

This is a repository copy of *Gravitymodels of interprovincial migration flows in Canada with hierarchical multifactor structure*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/168133/>

Version: Published Version

Article:

Shin, Yongcheol and Serlenga, Laura (2020) Gravitymodels of interprovincial migration flows in Canada with hierarchical multifactor structure. *Empirical Economics*. pp. 365-390. ISSN 0377-7332

<https://doi.org/10.1007/s00181-020-01938-9>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Gravity models of interprovincial migration flows in Canada with hierarchical multifactor structure

Laura Serlenga¹ · Yongcheol Shin²

Received: 6 February 2020 / Accepted: 3 September 2020
© The Author(s) 2020

Abstract

Following recent contributions on migration flows, we contribute to the literature by relaxing restrictions on how multilateral resistance to migration (MRM) may affect province-pair-specific migration flows. We follow recent advancements in the three-dimensional (3D) panel data models with a hierarchical multifactor structure and develop the more flexible specification for MRM. In addition to including unobserved global (country) factors with province-pair-specific coefficients, we can control for local origin (destination)-specific factors that have heterogeneous effects on destinations (origins). We apply the 3DCCE estimator advanced by Kapetanios et al. (J Econom, 2020) to an analysis of the determinants of interprovincial migration flows in Canada from 1976 to 2014. In particular, we find that the recent rise in the internal migration flows, registered in Canada from 2009 onwards, is more likely to be associated with the relative income inequality and network presence rather than the conventional long-run determinants such as income and unemployment differentials.

Keywords Multilateral resistance to migration cross section dependence · Multi-dimensional panel data models · Unobserved global and local factors · Interprovincial migration

JEL Classification C13 · C33 · F14 · F45

We are mostly grateful for the insightful comments by the editor and two anonymous referees for the Special Issue in honour of Badi Baltagi as well as George Kapetanios, Camilla Mastromarco and Chaowen Zheng. The usual disclaimer applies.

✉ Yongcheol Shin
yongcheol.shin@york.ac.uk
Laura Serlenga
lserlenga@gmail.com; laura.serlenga@uniba.it

¹ University of Bari, Bari, Italy

² University of York, York, UK

1 Introduction

A growing number of studies have analysed the importance of controlling cross section dependence (CSD, hereafter) in gravity model of migration flows. In particular, following the seminal study by Anderson and van Wincoop (2003), Bertoli and Fernández-Huertas Moraga (2013) defined multilateral resistance to migration (MRM, hereafter) as the influence that alternative destinations exert on the bilateral migration flows and developed a theoretical framework for the gravity equation through a random utility maximisation (RUM) model and showed that MRM can be controlled by a multifactor error structure. They applied the common correlated effects (CCE) methodology advanced by Pesaran (2006) to analysing international migration flows from 61 origin countries to Spain. The importance of properly taking into account MRM has been emphasised in a recent survey by Beine et al. (2016) such that the regression residuals should be cross-sectionally independent for the estimation to be consistent with the RUM model.

In the literature on gravity models, a number of studies have also highlighted the importance of explicitly developing the multi-dimensional models, e.g. Mitze (2016) and Mátyás (2017). Notice that the existing studies by Bertoli and Fernández-Huertas Moraga (2013) and Piras (2017) have attempted to control MRM through error components including bilateral pair fixed effects and unobserved (global) factors, but their approach is equivalent to the 2-dimensional multifactor model, ignoring the 3-dimensional characteristic of the data. Indeed, the main drawback of this approach lies in the fact that it does not account for the hierarchical factor structure in 3D panels. In this paper, we follow recent advancements in the three-dimensional (3D) panel data models with a hierarchical multifactor structure and relax restrictions on how MRM may affect country/province-pair-specific migration flows. Recently, Kapetanios et al. (2017, 2020) have extended the CCE methodology, mainly developed in the 2D panel data by Pesaran (2006), into the 3D panels. In particular, Kapetanios et al. (2020) propose a 3DCCE estimator, which explicitly accommodates cross section dependence (CSD), by employing the local and global cross section averages of dependent and independent variables as proxies for unobserved global and local factors, respectively. Kapetanios et al. (2020) show that the CCE estimator proposed by Pesaran (2006) is not effective in removing correlations between local factors and regressors within 3D panels; under these circumstances, it leads to biased estimates. In a 3D framework, in addition to including unobserved global factors with pair-specific coefficients, we can also control for origin (destination)-specific factors that have heterogeneous effects on destinations (origins), which can lead to the more flexible specification of MRM than the models employed in the existing empirical literature on migration flows. We apply these 3D estimation methodologies to the interprovincial migration flows in Canada from 1976 to 2014.

In Canada, there were slightly more than 300,000 interprovincial migrants in 2014, representing 0.85% of the Canadian population. Interprovincial migration flows can provide significant economic benefits by reallocating labour from the low-productivity regions with high unemployment to the high-productivity regions with low unemployment. Shannon (2015) highlights that the lack of immigration in some regions in Canada may raise concerns about the economic future of these regions, because

little immigration implies declining population inflows as well as an inability to use immigration to deal with regional skill shortages. Indeed, this is an important policy issue, intensely debated in other developed countries due to the worsening of the economic gap among regions within the same country (e.g. the urban–rural or the North–South divide in Italy, see Piras 2017). To deal with this issue, the Canadian government introduced the Provincial Nominees Program, which aims to create a balanced regional distribution of immigrants, see (Pandey and Townsend 2011) and (Baglay 2012).

The structural factors—such as regional differences in earnings, employment prospects and labour productivity—have been highlighted as the main drivers behind interprovincial migration in Canada, e.g. Coulombe (2006). Day and Winer (2006) apply a provincial fixed effect model for net migration flows (across age groups) among ten Canadian provinces and document evidence that interprovincial migration flows are driven mainly by the long-run regional differentials in unemployment rates and labour productivity, as well as the rural/urban differential structure of the provinces. Focusing on the migration flows from 9 provinces to Ontario, Basher and Fachin (2008) address the issue of CSD in examining the long-run determinants of interprovincial migrations over the period 1971–2004. They find a very strong link between migration flows and unemployment differentials and income in the sending province, in line with the earlier studies, e.g. Coulombe (2006) and Day and Winer (2006). The evidence on income effect is mixed, with the level of home income more important than relative income differentials.

In general, there is a strand of literature that examines regional income disparity across Canadian provinces. Convergence in income is likely to weaken the economic motivation for migration while a lack of convergence is likely to trigger migration from poorer to richer provinces. Though a complete convergence for all provinces may be impossible to achieve, there is a general consensus on the existence of a convergence process. The increased efforts by the federal government to establish interprovincial redistribution programmes (*i.e.* equalisation payments) may have facilitated such convergence process. Brown and Macdonald (2015) show that the episodes of divergence coincide with the largest economic shocks in the twentieth century. In the time span we consider, they explain the first episode of divergence coinciding with the first oil shock from 1973 to the mid-1980s while the second occurring at the beginning in 1996 and continuing through the resource boom of the 2000s.

We follow the pioneering work by Harris and Todaro (1970), who propose to explain the migration flows by the difference in expected income. In this regard, we examine the effects of the conventional determinants such as income and unemployment rate differentials. We also aim at investigating the role of networks (*i.e.*, a cluster of migrants of the same origin residing in the destination) and relative inequality, defined as the Gini coefficient in the origin divided by the Gini coefficient in the destination. Indeed, expat migrant networks play an important role on the migration decisions of potential migrants. Through the informational and financial support provided by the network, newcomers can lower their migration and assimilation costs, see Beine and Parsons (2015) and Beine et al. (2019). Further, as argued by Stark (1984), relative deprivation together with absolute income is an important factor to explain migration decisions. The relative deprivation hypothesis states that an improvement in an

agent's relative income improves her welfare. Unlike most existing studies, which may impose inconvenient restrictions on the type of unobserved heterogeneity that may lead to omitted variable bias, we analyse the determinants of interprovincial migration flows in Canada, using a 3D panel data model with the hierarchical multi-factor error structure, that accounts for CSD through unobserved global (country) and local (origin and destination provincial) factors, respectively.

Our main empirical findings are summarised as follows: an increase in income at the destination (at the origin) significantly increases (reduces) migration outflows whereas an increase in the unemployment rate at the destination (at the origin) significantly decreases (increases) migration flows. Further, we find evidence of a nonlinear effect of the relative inequality, consistent with the predictions of Borjas' (1987) selection model, and a significant effect of networks which relates to the positive effect of diasporas analysed by Beine et al. (2011). Our evidence is broadly consistent with previous studies on interprovincial Canadian migration, see Coulombe (2006) and Day and Winer (2006) for the positive effects of increasing employment opportunities and income differential. Furthermore, by conducting the rolling window-based time-varying estimation over the last 10 years of the sample, 2005–2014, we find that the effect of the unemployment differentials had been quite small but relatively stable, and the effect of the income differentials, after registering a peak in 2007, had steadily declined. By contrast the impacts of both relative inequality and network had displayed an increasing trend. These time varying patterns suggest that the recent rise in the internal migration flows, registered in Canada from 2009 onwards, is more likely to be associated with the relative income inequality and network presence rather than the conventional long-run determinants such as income and unemployment differentials.

Our empirical findings might provide a support for developing an appropriate Government policy such as the Canadian Provincial Nominees Program that aims at reallocating the labour supply across regions or provinces. Indeed, like many other developed countries, Canada faces declining population growths in certain regions. In an attempt to stem this phenomenon, fiscal and income redistribution policies should be designed to influence migration location choices on the basis of significant push and pull determinants.

The paper proceeds as follows: Sect. 2 overviews the literature on the gravity model of migration flows. Section 3 describes the 3D panel data modelling with hierarchical multifactor structure. Section 4 presents main empirical findings for Canadian interprovincial migration flows. Section 5 offers concluding remarks.

2 Overview on the gravity models of migration flows

The dependence across cross section units have been modelled explicitly in the 2D panels by two main approaches: the factor-based approach (e.g. Pesaran 2006; Bai, 2009) and the spatial modelling techniques (e.g. Behrens et al. 2012; Mastromarco et al. 2016). Chudik et al. (2011) show that the factor-based models exhibit the strong CSD, while the spatial-based models can deal with weak CSD. In the existing empirical literature, these approaches have been separately implemented to estimate gravity models of migration flows.

Bertoli and Fernández-Huertas Moraga (2013) follow the seminal study by Anderson and van Wincoop (2003) and control for the multilateral resistance to migration (MRM) through time-varying unobserved common factors. They apply the CCE approach to a gravity model of international migration flows from 61 origin countries to Spain. Migrants are faced with a discrete choice problem when selecting alternative destinations, as domiciles are imperfect substitutes and migration involves frictions. The importance of properly taking into account the MRM has been emphasised in a recent survey by Beine et al. (2016), where they highlight that the regression residuals should be cross-sectionally independent for the estimation to be consistent with the random utility maximisation (RUM) model. Crucially, Bertoli and Fernández-Huertas Moraga (2013) state that MRM introduces omitted variable bias in estimating the impacts of the determinants of bilateral migration flows. In this regard, it is widely acknowledged that the inclusion of destination-time dummies or origin-time dummies as proxies for MRM in receiving and sending countries, as proposed by Ortega and Peri (2013) and Beine and Parsons (2015), is likely to produce biased estimation results in the presence of strong CSD, see also Kapetanios et al. (2017).

Piras (2017) considers a triple-index panel gravity model of internal migration flows with the human capital across Italian regions during 1970–2005. Finding evidence of strong CSD by applying the cross section dependence (CD) test by Pesaran (2015), he applies the CCE estimator together with the mean group estimator by Pesaran and Smith (1995) and the augmented mean group estimator by Eberhardt and Teal (2010). The main results confirm that the macroeconomic variables (per capita GDP and unemployment) are the leading drivers behind migration flows across Italian regions. Furthermore, the human capital has no role at the destination while it works as a restraining factor at the origin. Piras (2017) justifies this finding with the severe territorial differences, which characterise the North-South divide in Italy. However, to control for strong CSD, he imposes that the error components contain the bilateral pair fixed effects and unobserved heterogeneous factors [see Eq. 9], which are essentially the same as the 2D specification.

Beenstock et al. (2015) highlight the importance of modelling the relationship between origin and destination country as well as the relationship between origin/destination and a third country, which might be connected to another destination. Since the migration flows are essentially multilateral, the standard approach using a bilateral specification, that ignores the multilateral or third-country effects, would lead to biased estimation and misleading inference. They suggest to control third country effects by applying the spatial econometric techniques, where the spatial effects are allowed to be different at the origin and at the destination. They apply the double-spatial lag ML estimator developed by Elhorst et al. (2012) to an analysis of the determinants of immigration from European neighbourhoods to the EU over the period 2000–2010. They find that unemployment rates, social spending per head and inequality affect migration. In the trade flows, Behrens et al. (2012) adapt the approach by LeSage and Pace (2008), and derive the spatial connectivity matrices, interpreted as a multilateral linkage. LeSage and Llano (2016) extend the (unilateral) spatial dependence into a multilateral case where the (latent) multilateral spatial effects are estimated for both origin and destination. In particular, they propose to combine the Bayesian hierarchical approach and the Spatial autoregressive (SAR) model to estimate the regional effects.

Mitze (2016) aims to introduce the ‘competing destinations’ effect by applying a spatial Durbin model to interregional migration flows in Germany. The basic intuition is to model the human behaviour as a spatial choice process such that an actual migration choice is made through the hierarchical information processing as migrants are supposed to evaluate a limited number of alternatives. Prospective migrants attempt to simplify these alternatives by categorising them into clusters, where the probability of one destination in a certain cluster being selected is related to the other regions in that cluster. This clustering process requires that the spatial proximity of destinations has an influence on the destination choice by migrants from one particular origin.

In the case of Canadian interprovincial migration flows, Day and Winer (2006) apply a triple index model to investigate the impacts of public policy on interprovincial migration in Canada from 1974 to 1996. They take into account provincial fixed effects and analyse the effect of regional variation in federal unemployment insurance, provincial social assistance, federal and provincial personal income taxes and public spending of different types on interprovincial migration. They find that the prime determinants of interprovincial migration are differentials in earnings, employment prospects and moving costs, whereas the impacts of public policies are relatively small. Basher and Fachin (2008) apply the bootstrap-based panel cointegrating gravity model, proposed by Fachin (2007) that are robust to the presence of CSD. They employ a 2D gravity model to analyse migration flows to Ontario from other 9 provinces from 1971 to 2004, with the main aim to analyse the long-term decline of internal migration in Canada. They propose a bootstrap panel cointegration analysis for a model where migration flows depend on income, unemployment and Federal transfers and find that the main determinants of migration are unemployment differentials and income in the sending province, whereas income and federal transfer differentials appear to play only a minor role. They also show that interprovincial migration flows have been significantly reduced by shrinking differentials in the labour market and income growth in the sending provinces.

In sum, we may conclude that income and unemployment are key drivers of the inter-regional migration process. Further, given the multidimensional structure of interprovincial migration data, recent empirical literature has also underlined the importance of properly modelling the general specification of MRM in order to avoid omitted variable bias.

Given the growing availability of dataset which contain information on the multiple dimensions, the recent literature on panel data has focused on extending the two-way error components specification to the multidimensional setting. The triple-index specification has been popularised by Mátyás (2017), where time, origin and destination fixed effects are specified, respectively, as unobservables. Baltagi et al. (2003) propose an extended specification with fixed origin-time, destination-time and country-pair effects. Balazsi et al. (2017b) generalise the 3D within estimator while Balazsi et al. (2017c) consider the random effects approach and propose a sequence of GLS estimators. This multi-dimensional approach is an essential tool for the analysis of complex interconnectedness of the data, which can be applied to the number of bilateral flows between countries or regions such as trade, FDI, capital or migration flows.

In this paper, we focus on controlling strong CSD by following recent developments on the multidimensional panel data models. We extend the microfounded specifications proposed by Bertoli and Fernández-Huertas Moraga (2013) and Piras (2017) to a 3D framework with a hierarchical multifactor structure, through which we can deal with two sources of cross section correlations, one driven by global factors and the other driven by local factors. Allowing for such hierarchical multifactor structure would be extremely relevant in the context of the provincial data where cross section entities within a single country are deeply integrated and interconnected such that interprovincial migration flows might be affected by both global (country) and local (origin and destination provincial) effects.

3 The 3D panel data estimator with hierarchical multifactor structure

Migration choices are often explained by using a random utility maximisation (RUM) model. The utility that individual q , who was located in country i at time $t - 1$, derives from opting for country j belonging to the choice set D at time t is:

$$U_{qijt} = w_{ijt} - c_{ijt} + \epsilon_{qijt} \quad (1)$$

where w_{ijt} is a deterministic component of utility, c_{ijt} is the cost of moving from i to j , and ϵ_{qijt} is an individual-specific stochastic term. Assuming that ϵ_{qijt} follows an independent and identically distributed extreme value type 1 distribution (McFadden, 1974), Beine et al. (2016) express the expected gross migration flow from country i to j as

$$E(M_{ijt}) = \frac{e^{w_{ijt} - c_{ijt}}}{\sum_{l \in D} e^{w_{ilt} - c_{ilt}}} s_{it} \quad (2)$$

where s_{it} is the ability of the origin i to send out migrants. Assuming that the deterministic component of utility does not vary with i , they write (2) as a gravity equation, such that

$$E(M_{ijt}) = \phi_{ijt} \frac{y_{jt}}{\Omega_{it}} s_{it}.$$

where $\Omega_{it} = \sum_{l \in D} e^{w_{ilt} - c_{ilt}}$ represents the exponentiated value of the expected utility of prospective migrants, $y_{jt} = e^{w_{jt}}$ proxies for the attractiveness of destination j and $\phi_{ijt} = e^{-c_{ijt}}$ is the accessibility of destination j for potential migrants from i . Taking the ratio between $E(M_{ijt})$ and $E(M_{iit})$ and normalising ϕ_{iit} to 1, we obtain:

$$\frac{E(M_{ijt})}{E(M_{iit})} = \phi_{ijt} \frac{y_{jt}}{y_{it}}. \quad (3)$$

Bertoli and Fernández-Huertas Moraga (2013) stress the importance of allowing correlations in the stochastic component of utility across different alternatives in the

migrant choice set. Using a more general distributional assumption on the stochastic component in (1), which does not rely on the independence from irrelevant alternatives property and allows for a correlation in the stochastic component of utility across different alternatives in the choice set, they show that (3) is generalised to

$$\frac{E(M_{ijt})}{E(M_{iit})} = \phi_{ijt}^{1/\tau} \frac{y_{jt}^{1/\tau}}{y_{it}} \frac{\Omega_{iit}}{\Omega_{ijt}} \quad (4)$$

where τ is a dissimilarity parameter and Ω_{ijt} is a resistance term which varies across destinations, see equation (9) in Beine et al. (2016). Specifically, an increase in the attractiveness of an alternative destination that is perceived as a close substitute to destination j , will reduce $E(M_{ijt})$ more than $E(M_{iit})$, thus inducing a decline in the ratio in (4).¹

Taking the logarithm of the odds of migrating to country j over staying in country i , (3) can be expressed as a linear function of the differential in the deterministic component of utility associated with the two countries

$$\ln \left(\frac{M_{ijt}}{M_{iit}} \right) = \frac{\ln y_{jt}}{\tau} - \ln y_{it} + \frac{\ln \phi_{ijt}}{\tau} + \ln \left(\frac{\Omega_{iit}}{\Omega_{ijt}} \right). \quad (5)$$

In the empirical macro analysis of migration flows, the first three terms in (5), which represent attractiveness of destination and/or origin and the accessibility of destination in terms of the time-specific cost of moving, are generally proxied by push and pull variables which might vary over the two indexes, (i, t) and/or (j, t) , and over the three indexes, (i, j, t) .² On the contrary, the importance of taking into account of the resistance term Ω_{ijt} has been mostly neglected in the empirical literature, with a few exceptions such as Bertoli and Fernández-Huertas Moraga (2013), Ortega and Peri (2013) and Piras (2017), who embody the resistance term in a compound unobserved error term, see the error components specifications (8) and (9) below.

In this paper, we propose to model the migration flows by means of a more general 3D error components specification, which can properly address the important issue of MRM. Consider the 3D heterogeneous panel data model:

$$y_{ijt} = \beta'_{1,ij} \mathbf{x}_{ijt} + \beta'_{2,ij} \mathbf{x}_{it} + \beta'_{3,ij} \mathbf{x}_{jt} + \delta'_{ij} \mathbf{d}_t + e_{ijt}, \quad i = 1, \dots, N_i, \\ j = 1, \dots, N_j, \quad t = 1, \dots, T, \quad (6)$$

where y_{ijt} is the dependent variable observed across three indices, the origin i , the destination j at period t , \mathbf{x}_{ijt} is the $m_1 \times 1$ vector of covariates observed across three indices, \mathbf{x}_{it} (\mathbf{x}_{jt}) is the $m_2 \times 1$ ($m_3 \times 1$) vector of push (pull) covariates observed across the origin i (j) at period t , and \mathbf{d}_t is the $m_d \times 1$ vector of observed common effects including deterministic components such as constant and trend. $\beta_{1,ij}$, $\beta_{2,ij}$,

¹ This, in turn, questions the longstanding tradition in empirical literature of estimating the determinants of bilateral migration rates as a function of the attractiveness of i and j only (Hanson 2010).

² Bertoli and Fernández-Huertas Moraga (2013) note that the aggregate migration data do not necessarily allow to separately identify all the parameters of the RUM model.

$\beta_{3,ij}$ and δ_{ij} are the $m_1 \times 1$, $m_2 \times 1$, $m_3 \times 1$, and $m_d \times 1$ vectors of corresponding parameters.

Beine and Parsons (2015) attempt to model MRM in receiving countries by imposing the following simple error components specification:

$$e_{ijt} = \alpha_i + \theta_{jt} + \varepsilon_{ijt} \quad (7)$$

where α_i is the origin individual effect and θ_{jt} is the destination-time dummy. Alternatively, Ortega and Peri (2013) suggest to capture MRM by including the origin-time dummies θ_{it} :

$$e_{ijt} = \alpha_i + \theta_{it} + \varepsilon_{ijt}. \quad (8)$$

Notice that the destination/origin-time dummies have been widely used in the literature as proxies for unmeasured origin-specific push factors and destination-specific pull factors of bilateral migration flows.

To explicitly address the important issue of modelling CSD in migration flows, Bertoli and Fernández-Huertas Moraga (2013) and Piras (2017) propose to model MRM through the following multifactor structure:

$$e_{ijt} = \alpha_{ij} + \gamma'_{ij} f_t + \varepsilon_{ijt} \quad (9)$$

where f_t is an $m_f \times 1$ vector of unobserved factors and γ_{ij} is an $m_f \times 1$ vector of heterogeneous loadings. Bertoli and Fernández-Huertas Moraga (2013) define MRM as the confounding influence that the attractiveness of alternative destinations exerts on the determinants of bilateral migration rate and show that the heterogeneous unobserved factors are expected to capture MRM. Notice, however, that the standard CCE estimator developed in 2D panels by Pesaran (2006) can be applied to the model (6) with the error components in (9).

In the context of the 3D models, Baltagi et al. (2003) propose the following triple error components specification:

$$e_{ijt} = \alpha_{ij} + \theta_{it} + \theta_{jt}^* + \varepsilon_{ijt}, \quad (10)$$

where α_{ij} is the bilateral pair-fixed effects, and θ_{it} (θ_{jt}^*) are the origin and destination country-time fixed effects (CTFE), respectively. Notice that (10) encompasses (7) and (8) as a special case.³ Balazsi et al. (2017b) develop the 3D within estimator and analyse their behaviour in the case of no self-flow data, unbalanced data, and dynamic autoregressive models. The CTFE specification has been popularly applied to control for MRM of exporters and importers in the structural gravity model of trade flows, e.g. Baltagi (2015). However, the main drawback of the CTFE approach lies in the underlying assumption that bilateral flows are independent of what happens to the

³ Baltagi et al. (2003) and Baier and Bergstrand (2007) consider several forms of fixed effects such as $e_{ijt} = \alpha_i + \gamma_j + \theta_t + \varepsilon_{ijt}$, $e_{ijt} = \alpha_{ij} + \theta_t + \varepsilon_{ijt}$, $e_{ijt} = \theta_{it} + \varepsilon_{ijt}$, and $e_{ijt} = \theta_{jt}^* + \varepsilon_{ijt}$, $e_{ijt} = \theta_{it} + \theta_{jt}^* + \varepsilon_{ijt}$.

rest of the world.⁴ Existing studies have neglected the important issue of properly controlling CSD in 3D models. In this regard, Kapetanios et al. (Kapetanios et al. 2017, KMSS) propose to combine unobserved heterogeneous global factors, θ_t with the CTFE specification by

$$e_{ijt} = \alpha_{ij} + \theta_{it} + \theta_{jt}^* + \pi_{ij}\theta_t + \varepsilon_{ijt} \quad (11)$$

KMSS develop the two-step consistent estimation procedure by approximating global factors with the double cross section averages of the dependent variable and regressors and then applying the 3D-within transformation.

The above discussions have at least two implications: (i) it is important to develop more proper specifications for accommodating CSD within multi-dimensional dataset and (ii) such specifications should involve several layers of components to account for effects across different dimensions. Following this research trend, Kapetanios et al. (2020, KSS) have developed the hierarchical multi-factor error components specification:

$$e_{ijt} = \gamma'_{ij}f_t + \gamma'_{oj}f_{iot} + \gamma'_{io}f_{ojt} + \varepsilon_{ijt}, \quad (12)$$

where f_t is an $m_f \times 1$ vectors of global factors, and f_{iot} and f_{ojt} are the $m_{\bullet o} \times 1$ and $m_{o\bullet} \times 1$ vectors of local factors with γ_{ij} , γ_{oj} and γ_{io} being the $m_f \times 1$, $m_{\bullet o} \times 1$ and $m_{o\bullet} \times 1$ vectors of corresponding heterogeneous loadings. The specification in (12) is more parsimonious than (10) and (11), and it is more structural by allowing the presence of global and local factors, explicitly. Furthermore, we can also estimate the parameters of push and pull regressors, say x_{it} and x_{jt} , that vary over (i, t) and over (j, t) in the model, (6), unlike the CTFE and KMSS approaches, which remove x_{it} and x_{jt} together with α_{ij} , θ_{it} and θ_{jt}^* by applying the 3D within transformation.⁵

f_t are the global factors such as common demographic shift that affects every migration flows between source i and destination j . For example, if the Canadian population increases substantially, then migration flows are likely to increase between any pair of provinces. The local factors, f_{iot} and f_{ojt} , represent the origin and destination provincial factors, respectively. f_{iot} are the factors of the i th sending province (say, Newfoundland) that affect the migration flows from Newfoundland to all destination provinces while f_{ojt} are the factors of the j th receiving province (say, British Columbia) that affect all migration flows to British Columbia from all other provinces. If unemployment insurance increases in Newfoundland, migration outflows from Newfoundland to all other countries may decrease, whereas if unemployment insurance in

⁴ To control for multilateral cross section correlations across trade flows, Behrens et al. (2012) develop the spatial econometric technique in cross section. A number of studies have established the importance of taking into account multilateral resistance, trade costs and bilateral heterogeneity in 2D panels, e.g. Serlenga and Shin (2007) and Mastromarco et al. (2016).

⁵ To recover those coefficients, it would be worthwhile to develop an extension of the Hausman and Taylor (1981) estimation, which has been popular in the 2D panel data models even in the presence of cross-sectionally correlated errors (e.g. Serlenga and Shin 2007). Balazsi et al. (2017a) develop an extended Hausman–Taylor estimator for multi-dimensional models, but the complexity of such estimator increases remarkably as the number of unobserved heterogeneities increase with the dimension of the data.

British Columbia decreases, inflows from all other countries may decrease.⁶ This factor structure allows to model the interconnection among cross section units; if Québec and Newfoundland are connected by business cycle synchronisation as they are both specialised in the fishery industry located in the Atlantic coast, a positive shock in Québec will also lead to a positive cycle in Newfoundland.

KSS allow unobserved factors to be correlated with the explanatory variables, x_{ijt} , and consider the following data generating process for x_{ijt} :

$$x_{ijt} = \mathcal{D}_{ij}d_t + \Gamma_{ij}f_t + \Gamma_{oj}f_{iot} + \Gamma_{io}f_{ojt} + v_{ijt}, \tag{13}$$

where \mathcal{D}_{ij} is the $(m_x \times m_d)$ parameter matrix on observed common effects, Γ_{ij} , Γ_{oj} and Γ_{io} are $(m_x \times m_f)$, $(m_x \times m_{\bullet o})$, $(m_x \times m_{o\bullet})$ factor loading matrices, and v_{ijt} are the idiosyncratic errors. Let $z_{ijt} = (y_{ijt}, x'_{ijt})'$ and $f_{ijt} = (f'_t, f'_{iot}, f'_{ojt})'$, and construct the global and local cross section averages of z_{ijt} , respectively, by

$$\bar{z}_t = \frac{1}{N^2} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} z_{ijt}, \quad \bar{z}_{iot} = \frac{1}{N} \sum_{j=1}^{N_j} z_{ijt}, \quad \bar{z}_{ojt} = \frac{1}{N} \sum_{i=1}^{N_i} z_{ijt}, \tag{14}$$

KSS propose to use $(d'_t, \bar{z}'_{ijt})'$ with $\bar{z}_{ijt} = (\bar{z}'_t, \bar{z}'_{iot}, \bar{z}'_{ojt})'$ as observable proxies for f_{ijt} . Then, we can estimate the individual slope coefficients, β_{ij} and their means β consistently by augmenting the regression, (6) with $(d'_t, \bar{z}'_{ijt})'$. This estimator is the 3D common correlated effect (3DCCE) estimator, which will be applied in the main empirical applications. This hierarchical multi-dimensional approach is expected to become an essential tool for the analysis of complex interconnectedness of the bilateral (origin-destination) flows such as trade, FDI, capital or migration flows.

We develop a diagnostic tool for detecting the presence of CSD as well as evaluating the goodness of the estimation strategy by jointly applying the modified cross section dependence (CD) test and estimating the CSD exponent. We first develop the CD test, which is a modified counterpart of an existing CD test advanced by Pesaran (2015). For convenience, we represent the vector of residuals $\hat{\epsilon}_{ij} = (\hat{\epsilon}_{ij1}, \dots, \hat{\epsilon}_{ijT})'$ as the (ij) pair using the single index $n = 1, \dots, N_1N_2$ - where N_1 and N_2 are the number of origin and destination countries, respectively—and compute the pair-wise residual correlations between n and n' cross section units by

$$\hat{\rho}_{nn'} (= \hat{\rho}_{n'n}) = \frac{\hat{\epsilon}'_n \hat{\epsilon}_{n'}}{\sqrt{(\hat{\epsilon}'_n \hat{\epsilon}_n) (\hat{\epsilon}'_{n'} \hat{\epsilon}_{n'})}}, \quad n, n' = 1, \dots, N_1N_2 \text{ and } n \neq n'.$$

⁶ Day and Winer (2006) describe that an unemployment insurance system in Canada is substantially more generous towards residents, who live in regions with above-average unemployment rates. Since individuals base their location decisions on a comparison of comprehensive incomes, and not just on earnings, such regional differentials in comprehensive incomes may lead to a misallocation of labour.

Then, we construct the CD statistic by

$$CD = \sqrt{\frac{2}{N_1 N_2 (N_1 N_2 - 1)}} \sum_{n=1}^{N_1 N_2 - 1} \sum_{n'=n+1}^{N_1 N_2} \sqrt{T} \hat{\rho}_{nn'} \quad (15)$$

which has the limiting $N(0, 1)$ distribution under the null hypothesis of (pairwise) residuals weak cross section dependence.

Next, we follow Bailey et al. (2016, BKP) and derive the exponent of CSD based on $\bar{\hat{e}}_{..t} = (N_1 N_2)^{-1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \hat{e}_{ijt}$. If \hat{e}_{ijt} 's are cross sectionally correlated across (ij) pairs, $Var(\bar{\hat{e}}_{..t})$ declines at a rate that is a function of α , where α is defined as

$$\lim_{N_1 N_2 \rightarrow \infty} (N_1 N_2)^{-\alpha} \lambda_{\max}(\Sigma_{\hat{e}})$$

and $\Sigma_{\hat{e}}$ is the $N_1 N_2 \times N_1 N_2$ covariance matrix of $\hat{\mathbf{e}}_t = (\hat{e}_{11t}, \dots, \hat{e}_{N_1 N_2 t})'$ with $\lambda_{\max}(\Sigma_{\hat{e}})$ denoting the largest eigenvalue. $Var(\bar{\hat{e}}_{..t})$ cannot decline at rate faster than $(N_1 N_2)^{-1}$ as well as it cannot decline at rate slower than $(N_1 N_2)^{\alpha-1}$ with $0 \leq \alpha \leq 1$. Furthermore, for $1/2 < \alpha \leq 1$, BKP propose the bias-adjusted estimator:

$$\hat{\alpha} = 1 + \frac{1}{2} \frac{\ln(\hat{\sigma}_{\bar{\hat{e}}_{..t}}^2)}{\ln(N_1 N_2)} - \frac{\ln(\hat{\mu}_{\pi}^2)}{2 \ln(N_1 N_2)} - \frac{\hat{c}_{N_1 N_2}}{2 \left[N_1 N_2 \ln(N_1 N_2) \hat{\sigma}_{\bar{\hat{e}}_{..t}}^2 \right]} \quad (16)$$

where $\hat{\sigma}_{\bar{\hat{e}}_{..t}}^2 = T^{-1} \sum_{t=1}^T \bar{\hat{e}}_{..t}^2$, $\hat{c}_{N_1 N_2} = \widehat{\bar{\sigma}_{N_1 N_2}^2} = (N_1 N_2)^{-1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \hat{\sigma}_{ij}^2$ and $\hat{\sigma}_{ij}^2 = T^{-1} \sum_{t=1}^T \hat{\xi}_{ijt}^2$ is the ij th diagonal element of the estimated covariance matrix, $\hat{\Sigma}_{\hat{e}}$ with $\hat{\xi}_{ijt} = \hat{e}_{ijt} - \hat{\delta}_{ij} \bar{\hat{e}}_{..t}$ and $\hat{\delta}_{ij}$ is the OLS estimator from the regression of \hat{e}_{ijt} on $\bar{\hat{e}}_{..t}$. We evaluate the confidence band for $\hat{\alpha}$ by employing the test statistic in (B47) in Supplementary Appendix VI in BKP.

The CD test statistic is increasingly applied to the residuals of regression models as an ex-post diagnostic tool. However, the CD test fails to reject the null hypothesis of weak error CSD when the factor loadings have zero means, implying that the CD test will display very poor power when it is applied to cross-sectionally demeaned data. Furthermore, the residual-based CD test has been shown to often reject the null hypothesis of no remaining CSD in the case of the CCE estimator (e.g. Mastromarco et al. 2016). Juodis and Reese (2018) prove that the application of the CD test to residuals obtained from interactive effects models introduces a bias term of order \sqrt{T} , rendering an erroneous rejection of the null. Nonetheless, we suggest to use the CD test in conjunction with the estimate of α as a model-selection tool, where a reduction in the absolute value of the CD test statistic and a moderate range of the estimate of α ,⁷ would be interpreted as an indication of an improved model specification.

⁷ BKP show that the values of $\alpha \in \left(\frac{1}{2}, \frac{3}{4}\right)$ represent a moderate degree of CSD.

4 The empirical application

4.1 The data

We collect the data over the period 1976–2014 from Statistics Canada, which records annual migration streams by province of origin and destination. We consider interprovincial migration flows for 10 provinces: Prince Edward Island, Nova Scotia, New Brunswick, Québec, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia and Newfoundland and Labrador. Then, we obtain the data for 90 province-pairs ($N = 90$, $T = 38$). The panel is balanced and interprovincial migration flows are always positive. Table 1 shows descriptive statistics. The richest provinces are Alberta, British Columbia and Ontario in terms of disposable income. They mostly attract migrants from other provinces and register relatively lower unemployment rates and higher Gini coefficients. On the other hand, Prince Edward Island, New Brunswick and Newfoundland and Labrador show the lowest income level and highest unemployment rate with higher migration outflows.

Figure 1 displays the time-varying aggregate inflows across provinces and provide an insight of how the pattern of interprovincial migration, as an aggregate, evolves over time. The interprovincial migrants flows have been on a declining trend till the mid 1990s, after which they became rather stable. Such overall declining patterns broadly correspond to the stylised facts observed by Shannon (2015) in Canada and by Molloy et al. (2011) in the US. Interprovincial migration flows in Canada were mostly flat since 2000, though it stepped up slightly in 2009 and remained substantially below the US interstate migration flows. Shannon (2015) explains that the decline in the interprovincial migrant share may be driven by an environment change (i.e. factors creating poorer conditions at destination), which renders the migration less attractive. Molloy et al. (2011) point out that the job opportunities have become more similar across the US regions, implying that fewer potential moves would generate gains large enough to offset the costs of moving, resulting in a decline in internal migration. If this concept were applied to the Canadian case, the decline in interprovincial migration would imply that immigrants are now more evenly spread across less unequal regions than in the past.

Figure 2 plots net interprovincial migration flows, defined as the sum of the net migrants (inflow minus outflow) for the 10 provinces. Notice that positive net flows have been concentrated in a small number of provinces. For example, in 2014, Alberta accounted for 76% of net positive migration, British Columbia for 23% and Saskatchewan for 1%. Over a longer time horizon (1987–2014), only two provinces had positive net migration; Alberta with 433,851 (57% of total net positive migration) and British Columbia with 331,083 (43%). Net migration flows have varied considerably over time, in terms of the number of net positive migrants and the destination provinces. These variations are closely related to provincial economic performance, but they are also caused by a couple of the major political events. In the early 1990s, most of the net internal migration was recorded in British Columbia. Since 1996, Alberta has been the major destination of interprovincial migrants. From 1997 to 2002, Ontario had the second highest net positive migration while British Columbia

Table 1 Descriptive statistics

	A	BC	M	NB	NLN	NS	O	PEI	Q	S
Inflows	0.437 (0.325)	0.275 (0.258)	0.076 (0.076)	0.097 (0.115)	0.047 (0.043)	0.157 (0.178)	0.504 (0.227)	0.021 (0.025)	0.094 (0.073)	0.083 (0.103)
Outflows	0.221 (0.286)	0.138 (0.198)	0.197 (0.221)	0.197 (0.186)	0.219 (0.284)	0.213 (0.212)	0.075 (0.072)	0.245 (0.224)	0.055 (0.106)	0.231 (0.339)
Income	11.15 (0.149)	11.08 (0.086)	10.99 (0.095)	10.87 (0.090)	10.83 (0.129)	10.9 (0.090)	11.12 (0.072)	10.88 (0.094)	10.94 (0.082)	10.98 (0.150)
Unemployment	6.109 (2.139)	8.368 (2.291)	6.138 (1.408)	10.92 (1.354)	15.64 (2.269)	10.02 (1.555)	7.092 (1.386)	12.19 (2.07)	9.698 (1.77)	5.521 (1.251)
Gini	0.303 (0.012)	0.301 (0.018)	0.291 (0.009)	0.281 (0.010)	0.29 (0.012)	0.285 (0.012)	0.3 (0.019)	0.266 (0.016)	0.285 (0.009)	0.302 (0.012)

The table shows descriptive statistics—average value and standard deviation in (.)—of interprovincial migration inflows and outflows, income, unemployment rate and the Gini index over 1976–2014 for 10 provinces. Inflows and outflows are expressed as percentage of population, whereas income is expressed as the logarithm of disposal income. In the columns A = Alberta; BC = British Columbia; M = Manitoba; NB = New Brunswick; NLN = Newfoundl and Lab Northwest; NS = Nova Scotia; O = Ontario; PEI = Prince Edward Island; Q = Québec; SA = Saskatchewan

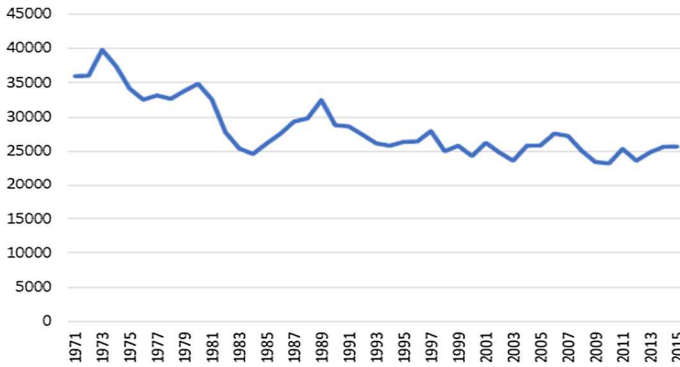


Fig. 1 Aggregate interprovincial migration flows in Canada

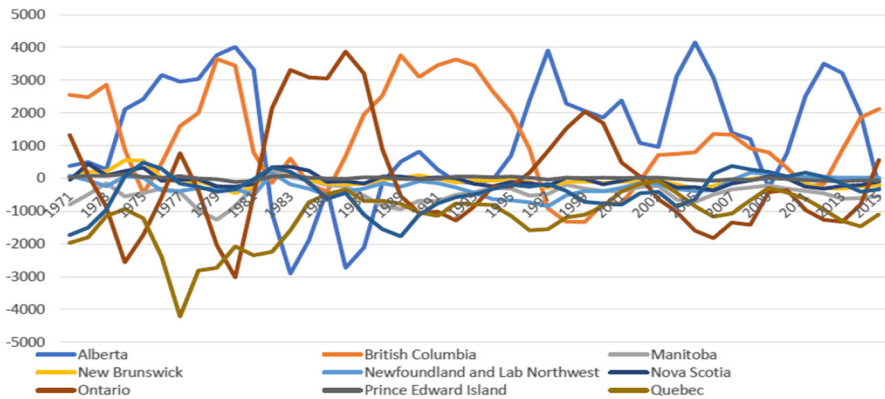


Fig. 2 Interprovincial migration net flows (inflows-outflows) in Canada

experienced net outflows. Since 2003, British Columbia has become a major destination province again. Not surprisingly, net migration flows to Alberta appear to be closely related to the oil industry boom. When the oil price collapsed in 2009, there was a sharp decline in net migration flows to Alberta, while when oil prices soared 2 years later, they increased sharply. The subsequent oil price drop and economic challenge resulted in substantial increase in outflows from Alberta in 2015. Further, as noticed by Day and Winer (2006), extraordinary policy changes could influence interprovincial migration flows. The election of the separatist Parti Québécois government in Québec in the second half of the 1970s, and the closing of the cod fishery in Atlantic Canada in 1992 appeared to have substantially altered interprovincial migration in Québec during 1977–1980 and in Newfoundland during 1993–1996, respectively.

In sum, we observe that interprovincial migration flows have been on a declining trend on average (Fig. 1), but we still find that a few provinces have experienced significant migration activities (Fig. 2). In what follows we aim at identifying the main determinants of these time-varying patterns.

4.2 Empirical results with regressors expressed as differentials

Following the income maximisation approach by Basher and Fachin (2008), Beine and Parsons (2015) and Beine et al. (2019), we first consider the following gravity regression model of interprovincial migration flows:

$$m_{ijt} = \beta_1 y_{ijt}^d + \beta_2 u_{ijt}^d + \beta_3 g_{ijt} + \beta_4 g_{ijt}^2 + \beta_5 n_{ijt} + e_{ijt}, \quad (17)$$

with origin $i = 1, \dots, N$, destination $j = 1, \dots, N$ and time $t = 1, \dots, T$, where m_{ijt} is bilateral migration flow expressed as $m_{ijt} = \ln \frac{M_{ijt}}{pop_{it}}$, M_{ijt} is the migration flow from i to j at time t , pop_{it} is the population in the origin i , which proxies for the number of stayers, M_{iit} in (5) as in most empirical studies. Here we construct the main regressors as differentials: $y_{ijt}^d = y_{it} - y_{jt}$ is the income differential and $u_{ijt}^d = u_{it} - u_{jt}$ is the unemployment rate differential, where y is the log of provincial disposable income per capita (in 2014 constant dollars) and u is the annual provincial unemployment rate (in percentage).⁸ We also investigate the impact of the second moment of the income distribution in origin and destination by introducing a measure of relative income inequality. Following Mayda (2010), we construct the origin province's relative inequality as a measure of the inequality in the origin relative to the destination (i.e. the Gini coefficient in the origin divided by the Gini coefficient in the destination, $g_{ijt} = \frac{g_{it}}{g_{jt}}$) and include both linear and quadratic terms; the total impact of relative inequality is measured at equal inequality degrees for origin and destination. This choice is motivated by Borjas' (1987) selection model, from which we infer that an increase in the origin's relative inequality has a nonlinear effect on the emigration rate: given low values of the origin country's relative inequality, as g_{ijt} rises, the emigration rate will increase. On the other hand, given high values of g_{ijt} , the emigration rate will decrease.⁹ Further, Beine and Parsons (2015) and Beine et al. (2019) stress the importance of taking into account of networks, i.e. the presence at destination of immigrants from the same origin. We measure networks, n_{ijt} by the accumulated (over time) net migration flow. The network presence lowers the migration costs of the next groups and the process continues as long as benefits exceed costs of migration, see Beine et al. (2011).¹⁰

⁸ In the income maximisation model, the migration flow, M_{ijt} is normalised by the total number of native workers of region i staying in their own region i at time t , and wage differentials are employed. Due to the data availability, we proxy the total number of native workers staying with the population in region i , pop_{it} , and wage differential with income differential. See Beine et al. (2019).

⁹ If income inequality in the origin is lower than in the destination, there is positive selection of immigrants: migrants are selected from the upper tail of the income distribution at home and end up in the upper tail of the income distribution abroad. Similarly, if income is more dispersed at home than abroad, there is negative selection: migrants are selected from the lower tail of the income distribution at home and end up in the lower tail of the income distribution abroad. Further, if relative inequality between i and j is low, an increase in inequality in the origin relative to the destination will give an incentive to migrate to those that are both at the higher and at the lower tail of the income curve. Instead, if the level of relative inequality is high, a further increase in relative inequality will not provide an incentive to move to those who did not migrate beforehand.

¹⁰ We acknowledge that this is a rather simple measure of migrants network which does not consider transit migration. The flow from the origin i and the destination j can include flows of i -born individuals, return

In what follows we apply four estimators to (17); namely the two-way fixed effects estimator (FE) with $e_{ijt} = \alpha_{ij} + \theta_t + \varepsilon_{ijt}$, the 2D-PCCE estimator ($PCCE_G$) with $u_{ijt} = \alpha_{ij} + \gamma'_{ij} f_t + \varepsilon_{ijt}$, the triple-way fixed effects estimator ($CTFE$) with $e_{ijt} = \alpha_{ij} + \theta_{it} + \theta_{jt}^* + \varepsilon_{ijt}$, and the 3D-PCCE estimator ($PCCE_{GL}$) with $u_{ijt} = \alpha_{ij} + \gamma'_{oj} f_{io} + \gamma'_{io} f_{oj} + \gamma'_{ij} f_t + \varepsilon_{ijt}$.¹¹ As an ex-post diagnostic tool, we report the CD test results applied to the residuals and estimates of the CSD exponent (denoted α).

Table 2 reports the panel gravity estimation results for the 90 province-pairs among the 10 Canadian provinces over the period 1976–2014 (39 years). As a benchmark, we first describe the results for FE . All the coefficients are statistically significant with expected signs. The impact of the income differentials (y_{ijt}^d) between origin and destination province is significant and negative at -0.40 , implying that the increase in y_{ijt}^d reduces the interprovincial migration flows. On the contrary, the impact of the unemployment differentials (u_{ijt}^d) is significant and slightly positive at 0.02 , implying that the increase in u_{ijt}^d tends to boost the interprovincial migration flows. Further, the coefficient on the linear inequality (g_{ijt}) is positive at 4.65 while the coefficient on the quadratic inequality (g_{ijt}^2) is negative at -2.01 . This is consistent with the selection model by Borjas (1987), suggesting that the effect of inequality depends on the size of emigration rates. In particular, an increase in the origin's relative inequality has a non-monotonic effect: the impact is positive (negative) if there is a positive (negative) selection. The impact of network (n_{ijt}) is significant and positive at 1.12 , showing that the presence of migrants of the same origin tends to encourage new inflows. Overall, these results are qualitatively consistent with *a priori* expectations, and with the previous studies, e.g. Coulombe (2006), Day and Winer (2006), Basher and Fachin (2008) and Beine and Parsons (2015). However, the CD test being applied to the residuals convincingly rejects the null of weak CSD, while α is estimated at 1.003 , implying the strong CSD. In this regard, we may conclude that the FE estimation results are likely to be less reliable.

Next, we find that the estimation results for $PCCE_G$ are similar to those of FE . All the coefficients display the expected signs and are significant with smaller standard errors. Still, the $PCCE_G$ results suffer from strong CSD in the residuals, though the estimate of α is 0.832 , substantially lower than the FE counterpart. A notable difference is that the impact of the network is now substantially reduced to 0.045 from 1.12 by FE .

Turning to the $CTFE$ estimation results, we find that the coefficients are mostly significant and with expected signs. Surprisingly, the impact of g_{ijt}^2 is slightly positive

Footnote 10 continued

migration flows of j -born individuals, and flows of individuals born in other Canadian provinces. The lack of more disaggregated data prevents the use of a even higher-dimensional model, and it may cast doubts on some of the variables, notably the one (past cumulated flows) which should capture network effects. Artuc and Ozden (2018) is an example of a paper in which the origin is indexed both by country of birth and by country of last residence. If we address such transit migration, then we need to extend the current 3D method to the 4D datasets of migration flows. In KMSS, we have discussed such methodological extensions though the development of the 4D models with the corresponding hierarchical multifactor structure would be much more challenging.

¹¹ Here we do not report the results for the mean group estimator, due to the fact that the number of individual estimation results tend to become unreliable.

Table 2 Alternative estimation results for the gravity regression in (17)

	Coeff FE	s.e.	<i>t</i> -ratio	Coeff CTFE	s.e.	<i>t</i> -ratio
y_{ijt}^d	-0.404	0.030	-13.38	-0.395	0.087	-4.541
u_{ijt}^d	0.022	0.001	21.57	0.021	0.003	7.161
g_{ijt}	4.653	0.362	12.85	0.547	0.273	2.004
g_{ijt}^2	-2.005	0.183	-10.96	0.037	0.122	0.305
n_{ijt}	1.115	0.003	449.2	1.209	0.007	167.5
CD stat	378.4			-3.81		
	$\alpha_{0.05}$	α	$\alpha_{0.95}$	$\alpha_{0.05}$	α	$\alpha_{0.95}$
CSD exponent	0.904	1.003	1.102	0.713	0.782	0.852
	<i>PCC_{EG}</i>			<i>PCC_{EGL}</i>		
y_{ijt}^d	-0.407	0.031	-13.06	-0.471	0.023	-20.14
u_{ijt}^d	0.044	0.001	58.47	0.073	0.001	73.18
g_{ijt}	3.93	0.168	23.42	3.915	0.192	20.38
g_{ijt}^2	-1.857	0.082	-22.67	-1.778	0.092	-19.26
n_{ijt}	0.045	0.002	20.79	0.017	0.002	9.336
CD stat	5.791			2.408		
	$\alpha_{0.05}$	α	$\alpha_{0.95}$	$\alpha_{0.05}$	α	$\alpha_{0.95}$
CSD exponent	0.763	0.832	0.884	0.762	0.823	0.884

Using annual data over 1976–2014 for 90 province-pairs, we estimate model (17), where the dependent variable is the logarithm of normalised migration flows and the regressors are differential of income, unemployment rate, relative income inequality, its squared value and network. FE stands for two-way fixed effects estimator that deals with pair and time fixed effect. CTFE represents the 3D within estimator. In *PCC_{EG}*, we implement global *PCC_E* presented in KSS where factors are $\mathbf{f}_t = \{\bar{y}_{..t}^d, \bar{u}_{..t}^d\}$. In *PCC_{EGL}*, we implement global and local *PCC_E* presented in KSS where factors are those used in *PCC_{EG}* plus $\mathbf{f}_{i,t} = \{\bar{u}_{i,t}^d\}$. CD test refers to testing the null hypothesis of residual cross-sectional error independence or weak dependence. α denotes the estimate of CSD exponent jointly with the 95% confidence bands

but insignificant. The impact of network is more positive at 1.21. Though the CD test still rejects the null of weak CSD in the residuals, α is estimated at 0.782 with the confidence band, (0.713,0.852), suggesting the moderate level of CSD.

Finally, the *PCC_{EGL}* estimation results demonstrate that the coefficients are all significant with expected signs. We still reject the null of weak residual CSD, but α is estimated at 0.823 with the confidence band, (0.762,0.884), suggesting the moderate level of CSD. As described in Sect. 3, the CD test has been used as an *ex-post* model-selection tool, with a reduction in the value of the CD test statistic typically interpreted as an indication of an improved model specification. On the basis of these results, we select the *PCC_{EGL}* as the preferred estimator. There are a few differences between *FE* and *PCC_{EGL}* results. First, the impact of the unemployment differentials is now more than 3 times larger while the impact of relative inequalities are substantially smaller. Next, we notice that the *PCC_{EGL}* estimators, that explicitly deal with CSD

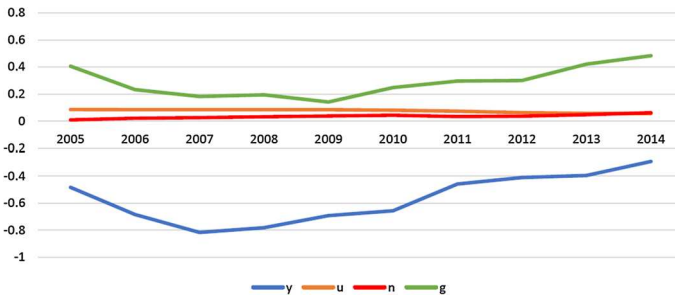


Fig. 3 Time varying coefficients of rolling window-based estimations of model (17) over 2005–2014

through unobserved heterogeneous factors, tend to produce a much smaller network effect than those in *FE* and *CTFE*, that ignore the presence of unobserved factors. This evidence may suggest that the network variable is endogenous due to unobserved MRM. As emphasised by Beine and Parsons (2015) and Özen et al. (2011), endogeneity might arise from the omission of unobserved factors that would be correlated with the network variable.¹²

Finally, we investigate the time varying patterns of the main coefficients by employing a rolling window-based estimation over the last ten years of the sample, 2005–2014, which are plotted in Fig. 3.¹³ The effect of the unemployment differentials has been relatively stable and quite small over the whole period. The effect of the income differentials shows a peak at 2007 and then steadily declining until 2014. On the contrary, the total impact of g_{ijt} —computed as $\beta_3 + 2\beta_4$ —has initially declined but started to increase from 2009 onwards, showing evidence in favour of a positive selection in more recent period. Further, the impact of network, albeit small, has shown a slightly increasing trend in the last few years. These time-varying patterns suggest that the rise in the internal migration flows, registered in Canada from 2009 onwards (see Fig. 1), is more likely to be associated with the relative income inequality and network presence rather than the conventional long-run determinants such as income and unemployment differentials.¹⁴

¹² We also evaluate the effect of an extraordinary policy event like the closing of the cod fishery on the east coast since 1992, which appears to have substantially affected interprovincial migrations, see Day and Winer (2006). We include two dummies, p_{it} equal to 1 if outflow from Newfoundland from 1993 to 1996 occurs and 0 otherwise, and p_{jt} equal to 1 if inflow from Newfoundland from 1993 to 1996 occurs and 0 otherwise. We find that this event exerted a positive significant effect on outflows from Newfoundland and a negative effect on inflows to Newfoundland. Interestingly, the impacts of this event are estimated at 0.34 and -0.09 by *FE*, whereas they are estimated at 0.03 and -0.11 by *PCC EGL*.

¹³ To track dynamic evolution of the slope parameters, we apply a rolling-window time-series regression and assign estimated β_t to the end date of each rolling sample. Fixing the window size at 10 years, we repeat a rolling-window regression until the end of the sample, and obtain 10 years time-varying estimates of β_t .

¹⁴ This new evidence deserves further investigations. If the data on skills or education of migrants were available at the interprovincial level, it would be interesting to further analyse the linkages between network and the structure of interprovincial migration flows. Only a few papers have analysed this issue employing the data on the education levels of international migrants and find that larger diasporas increase migration flows and lower their average educational level, see Beine et al. (2011).

Evidence on positive selection, in the spirit of Borjas' (1987) selection model, is broadly in line with the interprovincial skill redistribution hypothesis in Coulombe (2006) and Coulombe and Tremblay (2009), according to which (i) the typical interprovincial migrants in Canada has higher skills than the non-migrants and (ii) the most mobile component of the population is the educated young people who move from the relatively poor and rural provinces to the relatively rich and urban provinces. Also, given that the net interprovincial migration flows are typically oriented towards Ontario, Alberta, and British Columbia, the increasing importance of the network effect shows that migrants are attracted by the big cities where they can benefit from the various externalities of their diaspora (see Beine et al. 2011 and Beine and Coulombe 2018). Hence, migration location choice is no longer shown to be only guided by the highest probability of employment. It is often argued that immigrants tend to cluster together because the presence of established network facilitates assimilation of new arrivals, both in the labour market and in the social environment of the host. Analysing a survey on recent international immigrants to Canada, Goel and Lang (2019) find that relatives and friends, already living in the host country, help recent immigrants to find their first jobs. A similar mechanism might work for interprovincial migration and this issue deserves further investigations.

4.3 Empirical results with asymmetrical push and pull factors

Mayda (2010) and Piras (2017) emphasise that it is rather arbitrary to treat push and pull factors symmetrically, which does not properly disentangle the different structural impacts that the same variables could have as push or pull factor. In this regard, we consider the alternative gravity regression of the migration flows:

$$m_{ijt} = \beta_1 y_{it} + \beta_2 y_{jt} + \beta_3 u_{it} + \beta_4 u_{jt} + \beta_5 g_{ijt} + \beta_6 g_{ijt}^2 + \beta_7 n_{ijt} + e_{ijt}, \quad (18)$$

for $i, j = 1, \dots, N$ and $t = 1, \dots, T$. In particular, we define the income and unemployment variables separately for origin and destination such that y_{it} and y_{jt} are the logged per capita disposable income of the origin and destination while u_{it} and u_{jt} are the unemployment rates of the origin and destination.

Table 3 reports the panel gravity estimation and diagnostic results for the gravity regression in (18) for the 90 province-pairs among the 10 Canadian provinces over the period 1976-2014. The FE estimator suffers from strong CSD while α is estimated at 1.003. The CD test still rejects the null of weak CSD for both $PCCE_G$ and $PCCE_{GL}$, but the $PCCE_{GL}$ estimators display much lower CSD, as supported by the smaller CSD exponent estimate, $\hat{\alpha}$ at 0.904 for $PCCE_G$ and at 0.862 for $PCCE_{GL}$. On the basis of these results, we focus on the $PCCE_{GL}$ with the lowest CSD.¹⁵

All the coefficients are statistically significant with expected signs. We find that an income increase in the origin, y_{it} , will reduce the migration flows from i to j while an income increase in the destination, y_{jt} , will raise the migration flows. Next, the higher

¹⁵ We cannot report the CTFE estimation results for the gravity regression in (18), because 3D within transformation eliminates regressors varying over (i,t) and (j,t) together with origin/destination time fixed effects.

Table 3 Alternative estimation results for the gravity regression in (18)

	Coeff FE	s.e.	t-ratio	Coeff	s.e.	t-ratio
y_{it}	-0.306	0.020	-15.27			
y_{jt}	0.929	0.019	47.87			
u_{it}	0.009	0.001	8.689			
u_{jt}	-0.045	0.001	-40.21			
g_{ijt}	0.960	0.271	3.544			
g_{ijt}^2	-0.457	0.130	-3.515			
n_{ijt}	1.111	0.003	443.0			
CD stat	379.0					
	$\alpha_{0.05}$	α	$\alpha_{0.95}$			
CSD exponent	0.878	1.003	1.127			
	<i>PCCE_G</i>			<i>PCCE_{GL}</i>		
y_{it}	-0.347	0.026	-13.536	-0.416	0.020	-20.46
y_{jt}	0.970	0.019	52.015	0.484	0.020	24.62
u_{it}	0.042	0.001	54.314	0.026	0.001	34.55
u_{jt}	-0.056	0.001	-51.779	-0.035	0.001	-36.95
g_{ijt}	2.628	0.156	16.819	3.799	0.121	31.48
g_{ijt}^2	-1.313	0.076	-17.180	-1.845	0.057	-32.37
n_{ijt}	0.030	0.002	14.699	0.037	0.002	16.80
CD stat	8.61			16.29		
	$\alpha_{0.05}$	α	$\alpha_{0.95}$	$\alpha_{0.05}$	α	$\alpha_{0.95}$
CSD exponent	0.853	0.904	0.955	0.813	0.862	0.911

Using annual data over 1976–2014 for 90 province-pairs, we estimate model (18), where the dependent variable is the logarithm of normalised migration flows and the regressors are income at origin/destination, unemployment rate origin/destination, relative income inequality, its squared value and network. FE stands for two-way fixed-effects estimator that deals with pair and time fixed effect. In *PCCE_G*, we implement global *PCCE* presented in KSS where factors are $\mathbf{f}_t = \{\bar{y}_{..t}, \bar{u}_{..t}, \bar{g}_{..t}\}$. In *PCCE_{GL}*, we implement global and local *PCCE* presented in KSS where factors are those used in *PCCE_G* plus $\mathbf{f}_{i,t} = \{\bar{y}_{i..t}, \bar{u}_{i..t}\}$. CD test refers to testing the null hypothesis of residual cross-sectional error independence or weak dependence. α denotes the estimate of CSD exponent jointly with the 95% confidence bands

unemployment in the origin, u_{it} will increase the migration flows from i to j , whereas the higher unemployment in the destination, u_{jt} will reduce migration flows. In both cases, the pull impacts are slightly stronger than the push impacts,¹⁶ in line with the previous studies. Furthermore, as expected, the impact of the linear inequality (g_{ijt}) is positive, but the impact of the quadratic inequality (g_{ijt}^2) is negative, which is again consistent with the selection hypothesis by Borjas (1987).

¹⁶ We have conducted the t tests for the null hypothesis of the equal coefficient on y_{it} and y_{jt} as well as of the equal coefficient on u_{it} and u_{jt} . We find that the values of t-statistics are 2.4 and 6.4, respectively, which reject the null of symmetry. This may provide an empirical support in favour of the gravity specification (18) that considers push and pull factors, separately.

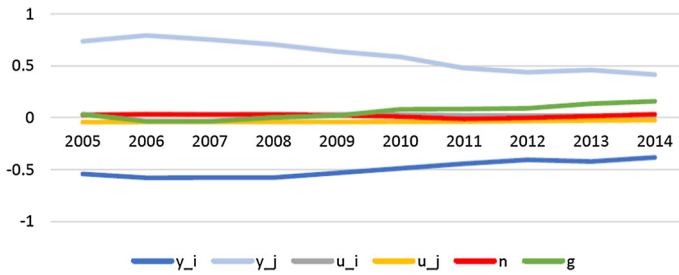


Fig. 4 Time-varying coefficients of rolling window-based estimations of model (18) over 2005–2014

The $PCCE_{GL}$ estimation results are somewhat different from the FE and $PCCE_G$ estimation results. First, the impacts of the income and unemployment at the origin are now larger in magnitude while those impacts at the destination are substantially smaller, such that the differences between them are much narrower. Importantly, we notice that the income coefficients by $PCCE_{GL}$ suggest that the pull effects no longer dominate the push effects, and they are close to be rather symmetric, in line with the recent view in favour of a convergence process. This evidence might suggest that global factors only do not capture the local MRM, leading to misleading results. Second, the impacts of the inequality measures are larger than the FE estimates, but close to those by $PCCE_G$. Third, both $PCCE_G$ and $PCCE_{GL}$ produce much smaller network impacts than FE . Overall, these estimation results are consistent with *a priori* expectations and with the previous evidence on the Canadian case (e.g. Coulombe 2006; Day and Winer 2006) and on the migration flows in general (e.g. Mayda 2010; Beine and Parsons 2015; Piras 2017).

Furthermore, we examine the time-varying slope coefficients through a rolling window-based $PCCE_{GL}$ estimation over 2005–2014 and plot them in Fig. 4. The time-varying effects of unemployment rate have been rather stable and small throughout the period with the coefficients on u_{jt} being generally larger than those on u_{it} . The impacts of both income at destination and income at origin have been much larger than those on unemployment. The time-varying effects of income at destination and income at the origin display an opposite sign, and the distance between them (in absolute terms) decreases from 2007 onwards. At the beginning of the period observed, y_{jt} has a larger impact on migration decision while in more recent years, y_{it} and y_{jt} have similar effects, implying that push and pull effects of per capita income rather approach symmetry. Similar to Fig. 3, the total impact of g_{ijt} initially declined but increased from 2007 onwards. The impact of network, albeit small, shows a slightly increasing pattern in the last few years. This confirms that the location choice of inter-provincial migration seems to be no longer only guided by the conventional long-run determinants such as income and unemployment differentials.

5 Conclusions

In this paper, we have applied recent 3D panel data techniques to an analysis of the interprovincial migration flows in Canada from 1976 to 2014. We have extended the microfounded specifications proposed by Bertoli and Fernández-Huertas Moraga (2013) and Piras (2017) to a 3D framework with a hierarchical multifactor structure. The specification proposed allows to deal with two sources of cross section correlations, one driven by global factors and the other driven by local (origin/destination) factors. This structure is shown to be particularly relevant in the context of provincial (or regional) data where cross sectional entities within a single country are deeply integrated and interconnected.

Our empirical findings suggest that the interprovincial migration flows in Canada have been mostly affected by average disposal income and relative income inequality, in line with Borjas' (1987) selection model. We also find that a smaller (albeit significant) effect is attributed to unemployment rate and network.

Further, the time-varying patterns of the slope coefficients derived through a rolling window-based estimation over the last ten years of the sample, 2005–2014, show that the rise in the internal migration flows, registered in Canada from 2009 onwards, is more likely to be associated with positive selection—in light of Borjas' (1987) selection model—and the existence of networks of migrants of the same origin rather than with the conventional long-run determinants such as income differentials. Evidence on positive selection is broadly in line with the interprovincial skill redistribution described in Coulombe (2006) and Coulombe and Tremblay (2009). Data on migrants' skills would be useful to further investigate the important nexus between self-selection of migrants and diaspora as discussed in Beine et al. (2011).

We conclude by noting an immediate avenue for future research. First, we might concern the application of the 3D hierarchical multifactor structure to international migration flows. Notably, migration patterns within country and across countries have different characteristics because domestic migration flows are generally less regulated by policy selection criteria or labour mobility. Next, if the data on transit migration were readily available as analysed in Artuc and Ozden (2018), then it is quite desirable to extend the 3D method to the 4D datasets of migration flows though the development of the 4D models with the hierarchical multifactor structure would be more challenging due to appropriately modelling several layers of components to account for effects across the higher dimensions.

Compliance with ethical standards

Conflict of interest Authors declare that they have no conflict of interest.

Human and animal rights They also declare that this research does not contain any studies with human participants or animals performed.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included

in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderson J, van Wincoop E. Gravity with gravitas: a solution to the border puzzle. *Am Econ Rev*. 2003;93:170–92.
- Artuc E, Ozden C. Transit migration: all roads lead to America. *Econ J*. 2018;128:306–34.
- Baglay S. Provincial nominee programs: a note on policy implication and future research needs. *Int Migr Integr*. 2012;13:121–41.
- Baier SL, Bergstrand JH. Do free trade agreements actually increase members' international trade? *J Int Econ*. 2007;71:72–95.
- Balazsi L, Bun MJG, Chan F, Harris MN. Models with endogenous regressors. In: Mátyás L, editor. *The econometrics of multi-dimensional panels*. Advanced studies in theoretical and applied econometrics. Berlin: Springer; 2017a.
- Balazsi L, Mátyás L, Wansbeek T. Fixed effects models. In: Mátyás L, editor. *The econometrics of multi-dimensional panels*. Advanced studies in theoretical and applied econometrics. Berlin: Springer; 2017b.
- Balazsi L, Baltagi BH, Mátyás L, Püs D. Random effects models. In: Mátyás L, editor. *The econometrics of multi-dimensional panels*. Advanced studies in theoretical and applied econometrics. Berlin: Springer; 2017c.
- Baltagi BH. Panel data gravity models of international trade. In: Baltagi BH, editor. *The Oxford handbook of panel data*. New York: Oxford University Press; 2015. p. 608–41.
- Baltagi BH, Egger P, Pfaffermayr M. A generalized design for bilateral trade flow models. *Econ Lett*. 2003;80:391–7.
- Basher B, Fachin S. The long-term decline of internal migration in Canada: the case of Ontario. *Let Spatial Resour Sci*. 2008;1:171–81.
- Bailey N, Kapetanios G, Pesaran MH. Exponent of cross-sectional dependence: estimation and inference. *J Appl Econ*. 2016;31:929–60.
- Beenstock M, Felsenstein D, Rubin Z. Immigration to the European Union from its neighborhoods: testing welfare-chasing and related hypotheses by spatial gravity. *Int J Manpower*. 2015;36:491–517.
- Beine M, Coulombe S. Immigration and internal mobility in Canada. *J Popul Econ*. 2018;31:69–106.
- Beine M, Parsons C. Climatic factors as determinants of international migration. *Scand J Econ*. 2015;117:723–67.
- Beine M, Docquier F, Özden C. Diasporas. *J Dev Econ*. 2011;95:30–41.
- Beine M, Bertoli S, Fernández-Huertas Moraga J. A practitioners' guide to gravity models of international migration. *World Econ*. 2016;39:496–512.
- Beine M, Bourgeon P, Bricongne JC. Aggregate fluctuations and international migration. *Scand J Econ*. 2019;121:117–52.
- Behrens K, Ertur C, Kock W. Dual gravity: using spatial econometrics to control for multilateral resistance. *J Appl Econ*. 2012;27:773–94.
- Bertoli S, Fernández-Huertas Moraga J. Multilateral resistance to migration. *J Dev Econ*. 2013;102:79–100.
- Borjas GJ. Self selection and the earnings of immigrants. *Am Econ Rev*. 1987;77:531–53.
- Brown M, Macdonald R. Provincial convergence and divergence in Canada, 1926 to 2011. *Economic Analysis (EA) Research Paper Series*, Statistics Canada; 2015.
- Coulombe S. Internal migration, asymmetric shocks, and interprovincial economic adjustments in Canada. *Int Reg Sci Rev*. 2006;29:199–223.
- Coulombe S, Tremblay J. Migration and skills disparities across the Canadian provinces. *Reg Stud*. 2009;43:5–18.
- Chudik A, Pesaran MH, Tosetti E. Weak and strong cross section dependence and estimation of large panels. *Econom J*. 2011;14:C45–90.
- Day KM, Winer SL. Policy-induced internal migration: an empirical investigation of the Canadian case. *Int Tax Public Finance*. 2006;13:535–64.

- Eberhardt M, Teal F. Productivity analysis in global manufacturing production. Department of Economics, University of Oxford; 2010.
- Elhorst JP, Lacombe DJ, Piras G. On model specifications and parameter space definitions in higher order spatial econometric models. *Reg Sci Urban Econ.* 2012;42:211–20.
- Fachin S. Long-run trends in internal migrations in Italy: a study in panel cointegration with dependent units. *J Appl Econom.* 2007;22:401–28.
- Finnie R. Who moves? A logit model analysis of inter-provincial migration in Canada. *Appl Econ.* 2004;36:1759–79.
- Goel D, Lang K. Social ties and the job search of recent immigrants. *ILR Rev.* 2019;72:355–81.
- Hanson GH. International migration and the developing world. In: Rodrik D, Rosenzweig M, editors. *Handbook of development economics*, vol. 5. Amsterdam: North-Holland; 2010. p. 363–414.
- Harris JR, Todaro MP. Migration, unemployment and development: a two-sector analysis. *Am Econ Rev.* 1970;60:126–42.
- Hausman J, Taylor W. Panel data and unobservable individual effects. *Econometrica.* 1981;49:1377–98.
- Kapetanios G, Mastromarco C, Serlenga L, Shin Y. Modelling in the presence of cross-sectional error dependence, Chap. 10. In: Mátyás L, editor. *The econometrics of multi-dimensional panels. Advanced studies in theoretical and applied econometrics.* Berlin: Springer; 2017.
- Kapetanios G, Serlenga L, Shin Y. Estimation and inference for multi-dimensional heterogeneous panel datasets with hierarchical multifactor error structure. *J Econom.* (forthcoming). 2020. Available online at <https://doi.org/10.1016/j.jeconom.2020.04.011>.
- Juodis A, Reese S. The incidental parameters problem in testing for remaining cross-section correlation. *Papers arXiv:1810.03715.* 2018.
- LeSage JP, Llano C. A spatial interaction model with spatially structured origin and destination effects. In: Arbia G, Patuelli R, editors. *Spatial econometric interaction modelling.* Berlin: Springer; 2016.
- LeSage JP, Pace RK. Spatial econometric modeling of origin-destination flows. *J Reg Sci.* 2008;48:941–67.
- Mayda AM. International migration: a panel data analysis of the determinants of bilateral flows. *J Popul Econ.* 2010;23:1249–74.
- Mastromarco C, Serlenga L, Shin Y. Multilateral resistance and Euro effects on trade flows, Chap. 12. In: Arbia G, Patuelli R, editors. *Spatial econometric interaction modelling.* Berlin: Springer; 2016.
- Mátyás L. *The econometrics of multi-dimensional panels. Advanced studies in theoretical and applied econometrics.* Berlin: Springer; 2017.
- McFadden D. Econometric analysis of qualitative response models. In: Griliches Z, Intriligator M, editors. *Handbook of econometrics*, vol. 2. Amsterdam: Elsevier; 1984.
- Mitze T. On the mutual dynamics of interregional gross migration flows in space and time in spatial econometric interaction modelling. In: Arbia G, Patuelli R, editors. *Spatial econometric interaction modelling.* Berlin: Springer; 2016.
- Molloy R, Smith C, Wozniak A. Internal migration in the United States. *J Econ Perspect.* 2011;25:173–96.
- Pandey M, Townsend J. Quantifying the effects of the Provincial Nominee Programs. *Can Public Policy.* 2011;37:495–512.
- Pesaran MH. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica.* 2006;74:967–1012.
- Pesaran MH. Testing weak cross-sectional dependence in large panels. *Econom Rev.* 2015;34:1089–117.
- Pesaran MH, Smith R. Estimating long-run relationships from dynamic heterogeneous panels. *J Econom.* 1995;68(1):79–113.
- Piras R. A long-run analysis of push and pull factors of internal migration in Italy. Estimation of a gravity model with human capital using homogeneous and heterogeneous approaches. *Pap Reg Sci.* 2017;96:571–602.
- Ortega F, Peri G. The effect of income and immigration policies on international migration. *Migr Stud.* 2013;1:47–74.
- Serlenga L, Shin Y. Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors. *J Appl Econom.* 2007;22:361–81.
- Shannon M. Canadian migration destinations of recent immigrants and interprovincial migrants: similarities, differences and explanations. *IZA J Migr.* 2015;4:4–32.
- Stark O. Rural-to-urban migration in LDCs: a relative deprivation approach. *Econ Dev Cult Change.* 1984;32:475–86.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.