

The World Is Not Yet Flat

Transport Costs Matter!

Kristian Behrens

W. Mark Brown

Théophile Bougna



WORLD BANK GROUP

Development Research Group

Environment and Energy Team

October 2016

Abstract

The paper provides evidence of the effects of changes in transport costs on the geographic concentration of industries. The analysis uses micro-level commodity flow data and micro-geographic plant-level data to construct industry-specific ad valorem trucking rates and continuous measures of geographic concentration. The findings show that, controlling

for international trade exposure and input-output links, increasing trucking rates are significantly associated with declining geographic concentration. The effect is large: changes in trucking rates explain around 20 percent of the observed decline in geographic concentration of Canadian manufacturing industries between 1992 and 2008.

This paper is a product of the Environment and Energy Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at tbougna@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

The world is not yet flat: Transport costs matter!*

Kristian Behrens[†]

W. Mark Brown[‡]

Théophile Bougna[§]

Keywords: transport costs; trucking rates; geographic concentration; international trade exposure; input-output links.

JEL classification: R12; C23; L60.

*We thank the editor, Gordon Hanson, two anonymous referees, Gabriel Ahlfeldt, John Baldwin, Antoine Bonnet, Pierre-Philippe Combes, Gilles Duranton, Wulong Gu, Julien Martin, Se-il Mun, Yasusada Murata, Frédéric Robert-Nicoud, Mathieu Parenti, Eugenia Shevtsova, Peter Warda, and conference and seminar participants in many places for helpful comments and suggestions. Tim Pendergast provided excellent technical assistance. Financial support from Statistics Canada's 'Trade Cost Analytical Projects Initiative' (API) is gratefully acknowledged. This project was funded by the Russian Academic Excellence Project '5-100'. Behrens and Bougna gratefully acknowledge financial support from the CRC Program of the Social Sciences and Humanities Research Council (SSHRC) of Canada for the funding of the *Canada Research Chair in Regional Impacts of Globalization*. Bougna gratefully acknowledges financial support from CIRPÉE. This work was carried out while Behrens and Bougna were Deemed Employees of Statistics Canada. The views expressed in this paper and all remaining errors are ours. This paper has been screened to ensure that no confidential data are revealed.

[†]*Corresponding author:* Dept. of Economics, Université du Québec à Montréal, Canada; National Research University Higher School of Economics, Russia; CIRPÉE, Canada; and CEPR, UK. E-mail: behrens.kristian@uqam.ca

[‡]Statistics Canada, Economic Analysis Division (EAD); Chief, Regional and Urban Economic Analysis. E-mail: mark.brown@canada.ca

[§]Development Research Group (DECRG), The World Bank, Washington DC. E-mail: tbougna@worldbank.org

1 Introduction

Transport and communication costs fell precipitously during the last century, leading many observers to posit that the world has ‘become flat’, that locational differences no longer matter. With the help of cheap transport and communication, business can be done almost anywhere, or so the story goes, leaving policy makers with the impression that we have entered a ‘brave new frictionless’ world. But, if this were true, the costs of trading and transporting goods should no longer influence firms’ location choices and, thereby, the spatial structure of economic activity. Why is then the tendency for economic activity to cluster in space still strong? Why do many industries nowadays still exhibit strong geographic patterns, including new entrants that should face little locational constraints?

We address this seeming contradiction head-on by identifying the causal effect of transport costs on the geographic concentration of industries. Using micro-level commodity flow data and micro-geographic plant-level data, we construct industry-specific ad valorem trucking rates and continuous measures of geographic concentration. We find that, controlling for international trade exposure and input-output links, increasing trucking rates are significantly associated with declining geographic concentration. The effect is large: had trucking rates not fallen between 1992 and 2008, the observed decline in geographic concentration of Canadian manufacturing industries would have been about 20% greater. Our results hold up to a large variety of robustness checks and to instrumental variables estimations that deal with potential endogeneity concerns of our key variables.

Assessing empirically the impact of transport costs on the geographic concentration of industries is important for several reasons. First, it is fair to say that, despite their fundamental theoretical role in spatial modeling, little is still known empirically on how transport costs drive the geographic structure of industries. Second, among the forces that drive the clustering of industries, transport costs have been less studied, much less than the ‘Marshallian’ determinants such as input-output links, labor market pooling, and knowledge spillovers (see Rosenthal and Strange, 2004; Combes and Gobillon, 2015). Consequently, we have little quantitative evidence on the impact of those costs on the geography of industries, and on how their magnitude compares to that of the other forces. Third, transport costs certainly strongly bear on the local composition of economic activity, which is of great policy importance. Changes in transport costs, driven by, e.g., infrastructure investments, inform us on how local economic structure may change (e.g., Duranton, Morrow, and Turner, 2014). These changes play a major role in understanding how regions can weather shocks in international trading environments that have direct repercussions in local labor markets (e.g., Autor, Dorn, and Hanson, 2013; Dauth, Findeisen, and Suedekum, 2014).

Assessing empirically the impact of transport costs on the geographic concentration of industries is also a difficult task. First, we need fine measures of geographic concentration

across time to assess changes therein. In this paper, we employ – for the first time to our knowledge – a long panel of continuous measures of geographic concentration, computed from micro-geographic plant-level data using the approach of Duranton and Overman (2005).¹ Using panel data allows us to look at the time-series variation over a nearly 20 year period and to go beyond existing studies that have mainly looked at the cross-sectional variation in the geographic concentration of industries.

Second, we need detailed measures of industry-specific transport costs. We devote substantial effort to the construction of domestic ad valorem trucking rates for 257 industries and 20 years. Most of the literature has looked at infrastructure as a source of variation in transport costs (see Redding and Turner, 2015). By contrast, we build our trucking rates time series from the microdata files on truck shipments within Canada. These measures capture time-changing domestic transport costs and are invariant to the spatial structure of industry, thereby side-stepping the often endogenous nature of standard transportation measures (e.g, transportation margins from input-output accounts).

Third, as suggested by a simple general equilibrium model that we construct, to identify the causal effect of transport costs on the geographic concentration of industries we need to control for both time-varying changes in international trade exposure and general equilibrium effects of access to intermediate input suppliers and customers in vertical production chains. We tackle this challenge by developing rich measures of trade exposure and microgeographic measures of plants’ access to potential suppliers and customers.

Last, we need to deal with the possible endogeneity of our main covariates. For example, it is well documented that productivity rises as an industry concentrates geographically (see, e.g., Rosenthal and Strange, 2004; Combes and Gobillon, 2015). If these productivity gains are passed on to consumers in the form of lower prices that affect also ad valorem trucking rates, the causality may actually run from agglomeration to transport costs and not the other way round. We deal with this issue directly by purging our measure of ad valorem trucking costs of the effect multi-factor productivity. We also use us industry price indices to construct external instruments for our trucking rate series.

The remainder of the paper is structured as follows. Section 2 presents the theoretical background that informs our empirical strategy. Section 3 briefly documents the evolutions of geographic concentration and of transport costs in Canada, including a detailed discussion of the construction of the transport cost variable. Section 4 describes our empirical strategy and discusses the identification issues. Section 5 presents our results and provides a large number of robustness checks and iv estimates. Finally, Section 6 concludes. Technical details are relegated to an extensive set of appendices and a Supplementary Online appendix.

¹See Holmes and Stevens (2004) for an exhaustive survey of location patterns in North America. They do not report results using continuous measures. Ellison, Glaeser, and Kerr (2010) use a ‘lumpy (county-level) approximation’ of the Duranton and Overman (2005) measure and apply it to us manufacturing data.

2 Theoretical background

Our aim is to empirically assess the causal effect of changes in domestic transport costs on the geographic concentration of industries within a country. Changes in transport costs alter firms' access to customers and suppliers, which affects many key dimensions of their economic environment and behavior: competition, input prices, plant size, output, and optimal location choices, to name but a few. It also affects potential entry in and exit from markets and industries. When aggregated across firms, these changes transform the geographic landscape of plants, employment, and output in different industries.

To formally investigate how changes in domestic transport costs affect geographic concentration, we require a framework that features costly trade between domestic and between domestic and foreign locations, vertical industry links, and geographically mobile industries. To paint a full picture also requires that changes in the competitive environment – brought about by falling transport costs – affect firms' sizes and outputs. There is, to our knowledge, no model to date that features all these ingredients simultaneously.²

We hence first develop a 'simple' model that allows us to evaluate the impacts of changes in domestic transport costs and international trade exposure on the geographic concentration of industries, taking into account input-output links and allowing for varying toughness of competition (see Appendix A). In our model, there are two domestic regions, labeled r and s , and a rest-of-the-world, labeled R . Region r has L_r workers, and region s has L_s workers. Workers are also consumers, so that L_r and L_s denote regional market sizes. There are two vertically linked industries: a horizontally differentiated final good industry, and a horizontally

²The *empirical literature* has documented that changes in transport and trade costs influence the geographic concentration and composition of economic activity through a variety of channels: market access broadly defined (e.g., Redding and Venables, 2004; Redding and Sturm, 2008; Brühlhart, Carrère, and Trionfetti, 2012); infrastructure (e.g., Chandra and Thompson, 2000; Duranton, Morrow, and Turner, 2014); access to intermediates and changes in the value chain (e.g., Hanson, 1996, 1997; Holmes, 1999); tougher competition in product markets and exit dynamics (e.g., Dumais, Ellison, and Glaeser, 2002; Holmes and Stevens, 2014; Behrens, Boualam, and Martin, 2016); or any combination of these. See Redding and Turner (2015) for a recent survey of that empirical literature focusing mainly on infrastructure investments. There are only few contributions that look at transport costs more broadly defined (see, e.g., Storeygard, 2016, who interacts infrastructure with oil prices to provide time-varying measures of transport costs). The *theoretical literature* that investigates how domestic transport and international trade costs affect the geographic concentration of industries is large and reaches ambiguous conclusions. Decreasing transport and trade costs are dispersive in the models by Krugman and Livas Elizondo (1996), Helpman (1998), and Behrens, Mion, Murata, and Südekum (2013), whereas they are agglomerative in the models by Krugman (1991), Krugman and Venables (1995), and Fujita, Krugman, and Venables (1999). Using a richer spatial structure with two countries and four regions, as well as variable markups for firms, Behrens, Gaigné, Ottaviano, and Thisse (2007) find that falling domestic transport costs are agglomerative, whereas increasing international trade exposure is dispersive within countries. The reasons underlying the diverging results in the literature are related to differences in the modeling of agglomeration and dispersion forces. See Brühlhart (2011) for a review of the effects of increased trade openness on the internal geography of countries.

differentiated intermediate good industry. Workers are geographically immobile but perfectly mobile across industries. The final industry uses labor and a continuum of intermediates. The latter are produced using labor only. Both final and intermediate goods can be shipped at a transport cost t between domestic regions, and at a trade cost τ between domestic regions and the rest-of-the-world. A mass f_R of intermediates are imported from abroad. As transport and trade costs change, so do the prices of final and intermediate goods, as well as the trade patterns and firm sizes. Furthermore, changes in t and τ affect the masses of final firms (N_r and N_s) and the masses of intermediate firms (n_r and n_s) located in each region. In other words, the spatial structure of economic activity changes as the costs of trading goods domestically and internationally change.

Although the model is highly stylized, it is rich enough to preclude a full analytical investigation. Solving it explicitly and deriving paper-and-pencil comparative static results is, therefore, beyond the scope of this paper. We hence simulate it to guide our empirical analysis. The simulations provide information on: (i) the expected sign of key coefficients; and (ii) what controls are required to isolate the key effect of interest, namely the relationship between changes in domestic transport costs and the spatial concentration of industries. One important question is how we can evaluate the latter. Because we use the count-based Duranton and Overman (2005) measure of geographic concentration in our empirical analysis – which is computed using the distribution of bilateral distances between plants (see the Supplementary Appendix S.1) – a natural counterpart in our simple model is the distribution of plants across regions. Since the larger market has more plants – as it has a size advantage – we measure the geographic concentration of industries by the share of plants located in the larger region in excess of that region’s share of total market size. Formally, we use the following measures:

$$\left(\frac{N_r}{N} - \frac{L_r}{L} \right) \times 100\%, \quad \left(\frac{n_r}{n} - \frac{L_r}{L} \right) \times 100\%, \quad \text{and} \quad \left(\frac{N_r + n_r}{N + n} - \frac{L_r}{L} \right) \times 100\%, \quad (1)$$

where $L = L_r + L_s$ is the total population; $N = N_r + N_s$ is the total mass of final firms; and $n = n_r + n_s$ is the total mass of intermediate firms. The first two measures in (1) capture the over-representation (positive values) or under-representation (negative values) of final and intermediate plants in the larger region. The last measure in (1) reports the same information for both industries jointly, i.e., the aggregate spatial distribution of economic activity.

The key question of interest is how the geographic concentration of both industries across regions, as measured by (1), changes with t and τ . Table 1 below summarizes a series of results for different sets of parameter values of the model.

As can be seen from the top panel (a) of Table 1, decreasing domestic transport costs t lead to the geographic concentration of the final goods sector (column 1), slightly more geographic dispersion of the intermediate sector (column 2), and more geographic concentration overall (column 3). As can be seen from the middle panel (b), decreasing international trade costs τ have the opposite effect. They lead to more geographic dispersion of the final good sector

Table 1: Changes in geographic concentration with respect to t , τ , and f_R .

Parameter values			(1)	(2)	(3)	Relative wage
t	τ	f_R	Agglomeration final $\left(\frac{N_r}{N} - \frac{L_r}{L}\right) \times 100\%$	Agglomeration interm. $\left(\frac{n_r}{n} - \frac{L_r}{L}\right) \times 100\%$	Agglomeration both $\left(\frac{N_r+n_r}{N+n} - \frac{L_r}{L}\right) \times 100\%$	w_s/w_r
<i>(a) Changes in domestic transport costs, t</i>						
1.5	1.5	0.2	0.279	-0.452	-0.037	0.980
1.4	1.5	0.2	1.010	-0.617	0.268	0.981
1.35	1.5	0.2	1.396	-0.731	0.408	0.981
1.3	1.5	0.2	1.823	-0.877	0.554	0.982
1.25	1.5	0.2	2.364	-1.082	0.731	0.983
<i>(b) Changes in international trade costs, τ</i>						
1.4	1.6	0.2	2.266	-0.912	0.814	0.975
1.4	1.5	0.2	1.010	-0.617	0.268	0.981
1.4	1.4	0.2	0.915	-0.600	0.226	0.981
1.4	1.3	0.2	0.851	-0.594	0.199	0.982
1.4	1.2	0.2	0.790	-0.590	0.173	0.982
<i>(c) Changes in imported intermediates, f_R</i>						
1.4	1.5	0.1	1.036	-0.616	0.275	0.981
1.4	1.5	0.2	1.010	-0.617	0.268	0.981
1.4	1.5	0.3	0.986	-0.616	0.259	0.981
1.4	1.5	0.4	0.961	-0.616	0.250	0.981
1.4	1.5	0.5	0.935	-0.615	0.241	0.981

Notes: Simulations for the baseline parameter values $\alpha = 0.6$, $\sigma = 5$, $\beta = 0.33$, $L_r = 11$, $L_s = 10$, $F = 1$, $L_R = 1$, $\lambda_R = 0.3$, $f_R = 0.2$, and $c_R = 1$. The baseline case with $t = 1.4$ and $\tau = 1.5$ is displayed in bold font. All simulations have been carried out using Mathematica 10. The notebook is available online as supporting material. See Appendix A for details on the model structure.

(column 1), slightly more geographic concentration of the intermediate sector (column 2), and less geographic concentration overall (column 3). As can further be seen from the bottom panel (c) of Table 1, changes in the mass of imported intermediates also have a dispersive effect, this time on both sectors, although that effect is clearly smaller in magnitude. Last, although we do not explicitly summarize it in Table 1, it should be clear that the equilibrium level of geographic concentration depends on the model's structural parameters – the strength of scale economies, consumers' preferences for the goods, and the cost structure in each sector.

Our key qualitative findings can be summarized as follows: (i) decreasing domestic transport costs lead to more geographic concentration; (ii) decreasing international trade costs lead to more geographic dispersion; (iii) changes in the trading environment of one industry induce changes in the geographic concentration of vertically linked industries; and (iv) industry-specific fixed parameters related to the degree of scale economies, costs, and consumers' preferences influence the degree of geographic concentration.³

What are the empirical implications of the foregoing results? Three conclusions can be drawn for our empirical exercise. First, given the opposite effects of domestic transport and

³Results (i), (ii), and (iv) are qualitatively in line with those obtained by Behrens *et al.* (2007) in a two-country four-region model with quadratic quasi-linear preferences. Contrary to our model, there are no vertically linked industries in their model.

international trade costs, changes in the geographic concentration of industries can go either way, depending on the relative magnitude of changes in these costs. Since domestic transport and international trade costs are partly determined by the same underlying forces – international trade still requires domestic transportation at some point – any analysis of the changes in transport costs on the geographic concentration of industries needs to control for changes in international trade costs broadly defined. We will control for these costs using measures of trade exposure – both for imports and exports and for different groups of countries – though these are only proxies for international trade costs since they are an amalgam of transport costs, foreign productivity, and other forces.

Second, changes in the trading environment of one sector, intermediates in this case, have a direct impact on the geographic structure of the other sector, final in this case. As the mass of imported intermediates increases, holding trade and transport costs fixed, the final sector disperses more although it is not directly but only indirectly affected (through the input-output links). This suggests that we need to control for general equilibrium effects that are channeled through the impacts of changes in the geographic structure of some sectors on the other sectors. We will control for these equilibrium effects by constructing novel microgeographic measures that proxy access to potential local inputs and outputs.

Last, a large set of parameters that are fairly stable over time influence the degree of geographic concentration across industries in any cross section. To control for these factors, we will use panel data and focus on changes and not on levels. Doing so obviously does not solve the problem of industry-specific time-varying unobservables, but it is a step forward compared to purely cross-sectional analyses of the determinants of geographic concentration that most of the literature (e.g., Rosenthal and Strange, 2001; Ellison *et al.*, 2010) has used until now (yet, see Storeygard, 2016, for an approach using time-varying proxies for transport costs).

3 Measurement and descriptive evidence

Our analysis requires two key pieces of information: (i) a measure of geographic concentration of industries; and (ii) a measure of transport costs of industries. To control for unobserved factors that drive cross-sectional differences in geographic concentration and transport costs, we also require a panel dimension. We now discuss the data and the procedures that allow us to construct these key empirical measures. We also take a first descriptive look at the data.

3.1 Geographic concentration

We draw on Statistics Canada’s Annual Survey of Manufacturers (ASM) Longitudinal Microdata file from 1990 to 2009. This file contains between 32,000 and 53,000 manufacturing plants per year, covering 257 NAICS 6-digit industries. For every plant, we have information about: its

primary NAICS industry; its employment; its sales; and its 6-digit postal code. The latter allows us to effectively geo-locate the plants using latitude and longitude coordinates of postal code centroids, which are spatially very fine-grained in Canada. A detailed description of the data and the data sources is provided in Appendix B.1.

We measure the geographic concentration of industries using the Duranton and Overman (2005; henceforth, DO) K -densities. Since this approach has become fairly standard by now (see, e.g., Murata, Nakajima, Okamoto, and Tamura, 2014; Behrens and Bougna, 2015; Kerr and Kominers, 2015), we relegate the technicalities to the Supplementary Appendix S.1. The DO K -densities are kernel-smoothed distributions of the bilateral distances between plants in an industry. We compute them year by year for all industries at the NAICS 6-digit level.⁴ Since K -densities are distribution functions, we can also compute their cumulatives (CDF) up to some distance d . The CDF of the K -density at distance d provides a measure of the share of plants in an industry that are located at most at distance d from each other. Kerr and Kominers (2015) develop a model where the clustering of firms is driven by spatial interactions between pairs of plants and is subject to distance decay and (fixed) interaction costs. They show that the K -densities can be interpreted within that framework and that they provide information on the ‘overall degree’ of geographic concentration of industries and on ‘cluster shapes’.

Table 2: Mean of the Duranton-Overman CDFs across industries, 1990 and 2009.

	(1) Unweighted				(2) Employment weighted				(3) Sales weighted			
					CDF at a distance of							
Year	10 km	50 km	100 km	500 km	10 km	50 km	100 km	500 km	10 km	50 km	100 km	500 km
1990	0.020	0.076	0.139	0.420	0.021	0.083	0.151	0.449	0.022	0.086	0.156	0.453
2009	0.013	0.056	0.107	0.373	0.015	0.063	0.121	0.397	0.017	0.068	0.126	0.403
Mean (1990–2009)	0.015	0.064	0.121	0.394	0.017	0.073	0.136	0.422	0.019	0.077	0.141	0.428
% change	-36.0%	-27.1%	-22.6%	-11.3%	-28.7%	-23.3%	-20.3%	-11.4%	-21.5%	-21.2%	-19.3%	-11.0%

Notes: Authors’ computations based on the Annual Survey of Manufacturers Longitudinal Microdata file, 1990–2009. We report the values for the starting and the end years only since the series in between decrease rather smoothly (see our CEPR discussion paper version for the full table). The means of the CDF are based on 257 industries and are not weighted (but the CDFs for each industry are weighted by either employment in the middle columns (2), or by sales in the right columns (3); see Supplementary Appendix S.1 for details). ‘Mean (1990–2009)’ refers to the mean of the K -densities over the whole 1990–2009 period, on a yearly basis. ‘% change’ is the percentage change between 1990 and 2009.

The standard K -densities are computed based on plant counts, i.e., distances between pairs of plants without any weighting scheme. This relates to the theoretical measures we have used in our model as given by (1). Yet, we can also compute weighted versions. In particular, we can weight pairs of plants by either plant-level employment or plant-level sales. For these weighted versions of the K -density the unit of observation is the employee or a dollar of sales. A value of 0.086 at 50 kilometers using sales weights (see panel (3) of Table 2 for 1990) means, for example, that 8.6% of the industry sales are generated in plants that are at most 50 kilometers

⁴Table 9 in the Supplementary Appendix S.2 summarizes the K -densities for the geographically most concentrated industries in 1990, 1999, and 2009, respectively.

from each other.

Table 3: Summary statistics for K -densities and transport costs.

Variable	Industry detail	Mean	Standard deviation		
			Overall	Between	Within
Duranton-Overman K -density cumulative (CDF) at 10 km	6-digit	0.015	0.031	0.023	0.021
Duranton-Overman K -density cumulative (CDF) at 50 km	6-digit	0.065	0.060	0.050	0.034
Duranton-Overman K -density cumulative (CDF) at 100 km	6-digit	0.120	0.085	0.074	0.042
<i>Ad valorem</i> trucking costs as a share of goods shipped (avg. load, shipped 500 km)	L -level	0.034	0.035	0.030	0.005

Notes: Based on the sample that we use in our regression analysis, which includes 4,369 observations covering 257 industries and 17 years. The standard deviation is decomposed into between and within components, which measure the cross sectional and the time series variation, respectively. Additional information regarding our data sources and the construction of our key variables is provided in Appendix B.

Tables 2 and 3 provide descriptive statistics for our measures of geographic concentration. As can be seen from panel (1) of Table 2, manufacturing has geographically dispersed in Canada between 1990 and 2009.⁵ The average value of the CDF at 50 kilometers distance has decreased by 27.1% over a twenty year period. In other words, while 7.6% of the bilateral distances between the plants in the ‘average industry’ were less than 50 kilometers in 1990, only 5.6% of those distances still remained in that distance range in 2009. Table 2 further shows that concentration has decreased more at shorter distances: plants are dispersing, but less so at longer distances.⁶ This finding suggests that the incentives for plants to locate in very close spatial proximity to each other are weakening over time. It also likely reflects the fact that manufacturing industries have been bid out of cities because of higher land and labor costs there, and that they are moving to smaller nearby urban, sub-urban, or rural areas as a consequence (see, e.g., Henderson, 1997). As can further be seen from Table 2, that trend also affects the employment-weighted and the sales-weighted measures, but to a lesser extent. This is consistent with the fact that geographic concentration is stronger in terms of employment than in terms of plants, and even stronger in terms of sales than in terms of employment, especially at short distances (see Figure 6 in the Supplementary Appendix S.2). Last, Table 3 shows that although the bulk of the variation in the K -density CDFs is cross sectional, there is also

⁵This finding is consistent with the evidence reported in Behrens and Bougna (2015) and Behrens, Boualam, and Martin (2016), who have shown – using a different dataset – that the geographic concentration of manufacturing industries has decreased over the first decade of the years 2000 in Canada.

⁶Whereas the CDF of the K -density is easily interpretable and provides a natural measure to track the changing geographic concentration of industries, it cannot tell us anything about whether or not industries are statistically significantly concentrated or not. Table 10 in the Supplementary Appendix S.2 summarizes location patterns by year, based on their statistical significance (see Duranton and Overman, 2005, and the Supplementary Appendix S.1 for more details). We find that the share of statistically significantly concentrated industries has decreased over our study period, thus confirming the downward trend in the K -density CDFs. In a nutshell, there is a clear trend towards less geographic concentration, and that trend is captured by both the CDFs and the statistical tests.

substantial time-series variation. This variation will allow us to identify the effect of changes in transport costs on changes in geographic concentration.

To summarize, the descriptive evidence points to a significant decrease in the geographic concentration of manufacturing industries in Canada over the last 20 years, no matter whether that geographic concentration is measured in terms of plant counts, employment, or sales. The pace of decline, however, likely differs across industries in systematic ways. Understanding which factors drive that decrease to what extent and for which industries – with a special focus on transport costs – is the key objective of the remainder of this paper.

3.2 Transport costs

The second key ingredient of our analysis is an industry-specific measure of transport costs. Contrary to most existing studies, we use *direct measures* constructed from detailed micro-data files on shipments within Canada.

To estimate ad valorem rates, we first use a model to predict trucking firm (carrier) revenues for a 500 kilometers trip by commodity for the average tonnage using shipment (waybill) data from Statistics Canada’s Trucking Commodity Origin-Destination Survey (see Brown, 2015, for details). We estimate the ‘prices’ charged by trucking firms as a function of distance shipped, tonnage, and a set of commodity and firm fixed effects. To begin, we assume firms set prices such that both fixed and variable (linehaul) costs are just covered. Firms are assumed to set prices based on a fixed component and kilometers shipped: $R_{m,kc} = \alpha + \beta d_k$, where $R_{m,kc}$ is the revenue earned by carrier m for shipment k composed of commodity c , α is the fixed price component, β is rate per kilometer, and d_k is the distance shipped. Of course, firms may also price on a per tonne-km basis and this needs to be taken into account. Assuming firms set prices based on an unknown average tonnage t^* shipped implies that the rate per tonne-km is $R_{m,kc} = \alpha + \beta(d_k/t^*)t^*$. For loads less (greater) than t^* the implicit price per tonne-km will be scaled upward (downward) to ensure that the price on a per-km basis is maintained, which is captured by the following function:

$$R_{m,kc} = \alpha + \left[\frac{\beta}{t^*} + \phi(t^* - t_k) \right] d_k t_k = \alpha + \left(\frac{\beta}{t^*} + \phi t^* \right) d_k t_k - \phi d_k t_k^2, \quad (2)$$

where t_k is the actual tonnage shipped and $\phi(t^* - t_k)$ is the scaling factor. Factoring out the known tonnage results in a flexible function that allows firms to price using either rule (or some hybrid of the two). Equation (2) can be estimated using a simple quadratic form

$$R_{m,kc} = \alpha + \delta d_k t_k + \omega d_k t_k^2, \quad \text{with} \quad \delta = \beta/t^* + \phi t^* \quad \text{and} \quad \omega = -\phi. \quad (3)$$

The model is further augmented with commodity and carrier fixed effects and a series of controls to account for the quarter the shipment was made, the effect of empty backhauls on

prices, and fuel prices. The model is estimated across three types of carriers – truck-load, less-than-truck load, and specialized – with the estimated rate being the weighted average by value of the three by commodity (see Brown, 2015, for a detailed discussion of the data and model). Because prices are measured on a quarterly basis, fixed and variable costs are permitted to vary through time.⁷ Prices charged to shippers are predicted by commodity using their average tonnage for a 500 kilometers trip.⁸ The final predicted price, $\hat{R}_{c,t}$, is the weighted average by value of the three carrier types by commodity and year. While, as will become apparent, we only use predictions from 2008, the entire period is used in order to bring as much information to bear on the cross-sectional estimates by commodity. Predictions from the model closely match observed annual prices on aggregate.

Finally, the predicted prices $\hat{R}_{c,t}$ are converted to ad valorem trucking rates $\hat{T}_{c,t}$ by measuring the value of shipments. Since the value of shipments are not reported, they have to be estimated by multiplying the average tonnage shipped for each commodity by their respective value per tonne derived from an ‘experiment export trade file’ produced only in 2008. The ad valorem estimate at the commodity level in 2008 is $T_{c,2008} = \hat{R}_{c,2008}/V_{c,2008}$, where $V_{c,2008}$ is the estimated value of the commodity for the average tonnage shipped in 2008. Using an industry-commodity concordance, the ad valorem transportation rates in 2008 for commodities $T_{c,2008}$ are aggregated to an industry basis $T_{i,2008}$.

Finally, to generate a time series, yearly trucking industry price indices $\mathbb{P}_{\text{trans},t}$ and manufacturing industry price indices $\mathbb{P}_{i,t}$ from Statistics Canada’s KLEMS database are used to project the ad valorem rates backwards and forwards in time, thereby creating an industry-specific ad valorem transportation rate time series:

$$T_{i,t} = \frac{\mathbb{P}_{\text{trans},t}}{\mathbb{P}_{i,t}} T_{i,2008}. \quad (4)$$

Observe that our measure (4) of transport costs is more direct and detailed than those usually used in the agglomeration literature.⁹ Yet, it is by construction unlikely to be fully exogenous to industrial location patterns since it depends on price indices. We return to this important point later. Note, however, that we estimate transport costs for a ‘representative shipment’ by truck, holding distance fixed at 500 kilometers. Hence, variable shipping distances that result from optimal location choices of plants in an industry have a priori no direct influence on our measure.

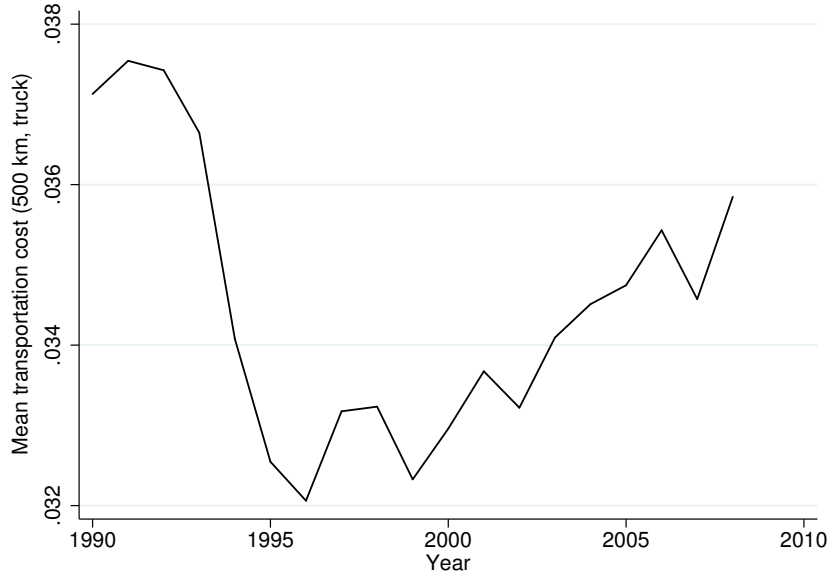
Figure 1 depicts the year-on-year changes in the (unweighted) cross-industry average transport costs for a representative 500 kilometers shipment. As can be seen, transport costs are

⁷Time enters as trends through a spline with knots set to reflect changing trends in a trucking price index generated from the same file.

⁸While we do not directly control for the time costs of transportation they will be, at least partially, embedded in the transport prices (which would capture quality of service for time-dependent trips).

⁹One exception is the work by Combes and Lafourcade (2005), who provide detailed estimates of transport costs for France. Their costs are, however, not industry specific.

Figure 1: Changes in average ad valorem trucking costs in Canada.



first decreasing – due, essentially, to reductions in labor costs at constant fuel prices – and then increasing – due, essentially, to increasing fuel prices at constant labor costs. They range from about 3.8% of the value of the shipment in the early nineties, to about 3.2% in the mid-nineties, with an average value of 3.4% (see Table 3). These figures are fairly close to the average ad valorem rates of 5% reported by Glaeser and Kohlhase (2004, p.206) using 2002 us data. As in their case, there is significant cross-industry variation in ad valorem trucking rates in our data. Table 11 in the Supplementary Appendix S.3 provides additional information on the highest and lowest ad valorem transport costs from 1992 to 2008. The highest ad valorem transport costs are for industries with low value to weight ratios (e.g., Cement Manufacturing and Breweries), with an average ad valorem transport cost across the top 10 industries in 2008 of 14%. The lowest are in industries with high value to weight ratios (e.g., Computer and Peripheral Equipment Manufacturing, and Medical Equipment and Supplies Manufacturing), with an average ad valorem transport cost in 2008 across the bottom 10 industries of 0.4%.

4 Empirical approach

We now discuss the empirical model that we estimate and the different identification concerns that we need to address.

4.1 Model

The patterns documented in Section 3.1 reveal that there is a clear trend toward the geographic dispersion of industries. As further documented in Section 3.2, transport costs have

also changed. To identify the causal effect of changes in the latter on changes in the former, we now turn to multivariate analysis. We work at the industry-year level and take advantage of the panel nature of our data. More precisely, we estimate the following model:

$$\gamma_{i,t}(d) = T_{i,t}\beta_T + \mathbf{X}_{i,t}\beta_X + \alpha_t + \mu_i + \varepsilon_{i,t}, \quad (5)$$

where $\gamma_{i,t}(d)$ is the K -density CDF for industry i in year t at distance d ; $T_{i,t}$ is our measure of ad valorem transport costs (4) of industry i in year t ; $\mathbf{X}_{i,t}$ is a vector of time-varying industry controls (including measures of international trade exposure and input-output links between industries; see Appendices B.2 and B.3 for details on how we construct those measures); α_t and μ_i are year and industry fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term.¹⁰ The latter is assumed to be independently and identically distributed with the usual properties for consistency of OLS. Our main coefficient of interest is β_T , which captures the impact of changes in transport costs on changes in the geographic concentration of industries.

4.2 Identification issues

In order for β_T to capture the causal effect of changes in transport costs on geographic concentration, we need to address a number of identification issues. In general, the two main problems that plague the identification of the determinants of geographic concentration are omitted variables and reverse causality (see Combes and Gobillon, 2015, for a recent survey). All studies based on cross-sectional industry data (e.g., Rosenthal and Strange, 2001; Ellison *et al.*, 2010) are potentially prone to these identification problems and use different strategies to overcome them. We first discuss the omitted variables problem and explain the large number of controls that we include in our panel estimations of (5) to deal with it. We then turn to the reverse causality problem and detail our instrumentation strategies. Table 8 in Appendix B.1 summarizes our main variables and provides descriptive statistics.

¹⁰One may be worried that identification in (5) comes from the *within* variation in the data. The latter could be small given yearly data, especially for the spatial variables. This point has been made in other studies (e.g., Ellison *et al.*, 2010, p.1200), but those studies usually use more aggregated measures of agglomeration. Those measures change much more slowly over time than the K -densities, especially at short distances. The reason is that our microgeographic measures are constructed from geocoded data, and that there is a lot of churning (entry and exit) at short distances. This is not picked up by spatially more aggregated measures, but it is by ours. Substantial plant-level churning creates a tension. On the one hand, there is *a lot of year-on-year variation*, which allows for identification using within variation (see Table 3). On the other hand, there is also *a lot of noise* at a small geographical scale, which makes the estimates less precise. The K -density CDFs are kernel-smoothed across distances, which reduces the noise due to plant-level churning (see our CEPR discussion paper version for additional details). Smoothing is also a standard procedure when working with the ASM longitudinal files (e.g., Statistics Canada usually uses three-year moving averages to filter noise in these series).

Omitted variables. We control for a wide range of time-varying industry-specific agglomeration factors. We know since Marshall (1890) that industries may concentrate for various reasons that are unrelated to the transport costs of goods, but linked to the costs of transporting people and ideas (see Duranton and Puga, 2004, for a review). Knowledge spillovers and labor market pooling are among the most important ‘Marshallian’ factors. Industries may also concentrate geographically because of localized natural advantages like natural resources or raw materials. The importance of controlling for these has been pointed out, among others, by Kim (1995), Ellison and Glaeser (1999), and Ellison *et al.* (2010). Furthermore, various other characteristics linked to the industrial organization of industries – industry size, the mean plant size, the size distribution of plants in the industry, the presence of multi-unit firms, or foreign ownership – are also likely to affect their spatial structure (see, e.g., Rosenthal and Strange, 2003). We account for these time-varying factors as follows. First, we control for knowledge spillovers using as a proxy an industry’s research and development (R&D) intensity, i.e., the ratio of R&D expenditure to total output of that industry (see Rosenthal and Strange, 2001). Second, we employ a proxy related to workers’ educational attainment to control for aspects of labor market pooling. More specifically, we use the share of hours worked by all workers with post-secondary education in the total number of hours worked in the industry.¹¹ Third, we control for the importance of natural advantages using the share of inputs from natural resource-based industries, and the sectoral energy inputs as a share of total sector output. Last, we control for basic industry structure and scale effects by including the following controls: total industry employment; mean plant size; the Herfindahl index of firm-level concentration; the share of plants controlled by multi-unit firms; and the share of plants controlled by foreign firms. These variables proxy for sectoral differences in the size distribution of firms, for potential differences in the location patterns of multi-unit multinationals, and differences in ‘business culture’.

As suggested by our theoretical model in Section 2 and Appendix A, controlling for input-output links and changes in international trade costs is important. We construct two controls that proxy for the costs of international trade and the proximity to customers and suppliers. Concerning the former, we construct measures of industry-level trade exposure (exports and imports), broken down by broad country groups – NAFTA, OECD excluding NAFTA, and low-cost countries. Changes in international trade exposure broadly capture changes in international trade costs, as well as changes in foreign productivity and the competitive environment in general. We provide details and descriptive statistics in Appendix B.2. We also discuss possible endogeneity concerns there. Turning to the proximity to customers and suppliers, we construct

¹¹We also constructed proxies for labor market conditions using the non-production to production workers ratio and other educational characteristics of the workforce. The latter are available at a more aggregated industry level (*L*-level) from Statistics Canada’s KLEMS database (e.g., the share of hours worked by all workers with a university degree, and the labor productivity index). These measures, however, are not statistically significant in the time series because they change quite slowly over time and are thus soaked up by the industry fixed effects.

novel microgeographic measures that are input-output share weighted distance measures to the closest potential customers and suppliers. These measures capture how close an industry is to other vertically linked industries from which it buys or to which it sells. We provide details on how we construct those measures, their descriptive statistics, and a discussion of their possible limitations and remaining endogeneity concerns, in Appendix B.3.¹²

Finally, the panel structure of our data allows us to control for industry-specific time-invariant factors and general macroeconomic trends. The former is important since, as suggested by the model in Section 2 and Appendix A, unobserved industry characteristics map into sizable cross-sectional differences in the geographic concentration of industries (see also Table 3, which shows that the major share of the variance in the K -densities is cross sectional). The latter captures, at least partly, general trends that affect the geographic concentration of industries (e.g., technological change due to improvements in $\pi\tau c$ that could make economic activity more footloose). Our broad range of time-varying controls, when combined with fixed effects, will capture many factors that may drive changes in geographic concentration of industries that are unrelated to changes in transport costs.

Reverse causality. Neither the panel structure of the data nor the controls help with potential problems of reverse causality. The latter may affect our key variable of interest, transport costs, which – as a price – reflect both demand and supply conditions in the market. A first problem is that productivity rises as an industry concentrates geographically (see, e.g., Rosenthal and Strange, 2004; Combes and Gobillon, 2015). Because our measure of transport costs is computed on an ad valorem basis and includes the industry price index, the causality may run from agglomeration to lower prices and, therefore, lower ad valorem transport costs. At the same time, agglomeration may lead to imbalances in shipping patterns, and the latter may increase the cost of transportation due to standard logistics problems like ‘backhaul’ of empty trucks (e.g., Jonkeren, Demirel, van Ommeren, and Rietveld, 2009; Behrens and Picard, 2011; Tanaka and Tsubota, 2016). Agglomeration would thus increase the transportation price index and bias our estimates. In a nutshell, $\mathbb{P}_{\text{trans},t}/\mathbb{P}_{i,t}$ in expression (4) is likely to be endogenous to the degree of geographic concentration of an industry, with stronger concentration increasing that ratio due to a combination of rising freight prices and lower output prices. Thus, the OLS estimate of β_T is likely to be upward biased in our model.¹³

¹²As explained in Appendix B.3, our proxies for input-output links are reduced form and not structural (unlike, e.g., the ‘structural supplier and market access’ in Redding and Venables, 2004). As pointed out by Combes and Gobillon (2015, p.274), there is generally no satisfying solution to control for supplier and market access in empirical estimation. Either we use a structural model – which requires a lot of assumptions and has its own limitations – or we use reduced-form proxies that aim at capturing those interactions.

¹³Industries that concentrate geographically are also likely to ship their output over different distances than industries that are less agglomerated. This problem does, however, not affect our estimates since our measure of transport costs is constructed for a representative shipment over a fixed distance of 500 kilometers.

To deal with those problems, we adopt three different strategies. First, we clear out the effect of productivity growth – one presumed source of endogeneity – on prices by regressing our transport cost series on industry multi-factor productivity indices (from the KLEMS database), as well as industry and year fixed effects. We then use the residual from that regression as a proxy for the transportation cost series. By definition, that residual is orthogonal to any productivity-driven price changes that could stem from the changing geographic concentration of industries. This strategy does not deal directly with the transportation price index.

Second, we use US manufacturing industry price indices as external instruments for the transport cost series. This instrumentation strategy is similar to that of Ellison *et al.* (2010), who instrument the US input-output tables and the US industry labor requirements with those of the UK. The underlying idea is the following. Assume that the geographic concentration of an industry increases over time because of unobserved factors that we cannot control for in our analysis. This increasing geographic concentration then raises ad valorem transport costs via price decreases of the industry’s output. Provided that the changes for the US are not driven by the same unobserved factors that affect the spatial concentration of the industry in Canada, but that the US price series $P_{trans,t}^{US}/P_{i,t}^{US}$ are correlated with the changes in $P_{trans,t}/P_{i,t}$, they will provide valid instruments for the Canadian transport cost series. Two potential limitations of these instruments are the following: (i) there may be common underlying unobserved factors that drive changes in the concentration of the same industries in Canada and in the US; or (ii) the geographic concentration of an industry in Canada affects directly the productivity – and, therefore, the price indices – in the US. While we cannot completely rule out those possibilities, neither strikes us as very plausible. First, the panel nature of our data and the extensive set of time-varying controls should pick up most of the unobserved factors that may drive the increasing concentration of the industry; and second, the Canadian economy is small compared to the US economy, so that changes in the geographic concentration in Canada are very unlikely to have substantial productivity impacts in the US.¹⁴

Last, as we have a large number of industries and a fairly long time series, our setting lends itself reasonably well to the construction of internal instruments. We implement the method suggested by Lewbel (2012), which exploits heteroscedasticity and variance-covariance restrictions to obtain identification with 2SLS when some variables are endogenous and when external instruments are either weak or not available. Appendix C provides more details on this method. We use this approach to deal with remaining potential endogeneity issues in our trade exposure and input-output distance measures (see Appendices B.2 and B.3 for a

¹⁴The elasticity of productivity to the density or size of economic activity is usually in the 3–8 percent range. Hence, huge changes in the geographic structure are needed to obtain large productivity changes. Furthermore, given that the Canadian economy is ten times smaller than the US economy, shocks to Canadian productivity have very limited effects on the US, save for a couple of states relatively close to the border or a couple of border-spanning industry networks (like the automotive industry). Excluding the automotive industries (NAICS 3361, 3362, and 3363) has no qualitative effect on the ‘ad valorem trucking cost (residual)’ coefficient.

discussion). We continue to use the external us instrument for transport costs. We view this approach as an additional robustness check – that takes care of possible endogeneity issues with our trade and input-output controls – on top of the previous two strategies that rely on either ‘filtering’ or the use of external instruments and standard 2SLS IV.

It should be clear that is very difficult to fully solve all endogeneity issues, given the sectoral level of aggregation at which we work. Yet, the panel nature of our data – which allows for the inclusion of industry fixed effects – our extensive set of time-varying controls, as well as the construction and instrumentation strategies for our main variable of interest, transport costs, all help us to be reasonably confident that we identify causal effects of changes in transport costs $T_{i,t}$ on our measure $\gamma_{i,t}(d)$ of geographic concentration.

5 Results

We first estimate several variants of equation (5), which differ by the sets of industry characteristics and controls they include.¹⁵ We only report results using the unweighted DO K -density CDF as our measure of geographic concentration. Using the employment or sales-weighted measures yields similar results (see Table 13 in the Supplementary Appendix S.4).

5.1 Baseline results

Table 4 summarizes our baseline results. **Model 1** reports the raw unconditional correlation without controls, whereas **Model 2** reports within estimates without controls. As can be seen, the coefficient on the trucking costs is negative and highly significant. In words, falling trucking costs are associated with the geographic concentration of industries. We then progressively add in **Models 3 to 6** the variables suggested by the model developed in Section 2 to control for international trade exposure, input-output links, and other industry characteristics.

Starting with **Model 3**, we add our trade variables (import and export exposure by country groups) to the baseline case. As can be seen, rising import shares are across the board associated with falling geographic concentration. The (non-OECD) Asian share of imports, which we use as a proxy for low-wage countries, has the largest estimated coefficient in absolute value and is the most statistically significant. One explanation for the dispersive effect of import competition is that firms become more footlose as they source a larger share of their intermediates from abroad and no longer rely on (possibly localized) domestic suppliers. Another explanation, for which Holmes and Stevens (2014) provide empirical evidence, is that import

¹⁵We performed a Hausman test for (5) to confirm that the appropriate estimator is a fixed-effects estimator and not a random-effects estimator. The result of the test strongly confirms (at the 1% level) that the fixed-effects estimator is the preferred specification. Note also that we work with the universe of manufacturing industries, so that there is no sampling variability in industries that would warrant the use of the random-effects estimator.

Table 4: Baseline estimation results for specification (5).

Variables	Dependent variable is the CDF at 50 km						
	(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)	(Model 6)	(Model 7)
Ad valorem trucking costs	-0.265 ^a (0.037)	-0.337 ^b (0.155)	-0.263 ^b (0.122)	-0.250 ^b (0.098)	-0.183 ^b (0.078)	-0.208 ^b (0.088)	
Ad valorem trucking costs (residual)							-0.261 ^a (0.078)
Asian share of imports			-1.639 ^a (0.413)		-1.309 ^a (0.379)	-1.132 ^a (0.380)	-1.118 ^a (0.383)
OECD share of imports			-1.103 ^a (0.400)		-0.638 ^c (0.341)	-0.491 (0.344)	-0.475 (0.345)
NAFTA share of imports			-1.188 ^a (0.363)		-0.715 ^b (0.327)	-0.562 ^c (0.327)	-0.549 ^c (0.327)
Asian share of exports			0.370 (0.542)		0.371 (0.422)	0.482 (0.405)	0.482 (0.412)
OECD share of exports			0.366 (0.274)		0.423 ^b (0.210)	0.440 ^b (0.189)	0.443 ^b (0.193)
NAFTA share of exports			0.298 (0.294)		0.279 (0.206)	0.319 (0.196)	0.318 (0.201)
Input distance				-0.359 ^a (0.064)	-0.343 ^a (0.059)	-0.361 ^a (0.055)	-0.358 ^a (0.055)
Output distance				-0.262 ^a (0.046)	-0.290 ^a (0.044)	-0.313 ^a (0.042)	-0.318 ^a (0.043)
Average minimum distance				-0.323 ^a (0.046)	-0.300 ^a (0.047)	-0.296 ^a (0.039)	-0.294 ^a (0.039)
Industry controls included	No	No	No	No	No	Yes	Yes
Number of NAICS industries	257	257	257	257	257	257	257
Number of years	17	17	17	17	17	17	17
Year dummies	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations (NAICS \times years)	4,369	4,369	4,369	4,369	4,369	4,369	4,369
R^2	0.165	0.054	0.096	0.442	0.473	0.516	0.518

Notes: The dependent variable is the unweighted (count based) Duranton-Overman K -density CDF at 50 kilometers distance. ^a, ^b and ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. We use simple OLS. Standard errors are clustered at the industry level and given in parentheses. Our measures of input and output distances, as well as average minimum distance, are computed using $N = 5$. 'Ad valorem trucking costs (residual)' denotes the residual of the regression of 'Ad valorem trucking costs' on industry multi-factor productivity. A constant term is included in all regressions but not reported. **Model 6** includes the following industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales). Last, our preferred specification, **Model 7**, uses the residual trucking rate as the main explanatory variable.

competition from low-wage countries leads to significant exit of large plants that produce standardized 'main segment' goods that are very exposed to competition.¹⁶ Since those plants are the ones that are predominantly clustered at short distances, their exit will significantly reduce the extent of measured geographic concentration.¹⁷ As can also be seen from **Model 3**

¹⁶We cannot disentangle the impact of exit vs relocation on the spatial structure. However, we control for the size of the industry, which at least partly picks up entry and exit dynamics. Note that relocations are quite rare and should have little impact on our results. The bulk of the variation is driven by entry and exit.

¹⁷This is a somewhat surprising result, because we would expect the productivity enhancing effects of geographic concentration to shelter firms from low-wage competition. Yet, one should keep in mind that clustering

in Table 4, rising export shares are across the board associated with increasing geographic concentration, although the effect is only significant for the share of exports to OECD countries. This pattern may be driven by the fact that more isolated non-exporting plants have a higher probability of exit, or that localization increases the export participation and performance of plants (e.g., Koenig, Mayneris, and Poncet, 2010).

Next, **Model 4** adds our input-output distances – the industry mean of the average minimum distance to a dollar of inputs or outputs computed using the five nearest plants in each industry – as well as our minimum distance (density) control.¹⁸ As can be seen, the estimated coefficients on the input and output distance measures are negative and highly significant, and they tend to be of similar magnitude. Industries tend to follow their suppliers and customers, i.e., industries where potential suppliers or clients disperse tend to also disperse. This shows that controlling for vertical industry links – as suggested by our model – is important in order to capture the general equilibrium effects of changing geographic industry structure. Note that this effect is not driven by changes in overall density which we control for (and the associated variable is highly significant). Note also that the coefficient on trucking costs barely changes when including our measures for access to suppliers and customers. **Model 5** shows that the joint inclusion of both trade and input-output controls, as suggested by theory, does not significantly change our baseline estimates. Although the coefficient on trucking cost drops, it remains negative and highly significant.

Next, **Model 6** adds a large number of time-varying industry controls that may influence the geographic concentration of industries. We include a measure of industry size, several controls for industry structure (the Herfindahl index of the firm-size distribution, mean plant size, the share of plants controlled by multiplant firms, and the share of plants controlled by foreign-owned firms), and several proxies for natural advantages (the share of inputs from natural resource-based industries, and the share of energy inputs in total output). We also include proxies for the skill composition of the workforce and for knowledge spillovers to control for ‘Marshallian’ forces. Our results are robust to the inclusion of these controls.

As discussed in Section 4.2, our measure of transport costs is potentially endogenous. We return to this point in more detail below. As a first step to address this problem, we use in **Model 7** – our preferred specification – the residual transport cost obtained from a first-stage regression of that cost on industry multi-factor productivities and a set of industry and year

provides firms with benefits as long as clusters grow (positive shocks), but that the unravelling of clusters (negative shocks) may lead to a domino effect as the agglomeration benefits dissipate with the exit of firms. Also, as shown by Holmes and Stevens (2014), plants in clusters operate on different market segments than non-clustered plants, and they are more vulnerable to import competition. See also Behrens, Boualam, and Martin (2016) for recent evidence on the dispersive effect of import competition on Canadian textile clusters between 2001 and 2013.

¹⁸See Appendix B.3 for a discussion of the need to include the minimum distance control. Using $N = 3, 7, 10$ plants to construct these distances yields qualitatively very similar results.

fixed effects.¹⁹ Consistent with our prior, the coefficient on transport costs becomes larger in absolute value when using the productivity-purged residual (compare **Models 6** and **7**). This is in line with our expectations discussed in Section 4.2, where we have argued that endogeneity concerns due to reverse causality are likely to bias the coefficient upwards (towards zero in this case). In what follows, we hence systematically use the residual measure of ad valorem trucking costs in all our regressions. We more fully deal with remaining endogeneity issues later. Last, note that in our preferred specification about half of the time-series variation in geographic concentration is explained by the model (5).

How large are the effects of transport costs on the geographic concentration of industries? As shown in Section 3.1, the degree of geographic concentration of manufacturing industries has significantly fallen in Canada between 1990 and 2009. How much of that change is explained by changes in transport costs? To gauge the strength of the effect, we compute the predicted change in the CDFs by holding the ad valorem trucking costs constant at their 1992 values, while still allowing the other variables to change through time. The observed change in the cross-industry average CDF between 1992 and 2008 at a distance of 50 kilometers is -23.37%. Holding the ad valorem trucking rate fixed at its 1992 level, the change would have been -28.36%. Thus, had ad valorem trucking costs not decreased by about 4% over the 1992–2008 period, the average geographic concentration of industries would have fallen by about 5 percentage points more (about 20% of the overall change).²⁰

5.2 Spatial scale

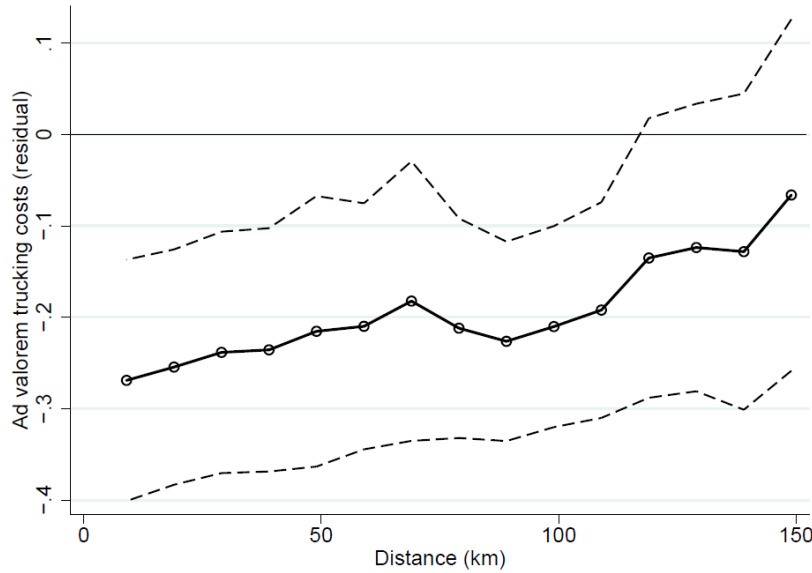
One of the strengths of the DO measure of geographic concentration is that we can vary it continuously across distances. We hence next investigate the spatial scale at which changes in transport costs operate by changing the distance d at which the K -density CDF is evaluated. Doing so allows us to highlight how our key variable of interest influences the geographic concentration of industries at different spatial scales. Furthermore, we can provide plots of the marginal effects over the whole distance range, and thus allowing for a fine analysis of the spatial dimensions of the changes in agglomeration due to changes in transport costs. The left half of Table 5 summarizes our results for different distances. To save space, we only report results for **Model 7** at three selected distances: 10, 100, and 500 kilometers. As can be seen, the qualitative results do not depend on the distance threshold d . Yet, there is a general tendency

¹⁹When using the ‘ad valorem trucking cost residual’ from the first-stage regression, we need to bootstrap the standard errors to control for the presence of an estimated regressor. We did this for the baseline specification (see **Model 7** in Table 7), and it makes virtually no difference. We hence report non-bootstrapped standard errors (yet clustered by industry) in all specifications.

²⁰We can repeat the exercise for import shares. Holding all import shares fixed at their 1992 level, the change in the CDF would have been -14.63%. In words, had imports remained at their 1992 levels, the geographic concentration would have fallen by about 9 percentage points (i.e., 60%) less than what we observed.

for the values and significance of the covariates to attenuate as the CDF increases in distance.

Figure 2: Estimated ‘ad valorem trucking costs (residual)’ coefficient by distance.



This can clearly be seen from Figure 2 and from the right half of Table 5, where we define the incremental distance of the CDF between distance d_1 and distance $d_2 > d_1$ as follows: $\Delta\gamma_i(d_1, d_2) = \gamma_i(d_1) - \gamma_i(d_2)$. We estimate the marginal effects of our variables by ‘distance bands’. As one can see, there is basically no more additional effect of ad valorem trucking costs or the Asian import share on the degree of geographic concentration beyond about 100 kilometers. Furthermore, the largest (and statistically most significant results) occur in the distance bands between either 10 and 25 kilometers, or between 25 and 50 kilometers. This result suggests that many of the agglomeration mechanisms linked to transportation operate at the scale of metropolitan areas, either by influencing within-metro patterns or between-metro specialization (see Duranton, Morrow, and Turner, 2014). At longer distances – beyond about 200 kilometers – other factors that do not figure in our model drive the clustering of firms, or incremental clustering becomes weak and fairly unimportant.²¹

Figure 2 depicts the incremental change in the ‘ad valorem trucking costs (residual)’ coefficient by 10 kilometers steps and also reports the 90% confidence intervals (dashed lines). Since all marginal coefficient changes are statistically zero after 110 kilometers, we limit the plot to a range of 150 kilometers. As can be seen from Figure 2, globally, the effects are significant up to about 110 kilometers.²² They are no longer significantly different from zero above 120

²¹The CDFs naturally display less variability across industries the greater is the distance d . The reason is that they are bounded from above by unity, and we converge by construction to that value for all industries if we compute them over sufficiently long distances.

²²That distance roughly encompasses the scale of the large Canadian metropolitan regions (e.g., the island of Montreal is about 50 kilometers long).

Table 5: Estimation of (5) by distance and by incremental change in the CDF.

Variables	Model (7), by distance			Model (7), by incremental CDF			
	CDF 10km	CDF 100km	CDF 500km	$\Delta\gamma_i(10, 25)$	$\Delta\gamma_i(25, 50)$	$\Delta\gamma_i(50, 100)$	$\Delta\gamma_i(100, 500)$
Ad valorem trucking costs (residual)	-0.269 ^a (0.080)	-0.250 ^a (0.073)	-0.212 ^a (0.048)	-0.253 ^a (0.079)	-0.238 ^a (0.080)	-0.229 ^a (0.069)	-0.105 (0.090)
Asian share of imports	-1.359 ^a (0.467)	-0.923 ^a (0.299)	-0.307 ^b (0.139)	-1.029 ^b (0.433)	-0.724 ^b (0.337)	-0.352 (0.235)	0.583 (0.429)
Input distance	-0.382 ^a (0.063)	-0.340 ^a (0.049)	-0.242 ^a (0.033)	-0.332 ^a (0.061)	-0.322 ^a (0.055)	-0.315 ^a (0.054)	-0.193 ^a (0.041)
Output distance	-0.307 ^a (0.046)	-0.307 ^a (0.040)	-0.197 ^a (0.027)	-0.341 ^a (0.045)	-0.340 ^a (0.045)	-0.302 ^a (0.045)	-0.122 ^a (0.039)
Average minimum distance	-0.322 ^a (0.046)	-0.268 ^a (0.035)	-0.137 ^a (0.024)	-0.298 ^a (0.041)	-0.243 ^a (0.043)	-0.204 ^a (0.038)	-0.038 (0.036)
Other trade shares included	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry controls included	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.473	0.540	0.545	0.481	0.417	0.436	0.168

Notes: All estimations for 257 industries and 17 years (4,369 observations). The dependent variable is the unweighted (count based) Duranton-Overman K -density CDF at the reported distance. ^a, ^b and ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. We use simple OLS. All specifications include industry and year fixed effects. Standard errors, given in parentheses, are clustered at the industry level. Our measures of input and output distances are computed using $N = 5$. 'Ad valorem trucking costs (residual)' denotes the residual of the regression of 'Ad valorem trucking costs' on industry multi-factor productivity. A constant term is included in all regressions but not reported. All industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales) are included but not reported. All trade share variables are included in the regressions, but we do not report detailed results to save space.

kilometers. Note also that the standard errors start to increase more quickly after about 110 to 120 kilometers, i.e., the estimated coefficients on transport costs become less precise.

5.3 Robustness checks

We now provide additional evidence on the robustness of our key findings. First, we investigate the robustness of our results to the choice of the dependent variable. Table 13 in the Supplementary Appendix S.4 shows that the effect of transport costs on geographic concentration is weaker – and the explanatory power of the model slightly lower – when the latter is measured using either employment- or sales-weighted CDFs. Although the key qualitative flavor of the results and the sign and significance of our key coefficient remain unchanged, the estimates using employment- or sales-weighted K -densities are slightly less sharp. Furthermore, the effect of import competition tends to be more limited to imports from Asia, and the coefficient tends to be smaller too. This suggests that much of the adaptation to import competition, particularly from low-wage countries which are responsible for the bulk of exit in Canadian manufacturing, occurs for smaller plants and firms (Behrens *et al.*, 2016). The residual transport cost variable remains significantly negative in all specifications that we estimate, irrespective of how we construct the dependent variable. In a nutshell, changes in transport

costs have a significant effect on the geographic concentration of economic activity, no matter whether we consider plants, employment, or sales to measure that concentration.

Table 6: Estimation results for specification (5), robustness checks.

Variables	Dependent variable is the CDF at 50 km		
	(Model 8)	(Model 9)	(Model 10)
Ad valorem trucking costs (residual)	-0.245 ^a (0.082)	-0.198 ^b (0.083)	-0.192 ^b (0.086)
Asian share of imports	-1.119 ^a (0.387)	-1.035 ^a (0.398)	-1.041 ^a (0.401)
Input distance	-0.337 ^a (0.056)	-0.238 ^a (0.055)	-0.232 ^a (0.056)
Output distance	-0.324 ^a (0.041)	-0.292 ^a (0.039)	-0.298 ^a (0.038)
Average minimum distance	-0.264 ^a (0.040)	-0.290 ^a (0.037)	-0.272 ^a (0.037)
Other trade shares included	Yes	Yes	Yes
Industry controls included	Yes	Yes	Yes
<i>Additional controls:</i>			
Distance to major container ports	-0.235 ^b (0.098)		-0.140 (0.104)
Eastern share of plants		0.758 ^a (0.159)	0.715 ^a (0.159)
Number of NAICS industries	257	257	257
Number of years	17	17	17
Year dummies	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes
Observations (NAICS \times years)	4,369	4,369	4,369
R^2	0.525	0.551	0.553

Notes: The dependent variable is the unweighted (count based) Duranton-Overman K -density CDF at 50 kilometers distance. ^a, ^b and ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. We use simple OLS. Standard errors are clustered at the industry level and given in parentheses. Our measures of input and output distances, as well as average minimum distance, are computed using $N = 5$. 'Ad valorem trucking costs (residual)' denotes the residual of the regression of 'Ad valorem trucking costs' on industry multi-factor productivity. A constant term is included in all regressions but not reported. All industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales) are included but not reported. All trade share variables are included in the regressions, but we do not report detailed results to save space.

Second, one may worry that import entry points are geographically localized, so that increasing imports could lead to the geographic concentration of industries around those entry points. This may, in turn, bias the transport cost coefficient. To address this potential concern, we include as a control the average minimum distance of plants in each industry from one of the five major Canadian container ports.²³ As can be seen from **Model 8** in Table 6, although

²³Those ports are Halifax (NS), Saint John (NB), Montreal (QC), Vancouver (BC), and Prince Rupert (BC).

industries that do move closer to import entry points tend to concentrate more geographically, the coefficients on the other variables are basically unchanged. Another concern we need to address is that the above-average economic growth of the western Canadian provinces might be driving the observed dispersion of manufacturing. To address that concern, we include in our regressions as an additional control the share of eastern plants in each industry.²⁴ As can be seen from **Model 9** in Table 6, this reduces somewhat the coefficients on trucking costs and Asian share of imports, yet does not change our qualitative results. Although the westward shift of the Canadian economy over the 1992–2008 period partly accounts for increasing geographic dispersion of industries, that effect is limited and does not substantially reduce the effect of trucking costs and import competition. Last, **Model 10** shows that the joint inclusion of both controls does not change our results either.

We also run a large number of robustness checks that we do not report in detail. First, we re-estimate the model by averaging all variables over five year periods. Doing so reduces the year-on-year volatility of some variables (e.g., the trade variables), and allows for slowly moving variables like R&D expenditures to be potentially better identified in the regressions. It also deals with business cycle aspects that may drive the changes in the geographic concentration of industries. The last three columns of Table 13 in the Supplementary Appendix S.4 show that our basic findings are unchanged when replacing year-on-year variations with five-year averages. Second, our results may be partly driven by a small number of sectors that were subject to major changes over our study period. For example, as documented by Behrens *et al.* (2016), the Canadian textile and clothing industry experienced a remarkable downward trend in the number of plants and in its geographic concentration in the wake of the end of the Multi-Fibre Arrangement in 2005. Given that the textile and clothing industry contains some of the initially most strongly agglomerated sectors in Canada (see Table 9 in the Supplementary Appendix S.2), the large changes in those sectors may drive some of our key results. That this is not the case, and that all of our main findings are robust to the exclusion of those sectors, is shown in Table 12 in the Supplementary Appendix S.2. We also run our regressions by excluding the ‘high-tech’ sectors, and the results are qualitatively unchanged. Third, we compute measures of ‘upstreamness’ of industries following Antràs, Chor, Fally, and Hillberry (2012). Using those measures, we split industries into the top quintile Q5 (most upstream industries) and the bottom quintile Q1 (most downstream industries). We then reestimate the model by interacting the transport and trade variables with those upstream-downstream dummies to capture potentially different impacts on different industries in the vertical production chain. The results are summarized in Table 14 in the Supplementary Appendix S.4. When including our input-output measures, splitting by upstreamness has virtually no effects on our main coefficients, which suggests that our input-output measures capture quite well vertical industry links. When excluding those measures, we find that more downstream industries are more

²⁴We compute each industry’s share of plants in Ontario, Quebec, and the Atlantic provinces.

sensitive to both transport costs and import competition, although the differential effects are quite imprecisely estimated.

We provide additional details on many other robustness checks that we run (controlling for sectoral changes in ICT; heterogeneous transport coefficients across industries; asymmetric coefficients for rising and falling costs; non-linear transport cost specifications etc.) in the Supplementary Appendix S.4. To summarize, our key findings are robust to a large number of robustness checks and continue to hold true in a variety of alternative specifications. Sectors that see their transport costs increase or that experience greater import competition tend to disperse more than other sectors.

5.4 Controlling for endogeneity

We finally address additional potential endogeneity concerns that we discussed in Section 4.2. The results of the different estimations are summarized in Table 7. **Model 11** replicates **Model 7** of Table 4. As explained previously, we use the residual of a regression of ad valorem trucking costs on sectoral multifactor productivity – including a set of industry and year fixed effects – in that specification. The residual from that regression is, by construction, orthogonal to multifactor productivity. **Model 11** in Table 7 differs from **Model 7** in Table 4 only by the standard errors, which are bootstrapped using 200 replications. The results do not change.

Although the residual transport cost is purged from productivity effects, endogeneity concerns linked to, e.g., backhaul, remain. Hence, we run some instrumental variable (IV) regressions to check the validity of our results. **Model 12** summarizes our IV-2SLS results where we instrument the ad valorem trucking rate residual by replacing the Canadian price indices with their US counterparts. The rationale underlying this instrumentation strategy was explained in Section 4.2. The first-stage results are summarized in column 3 of Table 7. As can be seen, the instrument is strong (with a first-stage F -test value of 19.07 and a first-stage R^2 of 0.62). Column 4 of Table 7 shows that, in line with our expectations, the instrumented coefficient is substantially more negative than the coefficient for the residual ad valorem trucking rate, which is itself already more negative than the coefficient using the unpurged trucking rate. The direction of the bias in the estimated coefficients is the same as in **Model 7**, which suggests that OLS estimates underestimate the impact of changes in ad valorem transport costs on the geographic concentration of industries.

Last, **Models 13** and **14** in Table 7 address remaining endogeneity concerns that may affect the trade variables and the input-output distance measures (see our discussions in Appendices B.2 and B.3, respectively).²⁵ Since we do not have any good external instruments for those

²⁵For example, plants are identified in our data by their primary industry only. Hence, although we exclude the industry itself when constructing the input-output distance measures, multi-product plants whose secondary activities belong to the same industry might be included in those measures.

Table 7: Controlling for the potential endogeneity of $T_{i,t}$ in specification (5).

Variables	Dependent variable is the CDF at 50 kilometers				
	(Model 11) Base	(Model 12) First stage	(Model 13) IV-2SLS	(Model 13) Lewbel 1	(Model 14) Lewbel 2
Ad valorem trucking costs (residual)	-0.261 ^a (0.0781)		-0.393 ^a (0.0956)	-0.197 ^b (0.0929)	-0.218 ^b (0.0915)
Ad valorem trucking costs us (instrument)		0.485 ^a (0.111)			
Asian share of imports	-1.118 ^a (0.383)	-0.056 (0.107)	-1.095 ^a (0.381)	-1.592 ^a (0.533)	-1.618 ^a (0.501)
Input distance	-0.358 ^a (0.0545)	0.035 ^c (0.020)	-0.356 ^a (0.055)	-0.143 ^c (0.075)	-0.224 ^a (0.075)
Output distance	-0.318 ^a (0.043)	-0.011 (0.015)	-0.322 ^a (0.042)	-0.386 ^a (0.083)	-0.360 ^a (0.087)
Average minimum distance	-0.294 ^a (0.039)	0.005 (0.014)	-0.290 ^a (0.039)		
Other trade shares included	Yes	Yes	Yes	Yes	Yes
Industry controls included	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.518		0.516	0.319	0.330
First-stage R^2		0.628			
First-stage F test of excluded instruments		19.07			

Notes: The dependent variable is the unweighted (count based) Duranton-Overman K -density CDF at 50 kilometers distance. ^a, ^b and ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. Our measures of input and output distances are computed using $N = 5$. 'Ad valorem trucking costs (residual)' denotes the residual of the regression of 'Ad valorem trucking costs' on industry multi-factor productivity. Model 11 replicates our preferred Model 7 but the standard errors are bootstrapped because of the generated regressor. Model 12 instruments the 'Ad valorem trucking costs' using costs constructed from us price indices. Models 13 and 14 use the Lewbel (2012) methodology to instrument input-output distances and trade shares. In Model 13, only a subset of the import shares is instrumented, while all trade shares are instrumented in Model 14. See Appendix C for additional details. A constant term is included in all regressions but not reported. All industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales) are included but not reported. All trade share variables are included in the regressions, but we do not report detailed results to save space.

variables, we use the Lewbel (2012) estimator with internal instruments for the input-output distances and the trade shares (see Lewbel, 2012, and Appendix C for more details on the implementation).²⁶ The excluded external instrument is the us price-based ad valorem trucking costs as before. As can be seen from the results in Table 7, the instrumented coefficient on the Asian share of imports increases, as do most of the other trade share coefficients. At the same time, both the magnitude of transport costs and of the input and output distances decreases slightly. However, these variables remain significant and their magnitude is in the same ballpark than in the case of OLS. Thus, our results appear to be robust. Decreases in ad val-

²⁶Since there is an insignificant correlation between the OECD export share and the squared residuals, we did not include it. We substituted instead the NAFTA import share because it is consistently significant in the baseline set of models and it meets the criteria for being internally instrumented (see Appendix C).

orem trucking costs lead to more geographic concentration of manufacturing industries, even when potential endogeneity concerns in transport costs, trade exposure, and buyer-supplier relationships are taken into account.

6 Conclusions

We develop a long industry panel of continuous measures of geographic concentration and combine them with measures of ad valorem trucking rates, international trade exposure, upstream and downstream links, and a large number of controls, to assess whether the ‘world is flat.’ Our answer is an emphatic, not yet! That is, the key message of our findings is that changes in the geographic concentration of industries due to changes in transport costs are sizable. This is a robust finding that survives a battery of checks, including extensive efforts to address inherent endogeneity issues that plague such estimations. We should also add that, to the best of our knowledge, this is the first instance where direct measures of transport costs have been used to assess their effects on the geographic concentration of industries.

The lessons for researchers from this work are twofold. The first is that it is difficult to contemplate investigating industry location (or co-location) without taking transport costs explicitly into account (see, e.g., Combes and Gobillon, 2015). In a nutshell, investing in better measures of transport costs is important and likely to pay substantial dividends. The second is that it is equally difficult to consider the effects of transport costs in isolation. Their general equilibrium effects on input-output links and competition, and more generally their endogenous nature as market prices, have to be grappled with. This involves challenges – both theoretical and empirical – with large investments required for both. While we believe we have made some strides developing the necessary empirics, theoretical work in moving from simulations to full-blown analytical results is still called for and needed.

The lesson for policy makers is simple: small changes in transport costs – e.g., due to infrastructure projects or any other policy that changes the costs of trading goods across nations and regions – still impact the economic geography of industries. Contrary to what seems a received wisdom in many policy circles, the world is not yet a flat featureless plain. Even small changes in transport costs – combined with historically low levels of these costs – can strongly affect geography because firms compete globally and their slim profit margins are driven to a large extent by locational advantage. In the end, the debate surrounding the ‘flat world’ is a classical instance of the fallacy consisting in equating ‘low’ with ‘unimportant’.²⁷

²⁷See, e.g., the ‘kaleidoscopic comparative advantage’ debate in international trade (Jagdish Bhagwati, “Why the world is not flat”, 2010; available at <http://www.worldaffairsjournal.org/blog/jagdish-bhagwati/why-world-not-flat>). Last accessed on July 11, 2016.

References

- [1] Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry. 2012. "Measuring the upstreamness of production and trade flows." *American Economic Review Papers and Proceedings* 102(3): 412–416.
- [2] Brown, W. Mark. 2015. "How much thicker is the Canada-U.S. border? The cost of crossing the border by truck in the pre- and post-9/11 eras." *Research in Transportation and Business Management* 16: 50–66.
- [3] Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the United States." *American Economic Review* 103(6): 2121–2168.
- [4] Behrens, Kristian, Brahim Boualam, and Julien Martin. 2016. "The resilience of the Canadian textile industries and clusters to shocks, 2001–2013." CIRANO Project Report #2016-RP05, available at <https://www.cirano.qc.ca/files/publications/2016RP-05.pdf>.
- [5] Behrens, Kristian, and Théophile Bougna. 2015. "An anatomy of the geographical concentration of Canadian manufacturing industries." *Regional Science and Urban Economics* 51(C): 47–69.
- [6] Behrens, Kristian, Théophile Bougna, and W. Mark Brown. 2015. "The world is not yet flat: Transport costs matter!" CEPR Discussion Paper #10356. Center for Economic Policy Research, London, UK.
- [7] Behrens, Kristian, Carl Gaigné, Gianmarco I.P. Ottaviano, and Jacques-François Thisse. 2007. "Countries, regions, and trade: On the welfare impacts of economic intergration." *European Economic Review* 51(5), 1277–1301.
- [8] Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Südekum. 2012. "Spatial frictions." CEPR Discussion Paper #8572, Center for Economic Policy Research, London, UK.
- [9] Behrens, Kristian, and Pierre M. Picard. 2011. "Transportation, freight rates, and economic geography." *Journal of International Economics* 85(2): 280–291.
- [10] Brülhart, Marius (2011). "The spatial effects of trade openness: a survey." *Review of World Economics* 147(1): 59–83.
- [11] Brülhart, Marius, Céline Carrère, and Federico Trionfetti. 2012. "How wages and employment adjust to trade liberalization: Quasi-experimental evidence from Austria." *Journal of International Economics* 86(1): 68–81.

- [12] Chandra, A., Thompson, E. 2000. "Does public infrastructure affect economic activity? Evidence from the rural interstate highway system." *Regional Science and Urban Economics* 30(4): 457–490.
- [13] Combes, Pierre-Philippe, and Laurent Gobillon. 2015. "The empirics of agglomeration economies." In: G. Duranton, J.V. Henderson, and W.C. Strange (eds.), *Handbook of Regional and Urban Economics*, vol. 5. North-Holland: Elsevier, pp.247–348.
- [14] Combes Pierre-Philippe, and Miren Lafourcade. 2005. "Transport costs : Measures, determinants, and regional policy implications for France." *Journal of Economic Geography* 5(3): 319–349.
- [15] Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum. 2014. "The rise of the east and the far east: German labor markets and trade integration." *Journal of the European Economic Association* 12(6): 1643–1675.
- [16] Dumais, Guy, Glenn D. Ellison, and Edward L. Glaeser. 2002. "Geographic concentration as a dynamic process." *Review of Economics and Statistics* 84(2): 193–204.
- [17] Duranton, Gilles, Peter M. Morrow, and Matthew A. Turner. 2014. "Roads and Trade: Evidence from the US." *Review of Economic Studies* 81(2): 681–724.
- [18] Duranton, Gilles, and Henry G. Overman. 2005. "Testing for localisation using micro-geographic data." *Review of Economic Studies* 72(4): 1077–1106.
- [19] Duranton, Gilles, and Diego Puga. 2004. "Micro-foundations of urban agglomeration economies." In: J. Vernon Henderson, and Jacques-François Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4. North-Holland: Elsevier B.V., pp. 2063–2117.
- [20] Ellison, Glenn D., and Edward L. Glaeser. 1999. "The geographic concentration of industry: Does natural advantage explain agglomeration?" *American Economic Review*, 89(2): 311–316.
- [21] Ellison, Glenn D., Edward L. Glaeser, and William R. Kerr. 2010. "What causes industry agglomeration? Evidence from coagglomeration patterns." *American Economic Review* 100(3): 1195–1213.
- [22] Fujita, Masahisa, Krugman, Paul R., and Anthony J. Venables (1999). *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, Cambridge, MA.
- [23] Glaeser, Edward L., and Janet E. Kohlhase. 2004. "Cities, regions and the decline of transport costs." *Papers in Regional Science* 83(1): 197–228.

- [24] Hanson, Gordon H. 1997. "Increasing returns, trade and the regional structure of wages." *Economic Journal* 107(440): 113–133.
- [25] Hanson, Gordon H. 1996. "Economic integration, intraindustry trade, and frontier regions." *Regional Science and Urban Economics* 40(3-5): 941–949.
- [26] Helpman, Elhanan. 1998. "The size of regions." In: D. Pines, E. Sadka, I. Zilcha (eds.), *Topics in Public Economics. Theoretical and Empirical Analysis*, Cambridge University Press, pp. 33–54.
- [27] Henderson, J. Vernon. 1997. "Medium-sized cities." *Regional Science and Urban Economics* 27(6): 583–612.
- [28] Holmes, Thomas J. 1999. "Localization of industry and vertical disintegration." *Review of Economics and Statistics* 81(2): 314–325.
- [29] Holmes, Thomas J., and John J. Stevens. 2014. "An alternative theory of the plant size distribution, with geography and intra- and international trade." *Journal of Political Economy* 122(2): 369–421.
- [30] Holmes, Thomas J. and John J. Stevens (2004). "Spatial distribution of economic activities in North America." In: J.V. Henderson, J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland: Elsevier B.V., pp. 2797–2843.
- [31] Jonkeren, O., Erhan Demirel, Jos van Ommeren, and Piet Rietveld. 2009. "Endogenous transport prices and trade imbalances." *Journal of Economic Geography* 11(3): 509–527.
- [32] Kerr, William R. and Scott D. Kominers. 2015. "Agglomerative forces and cluster shapes." *Review of Economics and Statistics* 97(4): 877–899.
- [33] Kim, Sukkoo. 1995. "Expansion of markets and the geographic distribution of economic activities: The trends in U.S. regional manufacturing structure, 1860-1987." *Quarterly Journal of Economics* 110(4): 881–908.
- [34] Koenig, Pamina, Florian Mayneris, and Sandra Poncet. 2010. "Local export spillovers in France." *European Economic Review* 54(4): 622–641.
- [35] Krugman, Paul R. 1991. "Increasing returns and economic geography." *Journal of Political Economy* 99(3): 483–499.
- [36] Krugman, Paul R., and R. Livas Elizondo. 1996. "Trade policy and the third world metropolis." *Journal of Development Economics* 49(1): 137–150.

- [37] Krugman, Paul R., and Anthony J. Venables (1995). "Globalization and the inequality of nations." *Quarterly Journal of Economics* 110(4): 857–880.
- [38] Lewbel, Arthur. 2012. "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models." *Journal of Business and Economic Statistics* 30(1): 67–80.
- [39] Murata, Yasusada, Ryo Nakajima, Ryosuke Okamoto, and Ryuichi Tamura. 2014. "Localized knowledge spillovers and patent citations: A distance-based approach." *Review of Economics and Statistics* 96(5): 967–985.
- [40] Redding, Stephen J., and Daniel M. Sturm. 2008. "The Costs of Remoteness: Evidence from German Division and Reunification." *American Economic Review* 98(5), 1766–1797.
- [41] Redding, Stephen J., and Anthony J. Venables. 2004. "Economic geography and international inequality." *Journal of International Economics* 62(1), 63–82.
- [42] Redding, Stephen J., and Matthew A. Turner. 2015. "Transportation costs and the spatial organization of economic activity." In: Duranton, Gilles, J. Vernon Henderson, and William C. Strange (eds.) *Handbook of Regional and Urban Economics*, vol. 5. North-Holland: Elsevier B.V., pp. 1339–1398.
- [43] Rosenthal, Stuart S., and William C. Strange. 2004. "Evidence on the nature and sources of agglomeration economies." In: J.V. Henderson, J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland: Elsevier B.V., pp. 2119–2171.
- [44] Rosenthal, Stuart S., and William C. Strange. 2003. "Geography, industrial organization, and agglomeration." *Review of Economics and Statistics* 85(2): 377–393.
- [45] Rosenthal, Stuart S., and William C. Strange. 2001. "The determinants of agglomeration." *Journal of Urban Economics* 50(2): 191–229.
- [46] Storeygard, Adam. 2016. "Farther on down the road: Transport costs, trade, and urban growth." Forthcoming, *Review of Economic Studies*.
- [47] Tanaka, Kiyoyasu, and Kenmei Tsubota. 2016. "Directional imbalance in freight rates: evidence from Japanese inter-prefectural data." Forthcoming, *Journal of Economic Geography*.

Appendix material

This set of Appendices is structured as follows. Appendix A presents a ‘simple’ general equilibrium model that encapsulates the key aspects of our empirical analysis. This model blends several features from the international trade and economic geography literatures and lends

itself readily to numerical analysis. Appendix B documents our data sources and explains the construction of our controls. We explain in particular the construction of our trade exposure and input-output distance controls. Last, Appendix C presents details on the Lewbel (2012) estimator and its implementation.

Appendix A. Model

Consider a setting with one final good industry, one intermediate good industry, two regions r and s within the same country, and a rest-of-the-world (henceforth, row, denoted by R). Each industry produces a continuum of differentiated varieties under monopolistic competition. There are L_r and L_s workers in regions r and s , respectively. The intermediate industry uses labor only. The final industry uses both labor and a CES aggregate of inputs from the intermediate industry. Labor is hired locally, whereas intermediate inputs can be sourced locally, from the other region, or from the row. Labor is immobile between regions. As in Krugman and Venables (1995) it is, however, mobile across industries. The intersectoral labor mobility can lead industries to cluster geographically through regional specialization.²⁸ Trading the intermediates and the final good across regions within the country incurs an iceberg *transport cost* of $t \geq 1$ (paid in units of the good shipped). Trading the intermediates and the final good internationally incurs an iceberg *trade cost* of $\tau \geq 1$ (paid in units of the good shipped). In what follows, we consider that t and τ are independent parameters, but it is easy to rewrite the model with $\tau = \tilde{\tau} \times t$, where $\tilde{\tau}$ is the international portion of trade costs and t the domestic portion (transport costs).

A.1. Final good industry. The unit cost function in region r for producing a variety of the final good is given by $c_r = w_r^\beta \mathbb{P}_r^{1-\beta}$, where $\beta \in (0,1)$ is the labor share, w_r is the wage, and \mathbb{P}_r is the CES price index for intermediates in region r . Hence, the conditional labor demand per unit of final good is given by $\partial c_r / \partial w_r = \beta w_r^{\beta-1} \mathbb{P}_r^{1-\beta}$. Let q_{rr}^F and q_{rs}^F denote the per capita demand for final goods in regions r and s , respectively, for a variety produced in region r . Let q_{Rr}^F and q_{Rs}^F denote those same demands for varieties of the final good produced in the row. Normalizing the variable input requirement to one and assuming that each firm incurs a fixed cost F , the total demand for the composite input (of labor and intermediates) is $F + L_r q_{rr}^F + L_s t q_{rs}^F + L_R \tau q_{Rr}^F$, where L_R denotes the ‘size’ of the demand emanating from the row. Hence, letting N_r stand for the mass of final producers in r , the total labor demand from

²⁸We abstract from interregional labor mobility since this makes the model more involved to simulate. Also, the Duranton-Overman metric that we use to measure the geographic concentration of industries takes the overall distribution of economic activity as the benchmark, i.e., works as if the geographic structure of manufacturing taken as a whole were fixed.

the final goods industry in region r is

$$N_r(F + L_r q_{rr}^F + L_s t q_{rs}^F + L_R \tau q_{rR}^F) \frac{\partial c_r}{\partial w_r}.$$

We assume that preferences are additively separable across the continuum of final good varieties as in Zhelobodko, Kokovin, Parenti, and Thisse (2012). Denote by $u(\cdot)$ the subutility function that is common to all consumers in all regions and for all varieties. The first-order conditions for utility maximization imply that

$$u'(q_{rr}^F) = \lambda_r p_{rr}^F \quad \text{and} \quad u'(q_{sr}^F) = \lambda_r p_{sr}^F,$$

where p_{sr}^F is the price of final goods produced in region s and consumed in region r ; and where λ_r is the multiplier associated with the budget constraint in region r , which is given for each firm because of the continuum assumption. Analogously, for the ROW, we have

$$q_{rR}^F = (u')^{-1}(\lambda_R p_{rR}^F) \quad \text{and} \quad q_{sR}^F = (u')^{-1}(\lambda_R p_{sR}^F).$$

In what follows, we take the ROW multiplier λ_R as exogenously given (small country assumption). The domestic multipliers λ_r and λ_s are, however, solved in the equilibrium of the model.

Given the above inverse demand schedules, final good producers maximize profits

$$\begin{aligned} \Pi_r &= (p_{rr}^F - c_r) L_r q_{rr}^F + (p_{rs}^F - t c_r) L_s q_{rs}^F + (p_{rR}^F - \tau c_r) L_R q_{rR}^F - F c_r \\ &= \left[\frac{u'(q_{rr}^F)}{\lambda_r} - c_r \right] L_r q_{rr}^F + \left[\frac{u'(q_{rs}^F)}{\lambda_s} - t c_r \right] L_s q_{rs}^F + \left[\frac{u'(q_{rR}^F)}{\lambda_R} - \tau c_r \right] L_R q_{rR}^F - F c_r, \end{aligned} \quad (\text{A.1})$$

with respect to the quantities q_{rr}^F , q_{rs}^F , and q_{rR}^F .²⁹ Firms cannot influence the market aggregates λ_r , λ_s , and λ_R . As in Zhelobodko *et al.* (2012), this yields profit-maximizing prices and quantities that satisfy

$$p_{rr}^F = \frac{c_r}{1 - r_u(q_{rr}^F)}, \quad p_{rs}^F = \frac{t c_r}{1 - r_u(q_{rs}^F)}, \quad \text{and} \quad p_{rR}^F = \frac{\tau c_r}{1 - r_u(q_{rR}^F)}, \quad (\text{A.2})$$

where $r_u(q) \equiv -u''(q)q/u'(q)$ is the relative love-of-variety (or relative risk aversion). This term captures how markups change with the competitive environment of the firms. Free entry drives profits to zero, which implies from (A.1) and (A.2) that

$$c_r \frac{r_u(q_{rr}^F)}{1 - r_u(q_{rr}^F)} L_r q_{rr}^F + t c_r \frac{r_u(q_{rs}^F)}{1 - r_u(q_{rs}^F)} L_s q_{rs}^F + \tau c_r \frac{r_u(q_{rR}^F)}{1 - r_u(q_{rR}^F)} L_R q_{rR}^F = F c_r.$$

Last, we need to determine the spending, E_r and E_s , of final good producers on intermediates. Because final goods producers spend a constant share $1 - \beta$ of their total costs on intermediates, it follows that

$$\begin{aligned} E_r &= (1 - \beta) w_r^\beta \mathbb{P}_r^{1-\beta} (F + L_r q_{rr}^F + L_s t q_{rs}^F + L_R \tau q_{rR}^F) \\ E_s &= (1 - \beta) w_s^\beta \mathbb{P}_s^{1-\beta} (F + L_s q_{ss}^F + L_r t q_{sr}^F + L_R \tau q_{sR}^F). \end{aligned} \quad (\text{A.3})$$

This completes the description of the final industry.

²⁹Price and quantity maximization are identical with a continuum of firms (see Vives, 1999).

A.2. Intermediate good industry. Let us next consider the intermediate good producers. Normalizing the unit input requirement to one for simplicity, and assuming that there are fixed costs Fw_r , the profit of an intermediate producer in region r is

$$\pi_r = (p_{rr}^I - w_r)N_r q_{rr}^I + (p_{rs}^I - tw_r)N_s q_{rs}^I + (p_R^I - \tau w_r)Q_R^I - Fw_r, \quad (\text{A.4})$$

where p_{rr}^I , q_{rr}^I , p_{rs}^I , and q_{rs}^I are the prices and quantities sold domestically, and p_R^I and Q_R^I are world prices and demands for intermediates. We take the latter as exogenously given (small country assumption). Intermediate good producers face CES demand functions emanating from the final good producers – recall that intermediates are assembled using a CES aggregator. Hence, the demands for intermediate goods are:

$$q_{rr}^I = \frac{(p_{rr}^I)^{-\sigma} E_r}{\mathbb{P}_r^{1-\sigma}}, \quad q_{rs}^I = \frac{(p_{rs}^I)^{-\sigma} E_s}{\mathbb{P}_s^{1-\sigma}} \quad \text{and} \quad Q_R^I = (p_R^I)^{-\sigma} E_R,$$

where E_r and E_s denote spending by final good producers in regions r and s on intermediates, and where E_R is spending on intermediate inputs from the ROW (taken, again, as exogenously given). With monopolistic competition and isoelastic demands, the profit-maximizing prices are a constant markup over marginal cost:

$$p_{rr}^I = \frac{\sigma}{\sigma-1} w_r, \quad p_{rs}^I = \frac{\sigma}{\sigma-1} tw_r \quad \text{and} \quad p_{rR}^I = \frac{\sigma}{\sigma-1} \tau w_r, \quad (\text{A.5})$$

where p_{rr}^I , p_{sr}^I , and p_{rR}^I are the prices of intermediates produced in regions r and sold locally, to the other region, and to the ROW. Let n_r and n_s denote the endogenously determined masses of intermediate goods producers in regions r and s , respectively. The price aggregate \mathbb{P}_r for intermediates in region r is then given by

$$\mathbb{P}_r = \left(n_r p_{rr}^{1-\sigma} + n_s p_{sr}^{1-\sigma} + n_R p_{rR}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \frac{\sigma}{\sigma-1} \left[n_r w_r^{1-\sigma} + n_s (tw_s)^{1-\sigma} + \tau^{1-\sigma} f_R \right]^{\frac{1}{1-\sigma}}, \quad (\text{A.6})$$

where $f_R \equiv n_R p_{rR}^{1-\sigma}$ denotes the impact of intermediate imports from the ROW on the price index in region r . Note that f_R is an amalgam of the mass of international competitors and productivity (costs) in the ROW. We take it as given. The second equality in (A.6) uses the profit-maximizing prices (A.5).

Substituting (A.5) into (A.4), we have $\pi_r = w_r Q_r / (\sigma - 1) - Fw_r$, where $Q_r = N_r q_{rr}^I + N_s t q_{rs}^I + \tau Q_R^I$ is total output produced by an intermediate producer in r . Hence, zero profits in the intermediate sector require that $Q_r = F(\sigma - 1)$, which implies that each intermediate producer hires a quantity of labor given by $Q_r + F = F\sigma$. Given n_r producers in region r , intermediate labor demand in region r is then given by $n_r F\sigma$.

Last, free entry into the intermediate industry drives profits to zero:

$$\begin{aligned} \pi_r &= \frac{w_r}{\sigma-1} \left[N_r \frac{(p_{rr}^I)^{-\sigma} E_r}{\mathbb{P}_r^{1-\sigma}} + N_s t \frac{(p_{rs}^I)^{-\sigma} E_s}{\mathbb{P}_s^{1-\sigma}} + \tau Q_R^I \right] - Fw_r \\ &= \frac{w_r}{\sigma-1} \left[\frac{N_r E_r}{\mathbb{P}_r^{1-\sigma}} \left(\frac{\sigma}{\sigma-1} w_r \right)^{-\sigma} + t \frac{N_s E_s}{\mathbb{P}_s^{1-\sigma}} \left(\frac{\sigma}{\sigma-1} tw_r \right)^{-\sigma} + \tau E_R \left(\frac{\sigma}{\sigma-1} \tau w_r \right)^{-\sigma} \right] - Fw_r = 0. \end{aligned}$$

This completes the description of the intermediate industry.

A.3. Equilibrium. To solve the equilibrium of the model, we need to specify the subutility function $u(\cdot)$ for consumers' preferences. Following previous work by Behrens and Murata (2007), we let $u(q) = 1 - e^{-\alpha q}$, where $\alpha > 0$ parametrizes firms' market power. The relative love-of-variety is then given by $r_u(q) = \alpha q$, which is increasing in quantity. This implies the existence of pro-competitive effects, both between regions and between each region and the row. The equilibrium conditions can be derived as follows. Using the definitions of expenditure on intermediates and the price indices, given by (A.3) and (A.6), as well as the unit cost function $c_r = w_r^\beta \mathbb{P}_r^{1-\beta}$, an equilibrium in our model satisfies the following five sets of equilibrium conditions:

(i) labor market clearing in each region:

$$\begin{aligned} L_r &= N_r(F + L_r q_{rr}^F + L_s t q_{rs}^F + L_R \tau q_{rR}^F) \beta w_r^{\beta-1} \mathbb{P}_r^{1-\beta} + n_r F \sigma \\ L_s &= N_s(F + L_s q_{ss}^F + L_r t q_{sr}^F + L_R \tau q_{sR}^F) \beta w_s^{\beta-1} \mathbb{P}_s^{1-\beta} + n_s F \sigma \end{aligned}$$

(ii) zero profits for final goods producers:

$$\begin{aligned} F c_r &= c_r \frac{r_u(q_{rr}^F)}{1 - r_u(q_{rr}^F)} L_r q_{rr}^F + t c_r \frac{r_u(q_{rs}^F)}{1 - r_u(q_{rs}^F)} L_s q_{rs}^F + \tau c_r \frac{r_u(q_{rR}^F)}{1 - r_u(q_{rR}^F)} L_R q_{rR}^F \\ F c_s &= c_s \frac{r_u(q_{ss}^F)}{1 - r_u(q_{ss}^F)} L_s q_{ss}^F + t c_s \frac{r_u(q_{sr}^F)}{1 - r_u(q_{sr}^F)} L_r q_{sr}^F + \tau c_s \frac{r_u(q_{sR}^F)}{1 - r_u(q_{sR}^F)} L_R q_{sR}^F \end{aligned}$$

(iii) zero profits for intermediate goods producers:

$$\begin{aligned} F(\sigma - 1) &= \left(\frac{\sigma}{\sigma - 1} w_r \right)^{-\sigma} \left(\frac{N_r E_r}{\mathbb{P}_r^{1-\sigma}} + t^{1-\sigma} \frac{N_s E_s}{\mathbb{P}_s^{1-\sigma}} + \tau^{1-\sigma} E_R \right) \\ F(\sigma - 1) &= \left(\frac{\sigma}{\sigma - 1} w_s \right)^{-\sigma} \left(\frac{N_s E_s}{\mathbb{P}_s^{1-\sigma}} + t^{1-\sigma} \frac{N_r E_r}{\mathbb{P}_r^{1-\sigma}} + \tau^{1-\sigma} E_R \right) \end{aligned}$$

(iv) regional budget constraints:

$$\begin{aligned} w_r &= N_r \frac{c_r}{1 - r_u(q_{rr}^F)} q_{rr} + N_s \frac{t c_s}{1 - r_u(q_{sr}^F)} q_{sr} + N_R \frac{\tau c_R}{1 - r_u(q_{rR}^F)} q_{rR} \\ w_s &= N_s \frac{c_s}{1 - r_u(q_{ss}^F)} q_{ss} + N_r \frac{t c_r}{1 - r_u(q_{rs}^F)} q_{rs} + N_R \frac{\tau c_R}{1 - r_u(q_{sR}^F)} q_{sR} \end{aligned}$$

(v) first-order conditions for utility maximization:

$$\begin{aligned} \frac{u'(q_{rr}^F)}{u'(q_{sr}^F)} &= \frac{\frac{c_r}{1 - r_u(q_{rr}^F)}}{\frac{t c_s}{1 - r_u(q_{sr}^F)}}, & \frac{u'(q_{ss}^F)}{u'(q_{rs}^F)} &= \frac{\frac{c_s}{1 - r_u(q_{ss}^F)}}{\frac{t c_r}{1 - r_u(q_{rs}^F)}} \\ \frac{u'(q_{rr}^F)}{u'(q_{rR}^F)} &= \frac{\frac{c_r}{1 - r_u(q_{rr}^F)}}{\frac{\tau c_R}{1 - r_u(q_{rR}^F)}}, & \frac{u'(q_{ss}^F)}{u'(q_{sR}^F)} &= \frac{\frac{c_s}{1 - r_u(q_{ss}^F)}}{\frac{\tau c_R}{1 - r_u(q_{sR}^F)}} \\ u'(q_{rR}^F) &= \lambda_R p_{rR}^F, & u'(q_{sR}^F) &= \lambda_R p_{sR}^F \end{aligned}$$

Letting $w_r \equiv 1$ by choice of numeraire, we can solve the set of 14 equations (i)–(v) for the 14 unknowns: $q_{rr}^F, q_{sr}^F, q_{ss}^F, q_{rs}^F, q_{rR}^F, q_{sR}^F, q_{Rr}^F, q_{Rs}^F, N_r, N_s, n_r, n_s, w_s$, and N_R . By Walras’ law, the last equilibrium condition – the trade balance condition, which we do not spell out here – then also holds. One comment is in order. As can be seen from above, we solve for N_R so that the budget constraints in each region hold. Without this condition, regions can run a budget surplus or deficit, because we do not close the model fully by specifying all equilibrium conditions for the ROW. We thus impose trade balance by assumption – balancing regional budget constraints. Alternatively, we could pin down wages for some given deficit or surplus.

Appendix B. Data sources and construction of controls

This appendix provides details on the data used and the data sources. A summary of the key variables and the associated descriptive statistics are given in Table 3 in the main text and in Table 8 below.

B.1. Data sources.

Plant-level data and industries. Our analysis is based on the Annual Survey of Manufacturers (ASM) Longitudinal Microdata file. This data cover the years from 1990 to 2010. Our focus is on manufacturing plants only. For every plant we have information on: its primary 6-digit NAICS code (the codes are consistent over the 20 year period); its year of establishment; its total employment; whether or not it is an exporter in selected years; its sales; the number of non-production and production workers; and its 6-digit postal code. The latter, in combination with the Postal Code Conversion files (PCCF), allow us to effectively geo-locate the plants by associating them with the geographical coordinate of their postal code centroids.

The survey frame of the ASM has evolved over time. Early in the period, it was relatively stable with, on average, about 32,000 plants per sample year. The sample of plants was restricted to those with total employment (production plus non-production workers) above zero, and plants must have sales in excess of \$30,000. Also, aggregate records were excluded. These records represent multiple (typically small) plants without latitudes and longitudes. In 2000, however, the number of plants in the survey increased substantially as the ASM moved from its own frame to Statistics Canada’s centralized Business Register, increasing the sample to an average of 53,000 plants. In 2004, however, the number of plants in the frame was once again restricted, with many of the small plants once again excluded, or included in aggregate records. With this in place, the sample returned to near previous levels, averaging about 33,000 plants between 2004 and 2009. The expanded survey scope in the early 2000s had little effect on trends in the CDFs, but there was an effect on the number of industries found to be localized or dispersed (see Table 10 in the Supplementary Appendix S.4). Our econometric analysis deals with the change in the sample frame through the inclusion of year fixed effects.

Table 8: Summary statistics for control variables and instruments.

Variable names and descriptions	Industry detail	Mean	Standard deviation		
			Overall	Between	Within
Share of industry imports from Asian countries (excluding OECD members)	NAICS6	0.120	0.183	0.172	0.062
Share of imports from OECD member countries (excluding U.S. and Mexico)	NAICS6	0.157	0.141	0.131	0.053
Share of imports from NAFTA countries (U.S. and Mexico)	NAICS6	0.662	0.273	0.263	0.074
Share of industry exports to Asian countries (excluding OECD members)	NAICS6	0.029	0.058	0.047	0.035
Share of exports to OECD member countries (excluding U.S. and Mexico)	NAICS6	0.086	0.101	0.085	0.054
Share of exports to NAFTA countries (U.S. and Mexico)	NAICS6	0.833	0.198	0.184	0.073
Industry mean of the avg. distance to a dollar of inputs from the 5 nearest plants (km)	NAICS6	241	111	95	57
Industry mean of the avg. distance to ship a dollar of output to the 5 nearest plants (km)	NAICS6	243	124	103	69.1
Minimum average distance to 5×257 closest plants	NAICS6	64.7	44.5	42.1	14.4
Share of inputs from natural resource-based industries	<i>L</i> -level	0.113	0.171	0.170	0.026
Sectoral energy inputs as a share of total sector output	<i>L</i> -level	0.032	0.046	0.045	0.013
Total industry employment	NAICS6	7,038	8,060	7,858.11	1,856
Herfindahl index of enterprise-level employment concentration	NAICS6	0.101	0.097	0.092	0.032
Mean plant size	NAICS6	73.7	145	139	41.8
Share of plants controlled by multi-plant firms	NAICS6	0.212	0.193	0.183	0.061
Share of foreign controlled plants	NAICS6	0.153	0.157	0.146	0.059
Share of hours worked by all workers with post-secondary education	NAICS6	0.401	0.082	0.071	0.041
Intramural research and development expenditures as a share of industry sales	<i>L</i> -level	0.011	0.039	0.027	0.005
Minimum distance from major container ports (km)	NAICS6	414	110	103	48
Eastern share of plants	NAICS6	0.749	0.133	0.124	0.047
Ad valorem trucking costs us (instrument)	NAICS6	0.038	0.034	0.034	0.006

Notes: All descriptive statistics are based on the sample we use in the regression analysis, which includes 4,369 observations covering 257 industries and 17 years. The standard deviation is decomposed into between and within components, which measure the cross sectional and the time series variation, respectively. Some industry-level data are available at the *L*-level only, which is the finest level of data for public release in Canada (between the NAICS 3- and 4-digit levels of aggregation).

We also use the ASM to construct controls for the labor market variables, for some natural advantage proxies, and for industry ownership structure variables that we include in the regressions. All variables are constructed by aggregating plant-level data to the industry level.

***L*-level input-output tables.** We use the *L*-level national input-output tables from Statistics Canada at buyers' prices to construct our plant-level proxies for the importance of input and output linkages. These tables – which constitute the finest sectoral public release – feature 42 sectors that are somewhere in between the NAICS 3- and NAICS 4-digit levels. We break them down to the 6-digit level based on industries' weights in terms of sales. For each industry, i , we allocate total inputs purchased or outputs sold in the *L*-level matrix to the corresponding NAICS 6-digit sectors. We allocate total sales to each subsector in proportion to that sector's sales in the total sales to obtain a 257×257 matrix of NAICS 6-digit inputs and outputs, which we use to construct the linkages.³⁰ From that table, we compute the share α_{ij} that sector i sells

³⁰Because of confidentiality reasons, we do not use the finer *W*-level matrices since this would make disclosure of results more problematic. However, the tests we ran using those matrices yield very similar results to the ones we report in this paper. Using the *L*-level matrix provides smoother series of input-output linkages than those obtained using the confidential *W*-level national input-output tables (which are directly in the 257×257 industries format).

to sector j . We also compute the share β_{ij} that sector i buys from sector j . We scale shares so that they sum to one.

KLEMS database. This database, which covers the period from 1961 to 2008, contains various industry-level informations useful for constructing proxies for natural advantage (e.g., energy intensity, water usage, etc.) or an industry’s labor force composition (educated vs non-educated workers).

Trucking micro-data. The trucking micro-data comes from Statistics Canada’s Trucking Commodity Origin-Destination Survey and from the ‘experiment export trade file’ produced in 2008 (see Brown, 2015; and Brown and Anderson, 2015, for details). Section 3.2 provides details on the methodology used to estimate ad valorem trucking rates by industry and year.

Geographical data. To geolocate firms, we use latitude and longitude data of postal code centroids obtained from Statistics Canada’s Postal Code Conversion files (PCCF). These files associate each postal code with different Standard Geographical Classifications (SGC) that are used for reporting census data in Canada. We match firm-level postal code information with geographical coordinates from the PCCF. Postal codes are less fine grained in rural areas, but the kernel smoothing of the K -density takes care of these variations (see Duranton and Overman, 2005).

Trade data. The industry-level trade data come from Innovation, Science and Economic Development Canada and cover the years 1992 to 2009. The dataset reports imports and exports at the NAICS 6-digit level by province and by country of origin and destination. We aggregate the data across provinces and compute the shares of exports and imports that go to or originate from a set of country groups: Asian countries (excluding OECD), OECD countries (excluding NAFTA), and NAFTA countries. Since the trade data is only available from 1992 on, whereas the KLEMS data is only available until 2008, we restrict our sample to the 1992–2008 period in all estimations to maintain comparability of results.

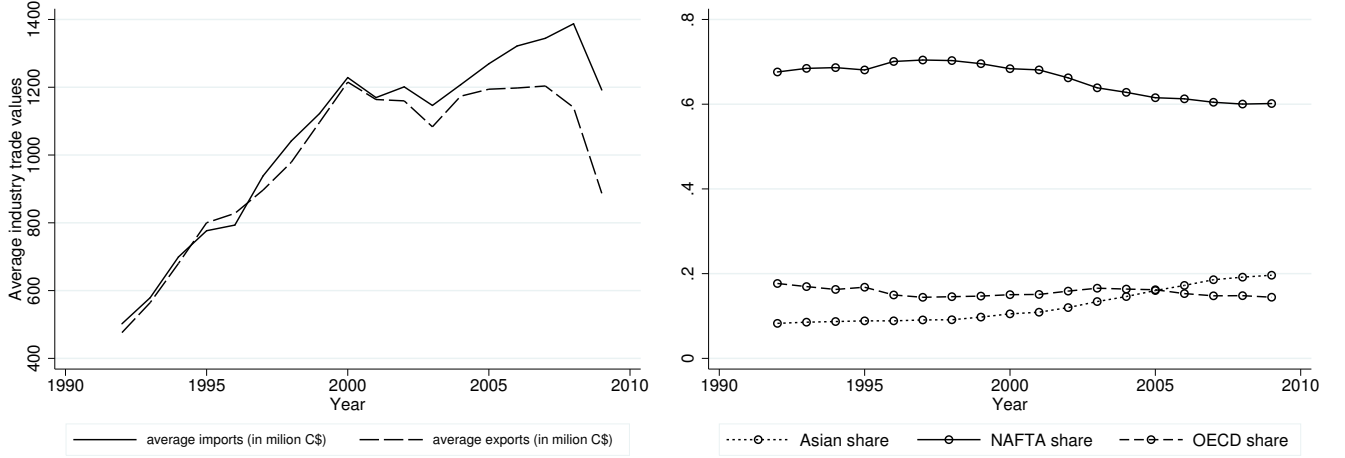
US price indices. We use detailed year-by-year NAICS 6-digit price indices from the NBER-CES Manufacturing Productivity Databas (<http://nber.org/data/nberces5809.html>) to construct instruments for Canadian industry-level transportation costs. Additional methodological details are provided in Section 4.2.

B.2. International trade exposure.

We use detailed yearly data on imports and exports by industry and country of origin and destination to control for industries’ import and export exposure (the ratio of industry imports

or exports to industry sales). To disentangle the different effects that depend on whether trade is in intermediates or final goods (on which we have unfortunately no information in our data), and on whether trade is ‘North-North’ or ‘North-South’, we break these measures down by countries of origin: low-cost Asian countries; OECD countries; and NAFTA countries.

Figure 3: Changes in import- and export trade values (left), and import shares (right).



The left panel of Figure 3 depicts the changes in the average import and export values by industry over our study period. The right panel provides a snapshot of how import and export shares change across broad groups of trading partners. As one can see, the importance of international trade has increased – up to the trade collapse starting 2008 – and there has been an increasing re-orientation of trade towards Asian countries (especially for imports).

Endogeneity concerns. The geographic concentration of plants increases productivity and, therefore, may increase the propensity of an industry to export or to import. For example, the agglomeration of an industry may reduce prices, which makes import penetration harder. In that case, the dispersion of an industry may be associated with increasing imports since productivity falls. Also, the agglomeration of an industry may be associated with rising exports due to productivity gains – although the productivity gains reduce unit export values, the total value of exports may increase. Since external instruments for import and export exposure of industries are difficult to find, we deal with potential endogeneity issues of trade exposure using the Lewbel (2012) estimator that relies on internal instruments.

B.3. Input-output distances.

We construct novel proxies for distance to inputs and outputs that make use of the microgeographic nature of our data. Consider a plant ℓ active in industry $\Omega(\ell)$. Let Ω denote the set of industries and Ω_i the set of plants in industry i . Let $k_i(r, \ell)$ denote the r -th closest industry- i

plant to plant ℓ . Our micro-geographic measures of input- and output linkages are constructed as weighted averages as follows:

$$\mathcal{I}\text{dist}(\ell) = \sum_{i \in \Omega \setminus \Omega(\ell)} \omega_{\Omega(\ell),i}^{\text{in}} \times \frac{1}{N} \sum_{j=1}^N d(\ell, k_i(j, \ell)), \quad (\text{B.1})$$

for inputs, and

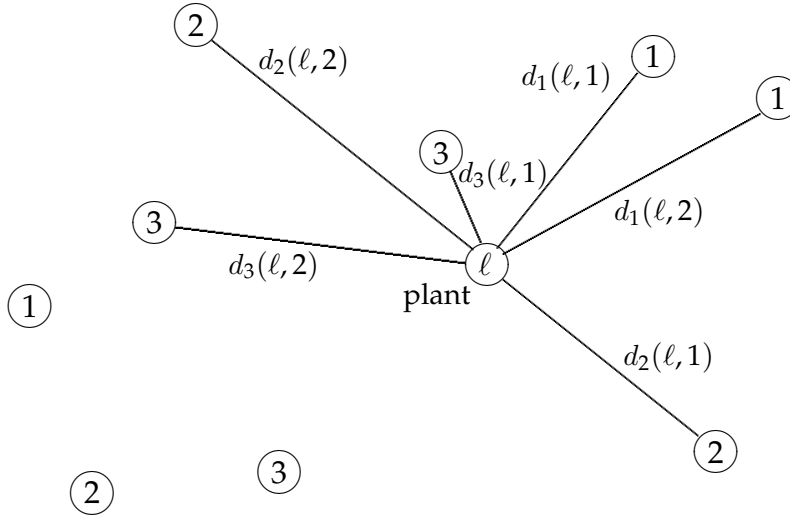
$$\mathcal{O}\text{dist}(\ell) = \sum_{i \in \Omega \setminus \Omega(\ell)} \omega_{\Omega(\ell),i}^{\text{out}} \times \frac{1}{N} \sum_{j=1}^N d(\ell, k_i(j, \ell)), \quad (\text{B.2})$$

for outputs, where $d(\cdot, \cdot)$ is the great circle distance between the plants' postal code centroids, and where $\omega_{\Omega(\ell),i}^{\text{in}}$ and $\omega_{\Omega(\ell),i}^{\text{out}}$ are sectoral input- and output shares. The latter are given by

$$\omega_{\Omega(\ell),i}^{\text{in}} \equiv \alpha_{\Omega(\ell),i} \quad \text{and} \quad \omega_{\Omega(\ell),i}^{\text{out}} \equiv \beta_{\Omega(\ell),i}, \quad (\text{B.3})$$

where α and β are the input and output shares computed using the L -level input-output tables (see Appendix B.1 for details). We exclude within-sector transactions where $\Omega(\ell) = i$. Figure 4 illustrates the construction of (B.1) and (B.2) for the case where $N = 2$ and with three industries.

Figure 4: Constructing input-output distances and ‘minimum distance’ measures.

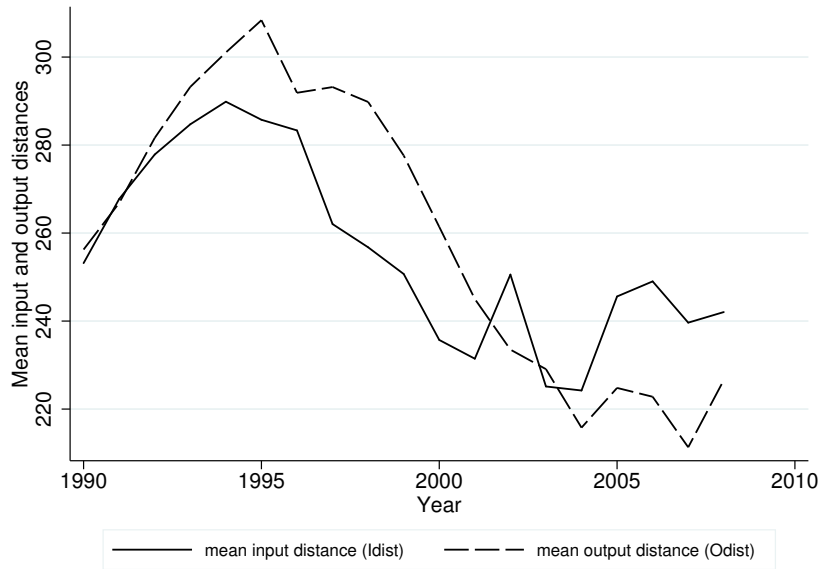


Observe that since $\sum_i \omega_{\Omega(\ell),i}^{\text{in}} = \sum_i \omega_{\Omega(\ell),i}^{\text{out}} = 1$ by construction, we can interpret $\mathcal{I}\text{dist}(\ell)$ as the minimum average distance of plant ℓ to a dollar of inputs from its N closest manufacturing suppliers. Analogously, $\mathcal{O}\text{dist}(\ell)$ is the minimum average distance plant ℓ has to ship a dollar of outputs to its N closest (industrial) customers.³¹ The larger are $\mathcal{I}\text{dist}(\ell)$ or $\mathcal{O}\text{dist}(\ell)$, the

³¹We have no micro-geographic information on final demand and thus cannot include it in our output linkage measures. Using a population-weighted market potential measure as a proxy is infeasible because of the very strong persistence through time. However, our industry fixed effects are likely to control for slow-changing final demand due to changes in the population distribution.

worse are plant ℓ 's input or output linkages – it is, on average, further away from a dollar of intermediate inputs or a dollar of demand emanating from the other industries. Note that our input and output linkages make use of *plant-level* location information, but only of *national* input and output shares. The latter is due to the fact that we do not directly observe input-output linkages at the plant level. Yet, given this, our procedure has the advantage to sidestep obvious problems of endogeneity of those plant-level linkages. Furthermore, our input-output measures are computed across all industries except the one the plant belongs to. Thus, our measures capture finely the whole cross-industry location patterns, but do not pick up industrial localization of the sector itself since it is excluded from the computation (see, however, the caveat in footnote 24). This is important to not confound input-output linkages with other drivers of geographical concentration.

Figure 5: Changes in average input-output distances, 1990–2009.



We compute (B.1) and (B.2) for all years and for all plants, using the $N = 3, 5, 7, 10$ nearest plants in each industry. We then average them across plants in each industry i and each year to get an industry-year specific measure of both input and output distances:

$$\mathcal{O}dist_i = \frac{1}{|\Omega_i|} \sum_{\ell \in \Omega_i} \mathcal{O}dist(\ell) \quad \text{and} \quad \mathcal{I}dist_i = \frac{1}{|\Omega_i|} \sum_{\ell \in \Omega_i} \mathcal{I}dist(\ell), \quad (\text{B.4})$$

where $|\Omega_i|$ denotes the number of plants in industry i . As expected, (B.1) and (B.2) are strongly correlated. Yet, despite that correlation we can include them simultaneously into our regressions and still identify their effect on industrial localization.

Figure 5 depicts the time-series changes in the (unweighted) average input and output measures across all industries. As one can see, in 2000 for example, plants were on average located about 235 kilometers from a dollar of inputs, and had to ship a dollar of their output

on average over a distance of 260 kilometers. Note that time-series changes in the input- and output-distance measures may reflect three things: (i) entry or exit of potential suppliers; (ii) changes in the geographic location of input suppliers and/or clients; and (iii) changes in the input-output coefficients, i.e., the technological relationships. We cannot dissociate the sources (i) and (ii) in our analysis, but entry and exit are more important than relocation when looking at plant-level data.

As can be seen from Figure 5, average input and output distances have fallen over the 1990–2009 period in Canada, from about 260 kilometers to about 240 kilometers. One may wonder how this result is compatible with our finding that industries have geographically dispersed, as documented in Section 3.1. To understand that result, one needs to keep in mind that the geographic dispersion we document in Section 3.1 is for *within-industry concentration*, whereas the measures of input-output distances are for *between-industry concentration*. Starting from a situation where industries are spatially segregated would yield a large value of within-industry geographic concentration, and large between-industry distances. As industries progressively disperse, the within-industry measure falls, whereas the between-industry distance can fall too if there is more ‘mixing’ of industries. In a nutshell, if there is less localization and more mixing between industries, the geographic concentration of industries would fall, but their distance to input suppliers and clients can decrease too. Hence, the two findings are not incompatible.

Note, finally, that one potential problem with the measures (B.1) and (B.2) is that they are mechanically smaller in denser areas. To control for this fact, we also compute a ‘minimum distance measure’, i.e., the distance of plant ℓ from the $M = N \times 257$ closest plants, regardless of their industry. Including that measure into our regressions then controls for the overall plant density in a location, which implies that our input-output linkage measures pick up the effect of being closer to a dollar of inputs or outputs conditional on the overall density of the area the plant is located in. Formally, we compute for each plant ℓ the following measure:

$$\mathcal{M}\text{dist}(\ell) = \frac{1}{M} \sum_{j=1}^M d(\ell, k_{\setminus \Omega(\ell)}(j, \ell)), \quad (\text{B.5})$$

where $d(\ell, k_{\setminus \Omega(\ell)}(j, \ell))$ denotes the distance to the j th closest plant in *any industry but* $\Omega(\ell)$. We then average this measure across all plants in the same industry as before and include it as an additional control into the regressions.

Endogeneity concerns. Our measures of input-output linkages are, by construction, reasonably exogenous to the spatial structure of a specific industry. First, observe that we compute those measures using national input-output shares instead of plant-level input-output shares. Hence, we do not pick up spuriously large values for inputs or outputs – due to substitution effects – when plants are located in close proximity to plants in related industries. Second, we exclude the own industry from the computation, so that the measures only pick up cross-

industry links and not the geographic concentration of the industry itself (which is on the left-hand side of our regressions).³² Last, for each plant, the input and output distance is computed using *all other 256 industries in Canadian manufacturing*. For the geographic concentration of one industry to drive the input-output linkage measure, that industry would need to substantially affect the whole location patterns of most other related industries, which strikes us as fairly unlikely (though we cannot completely rule out this possibility). Although the input- and output-measures should be reasonably exogenous, we deal with potential endogeneity issues of input-output links using the Lewbel (2012) estimator that relies on internal instruments. External instruments are hard to find for these measures.

Appendix C. Applying the Lewbel (2012) method.

To apply the Lewbel (2012) procedure, we need to verify two conditions: heteroscedasticity and correlation. First, we regress the potentially endogeneous variables (input and output distances, trade shares, and trucking costs residuals) on all other exogeneous variables of the model. We then predict the residuals of that regression and run a standard heteroscedasticity test. We need to reject the homoscedasticity assumption for the Lewbel method to be applicable. In our case, we strongly reject the null hypothesis of homoscedasticity for all series of residuals (the p -value is zero in all tests). Second, we take the square of the predict residuals from the foregoing regression, and check the correlation between the dependent variable of the regression (input distances, or output distances, or the different trade shares, or trucking costs residuals) and those squared residuals. The correlation needs to be ‘strong’ and statistically strongly significant for the instruments to work properly. In our case, this condition holds true for transport costs, the input and output distances, and for all import shares: the correlation of the squared residuals with the variable itself is significant at 1% in all cases. It is 0.067 for transportation costs, -0.081 for input distances, -0.089 for output distances, 0.130 for the Asian share of imports, and -0.079 for the NAFTA share of imports. We find no statistically significant correlation for the export shares.

Since the two conditions (heteroscedasticity of the residuals and correlation of the squared residuals with the variable) are met in our case, we can apply the Lewbel estimator. Since fixed effects cannot be included in the estimation (see `ivreg2h` in Stata), we de-mean all variables by industry first. The exogeneous variables are partialled-out for the Lewbel estimator and so their coefficients are not reported. Since we have an exogeneous instrument for transportation costs, we apply the Lewbel estimator only to deal with potential endogeneity concerns of trade shares and input-output distances.

³²Our dataset has the stand problem of reporting only a plant’s primary sector of activity. Hence, it is possible that a plant operates in multiple sectors, so that our measures still partly pick up own-industry location patters. There is not much we can do about this. We experimented with measures where we exclude all plants within the same 4-digit industry, and the results do not change qualitatively.

References

- [1] Behrens, Kristian, and Yasusada Murata. 2007. "General equilibrium models of monopolistic competition: A new approach." *Journal of Economic Theory* 136(1): 776–787.
- [2] Brown, W. Mark, and William P. Anderson. 2015. "How thick is the border: the relative cost of Canadian domestic and cross-border truck-borne trade, 2004-2009." *Journal of Transportation Geography* 42: 10–21.
- [3] Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and Jacques-François Thisse. 2012. "Monopolistic competition: Beyond the constant elasticity of substitution." *Econometrica* 80(6): 2765–2784.

Supplementary Online Appendix material

This set of supplementary appendices is structured as follows. Appendix S.1 describes the Duranton and Overman (2005) continuous spatial approach to measuring the geographic concentration of industries. Appendix S.2 provides additional descriptive evidence on the geographic concentration of industries, while Appendix S.3 provides additional descriptive evidence on transport costs. Appendix S.4 contains additional tables and summarizes robustness checks.

Appendix S.1. Measuring geographic concentration.

Following Duranton and Overman (2005), hereafter DO, the estimator of the kernel density (probability density function or PDF) of the distribution of bilateral distances between plants at some distance d , is given by:

$$\hat{K}(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=i+1}^n f\left(\frac{d - d_{ij}}{h}\right), \quad (\text{S.1})$$

where h is Silverman's optimal bandwidth and f is a Gaussian kernel function. The distance d_{ij} (in kilometers) between plants i and j is computed using the Great Circle distance formula:

$$d_{ij} = 6378.39 \cdot \text{acos}[\cos(|\text{lon}_i - \text{lon}_j|) \cos(\text{lat}_i) \cos(\text{lat}_j) + \sin(\text{lat}_i) \sin(\text{lat}_j)]. \quad (\text{S.2})$$

Alternatively, rather than using plant counts as the unit of observation in (S.1), we can characterize the geographic concentration of employment or sales at the industry level. This can be accommodated by adding weights to (S.1):

$$\hat{K}_W(d) = \frac{1}{h \sum_{i=1}^{n-1} \sum_{j=i+1}^n (e_i + e_j)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (e_i + e_j) f\left(\frac{d - d_{ij}}{h}\right), \quad (\text{S.3})$$

where e_i and e_j are the employment or sales levels of plants i and j , respectively.³³ The weighted K -density thus describes the distribution of bilateral distances between plants in a given industry, weighted by either employees or sales, whereas the unweighted K -density describes the distribution of bilateral distances between plants in that industry. When required, as in Table 10, we follow Duranton and Overman (2005) and implement a Monte Carlo approach for measuring the statistical significance of localization of industries.

To construct the K -densities, we need to fix a cutoff distance. Following Behrens and Bougna (2015), we choose a cutoff distance of 800 kilometers for computing the K -densities.

³³Contrary to Duranton and Overman (2005), who use a multiplicative weighting scheme, we use an additive one. The additive scheme gives less weight to pairs of large plants and more weight to pairs of smaller plants than the multiplicative scheme does (see Behrens and Bougna, 2015). Using a multiplicative scheme would imply that our results may be too strongly driven by a few very large firms in a given industry.

The interactions across ‘neighboring cities’ mostly fall into that range in Canada. In particular, a cutoff distance of 800 kilometers includes interactions within the ‘western cluster’ (Calgary, AB; Edmonton, AB; Saskatoon, SK; and Regina, SK); the ‘plains cluster’ (Winnipeg, MB; Regina, SK; Thunder Bay, ON); the ‘central cluster’ (Toronto, ON; Montréal, QC; Ottawa, ON; and Québec, QC); and the ‘Atlantic cluster’ (Halifax, NS; Fredericton, NB; and Charlottetown, PE). Setting the cutoff distance to 800 kilometers allows us to account for geographic concentration at both very small spatial scales, but also at larger interregional scales for which market-mediated input-output and demand linkages, as well as market size, might matter much more.

While the K -density PDF provides a snapshot of geographic concentration at every distance d , and while it allows for statistical testing, it is not well suited to capture globally the location patterns of industries up to some distance d . This can, however, be achieved by using the K -density cumulative distribution (CDF) up to distance d . In all our econometric estimations, we use as dependent variable the CDF of the K -densities. Those are given by:

$$\text{CDF}(d) = \sum_{\delta=1}^d \hat{K}(\delta) \quad \text{and} \quad \text{CDF}_W(d) = \sum_{\delta=1}^d \hat{K}_W(\delta), \quad (\text{S.4})$$

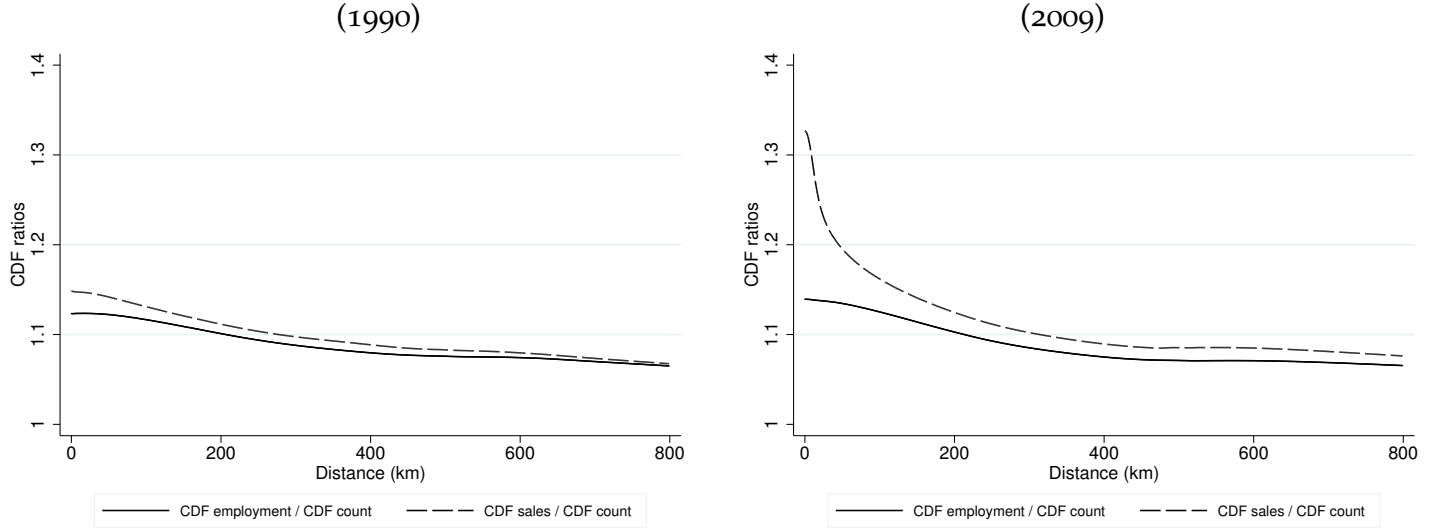
where the former is the unweighted (plant-count based) K -density and the latter is the employment (or sales) weighted K -density.

Appendix S.2. Additional tables and results for geographic concentration.

Table 2 in the main text summarizes the industry-average K -densities across years and for different weighting schemes. Comparing the unweighted to the employment- or sales-weighted K -densities reveals some interesting patterns. As can be seen from Figure 6, industries are on average always more concentrated in terms of employment than in terms of plant counts, and even more concentrated in terms of sales than in terms of employment. This is a manifestation of agglomeration economies, and it is consistent with the findings of Holmes and Stevens (2014) and others that more localized plants tend to be larger and more productive than less localized plants. Note that the ratios are increasing until about 2004, and slightly decreasing afterwards. In 2009, within 50 kilometers distance, the concentration of employment exceeds that of plant counts by about 13%, whereas the concentration of sales exceeds that of plant counts by about 20%. Note further that the de-concentration trend that we documented for the plant-count based measures also affects the employment-weighted and the sales-weighted measures of geographic concentration (see Table 2). Yet, as can be seen from Figure 6, although industries have in general become more geographically dispersed according to all three measures, the size of plant pairs in close proximity has tended to increase in relative terms regardless of whether size is measured by employment or by sales. Put differently, the process of dispersion is less pronounced when geographic concentration is measured by

either employment or sales, thus suggesting that smaller plants drive a substantial part of the dispersion process, either through entry and exit or through relocation. The additional results that we provide in Table 13 are consistent with that finding.

Figure 6: Ratios of mean employment- and sales-based CDFs to count-based CDF by distance.



We provide a number of additional results on the geographic concentration of industries in two tables. Table 9 below provides the (unweighted) K -density CDFs in 1990, 1999, and 2009 for the geographically most strongly concentrated industries in Canada. As can be seen, textile and clothing related industries rank very high in that table, which thus explains why we run robustness checks later to exclude them in Table 12. Table 10 summarizes the year-on-year location patterns of industries based on the formal significance test of Duranton and Overman (2005) that we have described in Appendix S.1. As can be seen, the number of significantly localized industries has fallen over time, whereas the number of industries that display location patterns that are not significantly different from random ones has increased.

Appendix S.3. Additional tables and results for transport costs.

Table 11 provides the average of the top- and the bottom-ten ad valorem trucking costs for each year from 1992 to 2008. The average top-ten rates are around 12.2% to 14.3%, whereas the average bottom-ten rates are around 0.3% to 0.4%. The ten industries with the highest ad valorem rates in 2008 are: Gypsum Product Manufacturing (NAICS 327420), Cement Manufacturing (NAICS 327310), Lime Manufacturing (NAICS 327410), Abrasive Product Manufacturing (NAICS 327910), All Other Non-Metallic Mineral Product Manufacturing (NAICS 327990), Sawmills (except Shingle and Shake Mills) (NAICS 321111), Alkali and Chlorine Manufacturing (NAICS 325181), All Other Miscellaneous Wood Product Manufacturing (NAICS 312999), Breweries (NAICS 312120), and Ready-Mix Concrete Manufacturing (NAICS 327320). The ten

Table 9: Ten most localized NAICS 6-digit industries (based on plant counts).

NAICS	Industry description	CDF
1990		
315231	Women's and Girls' Cut and Sew Lingerie, Loungewear and Nightwear Manufacturing	0.62
315233	Women's and Girls' Cut and Sew Dress Manufacturing	0.55
313240	Knit Fabric Mills	0.53
315292	Fur and Leather Clothing Manufacturing	0.42
315291	Infants' Cut and Sew Clothing Manufacturing	0.32
315210	Cut and Sew Clothing Contracting	0.30
337214	Office Furniture (except Wood) Manufacturing	0.21
332720	Turned Product and Screw, Nut and Bolt Manufacturing	0.21
313110	Fibre, Yarn and Thread Mills	0.19
333511	Industrial Mould Manufacturing	0.18
1999		
315231	Women's and Girls' Cut and Sew Lingerie, Loungewear and Nightwear Manufacturing	0.63
313240	Knit Fabric Mills	0.47
315210	Cut and Sew Clothing Contracting	0.22
333220	Rubber and Plastics Industry Machinery Manufacturing	0.20
336370	Motor Vehicle Metal Stamping	0.18
332720	Turned Product and Screw, Nut and Bolt Manufacturing	0.18
336330	Motor Vehicle Steering and Suspension Components (except Spring) Manufacturing	0.17
333519	Other Metalworking Machinery Manufacturing	0.16
337214	Office Furniture (except Wood) Manufacturing	0.15
315291	Infants' Cut and Sew Clothing Manufacturing	0.14
2009		
315231	Women's and Girls' Cut and Sew Lingerie, Loungewear and Nightwear Manufacturing	0.61
322299	All Other Converted Paper Product Manufacturing	0.29
337214	Office Furniture (except Wood) Manufacturing	0.17
336370	Motor Vehicle Metal Stamping	0.17
332720	Turned Product and Screw, Nut and Bolt Manufacturing	0.16
337215	Showcase, Partition, Shelving and Locker Manufacturing	0.15
321112	Shingle and Shake Mills	0.14
331420	Copper Rolling, Drawing, Extruding and Alloying	0.13
336360	Motor Vehicle Seating and Interior Trim Manufacturing	0.13
315110	Hosiery and Sock Mills	0.13

Notes: The CDF at distance d is the cumulative sum of the K -densities up to distance d . Results in this table are reported for a distance $d = 50$ kilometers. To understand how to read that table, take 'Women's and Girls' Cut and Sew Lingerie, Loungewear and Nightwear Manufacturing' (NAICS 315231) as an example. In 1990, 62 percent of the distances between plants in that industry are less than 50 kilometers. Put differently, if we draw two plants in that industry at random, the probability that these plants are less than 50 kilometers apart is 0.62. If we, however, draw two plants at random among *all* manufacturing plants, that same probability would only be about 0.08 (see Table 2). Clearly, this large difference suggests that the location patterns of plants in the 'Women's and Girls' Cut and Sew Lingerie, Loungewear and Nightwear Manufacturing' industry are very different from those of manufacturing in general. Plants in that industry are much closer than they 'should be' if they were distributed like overall manufacturing.

Table 10: Percentage of industries with random, localized, and dispersed point patterns, 1990 to 2009.

Year	Unweighted (plant counts)			Employment weighted			Sales weighted		
	Random	Localized	Dispersed	Random	Localized	Dispersed	Random	Localized	Dispersed
1990	52.53	34.63	12.84	52.53	36.96	10.51	54.86	37.35	7.78
1991	51.36	36.19	12.45	52.92	38.52	8.56	55.25	36.19	8.56
1992	53.70	36.19	10.12	56.42	35.02	8.56	58.37	33.46	8.17
1993	53.70	34.24	12.06	58.37	33.46	8.17	59.53	31.52	8.95
1994	49.81	36.96	13.23	57.20	33.07	9.73	60.70	30.74	8.56
1995	55.25	33.46	11.28	58.37	33.07	8.56	59.53	32.30	8.17
1996	54.09	35.41	10.51	56.03	35.41	8.56	59.53	33.46	7.00
1997	55.25	35.41	9.34	60.70	32.30	7.00	61.09	32.68	6.23
1998	55.64	34.24	10.12	58.37	35.02	6.61	61.87	32.68	5.45
1999	55.25	34.63	10.12	58.75	35.41	5.84	61.48	32.30	6.23
2000	47.86	37.74	14.40	51.75	40.47	7.78	53.31	40.47	6.23
2001	43.58	41.25	15.18	52.92	40.86	6.23	50.58	42.41	7.00
2002	45.91	39.69	14.40	50.97	41.63	7.39	54.86	37.35	7.78
2003	47.47	36.58	15.95	50.58	40.86	8.56	55.64	35.41	8.95
2004	60.31	30.35	9.34	60.31	33.07	6.61	60.70	32.30	7.00
2005	58.75	33.46	7.78	62.65	31.13	6.23	64.20	31.52	4.28
2006	60.31	30.35	9.34	60.31	33.46	6.23	62.26	33.85	3.89
2007	57.59	33.46	8.95	60.70	33.85	5.45	62.65	32.30	5.06
2008	56.03	34.24	9.73	61.48	31.91	6.61	64.59	29.96	5.45
2009	59.53	33.07	7.39	63.04	31.52	5.45	63.04	31.13	5.84

Source: Authors' computations using the *Annual Survey of Manufacturers* Longitudinal Microdata file. The statistical significance of the location patterns is computed using Monte Carlo simulations with 1,000 replications following the procedure developed by Duranton and Overman (2005).

industries with the lowest ad valorem rates in 2008 are: Telephone Apparatus Manufacturing (NAICS 334210), Navigational and Guidance Instruments Manufacturing (NAICS 334511), Radio and Television Broadcasting and Wireless Communications Equipment Manufacturing (NAICS 334220), Computer and Peripheral Equipment Manufacturing (NAICS 334110), Other Communications Equipment Manufacturing (NAICS 334290), Automobile and Light-Duty Motor Vehicle Manufacturing (NAICS 336110), Paper Industry Machinery Manufacturing (NAICS 333291), Medical Equipment and Supplies Manufacturing (NAICS 339110), Tobacco Product Manufacturing (NAICS 312220), and Other Wood Household Furniture Manufacturing (NAICS 337123).

Appendix S.4. Additional robustness checks and regression results.

We ran the following additional robustness checks which are not reported in the paper (see our discussion paper version for some of those results). First, we investigate whether changes in information and communication technologies (ICT) may lead to more geographic dispersion. To this end, we use the ICT investment variables from the KLEMS database, interacted with the other variables of the model, to check whether changes in communication costs have the same effect than changes in transport costs. We did not get any significant coefficients – neither for the direct effects, nor for the interaction terms. Second, we estimated models with heterogeneous coefficients since transport costs differ across industries. To this end, we split our sample

Table 11: Highest and lowest ad valorem transport costs.

Year	Average top-ten	Average bottom-ten
1992	14.3%	0.40%
1993	13.8%	0.39%
1994	12.8%	0.37%
1995	12.5%	0.37%
1996	12.3%	0.36%
1997	12.7%	0.36%
1998	12.6%	0.35%
1999	12.2%	0.34%
2000	12.6%	0.36%
2001	12.8%	0.36%
2002	12.5%	0.34%
2003	12.9%	0.35%
2004	13.2%	0.37%
2005	13.2%	0.37%
2006	13.5%	0.38%
2007	13.3%	0.37%
2008	14.0%	0.39%

Notes: Statistics Canada, Author's calculations.

into high versus low transport cost industries, using a 'below median'-'above median' criterion. The two coefficients were statistically identical. We also treated decreasing/increasing transport costs in an asymmetric way as they may have asymmetric impacts. Again, the two coefficients were fairly close. Third, we replaced our measures of input and output linkages with the industry 'material share to sales' ratio, a proxy for reliance on intermediate inputs. That variable turns out to be insignificant in our regressions, whereas the other coefficients are largely unaffected. Third, we also ran the model as a pooled cross-section and by year using a between estimator and found roughly the same signs and significant coefficients for transport costs. Although the levels of transport costs do seem to matter for the geographical concentration of industries, the time-series changes in those costs are much more strongly associated with changes in that concentration. Last, we also tried to control for the 'labor intensity' of an industry (not just high-skilled workers vs low-skilled workers). We constructed different measures using the quantity index of labor and the quantity index of capital from the KLEMS data, but these variables turned out again to be insignificant in our regressions.

Last, we also experimented with different non-linear transport cost specifications. More precisely, we estimated the effect of transportation costs with a spline, allowing the coefficients to vary between ad valorem rates of 0 to 0.05% (low), 0.05 to 15% (moderate), and 15% or greater (high). These are admittedly arbitrary categories, but ones that we believe make intuitive sense. The results are, by and large, consistent with the simpler specification that we use. Yet, we find that at low levels, the effect of transportation costs is positive or insignificant. At moderate levels, the coefficient is negative and always significant, and at high levels the coefficient is negative and insignificant. Transport costs thus seem to matter most strongly in the intermediate range.

Table 12: Estimation of specification (5) excluding textile and high-tech industries.

Variables	Excluding textiles industries			Excluding high-tech industries		
	CDF 10km	CDF 100km	CDF 500km	CDF 10km	CDF 100km	CDF 500km
Ad valorem trucking costs (residual)	-0.213 ^a (0.077)	-0.210 ^a (0.072)	-0.193 ^a (0.049)	-0.396 ^a (0.145)	-0.324 ^b (0.128)	-0.205 ^a (0.068)
Asian share of imports	-0.568 ^c (0.322)	-0.508 ^c (0.282)	-0.211 (0.174)	-1.517 ^a (0.554)	-1.035 ^a (0.350)	-0.380 ^b (0.155)
OECD share of imports	-0.035 (0.275)	0.007 (0.241)	0.137 (0.181)	-0.860 (0.530)	-0.474 (0.333)	-0.084 (0.177)
NAFTA share of imports	-0.097 (0.251)	-0.062 (0.221)	0.076 (0.156)	-0.878 ^c (0.499)	-0.531 ^c (0.317)	-0.133 (0.157)
Asian share of exports	0.627 (0.440)	0.505 (0.358)	0.096 (0.130)	0.468 (0.490)	0.469 (0.378)	0.111 (0.121)
OECD share of exports	0.471 ^b (0.186)	0.413 ^b (0.161)	0.249 ^b (0.097)	0.346 (0.236)	0.424 ^b (0.170)	0.271 ^a (0.098)
NAFTA share of exports	0.400 ^b (0.196)	0.348 ^b (0.170)	0.128 (0.080)	0.149 (0.246)	0.275 (0.179)	0.124 (0.085)
Input distance	-0.458 ^a (0.051)	-0.439 ^a (0.049)	-0.315 ^a (0.036)	-0.387 ^a (0.075)	-0.346 ^a (0.057)	-0.245 ^a (0.038)
Output distance	-0.265 ^a (0.043)	-0.245 ^a (0.040)	-0.155 ^a (0.029)	-0.333 ^a (0.051)	-0.336 ^a (0.044)	-0.216 ^a (0.030)
Average minimum distance	-0.289 ^a (0.041)	-0.265 ^a (0.038)	-0.142 ^a (0.026)	-0.321 ^a (0.053)	-0.257 ^a (0.038)	-0.128 ^a (0.026)
Industry controls included	Yes	Yes	Yes	Yes	Yes	Yes
Number of NAICS industries	229	229	229	198	198	198
Number of years	17	17	17	17	17	17
Industry dummies	yes	yes	yes	yes	yes	yes
Observations (NAICS \times years)	3,893	3,893	3,893	3,366	3,366	3,366
R^2	0.516	0.532	0.539	0.481	0.556	0.553

Notes: All estimations for 257 industries and 17 years (4,369 observations). ^a, ^b, ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. We use simple OLS. All specifications include industry and year fixed effects. Standard errors are clustered at the industry level and given in parentheses. Our measures of input and output distances are computed using $N = 5$. 'Ad valorem trucking costs (residual)' denotes the residual of the regression of 'Ad valorem trucking costs' on industry multi factor productivity. A constant term is included in all regressions but not reported. All industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales) are included but not reported. Our definition of high-tech sectors is based on the US Bureau of Labor Statistics classification by Hecker (2005). This definition of high-tech industries is 'input based'. An industry is 'high-tech' if it employs a high proportion of scientists, engineers or technicians. As shown by Hecker (2005), these industries are also usually associated with a high R&D-to-sales ratio, and they also largely – but not always – produce goods that are classified as 'high-tech' by the Bureau of Economic Analysis.

Table 13: Estimation of specification (5) using employment-weighted CDFs, sales-weighted CDFs, and five year averages.

Dependent variable	Employment weighted CDF			Sales weighted CDF			Unweighted CDF, five year averages		
Variables	CDF 10km	CDF 100km	CDF 500km	CDF 10km	CDF 100km	CDF 500km	CDF 10km	CDF 100km	CDF 500km
Ad valorem trucking costs (residual)	-0.158 ^b (0.077)	-0.150 ^b (0.072)	-0.148 ^a (0.053)	-0.134 ^c (0.076)	-0.127 ^c (0.070)	-0.137 ^a (0.045)	-0.377 ^a (0.085)	-0.361 ^a (0.076)	-0.315 ^a (0.060)
Asian share of imports	-0.684 ^b (0.312)	-0.531 ^b (0.252)	-0.241 ^c (0.145)	-0.713 ^b (0.349)	-0.604 ^b (0.276)	-0.285 ^c (0.162)	-1.463 ^b (0.579)	-1.012 ^a (0.357)	-0.383 ^c (0.202)
OECD share of imports	-0.377 (0.264)	-0.232 (0.217)	0.008 (0.164)	-0.305 (0.286)	-0.186 (0.236)	0.043 (0.176)	-0.770 (0.566)	-0.351 (0.336)	-0.006 (0.236)
NAFTA share of imports	-0.312 (0.244)	-0.208 (0.198)	-0.018 (0.141)	-0.262 (0.276)	-0.195 (0.226)	0.003 (0.159)	-0.821 (0.518)	-0.477 (0.317)	-0.104 (0.201)
Asian share of exports	0.264 (0.483)	0.368 (0.389)	0.065 (0.130)	0.217 (0.507)	0.299 (0.398)	0.082 (0.106)	0.322 (0.539)	0.366 (0.439)	0.051 (0.211)
OECD share of exports	0.212 (0.295)	0.330 (0.210)	0.181 ^c (0.094)	0.349 (0.288)	0.424 ^c (0.216)	0.280 ^a (0.096)	0.360 (0.386)	0.450 (0.314)	0.266 (0.191)
NAFTA share of exports	0.111 (0.310)	0.276 (0.206)	0.098 (0.075)	0.190 (0.303)	0.318 (0.213)	0.169 ^b (0.076)	0.265 (0.383)	0.442 (0.296)	0.180 (0.149)
Input distance	-0.256 ^a (0.063)	-0.238 ^a (0.054)	-0.186 ^a (0.032)	-0.256 ^a (0.064)	-0.239 ^a (0.056)	-0.180 ^a (0.033)	-0.258 ^a (0.073)	-0.246 ^a (0.059)	-0.221 ^a (0.043)
Output distance	-0.234 ^a (0.053)	-0.222 ^a (0.048)	-0.127 ^a (0.030)	-0.200 ^a (0.056)	-0.193 ^a (0.048)	-0.113 ^a (0.029)	-0.374 ^a (0.069)	-0.383 ^a (0.062)	-0.239 ^a (0.044)
Minimum distance	-0.312 ^a (0.050)	-0.246 ^a (0.039)	-0.119 ^a (0.026)	-0.327 ^a (0.054)	-0.249 ^a (0.039)	-0.131 ^a (0.026)	-0.400 ^a (0.067)	-0.297 ^a (0.043)	-0.141 ^a (0.032)
Industry controls included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of NAICS industries	257	257	257	257	257	257	257	257	257
Number of years	17	17	17	17	17	17	4	4	4
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations (NAICS × years)	4,369	4,369	4,369	4,369	4,369	4,369	1,028	1,028	1,028
R ²	0.318	0.371	0.381	0.294	0.359	0.376	0.517	0.599	0.598

Notes: ^a, ^b, ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. We use simple OLS. Standard errors, given in parentheses, are clustered at the industry level. Our measures of input and output distances are computed using $N = 5$. A constant term is included in all regressions but not reported. All industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales) are included but not reported.

Table 14: Estimation results for specification (5) with splits by upstreamness.

Variables	Dependent variable is the CDF at 50 km	
	(Model 11)	(Model 12)
Ad valorem trucking costs (residual)	-0.252 ^a (0.066)	-0.220 (0.135)
Ad valorem trucking costs (residual) Q1	-0.104 (0.378)	-0.543 (0.549)
Ad valorem trucking costs (residual) Q5	0.061 (0.142)	-0.083 (0.214)
Asian share of imports	-1.057 ^a (0.349)	-1.685 ^a (0.494)
Asian share of imports Q1	-0.583 (0.697)	-0.433 (0.737)
Asian share of imports Q5	1.473 ^a (0.539)	2.115 ^a (0.644)
OECD share of imports	-0.418 (0.314)	-0.866 ^b (0.423)
OECD share of imports Q1	-0.442 (0.692)	-0.893 (0.713)
OECD share of imports Q5	1.099 ^b (0.519)	1.602 ^a (0.530)
NAFTA share of imports	-0.382 (0.264)	-0.876 ^b (0.363)
NAFTA share of imports Q1	-0.787 (0.653)	-1.122 (0.686)
NAFTA share of imports Q5	0.810 ^c (0.447)	1.306 ^a (0.463)
Asian share of exports	0.482 (0.344)	0.406 (0.405)
OECD share of exports	0.478 ^a (0.178)	0.415 ^c (0.222)
NAFTA share of exports	0.338 ^b (0.171)	0.331 (0.232)
Input distance	-0.358 ^a (0.055)	
Output distance	-0.305 ^a (0.044)	
Average minimum distance	-0.300 ^a (0.039)	
Industry controls included	Yes	Yes
Industry dummies	Yes	Yes
Year dummies	Yes	Yes
Observations (NAICS \times years)	4,369	4,369
R^2	0.527	0.158

Notes: The dependent variable is the unweighted (count based) Duranton-Overman K -density CDF. ^a, ^b and ^c denote coefficients significant at the 1%, 5% and 10% levels, respectively. We use simple OLS. Standard errors are clustered at the industry level and given in parentheses. Our measures of input and output distances, as well as average minimum distance, are computed using $N = 5$. 'Ad valorem trucking costs (residual)' denotes the residual of the regression of 'Ad valorem trucking costs' on industry multi factor productivity. A constant term is included in all regressions but not reported. All industry controls (Total industry employment; Firm Herfindahl index (employment based); Mean plant size; Share of plants affiliated with multiplant firms; Share of plants controlled by foreign firms; Natural resource share of inputs; Energy share of inputs; Share of hours worked by all workers with post-secondary education; In-house R&D share of sales) are included but not reported.