DISCUSSION PAPER SERIES

No. 7369

TRADE, WAGES, AND PRODUCTIVITY

Kristian Behrens, Giordano Mion,
Yasusada Murata and Jens Südekum

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7369.asp**

ISSN 0265-8003

TRADE, WAGES, AND PRODUCTIVITY

Kristian Behrens, University of Quebec, Montreal, CIRPÉE and CEPR
Giordano Mion, London School of Economics, National Bank of Belgium, CEP, London and CEPR
Yasusada Murata, ARISH, Nihon University and NUPRI
Jens Südekum, University of Duisburg-Essen, Ruhr Graduate School of Economics and IZA, Germany

Discussion Paper No. 7369
July 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7369

July 2009

ABSTRACT

Trade, wages, and productivity

We develop a new general equilibrium model of monopolistic competition with heterogeneous firms, variable demand elasticity and multiple asymmetric regions, in which trade integration induces wage and productivity changes. Using Canada-US interregional trade data, we structurally estimate a theory-based gravity equation system featuring endogenous wages and productivity. Given the estimated parameter values, we first decompose 'border effects' into a 'pure' border effect, relative and absolute wage effects, and a selection effect. We then quantify the impacts of removing the trade distortions generated by the Canada-US border on regional market aggregates such as wages, productivity, markups, the mass of varieties produced and consumed, as well as welfare. Last, we extend the counterfactual analysis to the firm level by generating productivity distributions and their changes via simulation.

JEL Classification: F12, F15 and F17
Keywords: counterfactual analysis, endogenous markups, firm heterogeneity, general equilibrium, gravity equation system and monopolistic competition

Kristian Behrens
University of Quebec, Montreal
Department of Economics
Case postale 8888
Succursale Centre-Ville Montréal
Montréal, QC H3C 3P8
CANADA

Email: behrens.kristian@uqam.ca

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159512

Giordano Mion
Department of Geography
London School of Economics
Houghton Street
London WC2A 2AE
UK

Email: G.Mion@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157937

Yasusada Murata
Advanced Research Institute for the Sciences & Humanities,
Nihon University
12-5, Goban-cho
Chiyoda-ku
Tokyo 102-8251
JAPAN

Email:
murata.yasusada@nihon-u.ac.jp

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=158745

Jens Südekum
University of Duisburg-Essen
Mercator School of Management
Lotharstraße 65
47057 Duisburg
GERMANY

Email: jens.suedekum@uni-due.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=154217

Submitted 08 July 2009

1 Introduction

Over the last fifteen years, empirical research has uncovered a series of stylized facts about firms in international trade that hold across a large number of countries and industries. Firstly, only a small share of firms is engaged in foreign trade, and these firms differ along various dimensions from those operating entirely in domestic markets. Exporters are, in particular, larger and more productive than non-exporters (Bernard and Jensen, 1995; Clerides *et al.*, 1998; Bernard and Jensen, 1999; Aw *et al.*, 2000; Tybout, 2003; Bernard *et al.*, 2009). Secondly, most exporters earn only a small fraction of their revenue in foreign markets, and only a small share of exporting firms accounts for the lion's share of total export sales. Put differently, the bulk of exporters has a fairly low export intensity and the export sales distribution is skewed towards a small number of very large enterprises (Bernard and Jensen, 1995; Bernard *et al.*, 2003, 2007a). Thirdly, firm-level productivity differences are a channel through which trade liberalization brings about aggregate productivity gains. The reason is that trade liberalization increases competition, reduces markups, and forces the least efficient producers to leave the market, thereby reallocating market shares from low to high productivity firms (Aw *et al.*, 2000; Pavcnik, 2002; Trefler, 2004; Bernard *et al.*, 2007a).

Two recent strands of literature have moved firm-level productivity differences to the heart of trade analysis. The first strand, originating in the seminal contribution by Melitz (2003), extends an intra-industry trade model of monopolistic competition to allow for productivity differences across firms. That model can cope with the stylized facts mentioned above, namely that only the most productive firms operate in export markets and that trade liberalization induces aggregate productivity gains. The framework has, however, two rather restrictive features which make it difficult to take it to data: its reliance on symmetric countries with factor price equalization (FPE) and constant elasticity of substitution (CES). Though analytically convenient, the first feature is clearly at odds with the facts and neglects the general equilibrium impacts of differential factor price adjustments following trade liberalization across asymmetric countries. The second feature generates constant price-cost margins, which does not accord with recent empirical evidence that increasing trade integration and larger market size put downward pressure on markups (Syverson, 2004; Badinger, 2007).

To deal with these limitations, several alternatives have been put forward in the literature. Melitz and Ottaviano (2008) propose a monopolistic competition model with multiple asymmetric countries in which markups do depend on trade costs, firm-level productivity, and market size. However, FPE still holds in their trade equilibrium. Furthermore, their quasi-linear setup channels all income effects towards a numéraire good, which gives the analysis a partial equilibrium flavor. Bernard *et al.* (2007b) embed Melitz's model into a Heckscher-Ohlin framework, thereby allowing for factor price differences between two asymmetric countries. They rely, however, on the CES specification so that markups are again constant. In addition, extending their model beyond two countries, as required for taking it to data, is not straightforward.

The second strand of literature, rooted in the seminal contribution by Bernard *et al.* (2003), extends the Eaton and Kortum (2002) Ricardian model to allow for imperfect competition. That framework can accomodate multiple asymmetric countries, thus making it suitable as a basis for empirical work and counterfactual analysis. Furthermore, FPE is relaxed by assuming exogenous wage differences across countries.[1] The model by Bernard *et al.* (2003) can replicate several firm-level facts, in particular that the bulk of exporters has a fairly low export intensity. Despite its good empirical performance, their framework also displays three rather restrictive features: each market is served by a single firm per industry only (the lowest-cost supplier); the set of goods is fixed; and markups are independent of the number of competing firms and identically distributed across countries. These features are also at odds with the facts. There is indeed growing evidence that trade substantially expands the set of imported varieties, which is an important channel through which gains from trade materialize (Krugman, 1979, 1980; Broda and Weinstein, 2006). Furthermore, as argued before, price-cost margins are, in reality, affected by trade integration and by market size.

Our first contribution is to develop a unified theoretical framework that encompasses the key advantages of both the Melitz (2003) and the Bernard *et al.* (2003) types of approaches. More precisely, building on Behrens and Murata (2007), we propose a new general equilibrium model of monopolistic competition with heterogeneous firms, variable demand elasticity and multiple asymmetric regions, in which wages and markups are endogenous and need not be equalized. Our model is thus well equipped to trace out the impacts of trade integration on wages, productivity, and price-cost margins. Our framework is furthermore tractable enough to allow for clear-cut comparative static results with two asymmetric regions. We show, in particular, that if the two regions differ in size or technological possibilities, the larger or technologically more advanced region has the higher wage, productivity and welfare than the other region. Furthermore, trade integration favors convergence of regional wages and average productivities as bilateral trade barriers fall. These changes inevitably affect trade flows and welfare and must be taken into consideration when trying to assess the impacts of liberalization.

Our second contribution is to illustrate how our multi-region model can be taken to data by relating it to the vast literature on 'gravity equations' (McCallum, 1995; Anderson and van Wincoop, 2003). As is well known, gravity equations can be derived from trade models by using the value of exports between trading partners. The conventional method for estimating those equations is to rely on exporter and importer fixed effects. While such a procedure yields consis-

[1] Alvarez and Lucas (2007) extend the Eaton and Kortum (2002) model with perfect competition to include endogenously determined wages. To the best of our knowledge, such an extension is missing to date for the model by Bernard *et al.* (2003) with imperfect competition.

tent estimates of the trade friction parameters (Feenstra, 2004), it ignores labor market clearing, zero profit and trade balance conditions that must hold in general equilibrium. As shown by Anderson and van Wincoop (2003), disregarding equilibrium constraints places severe limitations on the usefulness of gravity equations for conducting counterfactual analysis. We hence derive a *gravity equation system* featuring both endogenous wages and productivity (firm selection) that subsume all the equilibrium information of our model.[2] Using Canada-US interregional trade data, we then structurally estimate the model's key parameters that are compatible with the equilibrium wages and productivity distributions.

Using the estimated parameter values, our last contribution is to conduct a counterfactual experiment in the spirit of Bernard *et al.* (2003), where we quantify the impacts of eliminating all trade distortions generated by the Canada-US border.[3] In so doing, we make use of the general equilibrium constraints to solve for regional wages and productivities that would prevail in a borderless world. This allows us to compute a series of *bilateral border effects* which illustrate how trade flows between any two regions would be affected by the trade integration and the induced wage and productivity responses. These bilateral border effects can be decomposed into a 'pure' border effect, relative and absolute wage effects, and a selection effect. We use this decomposition to provide a detailed account of each channel's contribution to changes in trade flows across Canadian provinces and US states. One key insight is that disregarding the endogenous wage and productivity responses leads to a substantial upward bias of Canadian bilateral border effects (by up to 50%).

What would be the regional effects for Canadian provinces and US states of our counterfactual trade integration scenario? Our analysis reveals that the border removal between Canada and the US would lead to higher average productivity, greater consumption diversity, and welfare gains for all provinces and states. Canadian labor productivity would rise by 5.71%, a figure that is roughly similar to the one estimated by Trefler (2004) for the 1988 Canada-US Free Trade Agreement.[4] The corresponding US figure is much smaller, with just a 0.3% labor productivity gain. The trade integration would favor wage convergence across the two countries. In fact, the

[2]Several recent contributions have derived gravity equations with heterogeneous firms (e.g., Chaney, 2008; Helpman *et al.*, 2008; Melitz and Ottaviano, 2008). In all these models, however, wages are either equalized or assumed to differ exogenously across regions.

[3]Bernard *et al.* (2003) consider the effects of a generalized 5-percent drop in trade costs and a 10-percent exogenous increase in the US relative wage. Whereas their Ricardian model is well suited to analyze trade among dissimilar countries, our monopolistic competition model is better suited to assess the impacts of trade integration among similar countries like Canada and the US.

[4]Trefler (2004, pp.880-881) estimates that Canadian labor productivity increased by 7.4%. It is worth emphasizing that he attributes the *sources* of these productivity gains to "market share shifts favoring high-productivity plants. Such share shifting would come about from the growth of high-productivity plants and the demise and/or exit of low-productivity plants [...]". These are precisely the key channels highlighted by our model.

relative wages in the Canadian provinces would rise by 1.79% to 5.84%.[5] Our counterfactual analysis also reveals that Canadian regional markups would fall by 1.78% to 5.19% and US regional markups would fall by 0.12% to 1.08%, thereby providing an idea of the magnitude of the pro-competitive effects of trade integration. Given those results, welfare gains are larger in Canada than in the US, averaging 10% and 3%, respectively. One important insight from our analysis is that endogenous wage and productivity responses are far from being uniform across Canadian provinces and US states. Investigating what drives these regional variations, we find that *geography* and *size* matter: less populous regions closer to the border tend to be more affected by the border removal.

We finally investigate the firm-level properties of our model by generating two large firm samples randomly drawn from the fitted productivity distributions of the different Canadian provinces and US states before and after the hypothetical border removal. This exercise firstly reveals that we can replicate fairly well several stylized facts, in particular that the productivity distribution of exporters lies to the right of that of non-exporters, that a very small share of exporters accounts for the bulk of export sales, and that the vast majority of exporters have a fairly low export intensity. Secondly, our analysis provides us with some novel explanations for current issues in firm-level studies. It has, for example, been shown that exporters' and domestic firms' productivity distributions exhibit a sizeable overlap (Bernard *et al.*, 2009). Our framework does deliver such an overlap because countries have an internal geography and because wages are not equalized across regions. Moreover, although selection due to trade integration works as a simple left-truncation of the regional productivity distributions in our model, the associated country-wise changes are more complex.

The remainder of the paper is organized as follows. Section 2 sets up the basic model and deals with the closed economy case. In Section 3 we extend it to a multi-region framework and prove analytical results for the case of two asymmetric regions. Section 4 derives the gravity equation system, describes the data, and presents the estimation procedure and results. In Section 5 we carry out the counterfactual experiment of removing the Canada-US border and discuss the associated regional and firm-level impacts. Section 6 concludes.

2 Closed economy

Consider a closed economy with a final consumption good, provided as a continuum of horizontally differentiated varieties. We denote by Ω the endogenously determined set of available

[5] We choose the wage in one US state (Alabama) as the numéraire. Though not directly comparable, our results for the relative wage increases in the Canadian provinces would therefore not be far away from the 3-5% wage increases reported by Trefler (2004, p.885).

varieties, with measure N. There are L consumers, each of whom supplies inelastically one unit of labor, which is the only factor of production.

2.1 Preferences and demands

All consumers have identical preferences which display 'love of variety' and give rise to demands with variable elasticity. Following Behrens and Murata (2007), the utility maximization problem of a representative consumer is given by:

$$\max_{q(j),\, j\in\Omega} \; U \equiv \int_\Omega \left[1 - \mathrm{e}^{-\alpha q(j)}\right] \mathrm{d}j \qquad \text{s.t.} \quad \int_\Omega p(j) q(j) \mathrm{d}j = E, \tag{1}$$

where E denotes expenditure; $p(j) > 0$ and $q(j) \geq 0$ stand for the price and the per capita consumption of variety j; and $\alpha > 0$ is a parameter. As shown by Behrens and Murata (2007), solving (1) yields the following demand functions:

$$q(i) = \frac{E}{N\overline{p}} - \frac{1}{\alpha} \left\{ \ln \left[\frac{p(i)}{N\overline{p}} \right] + h \right\}, \quad \forall i \in \Omega, \tag{2}$$

where

$$\overline{p} \equiv \frac{1}{N} \int_\Omega p(j) \mathrm{d}j \quad \text{and} \quad h \equiv - \int_\Omega \ln \left[\frac{p(j)}{N\overline{p}} \right] \frac{p(j)}{N\overline{p}} \mathrm{d}j$$

denote the average price and the differential entropy of the price distribution, respectively. Since marginal utility at zero consumption is bounded, the demand for a variety need not be positive. Indeed, as can be seen from (2), the demand for variety i is positive if and only if its price is lower than the reservation price p^d. Formally,

$$q(i) > 0 \quad \iff \quad p(i) < p^d \equiv N\overline{p}\, \mathrm{e}^{\frac{\alpha E}{N\overline{p}} - h}. \tag{3}$$

Note that the reservation price p^d is a function of the price aggregates $\overline{p}$ and h. Combining expressions (2) and (3) allows us to express the demand for variety i concisely as follows:

$$q(i) = \frac{1}{\alpha} \ln \left[\frac{p^d}{p(i)} \right]. \tag{4}$$

2.2 Technology and market structure

The labor market is assumed to be perfectly competitive so that all firms take the wage rate w as given. Prior to production, each firm engages in research and development, which requires a fixed amount F of labor paid at the market wage. Each firm discovers its marginal labor requirement $m(i) \geq 0$ only after making this irreversible investment. We assume that $m(i)$ is drawn from a common and known, continuously differentiable distribution G. Since research and development costs are sunk, a firm will survive (i.e., operate) in the market provided it can charge a price $p(i)$ above marginal cost $m(i)w$.

Each surviving firm sets its price to maximize operating profit

$$\pi(i) = L\big[p(i) - m(i)w\big]q(i), \tag{5}$$

where $q(i)$ is given by (4). Since there is a continuum of firms, no individual firm has any impact on p^d so that the first-order conditions for (operating) profit maximization are given by:

$$\ln\left[\frac{p^d}{p(i)}\right] = \frac{p(i) - m(i)w}{p(i)}, \quad \forall i \in \Omega. \tag{6}$$

A price distribution satisfying (6) is called a *price equilibrium*. Multiplying both sides of (6) by $p(i)$, integrating over Ω, and using (4) yield the average price as follows:

$$\overline{p} = \overline{m}w + \frac{\alpha E}{N}, \tag{7}$$

where $\overline{m} \equiv (1/N)\int_\Omega m(j)\mathrm{d}j$ denotes the average marginal labor requirement of the surviving firms. Observe that expression (7) displays pro-competitive effects, i.e., the average price is decreasing in the mass of surviving firms N.

Equations (4) and (6) imply that $q(i) = (1/\alpha)[1 - m(i)w/p(i)]$, which allows us to derive the upper and lower bounds for the marginal labor requirement. The maximum output is given by $q(i) = 1/\alpha$ at $m(i) = 0$. The minimum output is given by $q(i) = 0$ at $p(i) = m(i)w$, which by (6) implies that $p(i) = p^d$. Therefore, the cutoff marginal labor requirement is defined as $m^d \equiv p^d/w$. A firm that draws m^d is indifferent between producing and not producing, whereas all firms with a draw below (resp., above) m^d remain in (resp., exit from) the market.

Since firms differ only by their marginal labor requirement, we can express all firm-level variables in terms of m. Solving (6) by using the Lambert W function, defined as $\varphi = W(\varphi)\mathrm{e}^{W(\varphi)}$, the profit-maximizing prices and quantities, as well as operating profits, can be expressed as follows:

$$p(m) = \frac{mw}{W}, \quad q(m) = \frac{1}{\alpha}(1 - W), \quad \pi(m) = \frac{Lmw}{\alpha}\left(W^{-1} + W - 2\right), \tag{8}$$

where we suppress the argument $\mathrm{e}m/m^d$ of W to alleviate notation. It is readily verified that $W' > 0$ for all non-negative arguments and that $W(0) = 0$ and $W(\mathrm{e}) = 1$ (see Appendix A.1 for the derivation of (8) and the properties of W). Hence, $0 \leq W \leq 1$ if $0 \leq m \leq m^d$. The expressions in (8) show that a firm with a draw m^d charges a price equal to marginal cost, faces zero demand, and earns zero operating profit. Since $W' > 0$, we readily obtain $\partial p(m)/\partial m > 0$, $\partial q(m)/\partial m < 0$ and $\partial \pi(m)/\partial m < 0$. In words, firms with better draws charge lower prices, sell larger quantities, and earn higher operating profits than firms with worse draws.

2.3 Equilibrium

We now state the equilibrium conditions for the closed economy, which consist of zero expected profits and labor market clearing. First, given the mass of entrants N^E, the mass of surviving firms can be written as $N = N^E G(m^d)$. Using (5), the zero expected profit condition for each firm is given by:

$$L \int_0^{m^d} [p(m) - mw]\, q(m) \mathrm{d}G(m) = Fw, \tag{9}$$

which, combined with (8), can be rewritten as

$$\frac{L}{\alpha} \int_0^{m^d} m \left(W^{-1} + W - 2\right) \mathrm{d}G(m) = F. \tag{10}$$

As the left-hand side of (10) is strictly increasing in m^d from 0 to ∞, there always exists a unique equilibrium cutoff (see Appendix A.2). Furthermore, the labor market clearing condition is given by:[6]

$$N^E \left[L \int_0^{m^d} mq(m) \mathrm{d}G(m) + F \right] = L, \tag{11}$$

which, combined with (8), can be rewritten as

$$N^E \left[\frac{L}{\alpha} \int_0^{m^d} m \left(1 - W\right) \mathrm{d}G(m) + F \right] = L. \tag{12}$$

Given the equilibrium cutoff m^d, equation (12) can be uniquely solved for N^E.

How does population size affect entry and firms' survival probabilities? Using the equilibrium conditions (10) and (12), we can show that a larger L leads to more entrants N^E and a smaller cutoff m^d, respectively (see Appendix A.3). The effect of population size on the mass of surviving firms N is in general ambiguous. However, under the commonly made assumption that firms' productivity draws $1/m$ follow a Pareto distribution

$$G(m) = \left(\frac{m}{m^{\max}} \right)^k,$$

with upper bound $m^{\max} > 0$ and shape parameter $k \geq 1$, we can show that N is increasing in L.[7] Using this distributional assumption, we readily obtain closed-form solutions for the equilibrium cutoff and mass of entrants:

$$m^d = \left[\frac{\alpha F \left(m^{\max}\right)^k}{\kappa_2 L} \right]^{\frac{1}{k+1}} \quad \text{and} \quad N^E = \frac{\kappa_2}{\kappa_1 + \kappa_2} \frac{L}{F},$$

---

[6] Note that by using (9) and the budget constraint $N^E \int_0^{m^d} p(m)q(m)\mathrm{d}G(m) = E$, we obtain $EL/(wN^E) = L \int_0^{m^d} mq(m)\mathrm{d}G(m) + F$ which, together with (11), yields $E = w$ in equilibrium.

[7] The Pareto distribution has been extensively used in the previous literature on heterogeneous firms (e.g., Bernard *et al.*, 2007; Helpman *et al.*, 2008; Melitz and Ottaviano, 2008).

where κ_1 and κ_2 are positive constants that solely depend on k (see Appendices B.1 and B.2).[8] The mass of surviving firms is then given as follows:

$$N = \frac{\kappa_2^{\frac{1}{k+1}}}{\kappa_1 + \kappa_2} \left(\frac{\alpha}{m^{\max}}\right)^{\frac{k}{k+1}} \left(\frac{L}{F}\right)^{\frac{1}{k+1}},$$

which is increasing in population size L. One can further check that N is decreasing in the fixed labor requirement F and in the upper bound $m^{\max}$. Finally, since $\overline{m} = [k/(k+1)]m^d$ holds when productivity follows a Pareto distribution, a larger population also maps into higher average productivity $1/\overline{m}$.

3 Open economy

We now turn to the open economy case. As dealing with two regions only marginally alleviates the notational burden, we first derive the equilibrium conditions for the general case with K asymmetric regions that we use when taking our model to the data. We then present some clear-cut analytical results for the special case of two asymmetric regions in order to guide the intuition for the general case.

3.1 Preferences and demands

Preferences in the open economy case are analogous to the ones described in the previous section. Let $p_{sr}(i)$ and $q_{sr}(i)$ denote the price and the per capita consumption of variety i when it is produced in region s and consumed in region r. The utility maximization problem of a consumer in region r is given by:

$$\max_{q_{sr}(j),\, j\in\Omega_{sr}} \quad U_r \equiv \sum_s \int_{\Omega_{sr}} \left[1 - \mathrm{e}^{-\alpha q_{sr}(j)}\right] \mathrm{d}j \qquad \text{s.t.} \quad \sum_s \int_{\Omega_{sr}} p_{sr}(j) q_{sr}(j) \mathrm{d}j = E_r, \tag{13}$$

where Ω_{sr} denotes the set of varieties produced in region s and consumed in region r. It is readily verified that the demand functions are given as follows:

$$q_{sr}(i) = \frac{E_r}{N_r^c \overline{p}_r} - \frac{1}{\alpha}\left\{\ln\left[\frac{p_{sr}(i)}{N_r^c \overline{p}_r}\right] + h_r\right\}, \quad \forall i \in \Omega_{sr},$$

where N_r^c is the mass of varieties consumed in region r, and

$$\overline{p}_r \equiv \frac{1}{N_r^c} \sum_s \int_{\Omega_{sr}} p_{sr}(j)\mathrm{d}j \quad \text{and} \quad h_r \equiv -\sum_s \int_{\Omega_{sr}} \ln\left[\frac{p_{sr}(j)}{N_r^c \overline{p}_r}\right] \frac{p_{sr}(j)}{N_r^c \overline{p}_r} \mathrm{d}j$$

denote the average price and the differential entropy of the price distribution of all varieties consumed in region r. As in the closed economy case, the demand for domestic variety i (resp.,

[8] For this solution to be consistent, we must ensure that $m^d \leq m^{\max}$, i.e., $m^{\max} \geq \alpha F/(\kappa_2 L)$.

foreign variety j) is positive if and only if the price of variety i (resp., variety j) is lower than the reservation price p_r^d. Formally,

$$q_{rr}(i) > 0 \iff p_{rr}(i) < p_r^d \quad \text{and} \quad q_{sr}(j) > 0 \iff p_{sr}(j) < p_r^d,$$

where $p_r^d \equiv N_r^c \overline{p}_r \mathrm{e}^{\alpha E_r/(N_r^c \overline{p}_r) - h_r}$ is a function of the price aggregates $\overline{p}_r$ and h_r. The demands for domestic and foreign varieties can then be concisely expressed as follows:

$$q_{rr}(i) = \frac{1}{\alpha} \ln\left[\frac{p_r^d}{p_{rr}(i)}\right] \quad \text{and} \quad q_{sr}(j) = \frac{1}{\alpha} \ln\left[\frac{p_r^d}{p_{sr}(j)}\right]. \tag{14}$$

3.2 Technology and market structure

Technology and the entry process are identical to the ones described in Section 2. We assume that shipments from r to s are subject to trade costs $\tau_{rs} > 1$ for all r and s, that markets are segmented, and that firms are free to price discriminate.

Firms in region r independently draw their productivities from a region-specific distribution G_r. Assuming that firms incur trade costs in terms of labor, the operating profit of firm i in r is given by:

$$\pi_r(i) = \sum_s \pi_{rs}(i) = \sum_s L_s q_{rs}(i) \left[p_{rs}(i) - \tau_{rs} m_r(i) w_r\right]. \tag{15}$$

Each firm maximizes (15) with respect to its prices $p_{rs}(i)$ separately. Since it has no impact on the price aggregates and on the wages, the first-order conditions are given by:

$$\ln\left[\frac{p_s^d}{p_{rs}(i)}\right] = \frac{p_{rs}(i) - \tau_{rs} m_r(i) w_r}{p_{rs}(i)}, \quad \forall i \in \Omega_{rs}. \tag{16}$$

We first solve for the average price in region r. To do so, multiply (16) by $p_{rs}(i)$, use (14), integrate over Ω_{rs}, and finally sum the resulting expressions to obtain

$$\overline{p}_r \equiv \frac{1}{N_r^c} \sum_s \int_{\Omega_{sr}} p_{sr}(j) \mathrm{d}j = \frac{1}{N_r^c} \sum_s \tau_{sr} w_s \int_{\Omega_{sr}} m_s(j) \mathrm{d}j + \frac{\alpha E_r}{N_r^c}, \tag{17}$$

where the first term is the average of marginal delivered costs in region r. Expression (17) shows that $\overline{p}_r$ is decreasing in the mass N_r^c of firms competing in region r, which is similar to the result on pro-competitive effects established in the closed economy case.

Equations (14) and (16) imply that $q_{rs}(i) = (1/\alpha)[1 - \tau_{rs} m_r(i) w_r / p_{rs}(i)]$, which shows that $q_{rs}(i) = 0$ at $p_{rs}(i) = \tau_{rs} m_r(i) w_r$. It then follows from (16) that $p_{rs}(i) = p_s^d$. Hence, a firm located in r with draw $m_{rs}^x \equiv p_s^d/(\tau_{rs} w_r)$ is just indifferent between selling and not selling in region s. All firms in r with draws below m_{rs}^x are productive enough to sell to region s. In what follows, we refer to $m_{ss}^x \equiv m_s^d$ as the *domestic cutoff* in region s, whereas m_{rs}^x with $r \neq s$ is the *export cutoff*. Export and domestic cutoffs are linked as follows:

$$m_{rs}^x = \frac{\tau_{ss}}{\tau_{rs}} \frac{w_s}{w_r} m_s^d. \tag{18}$$

Expression (18) reveals how trade costs and wage differentials affect firms' ability to break into foreign markets. When wages are equalized ($w_r = w_s$) and internal trade is costless ($\tau_{ss} = 1$), all export cutoffs must fall short of the domestic cutoffs since $\tau_{rs} > 1$. In that case, breaking into any foreign market is always harder than selling domestically. However, in the presence of wage differentials and internal trade costs, the domestic and the foreign cutoffs can no longer be clearly ranked. The usual ranking, namely that exporting to s is more difficult than selling domestically in s, prevails only when $\tau_{ss}w_s < \tau_{rs}w_r$.[9]

The first-order conditions (16) can be solved as in the closed economy case. Switching to notation in terms of m, the profit-maximizing prices and quantities, as well as operating profits, are given by:

$$p_{rs}(m) = \frac{\tau_{rs}mw_r}{W}, \quad q_{rs}(m) = \frac{1}{\alpha}(1 - W), \quad \pi_{rs} = \frac{L_s\tau_{rs}mw_r}{\alpha}(W^{-1} + W - 2), \tag{19}$$

where W denotes the Lambert W function with argument $\mathrm{e}\tau_{rs}mw_r/p_s^d$, which we suppress to alleviate notation. It is readily verified that more productive firms again charge lower prices, sell larger quantities, and earn higher operating profits.

Observe that in an open economy, the masses of varieties consumed and produced in each region need not be the same. Given a mass of entrants N_r^E, only $N_r^p = N_r^E\, G_r\left(\max_s\{m_{rs}^x\}\right)$ firms survive, namely those which are productive enough to sell at least in one market. Finally, the mass of varieties consumed in region r is given by

$$N_r^c = \sum_s N_s^E G_s(m_{sr}^x). \tag{20}$$

3.3 Equilibrium

The zero expected profit condition for each firm in region r is given by

$$\sum_s L_s \int_0^{m_{rs}^x} [p_{rs}(m) - \tau_{rs}mw_r]\, q_{rs}(m)\mathrm{d}G_r(m) = F_r w_r, \tag{21}$$

where F_r is the region-specific fixed labor requirement. Furthermore, each labor market clears in equilibrium, which requires that

$$N_r^E \left[\sum_s L_s\tau_{rs}\int_0^{m_{rs}^x} mq_{rs}(m)\mathrm{d}G_r(m) + F_r\right] = L_r. \tag{22}$$

Last, trade is balanced for each region:

$$N_r^E \sum_{s\neq r} L_s \int_0^{m_{rs}^x} p_{rs}(m)q_{rs}(m)\mathrm{d}G_r(m) = L_r \sum_{s\neq r} N_s^E \int_0^{m_{sr}^x} p_{sr}(m)q_{sr}(m)\mathrm{d}G_s(m).$$

[9] As recently emphasized by Foster *et al.* (2008), firm selection and export performance depend more generally on profitability. In addition to physical productivity, locational heterogeneity matters in our model because firms face different wages and have different market access depending on which region they are located in.

As in the foregoing section, we can restate the equilibrium conditions using the Lambert W function (see Appendix C for details).

In what follows, we assume that productivity draws $1/m$ follow Pareto distributions with identical shape parameters $k \geq 1$. However, to capture differences in local technological possibilities, we allow the upper bounds to vary across regions, i.e., $G_r(m) = (m/m_r^{\max})^k$. A lower $m_r^{\max}$ implies that firms in region r have a higher probability of drawing a better productivity. Under the Pareto distributions, the equilibrium conditions can be greatly simplified. First, using the expressions in Appendices B.1 and C.1, labor market clearing requires that

$$N_r^E \left[\frac{\kappa_1}{\alpha \left(m_r^{\max}\right)^k} \sum_s L_s \tau_{rs} \left(\frac{\tau_{ss}}{\tau_{rs}} \frac{w_s}{w_r} m_s^d \right)^{k+1} + F_r \right] = L_r. \tag{23}$$

Second, using the expressions in Appendices B.2 and C.2, zero expected profits imply that

$$\mu_r^{\max} \equiv \frac{\alpha F_r \left(m_r^{\max}\right)^k}{\kappa_2} = \sum_s L_s \tau_{rs} \left(\frac{\tau_{ss}}{\tau_{rs}} \frac{w_s}{w_r} m_s^d \right)^{k+1}, \tag{24}$$

where $\mu_r^{\max}$ is a simple monotonic transformation of the upper bounds. Last, using the expressions in Appendices B.3 and C.3, balanced trade requires that

$$\frac{N_r^E w_r}{\left(m_r^{\max}\right)^k} \sum_{s \neq r} L_s \tau_{rs} \left(\frac{\tau_{ss}}{\tau_{rs}} \frac{w_s}{w_r} m_s^d \right)^{k+1} = L_r \sum_{s \neq r} \tau_{sr} \frac{N_s^E w_s}{\left(m_s^{\max}\right)^k} \left(\frac{\tau_{rr}}{\tau_{sr}} \frac{w_r}{w_s} m_r^d \right)^{k+1}. \tag{25}$$

The $3K$ conditions (23)–(25) depend on $3K$ unknowns: the wages w_r, the masses of entrants N_r^E, and the domestic cutoffs m_r^d. The export cutoffs m_{rs}^x can then be computed using (18). Combining (23) and (24) immediately shows that

$$N_r^E = \frac{\kappa_2}{\kappa_1 + \kappa_2} \frac{L_r}{F_r}. \tag{26}$$

The mass of entrants in region r therefore positively depends on that region's size L_r and negatively on its fixed labor requirement F_r.

Adding the term in r that is missing on both sides of (25), and using (24) and (26), we obtain the following equilibrium relationship:

$$\frac{1}{\left(m_r^d\right)^{k+1}} = \sum_s L_s \tau_{rr} \left(\frac{\tau_{rr}}{\tau_{sr}} \frac{w_r}{w_s} \right)^k \frac{1}{\mu_s^{\max}}. \tag{27}$$

Expressions (24) and (27) summarize how wages, upper bounds, cutoffs, trade costs and population sizes are related in general equilibrium.

3.4 Two-region case

Our model allows for clear-cut comparative static results with two asymmetric regions. Using (24)–(26), an equilibrium can be characterized by a system of three equations with three unknowns (the two cutoffs m_1^d and m_2^d, and the relative wage w_1/w_2) as follows:

$$\mu_1^{\max} = L_1\tau_{11}\left(m_1^d\right)^{k+1} + L_2\tau_{12}\left(\frac{\tau_{22}}{\tau_{12}}\frac{w_2}{w_1}m_2^d\right)^{k+1} \tag{28}$$

$$\mu_2^{\max} = L_2\tau_{22}\left(m_2^d\right)^{k+1} + L_1\tau_{21}\left(\frac{\tau_{11}}{\tau_{21}}\frac{w_1}{w_2}m_1^d\right)^{k+1} \tag{29}$$

$$\left(\frac{w_1}{w_2}\right)^{2k+1} = \left(\frac{\tau_{21}}{\tau_{12}}\right)^{k}\left(\frac{\tau_{22}}{\tau_{11}}\right)^{k+1}\left(\frac{m_2^d}{m_1^d}\right)^{k+1}\left(\frac{\mu_2^{\max}}{\mu_1^{\max}}\right). \tag{30}$$

Equations (28) and (29) can readily be solved for the cutoffs as a function of the relative wage $\omega \equiv w_1/w_2$:

$$(m_1^d)^{k+1} = \frac{\mu_1^{\max}}{L_1\tau_{11}}\frac{1-\rho\left(\frac{\tau_{22}}{\tau_{12}}\right)^k\omega^{-(k+1)}}{1-\left(\frac{\tau_{11}\tau_{22}}{\tau_{12}\tau_{21}}\right)^k} \quad \text{and} \quad (m_2^d)^{k+1} = \frac{\mu_2^{\max}}{L_2\tau_{22}}\frac{1-\rho^{-1}\left(\frac{\tau_{11}}{\tau_{21}}\right)^k\omega^{k+1}}{1-\left(\frac{\tau_{22}\tau_{11}}{\tau_{21}\tau_{12}}\right)^k}, \tag{31}$$

where $\rho \equiv \mu_2^{\max}/\mu_1^{\max}$ captures relative technological possibilities. A larger ρ (given F_r) implies, ceteris paribus, that firms in region 2 face a higher probability of drawing a worse productivity than those in region 1. Substituting the cutoffs (31) into (30) yields after some simplification

$$\text{LHS} \equiv \omega^k = \rho\frac{L_1}{L_2}\left(\frac{\tau_{21}}{\tau_{12}}\right)^k\frac{\rho\tau_{11}^{-k}-\tau_{21}^{-k}\omega^{k+1}}{\tau_{22}^{-k}\omega^{k+1}-\rho\tau_{12}^{-k}} \equiv \text{RHS}. \tag{32}$$

Assume that intraregional trade is less costly than interregional trade, i.e., $\tau_{11} < \tau_{21}$ and $\tau_{22} < \tau_{12}$. Then, the RHS of (32) is decreasing in ω on its relevant domain, whereas the LHS is increasing in ω. Hence, there exists a unique equilibrium such that the equilibrium relative wage ω^* is bounded by relative trade costs τ_{22}/τ_{12} and τ_{21}/τ_{11}, relative technological possibilities ρ, and the shape parameter k (see Appendix A.4).

Since the RHS of (32) is decreasing in ω, the comparative static results are straightforward to derive. In Appendix A.5 we show that, everything else equal: (i) the larger region has the higher wage, the lower cutoff and the higher welfare; (ii) the region with better technological possibilities has the higher wage, the lower cutoff, and the higher welfare; (iii) the region with the lower internal trade costs has the higher wage, the lower cutoff, and the higher welfare; (iv) if one region has better access to the other, that region has the higher wage, the lower cutoff and the higher welfare. Furthermore, we show that when the regions differ in size or technological possibilities, (v) wages, cutoffs and welfare converge between the regions as bilateral trade barriers fall.

4 Estimation

In this section we take the model with K asymmetric regions to the data. To this end, we first derive a theory-based gravity equation with general equilibrium constraints. Using the well-known Canada-US regional trade flow dataset by Anderson and van Wincoop (2003), we then structurally estimate trade friction parameters as well as other variables of the model. In the next section we turn to a counterfactual analysis, where we consider the impacts of a hypothetical decrease in trade frictions on various variables.

4.1 Gravity equation system

The value of exports from region r to region s is given by

$$X_{rs} = N_r^E L_s \int_0^{m_{rs}^x} p_{rs}(m) q_{rs}(m) \mathrm{d}G_r(m).$$

Using equations (19), (26), and the Pareto distribution for $G_r(m)$, we obtain the following gravity equation:[10]

$$\frac{X_{rs}}{L_r L_s} = \tau_{rs}^{-k} \tau_{ss}^{k+1} (w_s/w_r)^{k+1} w_r \left(m_s^d\right)^{k+1} (\mu_r^{\max})^{-1}. \tag{33}$$

As can be seen from (33), exports depend on bilateral trade costs τ_{rs}, internal trade costs in the destination τ_{ss}, origin and destination wages w_r and w_s, destination productivity m_s^d, and origin technological possibilities $\mu_r^{\max}$. A higher relative wage w_s/w_r raises the value of exports as firms in r face relatively lower production costs, whereas a higher absolute wage w_r raises the value of exports by increasing export prices p_{rs}. Furthermore, a larger m_s^d raises the value of exports since firms located in the destination are on average less productive. Last, a lower $\mu_r^{\max}$ implies that firms in region r have higher expected productivity, which raises the value of their exports. Conditions (24) and (27) give us the following general equilibrium constraints:

$$\mu_r^{\max} = \sum_v L_v \tau_{rv}^{-k} \tau_{vv}^{k+1} \left(\frac{w_v}{w_r}\right)^{k+1} \left(m_v^d\right)^{k+1} \qquad r = 1, 2. \dots K \tag{34}$$

$$\frac{1}{(m_s^d)^{k+1}} = \sum_v L_v \tau_{vs}^{-k} \tau_{ss}^{k+1} \left(\frac{w_s}{w_v}\right)^{k} (\mu_v^{\max})^{-1} \qquad s = 1, 2, \dots K \tag{35}$$

The *gravity equation system* consists of the gravity equation (33) and the $2K$ general equilibrium constraints (34) and (35) that summarize the interactions between the $2K$ endogenous variables, namely the wages and cutoffs.

[10] Contrary to standard practice in the gravity literature, we do not move the GDPs but instead move the population sizes to the left-hand side. Applying the former approach to our model would amount to assuming that wages are exogenous in the gravity estimation, which is not the case in general equilibrium (see Bergstrand, 1985, for an early contribution on this issue).

Expressions (34) and (35) are reminiscent of the constraints in Anderson and van Wincoop (2003), who argue that general equilibrium interdependencies need to be taken into account when conducting a counterfactual analysis based on the gravity equation.[11] One of our contributions is to go a step further by extending their approach to cope with endogenous wages and productivity. Note that expression (33) is similar to gravity equations that have been derived in previous models with heterogeneous firms. These models rely, however, on exogenous wages (Chaney, 2008; Melitz and Ottaviano, 2008) and also often disregard general equilibrium constraints when being estimated (Helpman *et al.*, 2008).[12]

4.2 Data and estimation procedure

To estimate the gravity equation system (33)–(35), we rely on aggregate bilateral trade flows X_{rs} and internal absorption X_{rr} for 10 Canadian provinces and 30 US states in 1993.[13] We further have geographical coordinates of the capitals, regional surface measures, and regional population sizes L_r in 1993 for these 40 regions. The latter are obtained from Statistics Canada and the US Census Bureau. For the specification of trade costs τ_{rs} we stick to standard practice by assuming that $\tau_{rs} \equiv d_{rs}^{\gamma} \mathrm{e}^{\theta b_{rs}}$, where d_{rs} stands for distance between r and s and is computed using the great circle formula. The internal distances are measured as $d_{rr} = (2/3)\sqrt{\text{surface}_r/\pi}$ like in Redding and Venables (2004).[14] The term b_{rs} is a border dummy valued 1 if r and s are not in the same country and 0 otherwise. With this specification of trade costs we relate our analysis to the vast literature on border effects (McCallum 1995), which has shown that regional trade flows are not only affected by physical distance, but also by the presence of the Canada-US border. The trade friction parameters γ and θ are to be estimated.

The estimation of the gravity equation system poses four main difficulties which we need to deal with. First, although we need to obtain the value of the shape parameter k, it cannot be structurally identified from the estimated parameters of the model. We thus proceed as follows. We choose an arbitrary initial value for k to estimate the gravity equation system as

[11] It is tempting to treat w_r, w_s, m_s^d and $\mu_r^{\max}$ as fixed effects in equation (33) by ignoring the general equilibrium constraints (34) and (35), as has been frequently done before. However, this approach cannot be used for a counterfactual analysis since the effect of the hypothetical decrease in trade frictions on the estimated fixed effects is not known. In our approach, the endogenous responses of wages and cutoffs are crucial when evaluating the counterfactual trade liberalization scenario below.

[12] One exception is Balistreri and Hillberry (2007), who allow regional incomes to adjust in response to trade liberalization. However, they do not consider firm heterogeneity in their analysis and take regional incomes as given in the gravity estimation.

[13] This publicly available data set has been widely used in the literature (see Anderson and van Wincoop, 2003; Feenstra 2004), which makes it easy to compare our results to existing ones.

[14] As a robustness check we also consider the alternative measure $d_{rr} = (1/4)\min_{s\neq r}\{d_{rs}\}$ as in Anderson and van Wincoop (2003). Results are little sensitive to that choice.

described below. Using the estimates thus obtained, as well as the chosen value of k, we then compute the productivity advantage of exporters from a random sample of 100,000 US firms drawn from the fitted productivity distributions of our model (see Section 5.3 and Appendix D for the procedure). We repeat this procedure for different values of k until our sample allows us to match the 33% productivity advantage of US exporters that is reported by Bernard *et al.* (2003) from 1992 Census of Manufactures data. This calibration yields $k = 9.5$, which we henceforth take as our benchmark.[15]

Second, there exists no data for $\mu_r^{\max}$ as it depends on the unobservables α, F_r and $m_r^{\max}$. To address this issue, we use the general equilibrium constraints (34) and (35). Ideally, we would plug data for $\mu_r^{\max}$ into these $2K$ constraints to solve for the $2K$ endogenous variables w_r and m_r^d. However, as $\mu_r^{\max}$ is unobservable, we rely instead on data for the K cutoffs m_s^d. This allows us to solve the $2K$ constraints (34) and (35) for theoretically consistent values of the $2K$ variables w_r and $\mu_r^{\max}$. Observe that, under the Pareto distribution, the domestic cutoff in each region is proportional to the inverse of the average productivity, i.e., $m_r^d = [(k+1)/k]\overline{m}_r$. We measure $\overline{m}_r$ by using the GDP per employee in Canadian dollars for each province and state in 1993, which is obtained from Statistics Canada and the US Census Bureau. Once we have computed the theory-consistent values of w_r from the general equilibrium constraints, we can evaluate the model fit by comparing computed with observed wages.[16]

Third, the estimates of the trade friction parameters γ and θ depend on w_r and $\mu_r^{\max}$, which depend themselves on the estimates of γ and θ. Put differently, the constraints (34) and (35) include the trade friction parameters, but to estimate the parameters of the gravity equation we need the solution to these constraints. We tackle this problem by estimating the gravity equation system iteratively. To summarize, our procedure consists of the following steps:

1. Given our specification of τ_{rs}, the gravity equation (33) can be rewritten in log-linear stochastic form as follows:

$$\ln\left(\frac{X_{rs}}{L_r L_s}\right) = -k\gamma \ln d_{rs} - k\theta b_{rs} + \zeta_r^1 + \zeta_s^2 + \varepsilon_{rs}, \tag{36}$$

   where all terms specific to the origin and the destination are collapsed into exporter and importer fixed effects ζ_r^1 and ζ_s^2; and where ε_{rs} is an error term with the usual properties. From (36), we obtain initial *unconstrained* estimates of the parameters $(\widehat{\gamma}', \widehat{\theta}')$.[17]

[15] As additional robustness checks we also consider $k = 3.6$, which is the value that has been used by Bernard *et al.* (2003), as well as $k = 8.5$ and 10.5. The qualitative results do not change with the value of k.

[16] We construct observed wages across provinces and states using hourly wage data from Statistics Canada and the Bureau of Labor Statistics. In order to match the unit of measurement of trade and GDP data, we compute average yearly wages in million Canadian dollars based on an average of 1930 hours worked yearly in Canada, and 2080 hours worked yearly in the US in 1993.

[17] Although we could choose the initial values for $(\widehat{\gamma}', \widehat{\theta}')$ arbitrarily, the fixed effects estimates provide a reason-

2. Using the initial estimates $(\widehat{\gamma}', \widehat{\theta}')$ and the observed cutoffs m_s^d in (34) and (35), we solve simultaneously for the equilibrium wages and the upper bounds $(\widehat{w}_r', \widehat{\mu}_r^{\max\prime})$.

3. We use the computed values $(\widehat{w}_r', \widehat{\mu}_r^{\max\prime})$ to estimate the gravity equation (33) as follows:

$$\ln\left(\frac{X_{rs}}{L_r L_s}\right) + k\ln\widehat{w}_r' - (k+1)\ln\widehat{w}_s' - \ln m_s^d + \ln\widehat{\mu}_r^{\max\prime}$$
$$= -\gamma k \ln d_{rs} + \gamma(k+1)\ln d_{rr} - k\theta b_{rs} + \varepsilon_{rs},$$

   which yields *constrained* estimates $(\widehat{\gamma}'', \widehat{\theta}'')$.

4. We iterate through steps 2 to 3 until convergence to obtain $(\widehat{\gamma}, \widehat{\theta})$ and $(\widehat{w}_r, \widehat{\mu}_r^{\max})$.

Last, we have to deal with the fact that bilateral trade flows among the 40 regions for which we have sufficient data are also affected by other out-of-sample regions and countries. This concern is particularly relevant in the context of a counterfactual analysis, since the trade creation and diversion effects of a hypothetical trade integration also feature general equilibrium repercussions with other trading partners. Fortunately, our gravity equation system allows us to take this issue into account, because we can include further regions and countries into the equilibrium constraints (34)–(35), even if we lack bilateral trade flow data for these observations.

Specifically, our full data set includes $K = 83$ areas, namely the 10 Canadian provinces, all 50 US states plus the District of Columbia, the 21 OECD members in 1993, and Mexico.[18] The distance, population and cutoff data for the 43 areas out of the gravity sample are defined in an analogous way as those for the areas in the gravity sample. For the rest of the world (ROW), we use OECD data (including for Mexico) to construct observed wages and average productivities. We use hourly wages and hours worked to construct the former, whereas the latter are obtained by converting 1993 hourly labor productivity in national currency into yearly figures (using hours worked) expressed in Canadian dollars.

In the iterative procedure, once we obtain initial unconstrained estimates for the structural parameters $(\widehat{\gamma}', \widehat{\theta}')$ using only the 40 regions from the 'in gravity sample', we can solve (34) and

able 'guess' for the starting values and allow for faster convergence of the iterative procedure. We experimented with different starting values and obtained the same estimates.

[18]See Tables 3 and 5 for the list of the 40 Canadian and US regions used in the gravity equation ('in gravity sample') and for the 21 regions used only in the general equilibrium constraints ('out of gravity sample'). Because of their extremely small population sizes we have excluded Yukon, Northwest Territories and Nunavut from the analysis. The rest of the world consists of Australia, Austria, Belgium-Luxembourg, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, UK and Mexico, which together with Canada account for the lion's share of US intra-industry trade (and 66.5% of total US exports and 64.7% of total US imports in 1993).

(35) for the wages and upper bounds $(\widehat{w}'_r, \widehat{\mu}_r^{\max\prime})$ for the full sample of areas.[19] The solutions for the 40 'in gravity' regions, which are affected by the general equilibrium interactions among *all* trading partners, are plugged back into (33) and the estimation proceeds as described above. The resulting final estimates of trade frictions $(\widehat{\gamma}, \widehat{\theta})$ and of wages and upper bounds $(\widehat{w}_r, \widehat{\mu}_r^{\max})$ for all 83 areas are consistent with theory as they take into account all the equilibrium information of the model. Using this information we can then retrieve the fitted values of bilateral trade flows $\widehat{X}_{rs}$ for *all* pairs of areas, even for those not in the gravity sample.

4.3 Estimation results

Our estimation results for the gravity equation system are summarized in Table 1. Column 1 presents the benchmark case with $K = 83$ areas and $k = 9.5$, whereas Columns 2-6 contain alternative specifications used as robustness checks.

Insert Table 1 about here.

As can be seen from Column 1 in Table 1, all coefficients have the correct sign and are precisely estimated. In our benchmark case, the estimated distance elasticity is -1.4195, which implies that $\widehat{\gamma} = 0.1494$. The border coefficient estimate is -1.5951, which implies that $\widehat{\theta} = 0.1679$. Note that our estimated border coefficient is virtually identical to the one of -1.59 obtained by Anderson and van Wincoop (2003). However, as shown later, the impacts of a border removal on trade flows differ substantially once endogenous wages and firm selection are taken into account.

Columns 2-4 report results for different values of k. Column 5 presents results when we use the alternative measure for internal distance as proposed by Anderson and van Wincoop (2003). In that case, we exclude the ROW as this distance measure is not really appropriate in an international context. Column 6 finally presents results obtained when we exclude the 49 zero trade flows in the sample. The coefficient of the border dummy varies only little across the different specifications, which mostly display adjusted R^2 values in excess of 0.9.

Last, as stated in the foregoing, we can get an additional idea of the goodness of fit of our model by comparing the relative wages $\widehat{w}_r$, obtained from the general equilibrium constraints, with observed relative wages. In our benchmark case, the correlation is 0.62 when including the ROW, and 0.76 when focusing only on Canadian provinces and US states. Thus, the predicted relative wages match observed ones fairly well.

[19] Even when focusing on the 40 regions, we still have to deal with 49 zero trade flows out of 1600 observations. Since there is no generally agreed-upon methodology to deal with this problem (see, e.g., Anderson and van Wincoop, 2004; Disdier and Head, 2008), we include a dummy variable for zero flows in the regressions. Note that our zeros are unlikely to be 'true zeros', as this would imply no aggregate manufacturing trade between several US states (see Helpman *et al.* 2008 on the treatment of 'true zeros'). As a robustness check, we estimate the system by dropping the 49 zero flows.

5 Counterfactual analysis

Having estimated the gravity equation system, we now turn to a counterfactual analysis in the spirit of Bernard *et al.* (2003) and Del Gatto *et al.* (2006) and investigate the potential impacts of fully eliminating all trade distortions generated by the Canada-US border. This hypothetical trade integration would induce various effects, both at the aggregated and at the disaggregated level. We start with the analysis of how regional trade flows would change, and how key regional economic aggregates such as wages, productivity, markups, as well as welfare would be affected. Afterwards we turn to a firm-level analysis which traces out some consequences of this trade integration for the productivity distributions, export sales distributions and export intensities.

5.1 The impacts on regional trade flows

To quantify the induced changes in regional trade flows we define *bilateral border effects* as the ratio of trade flows from r to s in a borderless world to those in a world with borders:

$$B_{rs} \equiv \frac{\widetilde{X}_{rs}}{\widehat{X}_{rs}} = \mathrm{e}^{k\widehat{\theta}b_{rs}} \left(\frac{\widetilde{w}_s/\widetilde{w}_r}{\widehat{w}_s/\widehat{w}_r}\right)^{k+1} \left(\frac{\widetilde{w}_r}{\widehat{w}_r}\right) \left(\frac{\widetilde{m}_s^d}{m_s^d}\right)^{k+1}, \tag{37}$$

where variables with a tilde refer to values in a borderless world. The value of B_{rs} can be computed as follows. First, we use the estimated wages $\widehat{w}_r$ and the observed cutoffs m_s^d in the presence of the border to obtain the relevant information for the initial fitted trade flows $\widehat{X}_{rs}$ in (37). Second, holding the shape parameter k, the estimated upper bounds $\widehat{\mu}_r^{\max}$, and trade frictions $(\widehat{\gamma}, \widehat{\theta})$ constant, we solve (34) and (35) by setting $b_{rs} = 0$ for all r and s. This yields the wages $\widetilde{w}_r$ and the cutoffs $\widetilde{m}_s^d$ that would prevail in a borderless world. Plugging these values into (37), we obtain $61 \times 61 = 3721$ bilateral border effects, each of which gives the change in the trade flows from r to s after the border removal.[20]

Insert Table 2 about here.

The bilateral border effects B_{rs} are typically greater than one when regions r and s are in different countries. The reason is that exports from region r to region s partly substitute for domestic sales as international trade frictions are reduced. For analogous reasons, the values of B_{rs} are typically less than one when r and s are in the same country. Table 2 provides some descriptive statistics on the series of computed bilateral border effects for the various specifications given in Table 1. One can see that the different specifications yield almost identically distributed and strongly correlated bilateral border effects.

[20]Strictly speaking, we could compute 83×83 bilateral border effects, but in the remainder of this paper we concentrate on the effects of the hypothetical border removal for the 61 regions in Canada and in the US. We do not report the induced changes of the border removal for the other countries, although the ROW is included in the underlying general equilibrium constraints.

5.1.1 Decomposing bilateral border effects

What drives bilateral border effects? As can be seen from expression (37), B_{rs} can be decomposed into four components: (i) a *pure border effect* $\mathrm{e}^{k\widehat{\theta} b_{rs}}$; (ii) a *relative wage effect* $\Delta(w_s/w_r) \equiv [(\widetilde{w}_s/\widetilde{w}_r)/(\widehat{w}_s/\widehat{w}_r)]^{k+1}$; (iii) an *absolute wage effect* $\Delta w_r \equiv \widetilde{w}_r/\widehat{w}_r$; and (iv) a *selection effect* $\Delta m_s^d \equiv \left(\widetilde{m}_s^d/m_s^d\right)^{k+1}$. Table 3 provides an example of this decomposition.

Insert Table 3 about here.

The left half of Table 3 lists the components of B_{rs} for exports from British Columbia (BC) to all Canadian provinces and US states. Consider, for example, the bilateral border effect with Washington (WA). The pure border effect corresponds to the predicted change in bilateral trade flows that would prevail if endogenous changes in wages and cutoffs were not taken into account. In this example, it states that the value of exports from BC to WA would rise by a factor of 4.9290. Yet, the wage in BC rises relative to that in WA after the border removal, and BC firms thus become less competitive in WA due to relatively higher production costs. This change is captured by the relative wage effect which decreases the export value by some 40% in this case. The absolute wage effect, on the contrary, raises BC exports by a factor of 1.0584 as the higher wage is reflected in the higher prices. When taken together, these two wage effects reduce the pure bilateral border effect from BC to WA by about 36.6% (as $0.5991 \times 1.0584 = 0.6340$). Put differently, neglecting the endogenous reaction of wages to the border removal leads to overstating the bilateral border effect by about 36.6%! Finally, there is a selection effect. The border removal reduces the cutoff marginal labor requirement that firms need to match to survive in WA. In other words, trade integration induces tougher selection and makes it harder for BC firms to sell in WA. This selection effect decreases the export value by a factor of 0.8359, i.e., it further reduces the bilateral border effect by 16.4%. Putting together the different components, the bilateral border effect is then given by $4.9290 \times 0.5991 \times 1.0584 \times 0.8359 = 2.6126$, which is almost half the size of the pure border effect without endogenous wage and productivity responses.

Trade flows between regions within the same country would also be affected by the border removal. Consider, for example, exports from BC to Ontario (ON). There is, of course, no pure border effect for this intranational trade flow, but due to the endogenous changes in wages and cutoffs we find a bilateral border effect equal to $1 \times 0.9350 \times 1.0584 \times 0.3581 = 0.3544$. The border removal thus reduces the value of exports from BC to ON by almost 65%. Note that wages in BC rise relative to those in ON, which provides BC firms with a cost disadvantage and per se decreases exports to ON by around 6.5%. The main effect at work in this case, however, is the tougher selection in ON due to the increased presence of more productive US firms. This makes it much harder for BC firms to sell in ON and reduces the export value by more than 60%. The induced selection effects are also visible in the bilateral border effect of BC with itself.

The value of local sales by BC firms drops by 67%, which is caused by the now much tougher selection in that market.

The right half of Table 3 provides the B_{rs} for exports from New York (NY) to all Canadian provinces and US states. Consider, for example, exports from NY to Québec (QC), which would rise by a factor of $4.9290 \times 1.5166 \times 1.0018 \times 0.4277 = 3.2026$. In this example wages in QC rise relative to those in NY, which gives NY firms a relative cost advantage and per se boosts export values, whereas the tougher selection in QC makes market penetration by NY firms more difficult, which per se reduces export values. Exports from NY to California (CA) change little after the border removal ($1 \times 0.9776 \times 1.0018 \times 0.9870 = 0.9667$). The explanation is that CA is large and far away from the border, so that little additional selection is induced there, while the wage in NY rises only slightly when compared to that in CA. On balance, these small changes affect trade flows only marginally. There is also only a slight reduction of local sales of NY firms (by about 3.4%) since the wage and the cutoff in NY are virtually unaffected by the relatively small Canadian economy.

To sum up, these examples show that any computation of border effects should take into account changes in wages and productivity. Neglecting these endogenous adjustments leads to biased predictions for changes in international and intranational trade flows in a borderless world, and that bias can be substantial.

5.1.2 Regional and national border effects

In his seminal work on border effects, McCallum (1995) finds that, conditional on regional GDP and distance, trade between Canadian provinces exceeds by roughly 22 times trade between Canadian provinces and US states. Anderson and van Wincoop (2003) argue that this estimate is substantially upward biased due to the omission of general equilibrium constraints. They find that, on average, the border increases trade between Canadian provinces 'only' by a factor of 10.7 when compared to trade with US states. The corresponding number for the US is 2.24.

What does our approach, which adds endogenous wages and firm selection to the analysis, predict for the impact of the border removal on *overall* Canadian and US trade flows? To evaluate this, we need to aggregate bilateral border effects first at the regional and then at the national level. We define the *regional border effect* for Canadian province r as follows:

$$B_r \equiv \frac{\sum_{s\in \text{US}} \widetilde{X}_{rs} / \sum_{s\in \text{US}} \widehat{X}_{rs}}{\sum_{s\in \text{CA}} \widetilde{X}_{rs} / \sum_{s\in \text{CA}} \widehat{X}_{rs}} = \frac{\sum_{s\in \text{US}} \lambda_{rs}^{\text{US}} B_{rs}}{\sum_{s\in \text{CA}} \lambda_{rs}^{\text{CA}} B_{rs}},$$

where $\lambda_{rs}^{\text{US}} = \widehat{X}_{rs} / \sum_{s\in \text{US}} \widehat{X}_{rs}$ and $\lambda_{rs}^{\text{CA}} = \widehat{X}_{rs} / \sum_{s\in \text{CA}} \widehat{X}_{rs}$ are the fitted trade shares. The numerator is the trade weighted average of international bilateral border effects, whereas the denominator is the trade weighted average of the intranational B_{rs}. The regional border effects

B_r thus summarize by how much cross-border trade would rise as compared to domestic trade for each Canadian province and US state. As can be seen from Column 2 in Table 4, regional border effects are naturally larger for Canadian provinces than for US states. As can be further seen, there is also some intranational variation, with border effects ranging from a low of 6.21 in Prince Edward Island to a high of 7.65 in Ontario for Canada; and from a low of 3.12 in Washington to a high of 3.46 in Maine for the US.

Insert Table 4 about here.

Turning to the national level, it can be easily verified that the *national border effect* for Canada can be simplified as follows:

$$B_{\mathrm{CA}} \equiv \frac{\sum_{r\in\mathrm{CA}}\sum_{s\in\mathrm{US}}\widetilde{X}_{rs}/\sum_{r\in\mathrm{CA}}\sum_{s\in\mathrm{US}}\widehat{X}_{rs}}{\sum_{r\in\mathrm{CA}}\sum_{s\in\mathrm{CA}}\widetilde{X}_{rs}/\sum_{r\in\mathrm{CA}}\sum_{s\in\mathrm{CA}}\widehat{X}_{rs}} = \frac{1}{K_{\mathrm{CA}}}\sum_{r\in\mathrm{CA}} B_r,$$

where $K_{\mathrm{CA}} = 10$ is the number of Canadian provinces. An analogous definition applies to the US. Using this approach we obtain $B_{\mathrm{CA}} = 6.8973$ and $B_{\mathrm{US}} = 3.2767$.[21]

Endogenous wages are crucial for explaining the difference between our results and those of Anderson and van Wincoop (2003). The border removal breaks the zero expected profit conditions in all regions. To recover zero expected profits, wages and productivity in Canada must rise relatively to those in the US, as will be shown in more detail in the next subsection. In a fixed-wage model such as Anderson and van Wincoop (2003), the measured Canadian border effect is therefore overstated, because the export dampening effects of the higher relative wage are not taken into account. The measured US border effect is understated for analogous reasons. This may explain why Anderson and van Wincoop (2003) find more dissimilar national border effects (10.7 for Canada and 2.24 for the US). The change in productivity also affects national border effects. The border removal intensifies competition in all markets and, therefore, bilateral border effects for all pairs of regions are reduced. A model that does not take such effects into account will therefore incorrectly assess the magnitude of the impact of the border removal on changes in regional trade flows.

[21] Strictly speaking, our definition of the national border effect differs slightly from that of Anderson and van Wincoop (2003). When using their definition in terms of geometric means (see Feenstra, 2004), we obtain 6.88 for Canada and 3.28 for the US. The advantage of our definition is that it precisely measures the (multiplicative) impact of the border removal on trade flows. Let us emphasize here that our primary objective is not to 'downsize the border effect' but rather to understand how it is affected by endogenous wages and cutoffs. There are many competing explanations for why measured border effects may be too large (see, e.g., Yi, 2009, for the impacts of trade in intermediate goods).

5.2 The impacts on key economic aggregates

We now investigate the predictions of our model on how trade integration affects other key economic aggregates at the regional level. More specifically, we describe the impacts of the border removal on wages, average productivity, average markups, the mass of varieties produced and consumed, and welfare.

Insert Table 5 about here.

In Table 5, we report the predicted regional changes for these variables after the border removal. It is clear that these endogenous changes must depend on the primitives of the model, in particular the vector of 1993 regional populations, and the matrix of pre- and post border-removal trade costs. Despite the fact that no explicit form can be obtained for the function that describes how exogenous regional characteristics map into changes in endogenous variables, we can, by using a simple OLS regression, approximate such a function by a linear relationship. Specifically, we regress endogenous regional changes on two crucial regional characteristics: *geography* and *size*. The former dimension is captured by the distance of region r to the closest foreign region, and the latter by the population size L_r. This simple regression allows us to understand if the border removal would mainly affect smaller regions and/or regions closer to the border. Furthermore, we investigate if regional changes differ systematically between Canada and the US by including a dummy for Canada. Table 6 reports the results. Note that the R^2 values consistently exceed 0.84, regardless of the specification and of the dependent variable. This suggests that our simple regression approach is in fact very useful for understanding the variation of the implied regional changes.[22]

Starting with the wage changes (listed in Column 1 of Table 5), the first estimated specification reveals (top part of Column 1 in Table 6) that the border removal favors wage convergence between the two countries. Pre-border wages in Canada are in fact lower than those in the US while the Canada dummy is positive, sizeable, and highly significant. As the gains from better market access are much larger for the Canadian provinces, wages generally rise relative to those in US states.[23] For Canadian provinces the wage changes range from 1.78% for Prince Edward Island to 5.84% for British Columbia, whereas those in the US are much smaller and can go either way. Investigating the overall variation of these changes, we find that regions further

[22] One could consider adding *technology* (local technological possibilities) as a third regional characteristic, which would be captured by $1/\widehat{\mu}_r^{\max}$. Recall, however, that $\widehat{\mu}_r^{\max}$ has been computed from the general equilibrium constraints, and may therefore not be strictly exogenous (in an econometric sense) to the predicted regional changes. Furthermore, when using $1/\widehat{\mu}_r^{\max}$ in the regressions we always obtain an insignificant coefficient for this variable, while the R^2 value does not improve. We hence drop it in what follows.

[23] This result is in line with result (v) derived analytically in our simple two-region setting in Section 3.4. Note that all wages are expressed relative to that in Alabama, which we set to one by choice of numéraire.

away from the border tend to experience smaller wage increases (the elasticity being -0.4323) and that there is no significant relationship with population size. These findings are likely to be driven by the fact that the 10 Canadian provinces experience the largest increases but display little variation in their distance to the border and in their size (at least when compared to US states). In a second specification (middle part of Column 1 in Table 6) we therefore consider interaction terms of our proxies for *geography* and *size* with a US-dummy, and focus interpretation on US states. Within the US, regions closer to the border also have stronger wage increases, but the elasticity is a bit smaller than before ($-1.0115 + 0.7734 = -0.2381$). This is consistent with the fact that for the US, being initially the much larger market, proximity to the new market opportunities matters less than for Canada. This second specification further shows that wage increases are stronger in smaller US states, thus favoring wage convergence across states (the elasticity with respect to size is $0.6001 - 0.6749 = -0.0748$). Finally, we explore whether *geography* or *size* contribute more to explaining regional variation in wage changes across the US. To do so we report (bottom part of Column 1 in Table 6) beta coefficients obtained from the second specification.[24] The beta coefficients reveal that distance to the border is the more important determinant of regional wage changes. In fact, this measure of *geography* is almost twice as important as that of *size*.

Insert Table 6 about here.

Looking at other endogenous changes reveals a similar pattern. Predicted cutoff changes are negative for *all* Canadian provinces and US states, which shows that removing the border induces tougher selection and increases average productivity.[25] The average productivity gain in Canada would be 5.71%, whereas in the US it is much smaller (0.3%). Clearly, since Canada is the smaller economy with less selection prior to the border removal, trade integration has more substantial consequences there. Still, across US states we find stronger productivity gains in smaller regions and in regions closer to the border, with *geography* adding more to the understanding of the regional variation than *size*. Average markups fall in all regions, but the reductions in Canada (where markups fall between 1.78% and 5.19%) are more substantial than in the US (0.12% to 1.08%). As we show in the next subsection, these pro-competitive effects arise because the border removal increases the share of firms engaged in cross-border transactions. More firms compete in each market after the border removal, which puts downward pressure on markups. Last, we find that the number of firms decreases everywhere due to tougher selection, but this exit is smaller than the increase in imported varieties so that consumption diversity expands in

[24] The value of -0.3390 for the beta coefficient on size means that a one standard deviation increase of regional size lowers the regional wage increase after the border removal by 33.90%.

[25] Quite naturally, the hypothetical border removal between Canada and the US hurts the ROW countries, who see their cutoffs marginally increase (results not reported here).

all Canadian provinces and US states.

To see that tougher selection or more diversity in consumption map into welfare gains, notice that since $\mathrm{e}^{-\alpha q_{sr}(m)} = p_{sr}(m)/p_r^d$ by (14), the indirect utility in region r is given by

$$U_r = \sum_s N_s^E \int_0^{m_{sr}^x} \left[1 - \mathrm{e}^{-\alpha q_{sr}(m)}\right] \mathrm{d}G_s(m) = N_r^c \left(1 - \frac{\overline{p}_r}{p_r^d}\right).$$

Using expression (17), one can verify that $\overline{p}_r = [k/(k+1)]p_r^d + \alpha w_r/N_r^c$, which allows us to express the indirect utility as $U_r = N_r^c/(k+1) - \alpha/(\tau_{rr}m_r^d)$. Since N_r^c is defined as in (20), and making use of the fact that expression (27) holds in equilibrium, we can rewrite the indirect utility as follows:

$$U_r = \left[\frac{1}{(k+1)\kappa_3} - 1\right] \frac{\alpha}{\tau_{rr}m_r^d}. \tag{38}$$

Hence, welfare is inversely proportional to the cutoff m_r^d. Alternatively, the equilibrium utility can be written as $U_r = [1/(k+1) - \kappa_3]N_r^c$., i.e., welfare changes in region r are proportional to changes in the mass of varieties available for consumption.[26] These effects are again more pronounced in Canada than in the US, and welfare gains in the US are strongest in small regions close to the border.

5.3 Firm-level analysis

In the foregoing section, we examined the impacts of trade integration at the regional level. Though fundamental for our understanding of the mechanisms at work, this approach does not provide enough insights on firm-level issues. We hence now investigate the impacts of our counterfactual trade liberalization scenario on firms' productivity distributions, their export sales distributions, and their export intensities. To this end, we generate two sets of random samples of 110,000 firms: 10,000 Canadian and 100,000 US firms, before and after removing the border. To make the samples representative, we draw firms in each region both before and after removing the border in proportion to its share of surviving firms (see Appendix D for additional details). We first use the generated samples to calibrate the shape parameter k of the Pareto distribution. We compute for each firm its 'empirical' productivity (based on revenue instead of the usually unavailable physical output). Our sample of Canadian exporters before removing the border delivers a revenue-based productivity advantage of exporters over non-exporters of 20.28%. The corresponding productivity advantage for US exporters in the presence of the border is 32.5%,

[26]Removing the border decreases all the domestic cutoffs, thus reducing domestic consumption diversity, yet increases all export cutoffs, thereby raising import consumption diversity. The net effect is to expand consumers' choice sets everywhere, which maps into welfare gains. Note that, as in Eaton and Kortum (2002), we report cardinal percentage changes in welfare. These figures are therefore sensitive to a monotonically increasing transformation of the utility function.

which matches the 33% advantage computed by Bernard *et al.* (2003) from 1992 US Census of Manufactures data. This vindicates our choice of $k = 9.5$.

Our Canadian sample delivers a share of exporters of about 7.08% before removing the border. The corresponding figure for the US is 1.95%. These are quite small numbers which convey the message that exporting firms are rare in the economy. Note that the US figure fits decently with the one reported by Bernard *et al.* (2009), who document that only 2.6% of all US firms reported exporting anything at all in 1993.[27]

5.3.1 Productivity distribution by export status and export sales distribution

Figure 1 depicts the (kernel-smoothed) density distributions of productivity $1/m$ for both exporters and non-exporters in Canada (left panel) and in the US (right panel) before the border removal.

Insert Figure 1 about here.

It should firstly be noted that the national productivity distributions of producers in our sample are clearly not Pareto, although all the underlying regional distributions are.[28] This results from the aggregation of heterogeneous regions, more specifically because of the positive relationship among size, productivity, and entry: smaller regions have in general higher cutoffs (less productive firms), and also fewer firms. Adding increasingly big regions with more productive firms to the national distribution then generates the initially upward sloping part. Secondly, as can be seen, exporters' productivity distribution lies to the right of the analogous distribution for domestic firms. This shows that high productivity firms, in general, tend to operate in export markets. However, there is some sizeable overlap in the two distributions, especially in the US case. The reasons for this overlap are regional heterogeneity in wages and market access, which implies that *a firm with a draw* $1/m$ *may be an exporter in some regions, but not in others.* Put differently, regional characteristics matter for the export status of a firm.[29] This applies more to the US as firms there face more heterogeneity in terms of accessibility to the Canadian market than the other way round. As can also be seen from Figure 1, both the Canadian and

[27]The share of exporters in a US state is defined as the share of firms selling to at least one Canadian province or one country in the ROW. Formally, it is given by $G_r(\max_{s\in \text{CA,ROW}}\{m^x_{rs}\})/G_r(\max_s\{m^x_{rs}\})$. The share of US exporters is then computed as the weighted average of the states' exporter shares, where the weights are proportional to the mass of surviving firms in each state (see Appendix D). All figures for the Canadian provinces are computed in an analogous way.

[28]The Kolmogorov-Smirnov test strongly rejects the null of a Pareto for both Canada and the US.

[29]It has been recognized that national distributions of exporters and domestic firms overlap even within narrowly defined industries. Eaton *et al.* (2009) introduce random firm-and-destination specific shocks to both demand and market entry costs to account for this overlap. As shown by our results, considerations related to wage differentials and the geography of trading partners are also able to cope with this feature.

the US distributions are quite 'spiky', which also stems from the regional heterogeneity that is aggregated into the overall distributions.

Insert Table 7 about here.

Table 7 shows that both in Canada and in the US, the large bulk of exporters sells little to nothing in the export markets, whereas a few firms have much higher export intensity and export sales. In Canada, export intensities of some firms exceed 90% (but fall short of 100%), whereas the US figures are lower, with the largest sellers earning slightly less than 70% of their revenue from exports. These different export intensities do, of course, reflect the size asymmetries between the two countries and the relative importance of the international market for their economies. As can also be seen from the top half of Table 7, our model can partly replicate the skewness of the empirical distribution of export intensities. The bottom half of Table 7 shows that our model delivers export sales distributions that are skewed towards the most productive firms. Although our model cannot replicate the amazing degree of skewness observed in the US distribution, as reported by Bernard *et al.* (2009) using 1993 US firm-level data, we still find that the top 25% of exporters account for about 92.65% of total US export sales, whereas the top 50% of exporters account for almost 98% of all export sales. The Canadian figures are of roughly similar magnitude.

5.3.2 Impacts of removing the border

Finally, we single out three important firm-level impacts of removing the border. First, as expected, *the share of exporters increases substantially.* The exporter share in Canada after the border removal is 33.51%, whereas the corresponding figure for the US is 5.89%. Clearly, removing the border increases the share of firms engaged in cross-border business by lowering trade frictions between the two countries. Furthermore, the productivity advantage of exporters shrinks to 9.53% for Canadian firms and to 21.56% for US firms.

Insert Figure 2 about here.

Second, *trade liberalization affects the productivity distributions in both countries.* Figure 2 depicts the (kernel smoothed) productivity density distributions of producers (top panel) and sellers (bottom panel) in both Canada (left panel) and in the US (right panel) before and after the border removal. Note that although selection works as a simple left-truncation of the productivity density in all regions, the observed country-wise effects are more complex. In fact, the Kolmogorov-Smirnov test rejects the null that the pre- and post border-removal distributions (for both Canada and the US) are a truncated version of each other. Also note from the bottom panel of Figure 2 that the productivity distribution of the firms selling in the Canadian market radically changes as the border is removed. The least productive Canadian firms go out of business, whereas the distribution of the surviving sellers shifts to the right in a non-trivial way. The US distribution barely changes. There is indeed some right-shift in the productivity distribution of surviving firms in the US, but that shift is very small. The reason is that Canadian firms are on average less productive and fewer in numbers, which implies that their entry into the US market will not spark substantial churning and market share reallocations across firms.

Third, *trade liberalization affects export intensities and polarization.* As can be seen from the bottom half of Table 7, removing the border slightly decreases the polarization of export sales of Canadian firms, whereas that of the most productive US firms increases. Firms' export intensities increase in both countries. Interestingly, some low productivity firms from British Columbia and Québec now earn a full 100% of revenue from exports. These firms can export to some US states in a borderless world, whereas they cannot sell anymore in their domestic market because of higher wages and increased competition. The border removal leads to wage convergence between Canada and the US (see above) but Canadian wages are still lower than in the US. As the firms from BC and QC substantially benefit from lower trade frictions, they now find it profitable to enter the high wage US market where they have a cost advantage. These less efficient Canadian sellers show up on the left-side of the productivity distribution of surviving firms in the US.[30] Such an effect can only be captured in a model featuring factor price differences that allow less productive firms to penetrate the export market and survive there.

6 Conclusions

We have developed a new and highly tractable trade model with monopolistic competition and heterogeneous firms. In contrast to previous contributions, it features variable demand elasticity and endogenous price-cost margins in a full-fledged general equilibrium setup with multiple asymmetric regions. Both wages and productivity are endogenous, need not be equalized across regions, and respond to trade integration. Given these properties, our framework provides a basis for empirical work and counterfactual analysis. To illustrate this, we have derived a gravity equation system and structurally estimated its parameters from 1993 regional trade data. Although our iterative procedure requires some customized programming, it has the advantage of yielding estimates that take into account all general equilibrium effects of the model. Contrary to the simpler fixed effects approach to gravity, we can take into account the endogenous responses of key regional variables when conducting policy experiments.

We have focused on one such counterfactual analysis and simulated the regional and firm-level effects of removing all trade distortions generated by the Canada-US border. The first message

[30] Although barely visible from the bottom right panel of Figure 2 due to the scaling, there is indeed entry of low productivity sellers in the US after removing the border.

from this simulation exercise is that endogenous wage and productivity responses are crucial for understanding the effects of trade integration: ignoring endogenous wages and selection effects systematically biases all bilateral border effects as well as the national border effects for Canada and the US. Our counterfactual analysis also reveals that the border removal would lead to non-negligible aggregate productivity gains, lower average markups, greater consumption diversity, and higher welfare in all provinces and states, but particularly so in Canada. Within the US there is substantial regional variation due to *geography* and *size*: smaller states closer to the border would gain more.

The second message is that our framework can replicate several stylized facts from firm-level studies in international trade, and that it can shed new light on some recent issues from that literature. Using a generated firm sample drawn from the estimated aggregate model, we have shown that the export sales and the export intensity distributions delivered by our model fit reasonably well with the empirical facts. Furthermore, the internal geography of countries provides a simple explanation for the observed overlap of domestic firms' and exporters' productivity distributions. Finally, we have found that at the country level there is no simple relationship between trade integration and its impact on the productivity distribution of firms. While trade integration leads to a simple left-truncation at the regional level, the aggregate implications at the national level are more complex.

To conclude, observe that our model allows for many extensions and alternative counterfactuals. A first extension would be to endogenize regional populations by taking into account interregional and international migration, which would nicely fit with our focus on North America. Taking this road would, as a by-product, partly bridge the gap between trade models with heterogeneous firms and the 'new economic geography' literature. A second extension would be to switch to a broader international setting including North-South trade. Given the absence of FPE, our model should be especially useful for this exercise once extended to cope with differential factor proportions and the presence of intermediate inputs. A third possible extension would be to show how revenue-based productivity depends upon firms' location, the wages they pay and the spatial structure of the demand they face. This may seem worthwhile since it has been recently emphasized that selection is likely to occur on profitability instead of narrowly defined physical productivity (Foster *et al.*, 2008). Locational heterogeneity matters in explaining firm selection and export status in our model because firms face different wages and have different market access depending on which region they are located in. A preliminary analysis using our generated sample of US firms and simple OLS regression reveals that idiosyncratic locational components explain about 13% of the variation of revenue-based firm productivity, whereas physical productivity accounts for the remaining 87%. This finding again emphasizes that location matters, and certainly calls for further research.

Acknowledgements. We thank Thomas Chaney, Robert C. Johnson, Alain Gabler, Marc Melitz, Gianmarco Ottaviano, Esteban Rossi-Hansberg, as well as participants at numerous seminars and conferences for helpful comments and suggestions. Kristian Behrens gratefully acknowledges financial support from the European Commission under the Marie Curie Fellowship MEIF-CT-2005-024266, from Université du Québec à Montréal (PAFARC), and from FQRSC Québec (Grant NP-127178). Yasusada Murata gratefully acknowledges financial support from the Japan Society for the Promotion of Science (17730165) and MEXT.ACADEMIC FRONTIER (2006-2010). Part of this paper was written while Kristian Behrens, Giordano Mion and Jens Südekum were visiting ARISH (Nihon University), and while Yasusada Murata and Jens Südekum were visiting CORE (Université catholique de Louvain) and UQAM. We gratefully acknowledge the hospitality of these institutions. The usual disclaimer applies.

References

[1] Alvarez, F.E. and R.E. Lucas (2007) General equilibrium analysis of the Eaton-Kortum model of international trade, *Journal of Monetary Economics* 54, 1726-1768.

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

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

[4] Aw, B., S. Chung and M. Roberts (2000) Productivity and turnover in the export market: Micro-level evidence from the Republic of Korea and Taiwan (China), *World Bank Economic Review* 14, 65-90.

[5] Badinger, H. (2007) Has the EU's Single Market Programme fostered competition? Testing for a decrease in mark-up ratios in EU industries, *Oxford Bulletin of Economics and Statistics* 69, 497-519.

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

[7] Behrens, K. and Y. Murata (2007) General equilibrium models of monopolistic competition: A new approach, *Journal of Economic Theory* 136, 776-787.

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

[9] Bernard, A.B. and J.B. Jensen (1995) Exporters, jobs and wages in US manufacturing, 1976-87, *Brookings Papers on Economic Activity, Microeconomics*. Washington DC.

[10] Bernard, A.B. and J.B. Jensen (1999) Exceptional exporter performance: Cause, effect, or both?, *Journal of International Economics* 47, 1-25.

[11] Bernard, A.B., J. Eaton, J.B. Jensen and S. Kortum (2003) Plants and productivity in international trade, *American Economic Review* 93, 1268-1290.

[12] Bernard, A.B., J.B. Jensen, S.J. Redding and P.K. Schott (2007a) Firms in international trade, *Journal of Economic Perspectives* 21, 105-130.

[13] Bernard, A.B., J.B. Jensen and P.K. Schott (2009) Importers, exporters and multinationals: A portrait of firms in the US that trade goods. In T. Dunne, J.B. Jensen and M.J. Roberts (eds.), *Producer Dynamics: New Evidence from Micro Data*, NBER Book Series Studies in Income and Wealth. University of Chicago Press.

[14] Bernard, A.B., S.J. Redding and P.K. Schott (2007b) Comparative advantage and heterogeneous firms, *Review of Economic Studies* 73, 31-66.

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

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

[17] Clerides, S.K., S. Lach and J.R. Tybout (1998) Is learning by exporting important? Microdynamic evidence from Colombia, Mexico, and Morocco, *Quarterly Journal of Economics* 113, 903-947.

[18] Corless, R.M., G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey and D.E. Knuth (1996) On the Lambert W function, *Advances in Computational Mathematics* 5, 329-359.

[19] Del Gatto, M., G. Mion and G.I.P. Ottaviano (2006) Trade integration, firm selection and the cost of non-Europe, CEPR Discussion Paper #5730.

[20] Disdier, A.-C. and K. Head (2008) The puzzling persistence of the distance effect on bilateral trade, *Review of Economics and Statistics* 90, 37-48.

[21] Eaton, J. and S. Kortum (2002) Technology, geography and trade, *Econometrica* 70, 1741-1779.

[22] Eaton, J., S. Kortum and F. Kramarz (2009) An anatomy of international trade: Evidence from French firms, CEPR Discussion Paper #7111.

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

[24] Foster, L., J.C. Haltiwanger and C. Syverson (2008) Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?, *American Economic Review* 98, 394-425.

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

[26] Krugman, P.R. (1979) Increasing returns, monopolistic competition, and international trade, *Journal of International Economics* 9, 469-479.

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

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

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

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

[31] Pavcnik, N. (2002) Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants, *Review of Economic Studies* 69, 245-276.

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

[33] Syverson, C. (2004) Market structure and productivity: A concrete example, *Journal of Political Economy* 112, 1181-1222.

[34] Trefler, D. (2004) The long and short of the Canada-US free trade agreement, *American Economic Review* 94, 870-895.

[35] Tybout, J. (2003) Plant and firm-level evidence on new trade theories. In: Choi, E.K. and J. Harrigan (eds.) *Handbook of International Trade*. Oxford: Basil-Blackwell, pp. 388-415.

[36] Yi, K.-F. (2009) Vertical specialization and the border effect puzzle, *American Economic Review*, forthcoming.

Appendix A: Proofs and computations

A.1. Derivation of (8) and properties of W. Using $p^d = m^d w$, the first-order conditions (6) can be rewritten as

$$\ln\left[\frac{m^d w}{p(m)}\right] = 1 - \frac{mw}{p(m)}.$$

Taking the exponential of both sides and rearranging terms, we have

$$\mathrm{e}\frac{m}{m^d} = \frac{mw}{p(m)}\mathrm{e}^{\frac{mw}{p(m)}}.$$

Noting that the Lambert W function is defined as $\varphi = W(\varphi)\mathrm{e}^{W(\varphi)}$ and setting $\varphi = \mathrm{e}m/m^d$, we obtain $W(\mathrm{e}m/m^d) = mw/p(m)$, which implies $p(m)$ as given in (8). The derivations of $q(m)$ and $\pi(m)$ then follow straightforwardly.

Turning to the properties of the Lambert W function, we clearly see that $\varphi = W(\varphi)\mathrm{e}^{W(\varphi)}$ implies that $W(\varphi) \geq 0$ for all $\varphi \geq 0$. Taking logarithms on both sides and differentiating yield

$$W'(\varphi) = \frac{W(\varphi)}{\varphi[W(\varphi)+1]} > 0$$

for all $\varphi > 0$. Finally, we have $0 = W(0)\mathrm{e}^{W(0)}$, which implies $W(0) = 0$; and $\mathrm{e} = W(\mathrm{e})\mathrm{e}^{W(\mathrm{e})}$, which implies $W(\mathrm{e}) = 1$.

A.2. Existence and uniqueness of the equilibrium cutoff m^d. We show that there exists a unique equilibrium cutoff m^d. To see this, apply the Leibnitz integral rule to the left-hand side of (10) and use $W(\mathrm{e}) = 1$ to obtain

$$\frac{\mathrm{e}L}{\alpha(m^d)^2}\int_0^{m^d} m^2\left(W^{-2}-1\right)W'\mathrm{d}G(m) > 0,$$

where the sign comes from $W' > 0$ and $W^{-2} \geq 1$ for $0 \leq m \leq m^d$. Hence, the left-hand side of (10) is strictly increasing. This uniquely determines the equilibrium cutoff m^d, because

$$\lim_{m^d\to 0}\int_0^{m^d} m\left(W^{-1}+W-2\right)\mathrm{d}G(m) = 0 \quad \text{and} \quad \lim_{m^d\to\infty}\int_0^{m^d} m\left(W^{-1}+W-2\right)\mathrm{d}G(m) = \infty.$$

A.3. Market size, the equilibrium cutoff, and the mass of entrants. Differentiating (10) and using the Leibniz integral rule, we readily obtain

$$\frac{\partial m^d}{\partial L} = -\frac{\alpha F\left(m^d\right)^2}{\mathrm{e}L^2}\left[\int_0^{m^d} m^2\left(W^{-2}-1\right)W'\mathrm{d}G(m)\right]^{-1} < 0,$$

because $W' > 0$ and $W^{-2} \geq 1$ for $0 \leq m \leq m^d$. Differentiating (12) with respect to L yields

$$\frac{\partial N^E}{\partial L} = \frac{F(N^E)^2}{L^2}\left\{1 - \frac{\mathrm{e}L^3}{\alpha F\left(m^d\right)^2}\left[\int_0^{m^d} m^2 W'\mathrm{d}G(m)\right]\frac{\partial m^d}{\partial L}\right\} > 0,$$

where the sign comes from $\partial m^d/\partial L < 0$ as established in the foregoing.

A.4. Existence and uniqueness in the two-region case. Under our assumptions on trade costs, the RHS of (32) is non-negative if and only if $\underline{\omega} < \omega < \overline{\omega}$, where $\underline{\omega} \equiv \rho^{1/(k+1)} (\tau_{22}/\tau_{12})^{k/(k+1)}$ and $\overline{\omega} \equiv \rho^{1/(k+1)} (\tau_{21}/\tau_{11})^{k/(k+1)}$. Furthermore, the RHS is strictly decreasing in $\omega \in (\underline{\omega}, \overline{\omega})$ with $\lim_{\omega \to \underline{\omega}+} \text{RHS} = \infty$ and $\lim_{\omega \to \overline{\omega}-} \text{RHS} = 0$. The LHS of (32) is, on the contrary, strictly increasing in $\omega \in (0, \infty)$. Hence, there exists a unique equilibrium $\omega^* \in (\underline{\omega}, \overline{\omega})$.

A.5. Comparative statics. (i) Assume that $\rho = 1$, $\tau_{12} = \tau_{21} = \tau$, and $\tau_{11} = \tau_{22} = t$. The equilibrium relative wage ω^* is increasing in L_1/L_2 as an increase in L_1/L_2 raises the RHS of (32) without affecting the LHS. This implies that if the two regions have equal technological possibilities and face symmetric trade costs, the larger region has the higher relative wage. Using (30), one can verify that $\omega^{2k+1} = \left(m_2^d/m_1^d\right)^{k+1}$ holds in that case. As $L_1 > L_2$ implies $\omega > 1$, it directly follows that $m_1^d < m_2^d$. Finally, we show in (38) that a lower cutoff maps into a higher welfare. Hence, the larger region has the higher welfare.

(ii) Assume that $L_1 = L_2$, $\tau_{12} = \tau_{21} = \tau$, and $\tau_{11} = \tau_{22} = t$. Since $t < \tau$ holds, the RHS of (32) shifts up as ρ increases, which then also increases ω^*. This implies that if the two regions are of equal size and face symmetric trade costs, the region with the better technological possibilities has the higher wage. Furthermore, evaluate (32) at $\omega = \rho^{1/(k+1)}$. The LHS is equal to $\rho^{k/(k+1)}$, which falls short of the RHS given by ρ (since $\rho > 1$ and $k \geq 1$). Since the LHS is increasing and the RHS is decreasing, it must be that $\omega^* > \rho^{1/(k+1)}$. It is then straightforward to see that $m_1^d < m_2^d$, because we can rewrite (30) as $\omega^{2k+1}/\rho = \left(m_2^d/m_1^d\right)^{k+1}$ and the LHS of this expression must be larger than one since $(\omega^*)^{2k+1} > (\omega^*)^{k+1} > \rho$. By (38), $U_1 > U_2$ follows again.

(iii) Assume that $\rho = 1$, $L_1 = L_2$ and $\tau_{12} = \tau_{21} = \tau$. Higher internal trade costs in one region reduce the relative wage of that region, because we can verify from (32) that

$$\frac{\partial(\text{RHS})}{\partial \tau_{11}} < 0 \quad \text{iff} \quad \omega^* > \underline{\omega} \quad \text{and} \quad \frac{\partial(\text{RHS})}{\partial \tau_{22}} > 0 \quad \text{iff} \quad \omega^* < \overline{\omega},$$

where the bounds $\underline{\omega}$ and $\overline{\omega}$ are given in Appendix A.4. Since $\omega^{2k+1} = \left(m_2^d/m_1^d\right)^{k+1}$, it follows that the region with the lower internal trade costs also has the lower cutoff and, thus, by (38) the higher welfare.

(iv) Assume that $\rho = 1$, $L_1 = L_2$ and $\tau_{11} = \tau_{22} = t$. Better access to the foreign market raises the domestic relative wage, whereas better access from the foreign to the domestic market reduces the domestic relative wage because (32) implies

$$\frac{\partial(\text{RHS})}{\partial \tau_{12}} < 0 \quad \text{iff} \quad \omega^* < \overline{\omega} \quad \text{and} \quad \frac{\partial(\text{RHS})}{\partial \tau_{21}} > 0 \quad \text{iff} \quad \omega^* > \underline{\omega}.$$

Since $\omega^{2k+1} = \left(m_2^d/m_1^d\right)^{k+1}$ holds, it follows that the region that ends up with the higher wage also ends up with the lower cutoff and, thus, with higher welfare.

(v) Finally, assuming that $\tau_{12} = \tau_{21} = \tau$ and that $\tau_{11} = \tau_{22} = t$, one can verify that

$$\frac{\partial(\text{RHS})}{\partial\tau} = -\frac{k\rho t^k}{\tau^{k+1}}\frac{L_1}{L_2}\frac{\rho^2 - \omega^{2(k+1)}}{[\omega^{k+1} - \rho(t/\tau)^k]^2}\left\{\begin{array}{c} > \\ = \\ < \end{array}\right\} 0 \quad \text{for} \quad \left\{\begin{array}{c} \underline{\omega} < \rho^{\frac{1}{k+1}} < \omega^* < \overline{\omega} \\ \underline{\omega} < \omega^* = \rho^{\frac{1}{k+1}} < \overline{\omega} \\ \underline{\omega} < \omega^* < \rho^{\frac{1}{k+1}} < \overline{\omega} \end{array}\right\}. \tag{39}$$

Note that when regions are of equal size, but have different upper bounds ($\rho > 1$), the first case of (39) applies since $\omega^* > \rho^{1/(k+1)}$ as shown in (ii). Hence, lower trade costs reduce the relative wage of the more productive region. Furthermore, when regions have the same upper bounds but different sizes ($L_1 > L_2$), we obtain $\omega^* > \rho^{k/(k+1)} = 1$, so that the first case of (39) applies again. In other words, when regions differ in size or technological possiblities, wages converge as bilateral trade barriers fall. Since $\omega^{2k+1} = \rho \left(m_2^d/m_1^d\right)^{k+1}$ always holds, this wage convergence directly implies (conditional) convergence of the regional cutoff productivities, and thus (conditional) convergence of regional welfare.

Appendix B: Integrals involving the Lambert W function

To derive closed-form solutions for various expressions throughout the paper we need to compute integrals involving the Lambert W function. This can be done by using the change in variables suggested by Corless *et al.* (1996, p.341). Let

$$z \equiv W\left(\mathrm{e}\frac{m}{I}\right), \quad \text{so that} \quad \mathrm{e}\frac{m}{I} = z\mathrm{e}^z, \quad \text{where} \quad I = m_r^d, m_{rs}^x,$$

where subscript r can be dropped in the closed economy. The change in variables then yields $\mathrm{d}m = (1+z)\mathrm{e}^{z-1}I\mathrm{d}z$, with the new integration bounds given by 0 and 1. Under our assumption of a Pareto distribution for productivity draws, the change in variables allows to rewrite integrals in simplified form.

B.1. First, consider the following expression, which appears when integrating firms' outputs:

$$\int_0^I m\left[1 - W\left(\mathrm{e}\frac{m}{I}\right)\right]\mathrm{d}G_r(m) = \kappa_1 \left(m_r^{\max}\right)^{-k} I^{k+1},$$

where $\kappa_1 \equiv k\mathrm{e}^{-(k+1)}\int_0^1 (1-z^2)\left(z\mathrm{e}^z\right)^k \mathrm{e}^z\mathrm{d}z > 0$ is a constant term which solely depends on the shape parameter k.

B.2. Second, the following expression appears when integrating firms' operating profits:

$$\int_0^I m\left[W\left(\mathrm{e}\frac{m}{I}\right)^{-1} + W\left(\mathrm{e}\frac{m}{I}\right) - 2\right]\mathrm{d}G_r(m) = \kappa_2 \left(m_r^{\max}\right)^{-k} I^{k+1},$$

where $\kappa_2 \equiv k\mathrm{e}^{-(k+1)}\int_0^1 (1+z)\left(z^{-1}+z-2\right)\left(z\mathrm{e}^z\right)^k \mathrm{e}^z\mathrm{d}z > 0$ is also a constant term which solely depends on the shape parameter k.

B.3. Finally, the following expression appears when integrating firms' revenues:

$$\int_0^I m\left[W\left(\mathrm{e}\frac{m}{I}\right)^{-1}-1\right]\mathrm{d}G_r(m)=\kappa_3\left(m_r^{\max}\right)^{-k}I^{k+1},$$

where $\kappa_3 \equiv k\mathrm{e}^{-(1+k)}\int_0^1 (z^{-1}-z)\,(z\mathrm{e}^z)^k\,\mathrm{e}^z\mathrm{d}z > 0$ is a constant term which solely depends on the shape parameter k. Using the expressions for κ_1 and κ_2, one can verify that $\kappa_3 = \kappa_1 + \kappa_2$.

Appendix C: Equilibrium in the open economy

In this appendix we restate the open economy equilibrium conditions of Section 3 using the Lambert W function.

C.1. Using (19), the labor market clearing condition can be rewritten as follows:

$$N_r^E\left\{\frac{1}{\alpha}\sum_s L_s\tau_{rs}\int_0^{m_{rs}^x} m\left[1-W\left(\mathrm{e}\frac{m}{m_{rs}^x}\right)\right]\mathrm{d}G_r(m)+F_r\right\}=L_r. \tag{40}$$

C.2. Plugging (19) into (21), zero expected profits require that

$$\frac{1}{\alpha}\sum_s L_s\tau_{rs}\int_0^{m_{rs}^x} m\left[W\left(\mathrm{e}\frac{m}{m_{rs}^x}\right)^{-1}+W\left(\mathrm{e}\frac{m}{m_{rs}^x}\right)-2\right]\mathrm{d}G_r(m)=F_r. \tag{41}$$

As in the closed economy case, the zero expected profit condition depends solely on the cutoffs m_{rs}^x and is independent of the mass of entrants.

C.3. Finally, the trade balance condition is given by

$$\begin{aligned}N_r^E w_r\sum_{s\neq r} L_s\tau_{rs}\int_0^{m_{rs}^x} m\left[W\left(\mathrm{e}\frac{m}{m_{rs}^x}\right)^{-1}-1\right]\mathrm{d}G_r(m)\\ = L_r\sum_{s\neq r} N_s^E\tau_{sr}w_s\int_0^{m_{sr}^x} m\left[W\left(\mathrm{e}\frac{m}{m_{sr}^x}\right)^{-1}-1\right]\mathrm{d}G_s(m).\end{aligned} \tag{42}$$

Applying the region-specific Pareto distributions $G_r(m) = (m/m_r^{\max})^k$ to (40)–(42) yields, after some algebra and using the results of Appendix B, expressions (23)–(25) given in the main text.

Appendix D: Firm-level simulation procedure

To simulate our model at the firm-level, we use the estimates from the gravity equation system: the relative wages ($\widehat{w}_r$), the upper bound transforms ($\widehat{\mu}_r^{\max}$), and the trade friction parameters

($\widehat{\theta}$ and $\widehat{\gamma}$). These estimates provide all the information required to construct the domestic and the export cutoffs in the economy ($\widehat{m}_s^d$ and $\widehat{m}_{rs}^x$). Note that even though we have no information on the preference parameter α and the region-specific entry costs F_r, we can still generate all relevant firm-level variables that we require for our analysis. Let $\widehat{\chi}_{rs} = 1$ if $m < \widehat{m}_{rs}^x$ (the firm is productive enough to export to market s when it is established in market r) and $\widehat{\chi}_{rs} = 0$ otherwise. For each firm with productivity m drawn in region r of country $I = \text{CA}, \text{US}$, we can compute the following firm-level quantities:

Export intensity. A firm's export intensity is defined as

$$\text{expint}_r(m) = \frac{\text{expsls}_r(m)}{\text{domsls}_r(m) + \text{expsls}_r(m)},$$

where total domestic sales and total export sales are respectively given by

$$\begin{aligned}
\text{domsls}_r(m) &= \sum_{s \in I} \widehat{\chi}_{rs} L_s p_{rs}(m) q_{rs}(m) \\
&= \frac{\widehat{w}_r m}{\alpha} \sum_{s \in I} \widehat{\chi}_{rs} L_s d_{rs}^{\widehat{\gamma}} [W(\text{e}m/\widehat{m}_{rs}^x)^{-1} - 1] \\
\text{expsls}_r(m) &= \sum_{s \notin I} \widehat{\chi}_{rs} L_s p_{rs}(m) q_{rs}(m) \\
&= \frac{\widehat{w}_r m}{\alpha} \sum_{s \notin I} \widehat{\chi}_{rs} L_s d_{rs}^{\widehat{\gamma}} \text{e}^{\widehat{\theta} b_{rs}} [W(\text{e}m/\widehat{m}_{rs}^x)^{-1} - 1].
\end{aligned}$$

Note that we can identify domestic sales and export sales up to a scaling parameter only (which depends on the unobservable α), but that this is immaterial for the firm's export intensity.

Revenue-based productivity. The revenue-based productivity, excluding the labor used for shipping goods, is given by:

$$\text{rbprod}_r(m) = \frac{\text{domsls}_r(m) + \text{expsls}_r(m)}{(m/\alpha) \sum_s \widehat{\chi}_{rs} L_s (1 - W(\text{e}m/\widehat{m}_{rs}^x))},$$

which is again independent of the scaling by α.

To calibrate the value of k and to obtain the productivity (both physical and revenue-based) and export intensity distributions, we generate two distinct samples of 110,000 firms each: one with the border, and one after removing the border (setting $b_{rs} = 0$ for all r, s). Comparing the two samples allows us to assess the distributional impacts of removing the border. To make the sample representative, we draw firms in all regions in proportion to that region's share of surviving firms in the national number of surviving firms. We know that

$$N_r^p = N_r^E G_r \left(\max_s m_{rs}^x \right) = \frac{\alpha}{\kappa_1 + \kappa_2} L_r \left(\mu_r^{\max} \right)^{-1} \left(\max_s m_{rs}^x \right)^k$$

so that each region's share of surviving firms in country $I = \mathrm{CA}, \mathrm{US}$ is given by

$$\widehat{\theta}_r = \frac{\widehat{N}_r^p}{\sum_{s \in I} \widehat{N}_s^p} = \frac{L_r \left(\widehat{\mu}_r^{\max}\right)^{-1} \left(\max_j \widehat{m}_{rj}^x\right)^k}{\sum_{s \in I} L_s \left(\widehat{\mu}_s^{\max}\right)^{-1} \left(\max_j \widehat{m}_{sj}^x\right)^k}, \quad r \in I.$$

Note that the foregoing expression is again independent of the unobservable parameter α. For a sample size $N_{CA} = 10,000$ and $N_{US} = 100,000$, we randomly draw $\mathrm{int}(\widehat{\theta}_s N_{CA})$ firms for each Canadian province and $\mathrm{int}(\widehat{\theta}_r N_{US})$ firms for each US state from the region-specific productivity distribution, where $\mathrm{int}(\cdot)$ stands for the integer part. This yields a representative sample for each country, while the overall sample respects country's relative sizes in 1993.

For the counterfactual case after the border removal, we proceed in the same way except that we use the simulated values $\widetilde{m}_{rs}^x$ for the various cutoffs. This yields counterfactual shares $\widetilde{\theta}_r$ which ensure that we still draw a representative firm sample for the scenario after the border removal. Note that we can keep the same overall sample size because all compositional changes due to trade integration are reflected by the new regional sample shares.

Table 1: Estimation of the gravity equation system

	Benchmark(1)	Robustness(2)	Robustness(3)	Robustness(4)	Robustness(5)	Robustness(6)
Regions	83 (40)	83 (40)	83 (40)	83 (40)	61 (40)	83 (40)
Flows	1560	1560	1560	1560	1560	1511
k	9.5	3.6	8.5	10.5	9.5	9.5
Internal dist.	Surface	Surface	Surface	Surface	AvW	Surface
Coefficients:						
constant	-4.2587	-4.1479	-4.2543	-4.2619	-4.4673	-4.2028
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\ln d_{rs}$	-1.4195	-1.5380	-1.4266	-1.4140	-1.1935	-1.4383
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\ln d_{rr}$	1.5690	1.9652	1.5944	1.5487	1.3191	1.5897
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
b_{rs}	-1.5951	-1.5803	-1.5957	-1.5945	-1.6773	-1.6927
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$0-$ dummy	-17.471	-17.546	-17.477	-17.466	-17.737	—
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	—
Adjusted R^2	0.9108	0.9076	0.9105	0.9108	0.8917	0.6828

Notes: p-values in parentheses. Benchmark(1) uses 1560 trade flows, excluding intraregional flows X_{rr} as in Anderson and van Wincoop (2003). 'Surface' refers to Redding and Venables' (2004) surface-based measure of internal distance, whereas AvW refers to Anderson and van Wincoop's (2003) measure. The convergence criterion for the iterative procedure is based on the difference of norms of the vector of regression coefficients between two successive iterations, with threshold 10^{-12}. Starting points for the iterative solver are obtained by OLS with importer-exporter fixed effects. We choose $w_{\text{Alabama}} = 1$ as numéraire and set starting wages to $w_r = 1$ for all r. Results are invariant to that choice.

Table 2: Descriptive statistics for bilateral border effect series

Descriptive statistics for bilateral border effect series:						
	Benchmark(1)	Robustness(2)	Robustness(3)	Robustness(4)	Robustness(5)	Robustness(6)
Minimum	0.3334	0.3504	0.3350	0.3322	0.3635	0.3332
Maximum	4.1121	4.3742	4.1381	4.0907	4.3323	4.1501
Mean	1.3144	1.3415	1.3171	1.3122	1.6642	1.3184
Std. dev.	0.9123	0.9594	0.9176	0.9079	1.2136	0.9216
Median	0.9772	0.9818	0.9774	0.9770	0.9624	0.9773
Skewness	2.0742	2.1059	2.0763	2.0725	1.0839	2.0775
Kurtosis	5.6195	5.7628	5.6278	5.6131	2.3343	5.6327
Correlation matrix for bilateral border effect series:						
	(1)	(2)	(3)	(4)	(5)	(6)
(1)	1	0.9997	1.0000	1.0000	0.9601	1.0000
(2)		1	0.9997	0.9997	0.9543	0.9997
(3)			1	1.0000	0.9598	1.0000
(4)				1	0.9603	0.9597
(5)					1	1.0000
(6)						1

Table 3: Bilateral border effects decomposition for the province of British Columbia and the state of New York

	Benchmark(1), $k = 9.5$									
	Exporter: British Columbia					**Exporter: New York**				
	Pure border $e^{\widehat{\theta} b_{rs}}$	Rel. wage $\Delta(w_s/w_r)$	Abs. wage Δw_r	Selection Δm_s^d	Bil. border B_{rs}	Pure border $e^{\widehat{\theta} b_{rs}}$	Rel. wage $\Delta(w_s/w_r)$	Abs. wage Δw_r	Selection Δm_s^d	Bil. border B_{rs}
Importer:	**In Gravity sample**									
Alberta	1.0000	0.7102	1.0584	0.6046	0.4545	4.9290	1.2644	1.0018	0.6046	3.7752
British Columbia	1.0000	1.0000	1.0584	0.3151	0.3334	4.9290	1.7805	1.0018	0.3151	2.7699
Manitoba	1.0000	0.8963	1.0584	0.3881	0.3682	4.9290	1.5958	1.0018	0.3881	3.0583
New Brunswick	1.0000	0.7188	1.0584	0.5909	0.4496	4.9290	1.2798	1.0018	0.5909	3.7343
Newfoundland	1.0000	0.7099	1.0584	0.6052	0.4547	4.9290	1.2639	1.0018	0.6052	3.7768
Nova Scotia	1.0000	0.6816	1.0584	0.6538	0.4717	4.9290	1.2136	1.0018	0.6538	3.9181
Ontario	1.0000	0.9350	1.0584	0.3581	0.3544	4.9290	1.6648	1.0018	0.3581	2.9435
Prince Edward Island	1.0000	0.6637	1.0584	0.6878	0.4831	4.9290	1.1818	1.0018	0.6878	4.0134
Quebec	1.0000	0.8518	1.0584	0.4277	0.3855	4.9290	1.5166	1.0018	0.4277	3.2026
Saskatchewan	1.0000	0.8337	1.0584	0.4455	0.3931	4.9290	1.4843	1.0018	0.4455	3.2655
Alabama	4.9290	0.5511	1.0584	0.9801	2.8178	1.0000	0.9812	1.0018	0.9801	0.9634
Arizona	4.9290	0.5527	1.0584	0.9747	2.8104	1.0000	0.9841	1.0018	0.9747	0.9609
California	4.9290	0.5491	1.0584	0.9870	2.8273	1.0000	0.9776	1.0018	0.9870	0.9667
Florida	4.9290	0.5491	1.0584	0.9868	2.8269	1.0000	0.9777	1.0018	0.9868	0.9666
Georgia	4.9290	0.5513	1.0584	0.9793	2.8167	1.0000	0.9816	1.0018	0.9793	0.9631
Idaho	4.9290	0.5670	1.0584	0.9283	2.7460	1.0000	1.0096	1.0018	0.9283	0.9389
Illinois	4.9290	0.5517	1.0584	0.9781	2.8151	1.0000	0.9823	1.0018	0.9781	0.9625
Indiana	4.9290	0.5550	1.0584	0.9671	2.7999	1.0000	0.9882	1.0018	0.9671	0.9573
Kentucky	4.9290	0.5552	1.0584	0.9664	2.7990	1.0000	0.9885	1.0018	0.9664	0.9570
Louisiana	4.9290	0.5501	1.0584	0.9837	2.8226	1.0000	0.9794	1.0018	0.9837	0.9651
Maine	4.9290	0.6113	1.0584	0.8045	2.5656	1.0000	1.0884	1.0018	0.8045	0.8772
Maryland	4.9290	0.5531	1.0584	0.9734	2.8086	1.0000	0.9848	1.0018	0.9734	0.9603
Massachusetts	4.9290	0.5561	1.0584	0.9635	2.7951	1.0000	0.9901	1.0018	0.9635	0.9557
Michigan	4.9290	0.5657	1.0584	0.9326	2.7520	1.0000	1.0072	1.0018	0.9326	0.9410
Minnesota	4.9290	0.5610	1.0584	0.9473	2.7726	1.0000	0.9989	1.0018	0.9473	0.9480
Missouri	4.9290	0.5529	1.0584	0.9740	2.8094	1.0000	0.9845	1.0018	0.9740	0.9606
Montana	4.9290	0.5872	1.0584	0.8684	2.6605	1.0000	1.0456	1.0018	0.8684	0.9096
New Hampshire	4.9290	0.5709	1.0584	0.9165	2.7294	1.0000	1.0164	1.0018	0.9165	0.9332
New Jersey	4.9290	0.5524	1.0584	0.9758	2.8119	1.0000	0.9835	1.0018	0.9758	0.9614
New York	4.9290	0.5616	1.0584	0.9454	2.7699	1.0000	1.0000	1.0018	0.9454	0.9471
North Carolina	4.9290	0.5541	1.0584	0.9702	2.8042	1.0000	0.9865	1.0018	0.9702	0.9588
North Dakota	4.9290	0.5852	1.0584	0.8743	2.6690	1.0000	1.0419	1.0018	0.8743	0.9126
Ohio	4.9290	0.5580	1.0584	0.9571	2.7862	1.0000	0.9935	1.0018	0.9571	0.9526
Pennsylvania	4.9290	0.5614	1.0584	0.9462	2.7710	1.0000	0.9995	1.0018	0.9462	0.9475
Tennessee	4.9290	0.5525	1.0584	0.9755	2.8115	1.0000	0.9837	1.0018	0.9755	0.9613
Texas	4.9290	0.5502	1.0584	0.9832	2.8220	1.0000	0.9796	1.0018	0.9832	0.9649
Vermont	4.9290	0.5959	1.0584	0.8445	2.6254	1.0000	1.0610	1.0018	0.8445	0.8977
Virginia	4.9290	0.5553	1.0584	0.9659	2.7984	1.0000	0.9888	1.0018	0.9659	0.9568
Washington	4.9290	0.5991	1.0584	0.8359	2.6126	1.0000	1.0668	1.0018	0.8359	0.8933
Wisconsin	4.9290	0.5584	1.0584	0.9559	2.7845	1.0000	0.9942	1.0018	0.9559	0.9521
Importer:	**Out of Gravity sample**									
Alaska	4.9290	0.5779	1.0584	0.8954	2.6993	1.0000	1.0289	1.0018	0.8954	0.9229
Arkansas	4.9290	0.5530	1.0584	0.9737	2.8090	1.0000	0.9846	1.0018	0.9737	0.9604
Colorado	4.9290	0.5556	1.0584	0.9650	2.7971	1.0000	0.9893	1.0018	0.9650	0.9564
Connecticut	4.9290	0.5544	1.0584	0.9692	2.8028	1.0000	0.9870	1.0018	0.9692	0.9583
Delaware	4.9290	0.5542	1.0584	0.9699	2.8038	1.0000	0.9867	1.0018	0.9699	0.9587
District of Columbia	4.9290	0.5468	1.0584	0.9950	2.8380	1.0000	0.9735	1.0018	0.9950	0.9704
Hawaii	4.9290	0.5562	1.0584	0.9632	2.7946	1.0000	0.9903	1.0018	0.9632	0.9555
Iowa	4.9290	0.5549	1.0584	0.9674	2.8003	1.0000	0.9880	1.0018	0.9674	0.9575
Kansas	4.9290	0.5517	1.0584	0.9780	2.8149	1.0000	0.9823	1.0018	0.9780	0.9625
Mississippi	4.9290	0.5576	1.0584	0.9585	2.7881	1.0000	0.9928	1.0018	0.9585	0.9533
Nebraska	4.9290	0.5500	1.0584	0.9837	2.8227	1.0000	0.9793	1.0018	0.9837	0.9651
Nevada	4.9290	0.5559	1.0584	0.9641	2.7958	1.0000	0.9897	1.0018	0.9641	0.9559
New Mexico	4.9290	0.5527	1.0584	0.9746	2.8102	1.0000	0.9841	1.0018	0.9746	0.9609
Oklahoma	4.9290	0.5668	1.0584	0.9292	2.7473	1.0000	1.0091	1.0018	0.9292	0.9393
Oregon	4.9290	0.5546	1.0584	0.9683	2.8017	1.0000	0.9875	1.0018	0.9683	0.9579
Rhode Island	4.9290	0.5530	1.0584	0.9738	2.8091	1.0000	0.9846	1.0018	0.9738	0.9605
South Carolina	4.9290	0.5701	1.0584	0.9190	2.7329	1.0000	1.0150	1.0018	0.9190	0.9344
South Dakota	4.9290	0.5594	1.0584	0.9526	2.7799	1.0000	0.9960	1.0018	0.9526	0.9505
Utah	4.9290	0.5599	1.0584	0.9509	2.7776	1.0000	0.9969	1.0018	0.9509	0.9497
West Virginia	4.9290	0.5593	1.0584	0.9530	2.7805	1.0000	0.9958	1.0018	0.9530	0.9507
Wyoming	4.9290	0.5479	1.0584	0.9911	2.8328	1.0000	0.9755	1.0018	0.9911	0.9686

Notes: Border effects are decomposed as indicated by (37).

Table 4: Regional border effects

	Benchmark (1), $k = 9.5$
	Regional BE B_r
State/province	**In Gravity sample**
Alberta	6.4824
British Columbia	7.5924
Manitoba	7.3196
New Brunswick	6.6774
Newfoundland	6.5406
Nova Scotia	6.2810
Ontario	7.6515
Prince Edward Island	6.2101
Quebec	7.1930
Saskatchewan	7.0252
Alabama	3.2538
Arizona	3.3001
California	3.2324
Florida	3.2529
Georgia	3.2448
Idaho	3.3087
Illinois	3.2387
Indiana	3.2186
Kentucky	3.2229
Louisiana	3.2669
Maine	3.4637
Maryland	3.2203
Massachusetts	3.3350
Michigan	3.2037
Minnesota	3.2962
Missouri	3.2626
Montana	3.4463
New Hampshire	3.3566
New Jersey	3.2291
New York	3.2761
North Carolina	3.2423
North Dakota	3.3872
Ohio	3.1975
Pennsylvania	3.2105
Tennessee	3.2401
Texas	3.2807
Vermont	3.3659
Virginia	3.2332
Washington	3.1195
Wisconsin	3.2499
State/province	**Out of Gravity sample**
Alaska	3.4068
Arkansas	3.2693
Colorado	3.3327
Connecticut	3.2860
Delaware	3.2330
Hawaii	3.2810
Iowa	3.2888
Kansas	3.2646
Mississippi	3.3020
Nebraska	3.2402
Nevada	3.3157
New Mexico	3.2885
Oklahoma	3.1705
Oregon	3.3157
Rhode Island	3.2479
South Carolina	3.3580
South Dakota	3.3340
Utah	3.2151
West Virginia	3.3429
Wyoming	3.1921
District of Columbia	3.2701

Notes: See Section 5 for details on computations.

Table 5: The impacts of fully removing the border, holding all other parameters fixed

	Benchmark (1), $k = 9.5$					
	Wage $\Delta(w_r)\%$	Cutoff $\Delta(m_r^d)\%$	Markups $\Delta(\alpha E_r/N_r^c)\%$	Consumed $\Delta(N_r^c)\%$	Produced $\Delta(N_r^p)\%$	Welfare $\Delta U_r^*\%$
State/province	**In Gravity sample**					
Alberta	2.4445	-4.6785	-2.3487	4.9085	-36.5674	4.9081
British Columbia	5.8387	-10.4167	-5.1864	11.6282	-31.8922	11.6280
Manitoba	4.7407	-8.6195	-4.2879	9.4331	-57.5273	9.4325
New Brunswick	2.5620	-4.8865	-2.4503	5.1382	-37.8699	5.1375
Newfoundland	2.4401	-4.6708	-2.3454	4.9004	-36.5184	4.8996
Nova Scotia	2.0448	-3.9661	-2.0030	4.1305	-31.9176	4.1299
Ontario	5.1635	-9.3179	-4.6357	10.2755	-60.5130	10.2754
Prince Edward Island	1.7870	-3.5021	-1.7788	3.6303	-28.7279	3.6292
Quebec	4.2339	-7.7714	-3.8666	8.4264	-43.9461	8.4262
Saskatchewan	4.0207	-7.4108	-3.6886	8.0046	-51.8805	8.0040
Alabama	0.0000	-0.1911	-0.1916	0.1920	-1.8010	0.1915
Arizona	0.0275	-0.2435	-0.2164	0.2445	-2.2891	0.2441
California	-0.0353	-0.1241	-0.1593	0.1242	-1.1727	0.1242
Florida	-0.0341	-0.1263	-0.1606	0.1267	-1.1937	0.1265
Georgia	0.0040	-0.1987	-0.1951	0.1995	-1.8718	0.1991
Idaho	0.2719	-0.7060	-0.4367	0.7117	-6.5094	0.7110
Illinois	0.0103	-0.2107	-0.2007	0.2115	-1.9838	0.2111
Indiana	0.0670	-0.3184	-0.2521	0.3199	-2.9845	0.3194
Kentucky	0.0706	-0.3252	-0.2553	0.3268	-3.0467	0.3262
Louisiana	-0.0180	-0.1569	-0.1753	0.1576	-1.4802	0.1571
Maine	0.9919	-2.0500	-1.0791	2.0936	-17.8625	2.0929
Maryland	0.0344	-0.2564	-0.2226	0.2575	-2.4096	0.2571
Massachusetts	0.0854	-0.3533	-0.2685	0.3549	-3.3060	0.3545
Michigan	0.2490	-0.6629	-0.4158	0.6676	-6.1226	0.6673
Minnesota	0.1705	-0.5144	-0.3453	0.5175	-4.7813	0.5171
Missouri	0.0313	-0.2507	-0.2199	0.2517	-2.3560	0.2513
Montana	0.6067	-1.3344	-0.7366	1.3532	-11.9816	1.3525
New Hampshire	0.3360	-0.8267	-0.4942	0.8343	-7.5835	0.8336
New Jersey	0.0221	-0.2331	-0.2114	0.2340	-2.1926	0.2336
New York	0.1807	-0.5337	-0.3542	0.5367	-4.9567	0.5366
North Carolina	0.0509	-0.2878	-0.2375	0.2891	-2.7013	0.2887
North Dakota	0.5729	-1.2712	-0.7065	1.2885	-11.4446	1.2876
Ohio	0.1189	-0.4168	-0.2987	0.4189	-3.8900	0.4185
Pennsylvania	0.1763	-0.5255	-0.3504	0.5286	-4.8824	0.5283
Tennessee	0.0235	-0.2358	-0.2129	0.2369	-2.2181	0.2364
Texas	-0.0157	-0.1613	-0.1772	0.1618	-1.5221	0.1616
Vermont	0.7474	-1.5968	-0.8621	1.6235	-14.1800	1.6227
Virginia	0.0729	-0.3296	-0.2573	0.3311	-3.0880	0.3307
Washington	0.7991	-1.6929	-0.9076	1.7223	-14.9732	1.7221
Wisconsin	0.1253	-0.4288	-0.3046	0.4311	-4.0003	0.4307
State/province	**Out of Gravity sample**					
Alaska	0.4531	-1.0470	-0.5992	1.0587	-9.5150	1.0580
Arkansas	0.0329	-0.2537	-0.2215	0.2550	-2.3844	0.2543
Colorado	0.0778	-0.3388	-0.2618	0.3404	-3.6919	0.3399
Connecticut	0.0563	-0.2979	-0.2423	0.2993	-2.7949	0.2988
Delaware	0.0527	-0.2911	-0.2394	0.2928	-2.7313	0.2919
Hawaii	-0.0753	-0.0478	-0.1232	0.0480	-0.4534	0.0478
Iowa	0.0872	-0.3568	-0.2704	0.3586	-3.3385	0.3581
Kansas	0.0655	-0.3155	-0.2508	0.3170	-2.9572	0.3165
Mississippi	0.0109	-0.2118	-0.2015	0.2128	-1.9937	0.2122
Nebraska	0.1117	-0.4030	-0.2925	0.4053	-3.7640	0.4047
Nevada	-0.0184	-0.1561	-0.1746	0.1565	-1.4733	0.1564
New Mexico	0.0824	-0.3475	-0.2661	0.3493	-3.2531	0.3487
Oklahoma	0.0283	-0.2449	-0.2172	0.2461	-2.3026	0.2455
Oregon	0.2671	-0.6969	-0.4321	0.7022	-6.4283	0.7018
Rhode Island	0.0604	-0.3058	-0.2463	0.3075	-2.8680	0.3068
South Carolina	0.0325	-0.2528	-0.2210	0.2540	-2.3758	0.2534
South Dakota	0.3225	-0.8013	-0.4822	0.8086	-7.3582	0.8078
Utah	0.1426	-0.4616	-0.3203	0.4644	-4.3000	0.4637
West Virginia	0.1514	-0.4782	-0.3283	0.4812	-4.4519	0.4805
Wyoming	0.1404	-0.4575	-0.3183	0.4602	-4.2628	0.4596
District of Columbia	-0.0558	-0.0849	-0.1413	0.0857	-0.8035	0.0850

Notes: See Section 5 for details on computations.

Table 6: Determinants of changes in regional variables

	Dependent variable (in logs)					
	Wage $\Delta(w_r)\%$	Cutoff $\Delta(m_r^d)\%$	Markup $\Delta(\alpha E_r/N_r^c)\%$	Consumed $\Delta(N_r^c)\%$	Produced $\Delta(N_r^p)\%$	Welfare $\Delta U_r^*\%$
Variable	**Estimated coefficients**, (All regions, $N = 61$ observations)					
CA-dummy	1.4521***	-2.5945***	-1.2588***	2.8260***	-15.9342***	2.8260***
DISTANCE TO BORDER (log)	-0.4323***	0.7740***	0.3752***	-0.8412***	3.6000***	-0.8412***
SIZE (log)	0.1030	-0.1659	-0.0834	0.2066	0.1876	0.2066
Constant	-23.4138***	41.3466***	20.0083***	-45.4731***	242.1900***	-45.473***
Adjusted R^2	0.8490	0.8584	0.8553	0.8460	0.8767	0.8460
Variable	**Estimated coefficients**, (All regions, $N = 61$ observations)					
DISTANCE TO BORDER (log)	-1.0115***	1.7066***	0.8445***	-2.0038***	-3.9498	-2.0038***
DISTANCE TO BORDER (log) × US-dummy	0.7734***	-1.2583***	-0.6308***	1.5478***	7.9061*	1.5478***
SIZE (log)	0.6001***	-1.0361***	-0.5085***	1.1802***	-4.9348***	1.1802***
SIZE (log) × US-dummy	-0.6749***	1.1771***	0.5758***	-1.3234***	6.1818***	-1.3234***
US-dummy	1.6685	-3.3316	-1.5584	3.1246	-102.0723***	3.1250
Constant	1.2010	-2.2619	-1.2102	2.5624	52.4719*	2.5592
Adjusted R^2	0.9288	0.9266	0.9253	0.9216	0.9063	0.9216
Variable	**Beta coefficients** (Only US States, $N = 51$ observations)					
DISTANCE TO BORDER (log)	-0.6205***	0.6212***	0.6208***	-0.6203***	0.6235***	-0.6203***
SIZE (log)	-0.3390***	0.3394***	0.3398***	-0.3391***	0.3418***	-0.3391***

Notes: See Section 5 for additional details on computations. Coefficients significant at 10% level (*), 5% level (**), and 1% level (***).

Table 7: Firm-level export intensities and export sales distributions

		Benchmark (1), $k = 9.5$			
Export intensity (percent)	US Percent of exporter (1992 Census, BEJK)	US Percent of exporters (BEJK simul.)	US Percent of exporters (our simul., with border)	Canada Percent of exporters (our simul., with border)	
0-10	66	76	65.16	41.81	
10-20	16	19	27.50	13.28	
20-30	7.7	4.2	4.05	15.96	
30-40	4.4	0.0	0.97	16.38	
40-50	2.4	0.0	0.78	5.65	
50-60	1.5	0.0	1.23	2.26	
60-70	1.0	0.0	0.31	1.41	
70-80	0.6	0.0	0.00	0.71	
80-90	0.5	0.0	0.00	1.27	
90-100	0.7	0.0	0.00	1.27	
Firm rank (exporters)	US Percent of exp. sales (1993 Census, BJS)	US Percent of exp. sales (with border)	US Percent of exp. sales (without border)	Canada Percent of exp. sales (with border)	Canada Percent of exp. sales (without border)
Top 1%	78.2	29.13	40.91	51.53	38.63
Top 5%	91.8	67.18	69.16	85.33	71.26
Top 10%	95.6	80.16	78.28	91.91	83.62
Top 25%	98.7	92.65	89.98	96.83	94.05
Top 50%	99.7	97.91	96.96	99.15	98.56

Notes: See Section 5 for additional details. Export intensity is defined as in Appendix D as the firm's share of export revenue in total revenue. Export sales are also defined as in Appendix D. Figures in Column 2 and 3 (top panel) are provided by Bernard *et al.* (2003). Column 2 reports the observed distribution using 1992 US Manufactuting Census data, whereas Column 3 provides the simulation results of Bernard *et al.* (2003). Columns 4 and 5 provide our own simulation results with $k = 9.5$. Figures in Column 2 (bottom panel) are provided by Bernard *et al.* (2009) using 1993 Census data. Columns 3-6 in the bottom half provide our own simulation results with $k = 9.5$.

Figure 1: National productivity distributions (CA and US, with border)

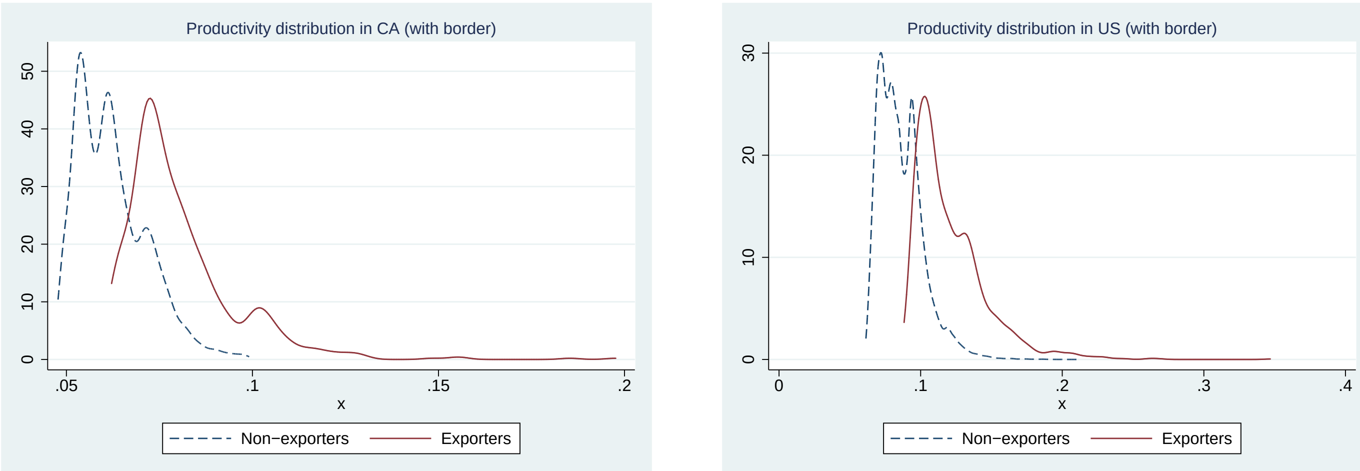

Figure 2: Producers' and sellers' distributions (with and without border)