria, respectively.

Table 4 — Homogeneous coefficients SARMA regressions (US-US flows ×1.3).

Model	SARMA(1)	SARMA(2)	SARMA(3)	SARMA(4)	SARMA(5)	SARMA(6)
Dependent variable	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln \frac{X_{ij}}{Y_i Y_j}$	$\ln \frac{X_{ij}}{Y_i Y_j}$	$\ln \frac{X_{ij}}{Y_i Y_j}$
Obs.	1600	1600	1600	1600	1600	1600
Weight matrix	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$	$\mathbf{W}-\mathbf{W_d}$
constant	−11.549	−11.718	−11.067	−18.633	−18.333	−18.639
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\ln Y_i$	0.898	0.894	0.901	—	—	—
	(0.000)	(0.000)	(0.000)			
$\ln Y_j$	2.031	2.073	1.950	—	—	—
	(0.000)	(0.000)	(0.000)			
$(\mathrm{I}-\mathbf{W})\ln w_i$ or $\ln w_i$	−1.184	−1.241	−1.164	−0.712	−0.785	−0.737
	(0.002)	(0.001)	(0.003)	(0.002)	(0.009)	(0.002)
$(\mathrm{I}-\mathbf{W})\ln d_{ij}$	−1.155	−1.218	−1.126	−1.213	−1.286	−1.179
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(\mathrm{I}-\mathbf{W})\mathrm{b}_{ij}$	−1.278	−1.295	−1.281	−1.226	−1.240	−1.232
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ρ	−0.839	−0.893	−0.767	−0.187	−0.180	−0.191
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)
λ	−4.392	−4.502	−3.985	−1.850	−1.739	−1.815
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
internal distance	$\frac{1}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$	$\frac{1}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{2}{3}\sqrt{\frac{\mathrm{surf}_i}{\pi}}$	$\frac{1}{4}\min_j\{d_{ij}\}$
AIC	−2.225	−2.252	−2.243	−2.262	−2.293	−2.272
BIC	−2.201	−2.228	−2.219	−2.245	−2.276	−2.256
Border effect† (total) (intra) / (inter)						
CA	9.643	9.935	9.701	8.797	9.012	8.893
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CA–CA / CA–US	3.105 / 0.322	3.152 / 0.317	3.115/0.321	2.966 / 0.337	3.002 / 0.333	2.982/0.335
	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)
US	1.336	1.341	1.337	1.320	1.324	1.322
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
US–US / US–CA	1.156 / 0.865	1.158 / 0.864	1.156 / 0.865	1.149 / 0.870	1.151 / 0.869	1.150 / 0.870
	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)	(0.000)/(0.000)
"Average border"	3.589	3.650	3.601	3.408	3.455	3.429
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are given in parentheses, those for border effect coefficients are computed using the Delta method. See Appendix B for an explanation of how to compute the border effects. Following Feenstra (2002, 2004), average border effects are computed as the geometric mean of the individual border effects. AIC and BIC stand for the Akaike and the Schwarz information criteria, respectively.

Table 5 — β-heterogeneous SARMA regressions.

Model				SARMA				
Dependent variable				$\ln X_{ij}$				
Obs.				1600				
Weight matrix				$\mathbf{W} - \mathbf{W_d}$				
Region Code	$\ln d_{ij}$	p-value	$\ln b_{ij}$	p-value	$\varepsilon_{d_{ij}}$	b_{ij} intra	b_{ij} inter	Total border
AB	−1.645	0.000	−1.278*	0.000	−1.645	3.163*	0.316*	10.006*
BC	−1.001	0.000	−1.023*	0.001	−1.002	2.525*	0.396*	6.376*
MN	−2.147	0.000	−1.339*	0.000	−2.147	3.314*	0.302*	10.981*
NB	−1.967	0.000	−1.602*	0.000	−1.967	4.183*	0.239*	17.494*
Nfld	−3.321	0.000	0.127	0.700	−3.321	0.898	1.113	0.807
NS	−1.622	0.000	−1.337*	0.000	−1.623	3.306*	0.302*	10.929*
ON	−1.381	0.000	−1.695*	0.000	−1.381	4.825*	0.207*	23.284*
PEI	−1.972	0.000	−1.031*	0.001	−1.972	2.511*	0.398*	6.303*
Que	−1.152	0.000	−1.459*	0.000	−1.153	3.808*	0.263*	14.504*
SK	−1.865	0.000	−1.386*	0.000	−1.865	3.454*	0.289*	11.933*
Ala	−1.582	0.000	−2.621	0.333	−1.591	1.344	0.744	1.806
Ari	−1.315	0.000	−3.898	0.145	−1.322	1.558	0.649	2.427
Cal	−0.703	0.000	−2.468	0.300	−0.728	1.384	0.722	1.916
Flo	−1.187	0.000	0.143	0.957	−1.209	0.974	1.026	0.950
Geo	−1.670	0.000	−2.419	0.370	−1.685	1.319	0.758	1.738
Ida	−1.255	0.000	−2.783	0.301	−1.257	1.362	0.734	1.856
Ill	−1.253	0.000	−0.195	0.941	−1.273	1.015	0.985	1.031
Ind	−1.572	0.000	−2.891	0.284	−1.584	1.391	0.719	1.934
Ken	−1.598	0.000	−1.429	0.600	−1.606	1.170	0.855	1.370
Lou	−1.840	0.000	−2.122	0.431	−1.851	1.269	0.788	1.610
Mai	−1.398	0.000	−7.486*	0.006	−1.401	2.330*	0.430*	5.426*
Mas	−0.847	0.000	−2.663	0.319	−0.854	1.355	0.738	1.836
Mic	−1.607	0.000	0.885	0.739	−1.627	0.893	1.120	0.797
Min	−1.744	0.000	1.624	0.544	−1.755	0.822	1.217	0.675
MO	−1.855	0.000	0.828	0.758	−1.869	0.901	1.110	0.811
Mon	−1.824	0.000	0.544	0.840	−1.826	0.933	1.072	0.870
Mry	−1.015	0.000	0.434	0.871	−1.021	0.943	1.060	0.890
Nca	−1.272	0.000	−2.713	0.311	−1.284	1.365	0.733	1.863
Nda	−2.397	0.000	−5.245	0.053	−2.399	1.801	0.555	3.243
NHm	−1.063	0.000	−1.299	0.629	−1.065	1.151	0.869	1.324
NJr	−0.856	0.000	−3.807	0.152	−0.865	1.556	0.643	2.421
Nyr	−0.689	0.000	−0.652	0.800	−0.705	1.076	0.929	1.158
Ohi	−1.327	0.000	1.226	0.645	−1.346	0.856	1.168	0.733
Pen	−0.771	0.000	−1.367	0.604	−0.784	1.170	0.855	1.369
Ten	−1.676	0.000	−0.821	0.763	−1.688	1.092	0.916	1.192
Tex	−1.149	0.000	−1.521	0.556	−1.176	1.200	0.834	1.439
Ver	−1.215	0.000	−4.496	0.095	−1.216	1.654	0.605	2.735
Vir	−1.079	0.000	−5.472*	0.041	−1.088	1.885*	0.531*	3.552*
Was	−1.267	0.000	−3.650	0.175	−1.275	1.518	0.659	2.303
Wis	−1.596	0.000	−2.208	0.409	−1.607	1.283	0.790	1.645
ρ_{CA}				−0.987				
				(0.000)				
ρ_{US}				−1.302				
				(0.000)				
λ_{CA}				−4.716				
				(0.000)				
λ_{US}				−6.195				
				(0.000)				

Notes: p-values are given in parentheses, all are computed using the Delta method. See Appendix B for an explanation of how to compute the border effects. Asterisks denote significantly positive border effects at the 5% level. Distance elasticities $\varepsilon_{d_{ij}}$ are computed from the explicit form of the model.

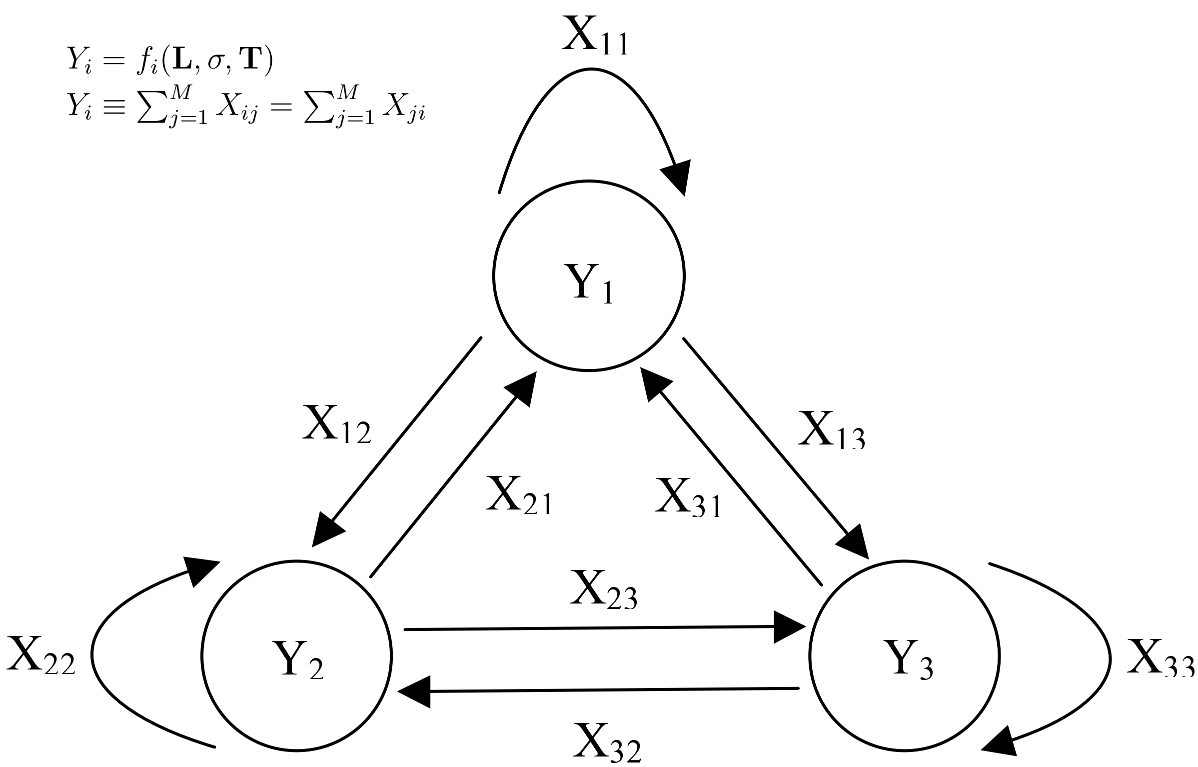

Figure 1. General equilibrium flows in a three-region trading network

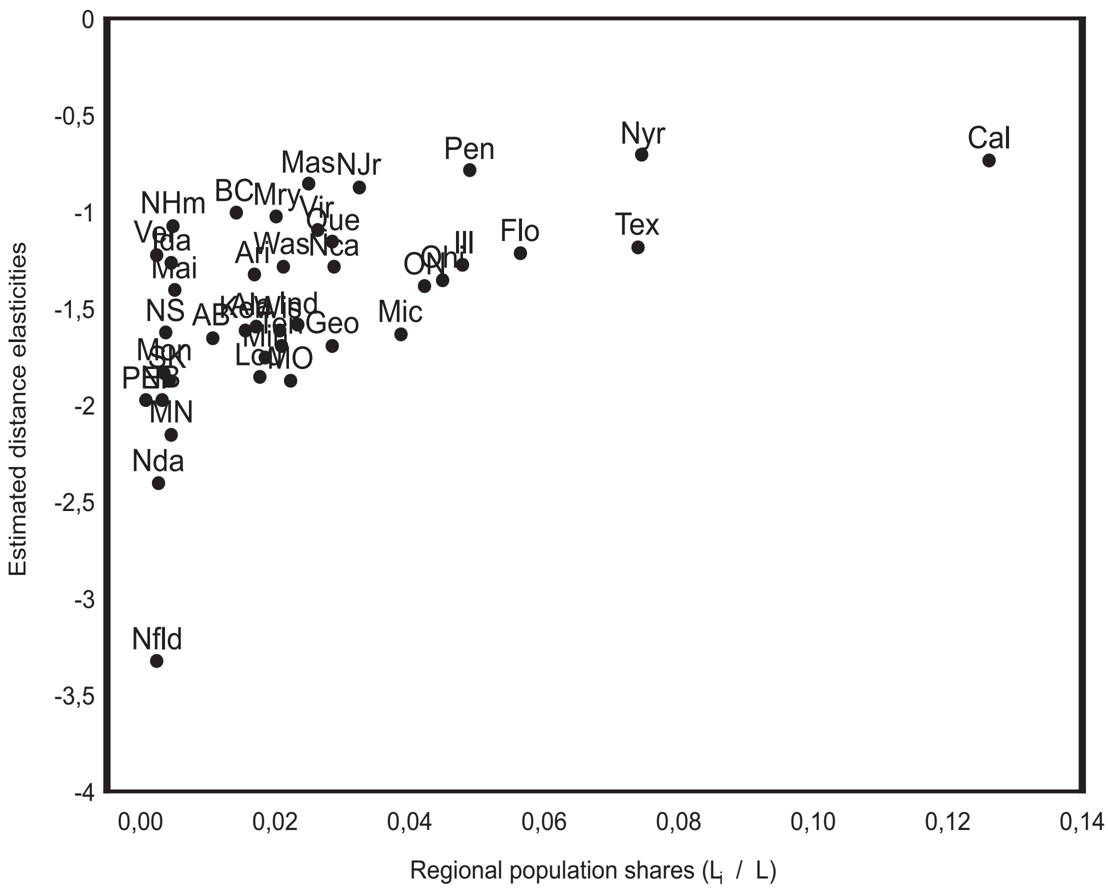

Figure 2. Distance elasticities and regional size (β-heterogeneous SARMA)

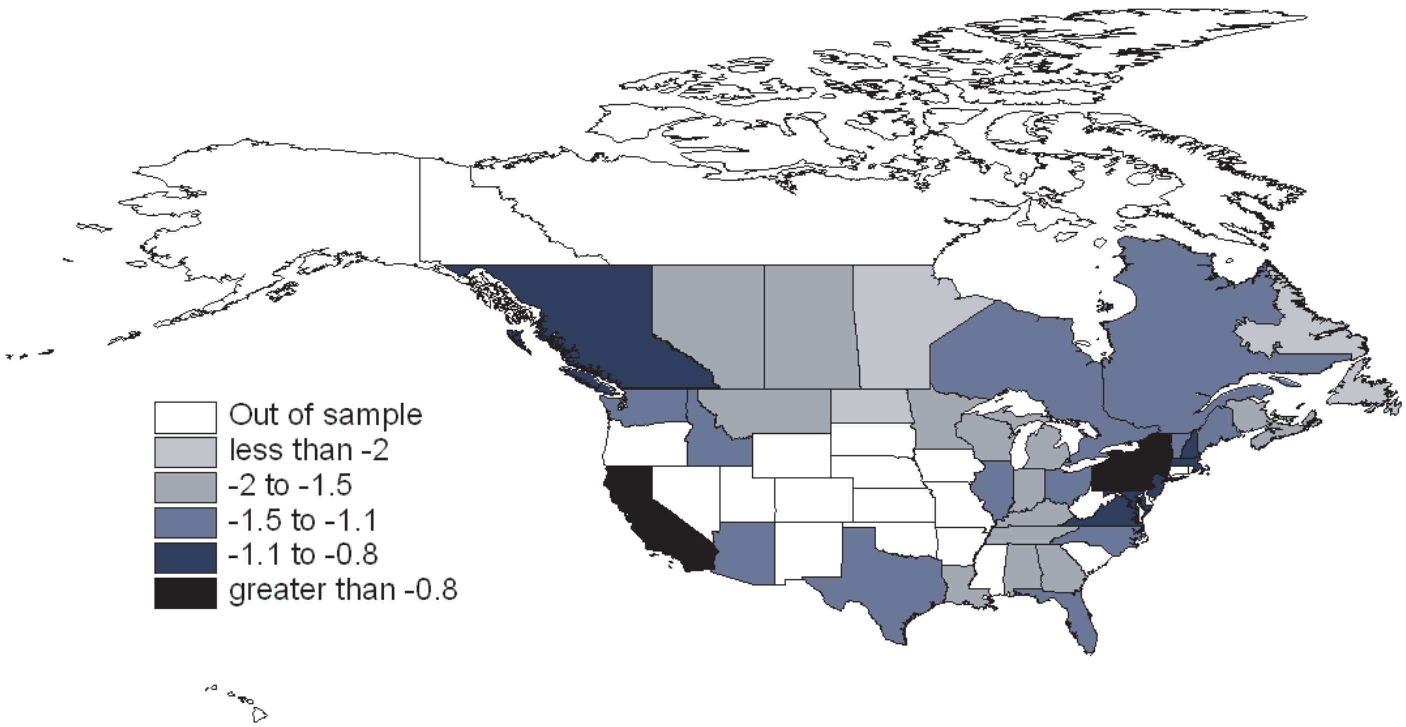

Figure 3. Regional structure of distance elasticities (β-heterogeneous SARMA)

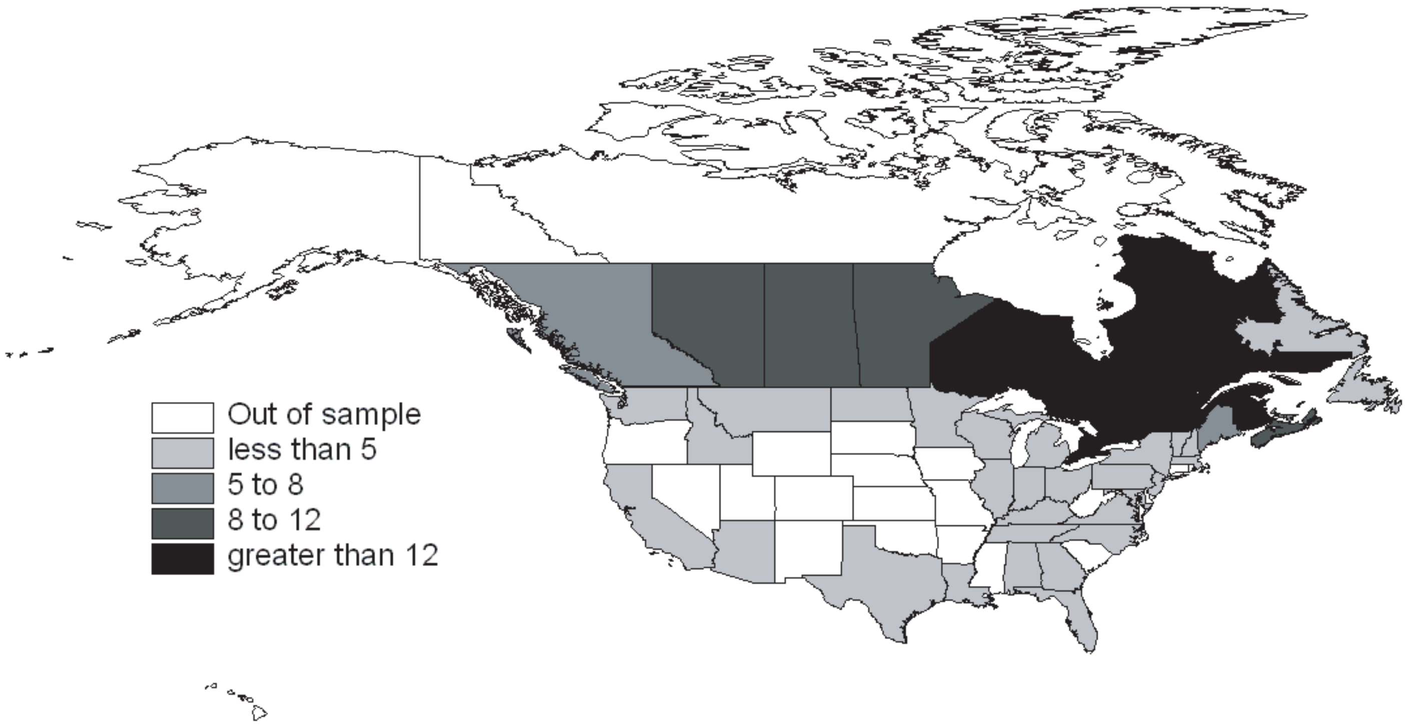

Figure 4. Regional structure of border effects (β-heterogeneous SARMA)

‘Dual’ gravity: Technical appendix (not intended for publication)

Kristian Behrens* Cem Ertur† Wilfried Koch‡

February 13, 2009

*Université du Québec à Montréal, Montréal, Canada; CIRPEE, Canada; and CORE, Université catholique de Louvain, Belgium. E-mail address: behrens.kristian@uqam.ca

†*Corresponding author:* LEO, Faculté de Droit d’Economie et de Gestion, Université d’Orléans, Rue de Blois – BP 6739, 45067 Orléans Cedex2, France. E-mail address: cem.ertur@univ-orleans.fr

‡LEG, Université de Bourgogne, France. E-mail address: kochwilfriedfr@aol.com

Appendix D: Log-likelihood and the information matrix

In this technical appendix, we derive the theoretical properties of the heterogeneous coefficients SARMA model with country-specific autoregressive parameters ($\overline{\rho}_j$ and $\overline{\lambda}_j$ for $j = 1, 2$) and region-specific non-autoregressive parameters ($\overline{\beta}_{0i}$, $\overline{\beta}_{1i}$, $\overline{\beta}_{2i}$, $\overline{\beta}_{3i}$, $\overline{\beta}_{4i}$ and $\overline{\theta}_i$ for $i = 1, ..., n$).

D.1. Model. To make notation as compact as possible, let $\mathbf{V}_i$ stand for the diagonal matrix defined by $\mathbf{V}_i \equiv \mathbf{E}_i \otimes \mathbf{I_n}$, where $\mathbf{E}_i = [\ 0 \mid 0 \mid \ \ldots e_i \ldots \ \mid 0 \mid 0\ ]$ with e_i (the i-th vector of the canonical base of $\mathbb{R}^n$) in position i and zero column vectors elsewhere. The diagonal matrix $\mathbf{V}_i$ is, therefore, a selection matrix with 1 on its main diagonal for the selected variables and 0 otherwise. Note that, by construction, $\sum_{i=1}^{n} \mathbf{V}_i = \mathbf{I_{n^2}}$. Analogously, let $\mathbf{D}_j$ stand for the diagonal selection matrix with 1 on its main diagonal for selecting canadian provinces or U.S. states, and 0 otherwise. Again, $\sum_{j=1}^{2} \mathbf{D}_j = \mathbf{I_{n^2}}$ by construction. Using the definitions of $\mathbf{V}_i$ and $\mathbf{D}_j$, the estimating equation (19) can be rewritten as follows:

$$\begin{aligned} \mathbf{X} &= \sum_i \mathbf{V_i} \left\{ \overline{\beta}_{0i} \mathbb{1} + \overline{\beta}_{1i} \mathbf{Y_d} + \overline{\beta}_{2i} \widetilde{\mathbf{Y}}_\mathbf{o} + \overline{\beta}_{3i} \widetilde{\mathbf{d}} + \overline{\beta}_{4i} \widetilde{\mathbf{w}} + \overline{\theta}_i \widetilde{\mathbf{b}} \right\} + \sum_j \mathbf{D}_j \, \overline{\rho}_j \mathbf{W_d} \mathbf{X} + \mathbf{u}, \\ &= \mathbf{Z} \overline{\beta} + \mathbf{D} \big(\overline{\rho} \otimes \mathbf{W_d} \big) \mathbf{X} + \mathbf{u}, \qquad \text{(D.1)} \end{aligned}$$

where

$$\mathbf{u} = \varepsilon + \mathbf{D}(\overline{\lambda} \otimes \mathbf{W_d})\varepsilon. \qquad \text{(D.2)}$$

In expressions (D.1) and (D.2), $\mathbf{W_d} \equiv \mathbf{W} - \mathbf{W_{diag}}$ denotes the spatial weight matrix without its diagonal elements; $\mathbf{Z} \equiv \mathbf{V}\,(\mathbf{I_n} \otimes \mathbf{M})$ denotes the $n^2 \times 6n$ block diagonal matrix of explanatory variables, with $\mathbf{M} \equiv [\ \mathbb{1} \mid \mathbf{Y_d} \mid \widetilde{\mathbf{Y}}_\mathbf{o} \mid \widetilde{\mathbf{d}} \mid \widetilde{\mathbf{w}} \mid \widetilde{\mathbf{b}}\ ]$; $\mathbf{V} \equiv [\ \mathbf{V_1} \mid \mathbf{V_2}\ \ldots\ \mathbf{V_i}\ \ldots\ \mathbf{V_n}\ ]$ stands for the $n^2 \times n^3$ selection matrix which extracts local subsamples from the full sample; $\overline{\beta}$ is the $6n \times 1$ vector of region-specific parameters; and $\overline{\rho}$ and $\overline{\lambda}$ are the 2×1 vectors of spatial autoregressive coefficients. Expressions (D.1) and (D.2) constitute the most compact and general specification of our model and will be useful for deriving the log-likelihood function and the information matrix.

Note that, in contrast to the SARMA model in the homogeneous case, we need to estimate two spatial autoregressive coefficients associated with different spatial weight matrices, the sum of which is equal to the spatial weight matrix that is used in the homogenous case ($\overline{\rho}_j = \rho$ and $\overline{\lambda}_j = \lambda$ for $j = 1, 2$). Letting $\mathbf{S}(\overline{\rho}) = \mathbf{I_{n^2}} - \mathbf{D}(\overline{\rho} \otimes \mathbf{W_d})$ and $\mathbf{S}(\overline{\lambda}) = \mathbf{I_{n^2}} - \mathbf{D}(\overline{\lambda} \otimes \mathbf{W_d})$, the equilibrium vector $\mathbf{X}$ is as follows:

$$\mathbf{X} = \mathbf{S}(\overline{\rho})^{-\mathbf{1}} \big[\mathbf{Z}\beta + \mathbf{S}(\overline{\lambda})\varepsilon \big], \qquad \text{(D.3)}$$

where $\mathbf{S}(\overline{\rho})$ and $\mathbf{S}(\overline{\lambda})$ are both non-singular. We propose to estimate this model by standard maximum likelihood techniques.

D.2. Log-likelihood. Let $\varepsilon(\theta) \equiv \mathbf{S}(\overline{\lambda})^{\mathbf{-1}} [\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{Z}\beta]$, where $\theta = \left[\ \beta' \mid \overline{\rho}' \mid \overline{\lambda}'\ \right]'$. The log-likelihood of (D.3) is then given by:

$$\ln L(\theta, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 + \ln|\mathbf{S}(\overline{\rho})| - \ln|\mathbf{S}(\overline{\lambda})| - \frac{1}{2\sigma^2}\varepsilon'(\theta)\varepsilon(\theta). \tag{D.4}$$

The Maximum Likelihood Estimators (MLE) $\hat{\theta}_{ML}$ and $\hat{\sigma}^2_{ML}$ are derived from the maximization of equation (D.4). In order to compute these estimators, it is convenient to work with the concentrated log-likelihood.

D.3. Estimators. The first-order conditions yield the following expressions for the estimators as a function of the autoregressive parameters:

$$\hat{\beta}_{ML}(\overline{\rho}, \overline{\lambda}) = \left[\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\right]^{-1} \mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X} \tag{D.5}$$

$$\hat{\sigma}^2_{ML}(\overline{\rho}, \overline{\lambda}) = \frac{1}{n}\mathbf{X}'\mathbf{S}'(\overline{\rho})\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{M}(\overline{\lambda})\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X}, \tag{D.6}$$

with $\mathbf{M}(\overline{\lambda}) \equiv \mathbf{I_{n^2}} - \mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\left[\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\right]^{-1}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}$ the $n^2 \times n^2$ projection matrix.

Proof. The first-order condition with respect to β is given by:

$$\nabla_\beta \ln L(\theta, \sigma^2) = 0 \iff \mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X} = \mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\beta,$$

which directly yields

$$\hat{\beta}_{ML}(\overline{\rho}, \overline{\lambda}) = \left[\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\right]^{-1} \mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X}.$$

The first-order condition with respect to σ^2 is given by:

$$\nabla_{\sigma^2} \ln L(\theta, \sigma^2) = 0 \iff -n + \frac{1}{\sigma^2}\varepsilon'(\theta)\varepsilon(\theta) = 0,$$

which directly yields

$$\hat{\sigma}^2_{ML}(\overline{\rho}, \overline{\lambda}) = \frac{1}{n}\varepsilon'(\theta)\varepsilon(\theta).$$

Using the definition of the projection matrix $\mathbf{M}(\overline{\lambda})$ we then obtain:

$$\hat{\sigma}^2_{ML}(\overline{\rho}, \overline{\lambda}) = \frac{1}{n}\mathbf{X}'\mathbf{S}'(\overline{\rho})\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{M}(\overline{\lambda})\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X},$$

which establishes the result. ■

D.4. Maximization of the concentrated log-likelihood. The concentrated log-likelihood can be rewritten as a function of the vectors $\overline{\rho}$ and $\overline{\lambda}$ as follows:

$$\begin{aligned} \ln L_c(\overline{\rho},\overline{\lambda}) &= -\frac{n}{2}(\ln(2\pi)+1) + \ln|\mathbf{S}(\overline{\rho})| + \ln|\mathbf{S}(\overline{\lambda})| \\ &- \frac{n}{2}\ln\left[\frac{\left(\mathbf{e}_0(\overline{\lambda}) - \sum_{i=1}^n \rho_i \mathbf{e}_i(\overline{\lambda})\right)'\left(\mathbf{e}_0(\overline{\lambda}) - \sum_{i=1}^n \rho_i \mathbf{e}_i(\overline{\lambda})\right)}{n}\right], \end{aligned} \tag{D.7}$$

where $\mathbf{e}_0(\overline{\lambda}) = \mathbf{M}(\overline{\lambda})\mathbf{S}(\overline{\lambda})^{-1}\mathbf{X}$, and where $\mathbf{e}_i(\overline{\lambda}) = \mathbf{M}(\overline{\lambda})\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d}\mathbf{X}$ for $i = 1,2$. Put differently, $\mathbf{e}_0(\overline{\lambda})$ is the vector of residuals of a regression of $\mathbf{X}$ on $\mathbf{Z}$, and $\mathbf{e}_i(\overline{\lambda})$ is the vector of residuals of a regression of $\mathbf{D}_i\mathbf{W_d}\mathbf{X}$ on $\mathbf{Z}$, for $i = 1,2$.

Proof. Note first that, using the expression for $\hat{\sigma}^2_{ML}(\overline{\rho},\overline{\lambda})$, we have the following relation: $\varepsilon'(\theta)\varepsilon(\theta) = n\hat{\sigma}^2_{ML}(\overline{\rho},\overline{\lambda})$. Moreover, using the expression of the projection matrix $\mathbf{M}(\overline{\lambda})$, it is straightforward to obtain the concentrated log-likelihood. ■

The MLEs of $\overline{\rho}$ and $\overline{\lambda}$, denoted respectively by $\widehat{\overline{\rho}}_{ML}$ and $\widehat{\overline{\lambda}}_{ML}$, maximize the concentrated log-likelihood (D.7). The MLEs of β and of σ^2 are then given by $\hat{\beta}_{ML} \equiv \beta_{ML}(\widehat{\overline{\rho}}_{ML},\widehat{\overline{\lambda}}_{ML})$ and by $\hat{\sigma}^2_{ML} \equiv \sigma^2_{ML}(\widehat{\overline{\rho}}_{ML},\widehat{\overline{\lambda}}_{ML})$, respectively.

D.5. Information matrix. The asymptotic covariance matrix of the maximum likelihood estimators is given by the inverse of the information matrix, which is defined as follows:

$$\mathbf{I}(\tilde{\theta}) = -E\left[\nabla^2_{\tilde{\theta},\tilde{\theta}'}\ln L(\tilde{\theta})\right] \tag{D.8}$$

with $\tilde{\theta} = (\theta',\sigma^2)'$. We can use the following estimator for this matrix:

$$\left[\hat{\mathbf{I}}(\hat{\tilde{\theta}})\right]^{-1} = \left[-\nabla^2_{\hat{\tilde{\theta}},\hat{\tilde{\theta}}'}\ln L(\hat{\tilde{\theta}})\right]^{-1} \tag{D.9}$$

To obtain this estimate, we need to compute that derivatives of the log-likelihood function.

D.6. First-order derivatives of the log-likelihood. We start with the first-order derivatives. By definition, $\varepsilon(\theta) = \mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\beta$. Because the transpose of a scalar is that scalar itself, we then obtain:

$$\begin{aligned} \varepsilon'(\theta)\varepsilon(\theta) &= \mathbf{X}'\mathbf{S}'(\overline{\rho})\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X} - 2\beta'\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X} \\ &+ \beta'\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\beta. \end{aligned} \tag{D.10}$$

The derivative of the log-likelihood with respect to β is given by:

$$\begin{aligned} \nabla_\beta \ln L(\theta,\sigma^2) &= -\frac{1}{2\sigma^2}\nabla_\beta\left[\varepsilon'(\theta)\varepsilon(\theta)\right] \\ &= -\frac{1}{2\sigma^2}\left[-2\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{S}(\overline{\rho})\mathbf{X} + 2\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}\beta\right] \\ &= \frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta). \end{aligned} \tag{D.11}$$

The derivative of the log-likelihood with respect to σ^2 is given by:

$$\nabla_{\sigma^2} \ln L(\theta, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{2}{4(\sigma^2)^2}\varepsilon'(\theta)\varepsilon(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\varepsilon'(\theta)\varepsilon(\theta). \tag{D.12}$$

The derivative of the log-likelihood with respect to $\overline{\rho}_i$, for $i = 1, 2$, is given by:

$$\nabla_{\overline{\rho}_i} \ln L(\theta, \sigma^2) = -\mathrm{tr}\left(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d}\right) + \frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}_\mathbf{d}'\mathbf{D}_i'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta). \tag{D.13}$$

Proof. To establish the expression for $\nabla_{\overline{\rho}_i} \ln L(\theta, \sigma^2)$, note that

$$\nabla_{\overline{\rho}} \ln L(\theta, \sigma^2) = \nabla_{\overline{\rho}} \ln |\mathbf{S}(\overline{\rho})| - \frac{1}{2\sigma^2}\nabla_{\overline{\rho}}\big(\varepsilon'(\theta)\varepsilon(\theta)\big)$$

Computation of the first term requires to apply the theorem for chain derivation of a matrix expression. Applying it for each element of the vector $\overline{\rho}$, we have:

$$\nabla_{\overline{\rho}_i} \ln |\mathbf{S}(\overline{\rho})| = \mathrm{tr}\left(\nabla_{\mathbf{S}(\overline{\rho})}(\ln |\mathbf{S}(\overline{\rho})|)'\nabla_{\overline{\rho}_i}\mathbf{S}(\overline{\rho})\right),$$

with $\nabla_{\mathbf{S}(\overline{\rho})} \ln |\mathbf{S}(\overline{\rho})| = (\mathbf{S}(\overline{\rho})')^{-1}$, and with

$$\nabla_{\overline{\rho}_i}\mathbf{S}(\overline{\rho}) = -\mathbf{D}\left[(\nabla_{\overline{\rho}_i}\overline{\rho}) \otimes \mathbf{W}_\mathbf{d} + \overline{\rho} \otimes (\nabla_{\overline{\rho}_i}\mathbf{W}_\mathbf{d})\right] = -\mathbf{D}(\mathbf{e}_i \otimes \mathbf{W}_\mathbf{d}) = -\mathbf{D}_i\mathbf{W}_\mathbf{d}.$$

As in the foregoing, $\mathbf{e}_i$ denotes the i-th vector of the canonical base, with 1 in position i and 0 otherwise. We then, therefore, obtain:

$$\nabla_{\overline{\rho}_i} \ln |\mathbf{S}(\overline{\rho})| = -\mathrm{tr}\left(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d}\right).$$

To compute the second term, note that

$$\begin{aligned}
\nabla_{\overline{\rho}_i}(\varepsilon'(\theta)\varepsilon(\theta)) &= \nabla_{\overline{\rho}_i}(\varepsilon'(\theta)\varepsilon(\theta)) + \varepsilon(\theta)\nabla_{\overline{\rho}_i}\varepsilon(\theta) \\
&= \mathbf{X}'\nabla_{\overline{\rho}_i}\mathbf{S}'(\overline{\rho})\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta) + \varepsilon'(\theta)\mathbf{S}(\overline{\lambda})^{-1}\nabla_{\overline{\rho}_i}\mathbf{S}(\overline{\rho})\mathbf{X} \\
&= 2\mathbf{X}'\nabla_{\overline{\rho}_i}\mathbf{S}'(\overline{\rho})\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta),
\end{aligned}$$

where we use the property that the transpose of a scalar is the scalar itself. We obtain:

$$\nabla_{\overline{\rho}_i}(\varepsilon'(\theta)\varepsilon(\theta)) = -2\mathbf{X}'(\mathbf{e}_i \otimes \mathbf{W}_\mathbf{d})'\mathbf{D}'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta) = -2\mathbf{X}'\mathbf{W}_\mathbf{d}'\mathbf{D}_i'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta)$$

Putting finally the expressions together, we have:

$$\nabla_{\overline{\rho}_i} \ln L(\theta, \sigma^2) = -\mathrm{tr}\left(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d}\right) + \frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}_\mathbf{d}'\mathbf{D}_i'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta)$$

for $i = 1, 2$, which establishes the result. ■

Next, the derivative of the log-likelihood with respect to the vector $\overline{\lambda}$ is given by:

$$\nabla_{\overline{\lambda}_i} \ln L(\theta, \sigma^2) = \mathrm{tr}\left(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d}\right) - \frac{1}{\sigma^2}\varepsilon'(\theta)\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta). \tag{D.14}$$

Proof. To begin with, note that

$$\nabla_{\overline{\lambda}} \ln L(\theta, \sigma^2) = -\nabla_{\overline{\lambda}} \ln |\mathbf{S}(\overline{\lambda})| - \frac{1}{2\sigma^2}\nabla_{\overline{\lambda}}(\varepsilon'(\theta)\varepsilon(\theta))$$

Computation of the first term requires to apply the theorem for chain derivation of a matrix expression. Applying it for each element of the vector $\overline{\lambda}$, we have:

$$\nabla_{\overline{\lambda}_i} \ln |\mathbf{S}(\overline{\lambda})| = \mathrm{tr}\left(\nabla_{\mathbf{S}(\overline{\lambda})} \ln |\mathbf{S}(\overline{\lambda})|'\nabla_{\overline{\lambda}_i}\mathbf{S}(\overline{\lambda})\right),$$

with $\nabla_{\mathbf{S}(\overline{\lambda})} \ln |\mathbf{S}(\overline{\lambda})| = (\mathbf{S}(\overline{\lambda})')^{-1}$, and with

$$\nabla_{\overline{\lambda}_i}\mathbf{S}(\overline{\lambda}) = -\mathbf{D}\left[(\nabla_{\overline{\lambda}_i}\overline{\lambda}) \otimes \mathbf{W}_\mathbf{d} + \overline{\lambda} \otimes (\nabla_{\overline{\lambda}_i}\mathbf{W}_\mathbf{d})\right] = -\mathbf{D}(\mathbf{e}_i \otimes \mathbf{W}_\mathbf{d}) = -\mathbf{D}_i\mathbf{W}_\mathbf{d}. \tag{D.15}$$

As in the foregoing, $\mathbf{e}_i$ denotes the i-th vector of the canonical base, with 1 in position i and 0 otherwise. We then, therefore, obtain:

$$\nabla_{\overline{\lambda}_i} \ln |\mathbf{S}(\overline{\lambda})| = -\mathrm{tr}\left(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d}\right).$$

To compute the second term, note that

$$\begin{aligned}
\nabla_{\overline{\lambda}_i}(\varepsilon'(\theta)\varepsilon(\theta)) &= \nabla_{\overline{\lambda}_i}\varepsilon'(\theta)\varepsilon(\theta) + \varepsilon'(\theta)\nabla_{\overline{\lambda}_i}\varepsilon(\theta) \\
&= \left[\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{Z}\beta\right]' \nabla_{\overline{\lambda}_i}\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta) + \varepsilon'(\theta)\nabla_{\overline{\lambda}_i}\mathbf{S}(\overline{\lambda})^{-1}\left[\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{Z}\beta\right] \\
&= 2\left[\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{Z}\beta\right]' \nabla_{\overline{\lambda}_i}\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta)
\end{aligned}$$

where we use the property that the transpose of a scalar is the scalar itself. We obtain:

$$\nabla_{\overline{\lambda}_i}\mathbf{S}'(\overline{\lambda})^{-1} = -\mathbf{S}'(\overline{\lambda})^{-1}\nabla_{\overline{\lambda}_i}\mathbf{S}'(\overline{\lambda})\mathbf{S}'(\overline{\lambda})^{-\mathbf{1}} = \mathbf{S}'(\overline{\lambda})^{-1}\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}.$$

Putting finally the expressions together, we have:

$$\nabla_{\overline{\lambda}_i}(\varepsilon'(\theta)\varepsilon(\theta)) = 2\left[\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{Z}\beta\right]' \mathbf{S}'(\overline{\lambda})^{-1}\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta)$$

and

$$\nabla_{\overline{\lambda}_i} \ln L(\theta, \sigma^2) = \mathrm{tr}\left(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d}\right) - \frac{1}{\sigma^2}\left[\mathbf{S}(\overline{\rho})\mathbf{X} - \mathbf{Z}\beta\right]' \mathbf{S}'(\overline{\lambda})^{-1}\mathbf{W}'_\mathbf{d}\mathbf{D_i}'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta),$$

for $i = 1, 2$, which establishes the result. ■

D.7. Second-order derivatives of the log-likelihood. We next turn to the second-order derivatives with respect to β. Deriving (D.11) with respect to β, we obtain:

$$\nabla^2_\beta \ln L(\theta,\sigma^2) = \frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\nabla_\beta\varepsilon(\theta) = -\frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{Z}. \tag{D.16}$$

Deriving (D.11) with respect to σ^2 yields:

$$\frac{\partial\big(\nabla_\beta \ln L(\theta,\sigma^2)\big)}{\partial\sigma^2} = -\frac{1}{(\sigma^2)^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta). \tag{D.17}$$

Taking the derivative of (D.11) with respect to $\overline{\rho}_j$ yields:

$$\begin{aligned}\frac{\partial\big(\nabla_\beta \ln L(\theta,\sigma^2)\big)}{\partial\overline{\rho}_j} &= \frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\nabla_{\overline{\rho}_j}\varepsilon(\theta)\\ &= \frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\nabla_{\overline{\rho}_j}\mathbf{S}(\overline{\rho})\mathbf{X}\\ &= -\frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W}_\mathbf{d}\mathbf{X}\end{aligned} \tag{D.18}$$

for $j=1,2$. Finally, the derivative of (D.11) with respect to $\overline{\lambda}_j$ is given by:

$$\begin{aligned}\frac{\partial\big(\nabla_\beta \ln L(\theta,\sigma^2)\big)}{\partial\overline{\lambda}_j} &= \frac{1}{\sigma^2}\mathbf{Z}'\left[\frac{\partial\mathbf{S}'(\overline{\lambda})^{-\mathbf{1}}}{\partial\overline{\lambda}_j}\varepsilon(\theta)+\mathbf{S}'(\overline{\lambda})^{-\mathbf{1}}\frac{\partial\varepsilon(\theta)}{\partial\overline{\lambda}_j}\right] \\ &= \frac{1}{\sigma^2}\mathbf{Z}'\big[\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta)\\ &\quad +\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W}_\mathbf{d}\mathbf{S}(\overline{\lambda})^{-1}(\mathbf{S}(\overline{\rho})\mathbf{X}-\mathbf{Z}\beta)\big]\\ &= \frac{1}{\sigma^2}\mathbf{Z}'\mathbf{S}'(\overline{\lambda})^{-1}\left[\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1}+\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W}_\mathbf{d}\right]\varepsilon(\theta)\end{aligned} \tag{D.19}$$

for $j=1,2$. We next derive (D.12) with respect to σ^2 to obtain the following second-order derivative:

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial(\sigma^2)^2} = \frac{n}{2(\sigma^2)^2}-\frac{1}{(\sigma^2)^3}\varepsilon'(\theta)\varepsilon(\theta). \tag{D.20}$$

The derivative of (D.12) with respect to $\overline{\rho}_j$ is computed as follows:

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\sigma^2\partial\overline{\rho}_j} = \frac{1}{2(\sigma^2)^2}\nabla_{\overline{\rho}_j}\left(\varepsilon'(\theta)\varepsilon(\theta)\right) = -\frac{1}{(\sigma^2)^2}\mathbf{X}'\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta), \tag{D.21}$$

for $j=1,2$. The derivative of (D.12) with respect to λ_j is given by:

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\sigma^2\partial\overline{\lambda}_j} = \frac{1}{2(\sigma^2)^2}\nabla_{\overline{\lambda}_j}\left(\varepsilon'(\theta)\varepsilon(\theta)\right) = \frac{1}{(\sigma^2)^2}\varepsilon'(\theta)\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta), \tag{D.22}$$

for $j=1,2$. We next derive (D.13) with respect to $\overline{\rho}_j$:

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\overline{\rho}_i\partial\overline{\rho}_j} = -\frac{\partial\big(\mathrm{tr}(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W}_\mathbf{d})\big)}{\partial\overline{\rho}_j}+\frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\nabla_{\overline{\rho}_j}\varepsilon(\theta) \tag{D.23}$$

for $j = 1, 2$. Since

$$\begin{aligned}
\frac{\partial\big(\mathrm{tr}(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W_d})\big)}{\partial\overline{\rho}_j} &= \mathrm{tr}\left(\nabla_{\overline{\rho}_j}\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W_d}\right) \\
&= \mathrm{tr}\left(-\mathbf{S}(\overline{\rho})^{-1}\nabla_{\overline{\rho}_j}\mathbf{S}(\overline{\rho})\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W_d}\right) \\
&= \mathrm{tr}\left(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_j\mathbf{W_d}\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W_d}\right),
\end{aligned}$$

and since

$$\nabla_{\overline{\rho}_j}\varepsilon(\theta) = \mathbf{S}(\overline{\lambda})^{-1}\nabla_{\overline{\rho}_j}\mathbf{S}(\overline{\rho})\mathbf{X} = -\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W_d}\mathbf{X},$$

we finally obtain:

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\overline{\rho}_i\partial\overline{\rho}_j} = -\mathrm{tr}\left(\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_j\mathbf{W_d}\mathbf{S}(\overline{\rho})^{-1}\mathbf{D}_i\mathbf{W_d}\right) - \frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W_d}\mathbf{X}. \tag{D.24}$$

We next derive (D.13) with respect to $\overline{\lambda}_j$, which yields:

$$\begin{aligned}
\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\overline{\rho}_i\partial\overline{\lambda}_j} &= \frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\left[\nabla_{\overline{\lambda}_j}\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta) + \mathbf{S}'(\overline{\lambda})^{-1}\nabla_{\overline{\lambda}_j}\varepsilon(\theta)\right] \\
&= \frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\left[\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1} + \mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W_d}\right]\varepsilon(\theta),
\end{aligned} \tag{D.25}$$

for $j = 1, 2$. Finally, the derivative of (D.14) with respect to $\overline{\lambda}_j$ is computed as follows:

$$\begin{aligned}
\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\overline{\lambda}_i\partial\overline{\lambda}_j} &= \frac{\partial\big(\mathrm{tr}(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d})\big)}{\partial\overline{\lambda}_j} - \frac{1}{\sigma^2}\Big[\nabla_{\overline{\lambda}_j}\varepsilon'(\theta)\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta) \\
&+ \varepsilon'(\theta)\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\left(\nabla_{\overline{\lambda}_j}\mathbf{S}'(\overline{\lambda})^{-1}\varepsilon(\theta) + \mathbf{S}'(\overline{\lambda})^{-1}\nabla_{\overline{\lambda}_j}\varepsilon(\theta)\right)\Big]
\end{aligned} \tag{D.26}$$

for $j = 1, 2$. We have:

$$\begin{aligned}
\frac{\partial\big(\mathrm{tr}(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d})\big)}{\partial\overline{\lambda}_j} &= \mathrm{tr}\left(\nabla_{\overline{\lambda}_j}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d}\right) \\
&= \mathrm{tr}\left(-\mathbf{S}(\overline{\lambda})^{-1}\nabla_{\overline{\lambda}_j}\mathbf{S}(\overline{\lambda})\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d}\right) \\
&= \mathrm{tr}\left(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W_d}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d}\right)
\end{aligned}$$

so that, using the foregoing results, we obtain:

$$\begin{aligned}
\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial\overline{\lambda}_i\partial\overline{\lambda}_j} &= \mathrm{tr}\left(\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W_d}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_i\mathbf{W_d}\right) - \frac{1}{\sigma^2}\varepsilon'(\theta)\Big[\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1} \\
&+ \mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{W}'_\mathbf{d}\mathbf{D}'_j\mathbf{S}'(\overline{\lambda})^{-1} + \mathbf{W}'_\mathbf{d}\mathbf{D}'_i\mathbf{S}'(\overline{\lambda})^{-1}\mathbf{S}(\overline{\lambda})^{-1}\mathbf{D}_j\mathbf{W_d}\Big]\varepsilon(\theta).
\end{aligned} \tag{D.27}$$