Gains from Trade with Flexible Extensive Margin Adjustment*

Chang-Tai Hsieh
University of Chicago and NBER

Nicholas Li
University of Toronto

Ralph Ossa
University of Zurich and CEPR

Mu-Jeung Yang
University of Utah

November, 2020

Abstract

We propose a new sufficient statistic to measure the gains from trade in models where the extensive margin trade elasticity is not necessarily constant. This statistic is a function of one data moment, the market share of continuing domestic products, and one parameter, the elasticity of substitution between products. It measures the gains from trade in a Melitz model with any productivity distribution and any pattern of selection into production and exporting. We apply our statistic to measure Canada's gains from the Canada-US Free Trade Agreement using actual and model-simulated data and find that they are smaller than suggested by statistics that assume a constant extensive margin trade elasticity.

*Emails: chsieh@chicagobooth.edu; nick.li@utoronto.ca; ralph.ossa@econ.uzh.ch; mjyang@eccles.utah.edu.
Financial support from IGM at Chicago Booth is gratefully acknowledged. The results of this paper have been screened by Statistics Canada and the US Census Bureau to ensure that no confidential data are revealed.

1 Introduction

Seminal models by Krugman (1980) and Melitz (2003) show that changes in the set of firms serving domestic and foreign markets constitute an important "extensive" margin of adjustment in response to changes in trade costs. A large subsequent empirical literature has quantified partial aspects of these extensive margin entry/exit effects with micro data. For example, Broda and Weinstein (2006) quantify the welfare gains from increases in US import variety, while Pavcnik (2002) and Melitz and Trefler (2012) quantify increases in average domestic productivity associated with domestic exit by less productive firms and re-allocation of labor to more productive firms.

While firm-level micro data is necessary to quantify these individual effects, an influential paper by Arkolakis et al. (2012) (henceforth ACR) shows that for many trade models, detailed quantification of these entry, exit and re-allocation effects using micro-data is unnecessary to calculate welfare changes. Instead, welfare gains from changes in trade costs can be quantified using a sufficient statistic based only on the aggregate trade share and the aggregate trade elasticity. Their result assumes an iso-elastic import demand system. When trade is only affected at the intensive margin, this follows immediately from the standard Dixit-Stiglitz preferences. However, with extensive margin adjustment, iso-elastic import demand also requires that selection into exporting be such that the extensive margin trade elasticity is constant. The supply-side restrictions on firm heterogeneity that yield a constant extensive margin trade elasticity are restrictive, e.g. in the standard Melitz (2003) model any deviation from Pareto-distributed firm productivity or even simple deviations from strict selection into exporting lead to a variable extensive margin trade elasticity and hence an import demand system that is not iso-elastic.

In this paper, we argue that plausible and data-consistent assumptions on the distribution of firm heterogeneity and the nature of firm selection imply significant biases when applying ACR and similar "welfare sufficient statistics" to data. We propose a new formula to measure welfare gains in trade models with an extensive margin. Our proposed formula is based on one data moment, the domestic market share of continuing domestic products, and one parameter, the elasticity of substitution between products. The key advantage of our formula is that it does not require knife-edge supply-side restrictions. It is therefore consistent with a much wider range of trade models - for example, a generalized Melitz (2003) model with any productivity distribution and any pattern of selection into production and exporting.

Our analysis is motivated by a key stylized fact: firm selection into export markets and exit does not follow strict sorting, as assumed in Melitz (2003). For example, Eaton et al. (2011) and Armenter and Koren (2015) show that there is a substantial overlap in the size distribution of exporters and non-exporters in France and the US. Figure 1 (top panel) replicates this evidence for U.S. and Canadian manufacturing. As can be seen, many exporters are smaller than non-exporters, and many non-exporters are larger than exporters. The bottom panel in Figure 1 show that the same is true for exiting firms vs. continuing firms. Specifically, many exiting plants are ex-ante substantially larger than continuing firms in Canada and the US.

We therefore derive a more robust sufficient welfare statistic to quantify the gains from trade. Our approach is essentially an application of the well-known Feenstra (1994) formula, albeit one that is different from what is typically done in the literature. While prior papers such as Broda and Weinstein (2006) use the Feenstra (1994) method to measure the welfare gains from increasing *import variety*, we apply it to measure the *overall* welfare gains from domestic and foreign extensive margin adjustment. For example, in a Melitz (2003) model, the welfare gains from a reduction in trade costs come from lower prices for *existing* imported products and access to new imported products net of the loss from domestic products that exit. The intuition behind our application of Feenstra (1994)'s formula is that new foreign products and lower prices of existing imported products take away the market of *incumbent* domestic firms, while the exit of domestic firms increases their market share. The change in the market share of domestic incumbent firms therefore captures the net effect of all three sources of welfare gains in the Melitz (2003) model. Domestic incumbent firms are a natural benchmark for calculating the value of domestic exit and changes in the price and variety of imports because the domestic real wage in terms of the output of these firms is fixed for a wide class of models and they make up a large share of all domestic consumption.[1]

[1]Models with no within-firm productivity changes or changes in domestic markups or other distortions have this implication. Empirically we recognize that there is growing evidence that trade affects within-firm innovation and productivity (see Melitz and Trefler (2012) for a review in the context of the CUSFTA or Kueng et al. (2020) for more evidence from Canadian manufacturing firms). We have little to add to the literature on estimation of firm-level productivity growth and markups, but we note that to the extent that the detailed

We first apply our statistic to Canadian and US data during the Canada-US Free Trade Agreement liberalization period (1988-1996), which has been argued to be an ideal natural experiment to analyze the effects of trade liberalizations for developed economies (Trefler (2004)). We show that our statistic implies lower welfare gains from trade liberalization than the ACR statistic, primarily due domestic exit effects that are larger than predicted by a Pareto parameter calibrated to match the domestic firm size distribution or the relative size of new US entrants to the Canadian market.

We then construct a two-country general equilibrium Melitz model that allows for deviations from iso-elastic import demand and deviations from both Pareto productivity and strict selection into exporting. The model is calibrated using data from Canada and the United States including macro moments on trade and micro moments on the overlap in size between exporters and non-exporters and continuing and exiting firms in both countries during the pre-CUSFTA period. Matching these data moments requires that we choose parameters that result in multiple deviations from a standard Melitz-Pareto model and a substantial deviation from iso-elastic import demand. Once our model is calibrated, it is an ideal laboratory to study the potential biases of existing welfare sufficient statistics and assess the causal effect of a CUSFTA-style trade shock on entry and exit into the Canadian market and welfare gains. We generate simulated data from the model and compare the true welfare change from trade liberalizations to the measured welfare change of different sufficient statistics, such as Arkolakis et al. (2012), Melitz and Redding (2015), Head et al. (2014). Our simulations confirm that under realistic deviations from iso-elastic demand, the ACR statistic substantially overestimates the welfare gains from trade liberalization relative to the true welfare gains captured by our statistic. Our analysis also shows that alternative welfare statistics proposed in the literature can generate large biases in both directions in a setting with multiple deviations from Pareto productivity and strict sorting into exporting and a large trade shock.

Our paper builds primarily on the Arkolakis et al. (2012) sufficient statistics approach to welfare gains from trade and the CES gains from variety approach developed by Feenstra

micro-data is available to estimate changes in firm markups and productivity, these will typically only be available for continuing domestic incumbent firms, precisely the firms that our statistic uses as a benchmark. We return to the question of changes in the value of domestic incumbent firms later in our analysis and consider a variant of our sufficient statistic that allows us to capture some of these effects by specifying a subset of domestic firms that are unaffected by CUSFTA liberalization.

(1994). Like Head et al. (2014) and Melitz and Redding (2015), we allow for departures from a Pareto distribution of firm productivity and therefore do not impose iso-elastic import demand. However, our welfare sufficient statistic and modeling exercise also allow for deviations from strict productivity-based sorting into production and export. These deviations from the standard Melitz model appear to be a common feature of trade data (Eaton et al. (2011), Armenter and Koren (2015)) as well as the CUSFTA setting, and they imply that productivity cutoffs and the firm size distribution are not enough to characterize the welfare implications of firm selection for the gains from trade. We differ from Eaton et al. (2011) and Armenter and Koren (2015) by focusing on how these deviations from standard Melitz selection affect domestic exit and welfare measurement.[2]

In contemporaneous work, Adao et al. (2020) also analyze welfare gains from trade with a similar minimal set of assumptions and allow for both an arbitrary distribution of firm productivity and deviations from productivity-based sorting into markets. Their approach involves estimating a general elasticity function for the extensive margin using cross-country data on trade, firm export shares and firm entry together with semi-parametric gravity equations. This has the advantage of allowing for estimates of counter-factual gains from trade for an arbitrary range of trade costs, including those consistent with the autarky equilibrium. However, estimating this general elasticity function for the extensive margin requires additional data and assumptions about how these data inform identification of the elasticity function. Our welfare statistic is much easier to implement, requiring only a measure of the change in market shares of continuing domestic firms, but since this can only be observed ex-post our statistic cannot be used to analyze counter-factuals. Our model calibration and simulation exercise allow us to consider counter-factuals, but they require more data and impose a parametric structure.

This paper also relates to our earlier work (Hsieh et al. (2020)) that explores the margins of domestic and foreign firm selection in the context of CUSFTA. Our focus here is on

[2] A recent contribution by Fernandes et al. (2020) suggests that whether deviations from iso-elastic import demand are quantitatively important depends on the details of the model. They calibrate a model with log-normal firm heterogeneity in productivity, destination specific tastes and fixed entry costs to match firm-level bilateral export data and find that the trade elasticity in their model is fairly constant. Head et al. (2014) also observe that when replacing Pareto with log-normal productivity in a symmetric Melitz model, the details of the calibration matter for whether the gains from trade vary substantially.

estimating the overall gains from trade liberalization rather decomposing the foreign and domestic extensive margins. Here we also adopt a simulation-based approach to quantifying the contribution of CUSFTA to firm entry and exit.

The paper proceeds as follows. We first derive a new sufficient statistic for the gains from trade that holds for all models of trade with Dixit-Stiglitz preferences, including ones where the import demand elasticity is not constant. We then use micro-data from Canadian and US manufacturing to calculate Canada's gains from trade in 1988-1996 after CUSFTA was signed using actual data. The last section calibrates the welfare effect of CUSFTA in a generalized two-country Melitz model that does not impose the assumption of iso-elastic import demand and allows for data-driven deviations from strict selection into exporting.

2 A New Sufficient Statistic for the Gains from Trade

This section derives our sufficient statistic for the gains from trade liberalization. In anticipation of our later application, we derive this statistic in the context of a generalized Melitz (2003) model where import demand is not necessarily iso-elastic. We do this by allowing the productivity distribution to take any form and by not imposing any restriction on selection into production and exporting. But it will become clear that our sufficient statistic also measures the gains from trade in a Ricardian model with an arbitrary distribution of productivity.

Consumers have constant elasticity of substitution preferences over differentiated varieties sourced from many countries. These varieties are produced by monopolistic firms with heterogeneous productivities at constant marginal costs using labor only and trade is subject to iceberg costs. We remain agnostic about the determinants of entry into production and exporting and simply say that M_{ij} firms from country i serve country j. Hence, there may or may not be fixed market access costs and firms may or may not sort into production and exporting according to productivity cutoffs.

In this environment, a country i firm with productivity φ faces a demand $q_{ij}(\varphi) = \frac{p_{ij}(\varphi)^{-\sigma}}{P_j^{1-\sigma}} Y_j$ in country j, where p_{ij} is the delivered price in country j, P_j is the price index in country j, Y_j is the income in country j, and $\sigma > 1$ is the elasticity of substitution. As a result, it adopts a constant markup pricing rule $p_{ij}(\varphi) = \frac{\sigma}{\sigma-1}\frac{w_i \tau_{ij}}{\varphi}$, where w_i is the wage in country i and $\tau_{ij} > 1$ are the iceberg trade costs. Bilateral trade flows can therefore be expressed as a function of average prices, $X_{ij} = M_{ij}\left(\frac{\tilde{p}_{ij}}{P_j}\right)^{1-\sigma} Y_j$, where average prices are in turn a function of average productivities, $\tilde{p}_{ij} = \frac{\sigma}{\sigma-1}\frac{w_i \tau_{ij}}{\tilde{\varphi}_{ij}}$, where $\tilde{\varphi}_{ij}$ is a weighted harmonic mean of productivity.[3]

Consider now a shock to the economy, which causes some firms to exit and others to enter. We focus on trade liberalization in our application but our method really applies to any shock. We denote by M^c_{ij} the subset of *continuing* firms, defined as firms which are active both before and after the shock. Bilateral trade flows associated with continuing firms can be written as $X^c_{ij} = M^c_{ij}\left(\frac{\tilde{p}^c_{ij}}{P_j}\right)^{1-\sigma} Y_j$, where average prices and average productivity are defined only over this subset of firms, $\tilde{p}^c_{ij} = \frac{\sigma}{\sigma-1}\frac{w_i \tau_{ij}}{\tilde{\varphi}^c_{ij}}$. By definition, there are no changes in the set of continuing firms so that M^c_{ij} remains unchanged and $\tilde{\varphi}^c_{ij}$ changes only if there are within-firm productivity effects (i.e. there are no Melitz-type selection effects on $\tilde{\varphi}^c_{ij}$).

We derive our sufficient statistic by focusing on the domestic market share of continuing domestic firms, $\lambda^c_{jj} \equiv \frac{X^c_{jj}}{Y_j}$. Using our expression for X^c_{ij} above, we can express price index changes as $\Delta \ln P_j = \Delta \ln \tilde{p}^c_{jj} + \frac{1}{\sigma-1}\Delta \ln \lambda^c_{jj}$. From our expression for $\tilde{p}^c_{ij}$ above, we know that $\Delta \ln \tilde{p}^c_{jj} = \Delta \ln w_j - \Delta \ln \tilde{\varphi}^c_{jj}$ so that we can write changes in the domestic real wage as $\Delta \ln \frac{w_j}{P_j} - \Delta \ln \tilde{\varphi}^c_{jj} = -\frac{1}{\sigma-1}\Delta \ln \lambda^c_{jj}$. Changes in the domestic real wage are equal to changes in per-capita welfare if labor income is proportional to total income since then $\Delta \ln \frac{w_j}{P_j} = \Delta \ln \frac{Y_j/L_j}{P_j} \equiv \Delta \ln W_j$. This holds, for example, under free entry and we impose this assumption henceforth. We can thus write:

$$\Delta \ln W_j - \Delta \ln \tilde{\varphi}^c_{jj} = -\frac{1}{\sigma-1}\Delta \ln \lambda^c_{jj} \tag{1}$$

This equation says that anything that affects welfare, other than the productivity of continuing domestic firms, shows up as changes in λ^c_{jj}. One implication of this is that the effect of changes in trade costs on welfare, including the effect of any reallocation and entry and exit induced by the change in trade costs, can be measured by one simple statistic, the change

[3]Specifically, $\tilde{\varphi}_{ij} \equiv \left(\int_{\varphi\in\Phi_{ij}} \varphi^{\sigma-1} dG_i\left(\varphi|\varphi\in\Phi_{ij}\right)\right)^{\frac{1}{\sigma-1}}$, where Φ_{ij} is the set of productivities corresponding to all country i firms serving country j and $G_i\left(\varphi|\varphi\in\Phi_{ij}\right)$ is their cumulative distribution.

in the domestic market share of continuing domestic firms $\Delta \ln \lambda^c_{jj}$, and one parameter, the elasticity of substitution σ.

We want to make clear two points about our proposed statistic for the gains from trade in equation (1). First, while it captures all welfare effects from domestic net entry and changes in the price or variety of foreign imports when comparing two equilibria, interpreting it as the welfare gains from trade assumes that all of these changes are brought about by trade. This may not be true. For example, in a closed economy our sufficient statistic boils down to $\frac{1}{\sigma-1}\Delta \ln \lambda^c = \frac{1}{\sigma-1}\Delta \ln M_{jj} + \Delta \ln \frac{\tilde{\varphi}_{jj}}{\tilde{\varphi}^c_{jj}}$, which is simply the gains from entry of new domestic varieties net of the losses from domestic exit.

Second, even if the changes underlying computation of 1 are driven by trade, we do not know whether they are due to changes in *trade costs*, e.g. import tariff cuts negotiated under CUSFTA. For example, differential productivity growth (domestic vs. foreign) or changes in the fixed cost of exporting can also affect welfare through domestic net entry and changes in the price and variety of foreign imports. Our approach to assessing the contribution of reductions in trade costs under CUSFTA to extensive margin welfare gains is to specify a full-blown model, which we do in section 5.

We end this section by comparing our sufficient statistic in equation (1) with two widely used statistics by Feenstra (1994) and ACR. First, our formula can be thought of as an application of Feenstra (1994), albeit one that is very different from what is typically done in the literature. While prior papers such as Broda and Weinstein (2006) use Feenstra (1994) to measure the *import variety* gains from trade, we apply it to measure the *overall* gains from trade. Feenstra (1994) decomposes price index changes (Δ ln P) into a term capturing changes in the prices of continuing goods ($\sum_{i\in I^c} \bar{\mu}^c_i \Delta$ ln p_i, where I^c is a subset of continuing goods and $\bar{\mu}^c_i$ are Sato-Vartia weights) and a residual commonly thought of as capturing changes in the set of available goods (the Feenstra ratio $\frac{1}{\sigma-1}\Delta \ln \lambda^c$). However, the set I^c can really include any subset of continuing goods in which case the Feenstra ratio then also captures changes in the prices of the remaining continuing goods.[4] Our approach essentially boils down

[4]Note that in principle this also allows us to choose a set of continuing goods that experience no change in productivity, eliminating the "continuing firm productivity term" in equation (1). Any changes in the productivity of other continuing firms, foreign or domestic, is then captured by the Feenstra ratio. We provide an illustrative example in the appendix (section A7) where we restrict the set of continuing domestic firms in I^c to only those in sectors that had initially low tariffs and hence experienced minimal tariff decreases due to

to choosing continuing domestic goods. Intuitively, the market share of continuing domestic goods measures the net effect of all the margins of adjustment triggered by a change in trade costs. In a model with only adjustment on the intensive margin, the share of continuing domestic goods falls when a reduction in trade costs lowers the prices of foreign goods. In models that also have adjustment on the extensive margin, the share of continuing domestic goods also falls with more and better entering foreign varieties and rises with more and better exiting domestic varieties.

Our sufficient statistic in equation (1) is also a generalization of the formula by ACR. While we derived this sufficient statistic in a generalized Melitz (2003) model, it should be clear from our derivations that it holds in all models satisfying $X_{ij} \propto M_{ij} \left(\frac{\tilde{p}_{ij}}{P_j}\right)^{1-\sigma} Y_j$, $\tilde{p}_{ij} \propto \frac{w_i \tau_{ij}}{\tilde{\varphi}_{ij}}$, and $w_j L_j \propto Y_j$. For example, it also holds in a generalized Eaton and Kortum (2002) model that allows for an arbitrary productivity distribution if M_{ij} is reinterpreted as the number of goods shipped from country i to country j. Recall that ACR require four "model primitives" - (i) Dixit-Stiglitz preferences, (ii) one factor of production, (iii) linear cost functions, and (iv) perfect or monopolistic competition - and three "macro-level restrictions" - (i) trade is balanced, (ii) aggregate profits are a constant share of aggregate revenues, and (iii) the import demand system is iso-elastic. Their model primitive (i) immediately implies our first key equation $X_{ij} \propto M_{ij} \left(\frac{\tilde{p}_{ij}}{P_j}\right)^{1-\sigma} Y_j$, while their model primitives (i)-(iv) together yield our second key equation $\tilde{p}_{ij} \propto \frac{w_i \tau_{ij}}{\tilde{\varphi}_{ij}}$. Our third key equation $w_j L_j \propto Y_j$ follows from their macro-level restrictions (i) and (ii) so that we effectively relax their macro-level restriction (iii). In the appendix (section A1), we elaborate further on the link between our formula and ACR's (as well as a related formula by Melitz and Redding (2015)). [5]

3 Data

The free trade agreement between Canada and the United States was signed on January 2, 1988. It mandated the elimination of bilateral import tariffs in manufacturing, phased-in over

CUSFTA.

[5]In an earlier paper (Hsieh et al. (2020)), we show how the Melitz framework analyzed here generalizes to richer economic environments. In particular, we consider non-traded and intermediate goods, endogenous markups, tariff revenue, multiproduct firms, and heterogeneous quality. The bottom line is that only minor (if any) adjustments are needed to accommodate these extensions so we do not elaborate on this here.

a ten-year period starting on January 1, 1989. By 1996, Canadian tariffs on US imports had fallen from an average of 8% (equivalent to a 16% effective tariff rate) to about 1%. US tariffs on Canadian imports fell from about 4% in 1988 to below 1% during this period. Bilateral manufacturing trade between the two countries roughly doubled in nominal terms during this period. For Canada, the free trade agreement was a large shock, as trade with the US accounts for about 70% of Canadian trade in manufacturing. In addition, as discussed in Trefler (2004), CUSFTA was not accompanied by other macroeconomic reforms or implemented in response to a macroeconomic crisis.

We need information on domestic sales of continuing firms in Canada before and after CUSFTA came into force. We use the micro-data from Canada's Annual Survey of Manufacturing Establishments.[6] This survey covers all but the very smallest Canadian manufacturing establishments with sales below $30,000 Canadian dollars. We focus on the 1978-1988 and 1988-1996 time periods. We consider the 1978-1988 period as the "pre-CUSFTA" and the 1988-1996 period as "post-CUSFTA" period.[7] The information we use from this data is the establishment's id, exports, and sales. In each of the two time periods, we use the establishment's id to identify firms as entrants, exiters, and continuing firms. We define an *entrant* as an establishment not in the data at the *beginning* of the time period, an *exiter* as an establishment not in the data at the *end* of the time period, and a *continuing* establishment as one that was present in the data at the beginning and at the end of a time period. We supplement these data on domestic sales by Canadian firms with data on total manufacturing exports to Canada from the United States.

Equation (1) says that all we need to measure the gains from trade is the change in the Canadian market share of continuing Canadian firms, which we can calculate using only the Canadian firm data together with data on aggregate US manufacturing exports to Canada. When we later calibrate a model to the pre-CUSFTA equilibrium, we also need micro-data on US firm exports to Canada. We use micro-data from the quinquennial US manufacturing census, which provides data on exports at the establishment level starting in 1987.[8] The last

[6]This survey was initially called the Census of Manufactures and is now known as Annual Survey of Manufactures.

[7]We also chose these time periods because Statistics Canada officials indicated to us that the years with the best sampling frame are 1978, 1988, and 1996.

[8]The US census does not report establishment level exports by destination country so we assume that all

thing we need is the elasticity of substitution, and we use the estimates from Oberfield and Raval (2014) based on firm markups in US manufacturing.[9] The elasticities at the two-digit level range from 3.3 to 4.4 and average to 3.7. Since we only use aggregate data, we simply work with the average elasticity of 3.7.[10]

4 Canada's Gain during CUSFTA

In this section, we apply our sufficient statistic (1) to measure Canada's gains from CUSFTA in a simple before-after analysis. As we discussed earlier, this analysis is only identified if the observed change in the sufficient statistic is also entirely driven by CUSFTA. We will investigate this question in detail in the next section with the help of a quantitative model. Our conclusion will be that around 90% of the gains from CUSFTA implied by our simple before-after analysis can be causally attributed to CUSFTA.

We begin by showing key summary statistics from the Canadian and US micro-data. Table 1 shows the domestic sales of exiting (row 1) and entering (row 2) Canadian firms as a share of domestic sales of all Canadian firms in 1978-1988 (column 1) and 1988-1996 (column 2). The share of exiting firms is defined as the ratio of revenues of firms that exit between t and $t+1$ to total revenues at time t. And the share of entrants is the ratio of revenues of firms that enter between t and $t+1$ to total revenues in period $t+1$. To compare the numbers across the two time periods we study, we convert the share of entrants and exiters in the 1978-1988 period into shares over an 8 year period.[11] A key fact is that the share of exiting firms increased, from 24.41% to 28.01% between 1978-1988 and 1988-1996, while the same share for entrants declined from 21.55% to 18.81%.

Row 3 in Table 1 shows the change in the share of continuing domestic firms as a share of all domestic firms implied by the change in the share of exiters in row 1 and entrants in row 2. Specifically, the third row shows that the change in the sales of continuing domestic firms

US exporters in manufacturing also export to Canada. Canadian customs collects transaction-level data on imports from the US, but this data is only available after 1992 and cannot be reliably matched to US firms. US customs does not separately collect transaction-level data on exports to Canada.

[9] See Table VII of Oberfield and Raval (2014)'s online appendix. We used the concordance from Peter Schott's website to match them to 2-digit Canadian SIC codes.

[10] This value is also consistent with mean elasticities estimated using panel import data for the US (Broda and Weinstein (2006)) and Canada (Chen and Jacks (2012)) at similar levels of aggregation.

[11] We multiply the share of entrants and exiters in 1978-1988 by 8/10.

as a share of sales of all domestic firms increased from 2.97% before CUSFTA (1978-1988) to 12.03% after CUSFTA (1988-1996).

Table 1: Revenue Shares of Canadian firms in Canadian Market

	1978–1988	1988–1996
Exiter Domestic Revenue Share [1]	24.41%	28.01%
Entrant Domestic Revenue Share [2]	21.55%	18.81%
$\Delta \ln X^c_{jj}/X_{jj}$ [3]	2.97%	12.03%
$\Delta \ln \lambda_{jj}$ [4]	-1.36%	-23.86%
$\Delta \ln \lambda^c_{jj}$ [5]	1.61%	-11.83%

[1] Total domestic revenues of *exiting* Canadian firms/total domestic revenues of all Canadian firms at beginning of each period.
[2] Total domestic revenues of *entrant* Canadian firms/total domestic revenues of all Canadian firms.
[3] Change in domestic revenues of *continuing* Canadian firms/all Canadian firms.
[4] Change in total domestic sales of *all* Canadian firms/total sales in Canadian market.
[5] Change in total domestic revenues of *continuing* Canadian firms/total sales in Canadian market.

Sources: Column 1 imputes the share or changes over 8 years based on the change over 10 years. Calculated from micro-data of Canada's Annual Survey of Manufacturing. See text for details.

Equation (1) says that the key statistic is the change in domestic sales of continuing domestic firms as a share of *total* sales in the domestic market. This is simply the sum of the change in sales of continuing domestic firms as a share of all domestic firms shown in row 3 and the share of domestic firms in total sales. The latter, shown in the fourth row, indicates that the market share of domestic Canadian firms fell massively after CUSFTA. The last row shows $\Delta \ln \lambda^c_{jj}$ as the sum of row 3 and 4. The share of continuing domestic firms in total sales in the Canadian market fell by 11.83% in 1988-1996, compared to an increase of 1.61% in the period prior to CUSFTA.

Table 2 provides further insight into what drives the change in revenue shares of incumbent domestic producers by presenting data on the number of plants ("variety" margin in the top panel) and in the average productivity of these plants ("productivity" margin, in the bottom panel). The first two rows in Table 2 show the number of exiting and entering plants as a share of all plants, and the third row shows the implied change (over an 8-year period) in the number of plants. So the key facts here are that over the 1988-1996 period, the number

Table 2: Variety and Productivity Margins of Entry and Exit

	Canadian Firms 1978–1988	Canadian Firms 1988–1996	US Exporters 1987–1997
Variety Margin			
Exiters (% of All Plants)[1]	51.68%	49.56%	54.69%
Entrants (% of All Plants)[2]	64.46%	43.76%	72.87%
$\Delta \ln M_{ij}$ [3]	24.57%	-10.90%	41.04%
Productivity Margin			
Exiters (Size relative to all plants)[4]	47.23%	56.52%	64.86%
Entrants (Size relative to all plants)[5]	33.44%	42.99%	53.14%
$\Delta \ln \tilde{\varphi}_{ij}/\tilde{\varphi}^c_{ij}$ [6]	-10.20%	-0.42%	-13.67%

[1] # exiting plants/total # of plants (columns 1-2) or exiting exporters/total # of exporters (column 3) at beginning of each period.

[2] # new plants/total # of plants (columns 1-2) or new exporters/total # of exporters (column 3) at end of each period.

[3] % change in total # of plants (columns 1-2) or exporters (column 3).

[4] Average domestic revenues of exiting plants/all plants (columns 1-2) or average exports of exiting exporters/all exporters (column 3) at beginning of period.

[5] Average domestic revenues of new plants/all plants (columns 1-2) or average exports of new exporters/all exporters (column 3) at end of period.

[6] Productivity growth of all plants/continuing plants (columns 1-2) or all exporters/continuing exporters (column 3), measured as $\Delta \ln \tilde{\varphi}_{ij}/\tilde{\varphi}^c_{ij} = -\frac{1}{\sigma-1}\Delta \ln \frac{\tilde{r}^c_{ij}}{\tilde{r}_{ij}}$ where $\tilde{r}$ denotes average revenues.

Notes: Columns 1 and 3 impute 8-year change from 10-year changes. Calculated from micro-data of Canada's Annual Survey of Manufacturing and US Manufacturing Census. See text for details.

of Canadian plants decreased, which pushes up the domestic revenue share of continuing Canadian plants relative to all Canadian plants, and the number of US exporters increased, which contributed to the lower share of the domestic market captured by Canadian plants.

But entry and exit also potentially change average productivity through selection effects. This effect is shown in the bottom panel in Table 2. The first two rows show the raw data, namely sales of exiting firms relative to all firms and sales of entering firms relative to all firms. As expected, exiting and entering firms are smaller than the average firm. The net effect of entry and exit on average productivity is then given by $\Delta \ln \left(\tilde{\varphi}_{ij}/\tilde{\varphi}^c_{ij}\right) = \frac{1}{\sigma-1}\Delta \ln \left(\tilde{r}^c_{ij}/\tilde{r}_{ij}\right)$ where $\tilde{r}$ denotes average revenues. This number is shown in the last row in bottom panel of Table 2. Net entry of new US exporters lowered the average productivity of US exporters.

Exit by Canadian firms during this period did eliminate firms that were less productive than average, but there was also entry by less productive firms, leading to only small effects of net exit on average productivity due to seletion during the 1988-1996 CUSFTA period. However, domestic firm selection during the CUSFTA period clearly increased average productivity of Canadian firms relative to the period before CUSFTA (in 1978-1988).

We re-emphasize the point that what matters for welfare is the net effect of the extensive and the intensive margins, and the sufficient statistic for this net effect is the share of exiting or entering firms in total revenues shown in Table 1. So in a sense, once we have data on shares of exiting and entering firms, we do not need to know the contribution of the number of plants or plant revenue to these shares. Put differently, exit is always welfare reducing and entry is always welfare enhancing, and how much they matter is measured by the revenue shares of the two types of firms relative to continuing firms.

Table 3 shows the welfare gains calculated as $-\frac{1}{\sigma-1}\Delta\ln\lambda^c_{jj}$, where the latter is measured as the difference between $\ln\lambda^c_{jj}$ before and after CUSFTA and assuming $\sigma = 3.7$. The first row in the table says that Canadian welfare rose by almost 5% over the 1988-1996 CUSFTA period from the net effect of domestic exit, US entry and lower prices for continuing US firms.[12]

The next two rows report welfare gains assuming iso-elastic import demand. In particular, we can use ACR's sufficient statistic if we assume that firm productivity follows a Pareto distribution and there is strict productivity-based sorting into production and exporting. With these two assumptions, the welfare gains are given by:

$$\Delta\ln W_j - \Delta\ln\tilde{\varphi}^c_{jj} = -\frac{1}{\theta}\Delta\ln\lambda_{jj} \tag{2}$$

where θ is the trade elasticity, which happens to be the shape parameter of the Pareto distribution in this particular model. There are two differences between (2) and (1). First, λ_{jj} in equation (2) is the share of *all* domestic firms while λ^c_{jj} in equation (1) is the share of

[12]Note that this underestimates the total welfare gains during this period if productivity for continuing Canadian firms grew, regardless of whether this within-firm productivity growth can be attributed to CUSFTA or other factors. Appendix Section A2 provides additional analysis of the potential magnitude of these effects using a similar framework under the assumption that only continuing firms in industries with high initial tariffs were affected by CUSFTA.

continuing domestic firms. The former is the number in row 4 in Table 1 and the latter in row 5. Second, the elasticity in equation (2) is a function of the shape parameter of the Pareto distribution, while the elasticity in equation (1) is a function of the elasticity of substitution across varieties.

So the only additional number we need to estimate equation (2) is θ. There are at least three ways to estimate this parameter. The first two involve data moments from a trade shock. First, we can use the fact that with an (untruncated) Pareto distribution, $\theta = -\Delta \ln \left(\frac{M_{ij}/M_{jj}}{\tilde{\varphi}_{ij}/\tilde{\varphi}_{jj}} \right)$. A second approach is to use the trade elasticity implied by the model. In particular, Melitz and Redding (2015) propose the "arc elasticity" $\theta = \Delta \ \ln \frac{(1-\lambda_{jj})/\lambda_{jj}}{(1+\tau_{ij})}$. A third way is to use the distribution of firm size in the steady state. Specifically, if the distribution of firm productivity follows a Pareto distribution, θ is the product of $\sigma - 1$ and the elasticity of the rank of firm size with respect to firm size.

We use our micro-data from Canada and the US to estimate θ using these three methods. For the first method, we use the $\Delta \ln M$ and $\Delta \ln \tilde{\varphi}$ from 1978-1988 to 1988-1996 (shown in Table 2), which gives us $\theta = 2.9$. This estimate for θ is almost identical to that obtained from the cross-sectional distribution of firm size. Specifically, the elasticity of firm rank with respect to firm size is 1.06, which combined with $\sigma = 3.7$ yields $\theta = 2.86$. Finally, we estimate θ from the "arc-elasticity" proposed by Melitz and Redding (2015) as $\theta = \Delta \ \ln \frac{(1-\lambda_{jj})/\lambda_{jj}}{(1+\tau_{ij})}$. As we will describe later, CUSFTA lowered the iceberg trade cost of shipping goods from the US to Canada by 23%, which combined with the change in λ_{jj} (shown in Table 1) yields $\theta = 2.63$.

The last two rows in Table 3 calculate the ACR statistic using these two values of θ. Using θ=2.9, ACR's statistic suggests that CUSFTA increased welfare by 7.76%. Using $\theta = 2.63$, ACR's gains are even larger, at 8.56%. These numbers are 50%-70% larger than our estimate of 4.98%, shown in the first row of the table.

5 Simulation of welfare effects of CUSFTA

In this section, we use a quantitative trade model to isolate the causal effect of CUSFTA on our sufficient statistic (1). This is meant to complement our simple before-after analysis

Table 3: Alternative Estimates of Canada's Welfare Gains from Trade after CUSFTA

	Δ Welfare
Our statistic: $-\frac{1}{\sigma-1}\Delta \ln \lambda_{jj}^{c}$	4.98%
ACR with $\theta = 2.9$ [1]	7.76%
ACR with $\theta = 2.63$ [2]	8.56%

[1] ACR welfare statistic (2) using $\theta = -\Delta \ln \left(\frac{M_{ij}/M_{jj}}{\tilde{\varphi}_{ij}/\tilde{\varphi}_{jj}}\right)$ or using Zipf's Law calculated as $\theta = \xi \cdot (\sigma - 1)$ where ξ is the elasticity of firm rank with respect to firm employment.
[2] ACR welfare statistic (2) with arc-elasticity proposed by Melitz and Redding (2015) and measured as $\theta = \Delta \ln \frac{(1-\lambda_{jj})/\lambda_{jj}}{(1+\tau_{ij})}$
Note: Table shows the welfare gains for Canada from 1988 to 1996 (relative to the 1978-1988 period) based on the data moments in Table 1.

from the previous section which simply attributed the entire change observed in the data to CUSFTA. We build on a standard Melitz-Pareto model but want to relax the strong supply side assumptions necessary to generate an iso-elastic import demand system. Therefore, we consider two generalized Melitz-Pareto models, both of which have as a special case the model with iso-elastic import demand. We then calibrate these models to the data, and let the data tell us about the extent of the departure from a model with iso-elastic import demand.

Our first model follows Melitz and Redding (2015). Specifically, we assume two countries populated by representative consumers with CES preferences, firms pay a common fixed cost to produce in each period and another fixed cost to export, and the steady-state distribution of firm productivity follows a *truncated* Pareto distribution.[13] In this model then the truncation parameter determines the extent of the departure from an iso-elastic import demand. With a large enough truncation parameter, the model will be quantitatively close to a Melitz-Pareto model with iso-elastic import demand.

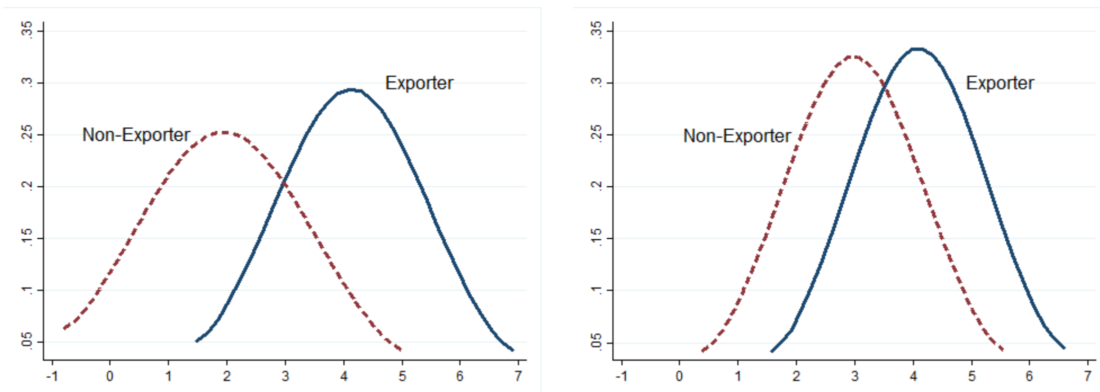
Our second generalization drops the assumption that all firms face the same fixed cost of production and exporting. We do this for two reasons. First, there is abundant evidence that strict sorting into exporting may not hold empirically. Eaton et al. (2011) and Armenter and Koren (2015) show that there is a substantial overlap in the size distribution between exporters and non-exporters in France and the US. Figure 1 (top panel) plots the distribution

[13] We also assume free entry (after paying the fixed cost of entry), balanced trade, and that firm productivity is subject to a shock that follows a mean zero log normal distribution.

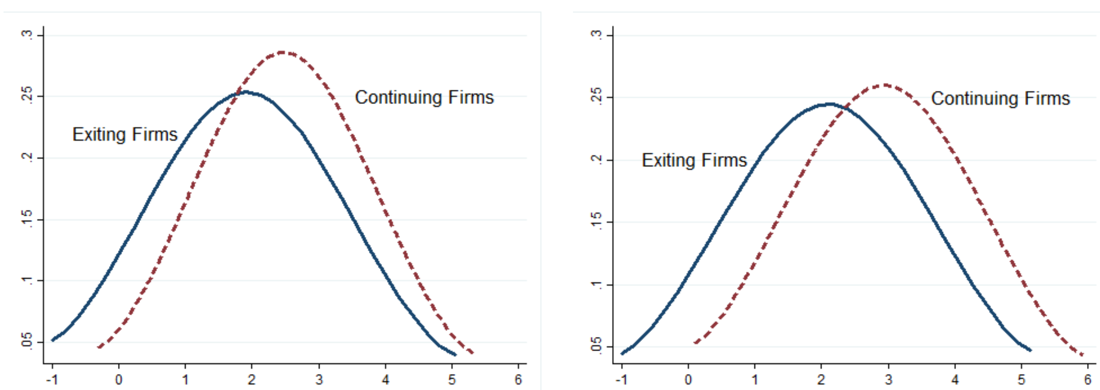
of employment for exporters and non-exporters in our Canadian and US data, and shows that a similar fact holds in Canada and the US We can capture the fact that many exporters are smaller than non-exporters, and vice versa, by allowing the fixed cost of exporting to differ across firms. Specifically, we assume that export fixed costs in each country are i.i.d. Pareto, as in Armenter and Koren (2015).

Figure 1: Distribution of Employment

US | Canada



Exporters vs. Non-Exporters



Exiting vs. Continuing Establishments

Note: Top panel shows the distribution of log employment of exporting and non-exporting establishments in Canada in 1996 and the US in 1997. Bottom panel shows distribution of log employment of exiting and continuing establishments in Canada in 1988 and US in 1987. Exiting plants leave the data between 1988 and 1996 (Canada) or between 1987 and 1997 (US). Continuing plants are in the data in the initial and final years. Statistics calculated from the micro-data of manufacturing firms in Canada and the US

Second, the bottom panel in Figure 1 plots the ex-ante distribution of plant size for exiting vs. continuing firms in US and Canadian manufacturing. As can be seen, many exiting plants are ex-ante substantially larger than continuing plants, and vice versa. We will also allow our model to reproduce this pattern, this time by relaxing the assumption that the fixed cost of domestic production are the same across firms. Specifically, we allow production fixed costs to vary with firm productivity according to $f_d + \beta \ln \varphi$ (note that we do not impose $\beta \neq 0$ but calibrate it using the data). The appendix (section A3) provides more detail on our models and our solution algorithm.

5.1 Model Calibration

Table 4 summarizes the key parameters of the two models. In both models, we assume $\sigma = 3.7$ and $\theta = 2.9$ and take as given employment in manufacturing in the two countries. In the model where all firms have common fixed production and exporting cost (Melitz and Redding (2015)), we then choose the parameters in the first column in Table 4 to fit the trade share, the number and relative employment of exporting and non-exporting firms, and the exit rate in the two countries.[14] The first column in Table 4 shows the resulting estimates of these parameters for this model.[15]

Our second model, which features heterogeneity in fixed production and export costs, has two more parameters: the elasticity of production fixed costs with respect to firm productivity β and the shape parameter of the Pareto distribution of export fixed cost α. We choose these additional parameters using data on the overlap in firm size between exporting and non-exporting firms and exiting vs. continuing firms in Canada and the US[16] The second column in Table 4 reports the parameters of the model that allows for overlap in the distribution of exporters and non-exporters and exiters and continuing firms. We need a very high level of dispersion in the fixed exporting costs to match the degree of overlap in employment between exporters and non-exporters in the data, with a (Pareto) shape parameter of $\alpha = 0.17$ for Canada. In addition, we need a positive elasticity of fixed operating cost with respect to productivity to match the overlap in the distribution of exiters vs. continuing firms. The

[14] Table 8 in the appendix (section A4) shows the values of the moments we target.

[15] Table 4 only shows the model parameters for Canada. Table 9 in the appendix (section A5) shows the parameters for the US

[16] Specifically, we target the 20th, 50th, and 80th percentiles of the employment distribution of exporting versus non-exporting firms, and continuing versus exiting firms. Table 8 in the appendix (section A4) shows the precise moments of the firm size distribution we target.

Table 4: Parameters for Canada

	Melitz-Redding *truncated Pareto only*	**Full model** *truncated Pareto + fixed cost heterogeneity*
Productivity Pareto truncation	3.8	4.3
Trade friction τ	2.15	2.15
Entry fixed cost	0.83	0.83
Production fixed cost intercept f_d	0.65	0.2
Production fixed cost slope β	-	3
Export fixed cost location f_x	0.6	0.09
Export fixed cost shape α	-	0.17
Exit shock standard deviation	0.07	0.051

Notes: Column 1 shows the model parameters for the model with a truncated Pareto distribution and common fixed costs (the Melitz and Redding (2015) model). Column 2 shows the calibrated parameters for the model a truncated Pareto distribution and heterogeneity in fixed costs of production and exporting. See Table 9 in the appendix (section A4) for US parameters.

appendix (sections A4 and A5) provides more detail on our calibration.

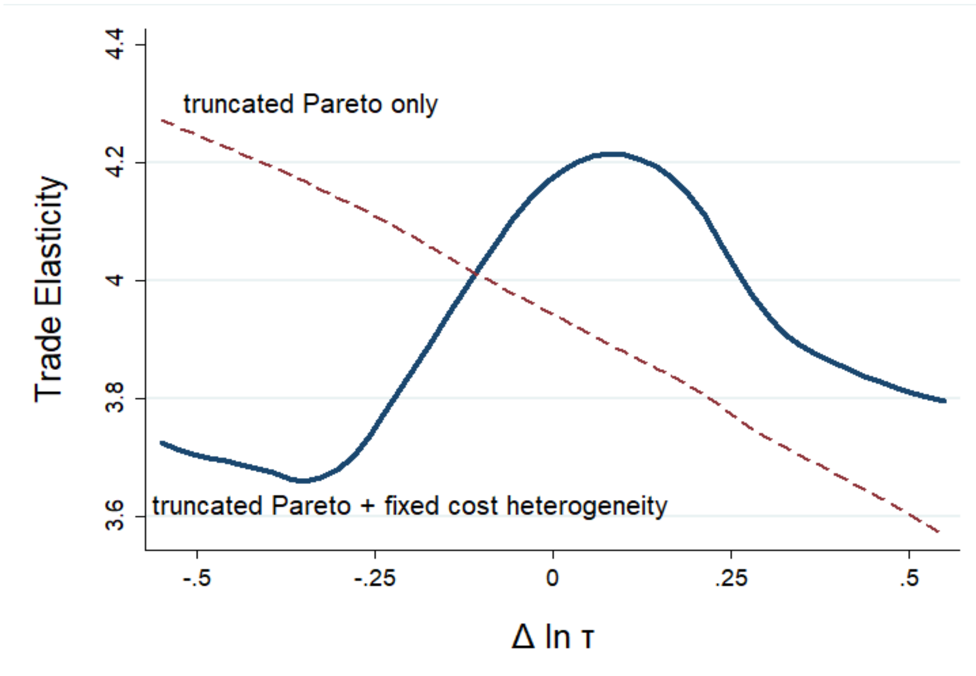
Figure 2 shows the trade elasticity for varying levels of changes in trade costs implied by the estimated parameters in the two models. As can be seen, the estimated model parameters imply significant departures from an iso-elastic import demand structure. So the gains from trade in our two models will not be measured accurately by ACR's formula. Of course, this observation alone does not tell us the degree of bias if we use ACR's statistic in these two models, which is a question we address in section 5.3.

5.2 Predicted welfare effects of CUSFTA

We now simulate the effect of the reduction in trade costs due to CUSFTA on Canada in the two models. We assume that Canadian trade costs for US imports fell by 23% due to CUSFTA. This includes a 16% decline in tariffs and the elimination of non-tariff barriers equivalent to about a 7% tariff.[17] On the US side, CUSFTA lowered US tariffs on Canadian imports by about 58% of the decline in Canadian tariffs on US imports. So we assume that

[17] See Trefler (2004) for the reduction in tariffs and Head and Ries (2001) for the tariff equivalent of non-tariff barriers.

Figure 2: Heterogeneous Trade Elasticities in Simulated Models



Note: Figure shows the simulated trade elasticity for different values of changes in trade cost in the model with a truncated Pareto distribution only and the model that also has heterogeneity in production and export costs. See text for details.

US trade costs for Canadian imports fell by 13.3% due to CUSFTA.[18]

The first panel in Table 5 shows the predicted change of the key data moments in response to the change in trade costs due to CUSFTA in the two models. For comparison, the table also replicates the same moments in the data in column 1.[19] Recall that we are using the models to isolate the effect of trade cost reductions so these are not moments we are trying to match. In the first row, we show that the growth in the domestic market share of incumbent Canadian firms is lower in both models than in the data. This is because the share of exiting firms is lower in both models than in the data. The second row shows the change in the share of all domestic firms. The model with only a truncated Pareto distribution predicts a

[18] In Table 10 in the appendix (section A6) we also consider a simulation where fixed exporting costs fall by a similar magnitude as tariffs, i.e. the fixed cost of exporting to Canada falls by 23% and the fixed cost of exporting to the US falls by 13%. Unsurprisingly, this results in larger changes in all of the table entries for both models except the "intensive margin" term. It has a much larger effect in the Melitz-Redding model than our full model because the dispersion of export fixed costs is so large in our model that there are many fewer marginal firms induced to export due to falling export fixed costs relative to a truncated Pareto model with strict sorting. We focus our analysis on changes in variable trade costs as we are not aware of direct evidence or previous studies that quantify changes in fixed exporting costs during this period.

[19] Specifically, "data" refers to the difference between 1988-1996 and 1978-1988 shown in Table 1.

Table 5: Simulated effects of CUSFTA: Data vs. Model

	Data 1988-1996	**Melitz-Redding** *Truncated Pareto Only*	**Full model** *Truncated Pareto + Fixed Cost Het.*
Market Shares:			
$\Delta \ln X_{jj}^c/X_{jj}$ [1]	9.04%	6.68%	5.91%
$\Delta \ln \lambda$ [2]	-22.50%	-22.55%	-18.04%
$\Delta \ln \lambda^c$ [3]	-13.44%	-15.87%	-12.13%
Total Welfare Change [4]	4.98%	5.88%	4.49%
Other moments:			
$\Delta \ln M_{jj}$ [5]	-10.9%	-17.39%	-14.19%
Exiter relative size [6]	0.57	0.37	0.57
$\Delta \ln M_{ij}$ [7]	41.0%	53.8%	36.05%
$\Delta \ln X_{ij}^c/X_{ij}$ [8]	4.16%	20.11%	6.16%

[1] Percentage change in domestic revenues of *continuing* Canadian firms/all Canadian firms. (Difference to 1978-1988 pre-trend.)
[2] Percentage change in total domestic sales of *all* Canadian firms/total sales in Canadian market. (Difference to 1978-1988 pre-trend.)
[3] Percentage change in total domestic revenues of *continuing* Canadian firms/total sales in Canadian market. (Difference to 1978-1988 pre-trend.)
[4] Gains from trade calculated from equation (1) from simulated data. Note: Simulated effect of a 23% (13%) reduction in Canadian (US) tariffs in model with a truncated Pareto distribution of productivity and common fixed costs (column 2) and heterogeneous fixed costs (column 3). "Data" in top panel is the difference between the annualized change in 1988-1996 compared to 1978-1988 shown in Table 1, accumulated over 8 years. See text for details. [5] Percentage change in count of Canadian plants 1988-1996.
[6] Relative size of exiting Canadian plants (1988-1996) relative to all plants in 1988.
[7] Percentage change in US exporting plants 1987-1997 from US Manufacturing Census.
[8] Percentage change in total revenues of *continuing* US exporting firms/total sales of exporting firms from US Manufacturing Census.

change in the domestic spending share almost exactly in line with the data, while the model that also allows for heterogeneity in export and production fixed costs predicts a lower fall in the domestic spending share. The third row, the share of *continuing* domestic firms in domestic spending, is simply the sum of the previous two rows and the key summary statistic in equation (1). So, the model with only the truncated Pareto distribution (column 2) predicts higher gains from CUSFTA than our earlier before-after comparison based on the raw data moments, while the model that also allows for heterogeneity in fixed costs (column 3) predicts slightly lower welfare gains.

The fourth row shows the change in welfare. In the data, this is about 5%. The Melitz-

Redding model predicts that CUSFTA increased welfare by 5.9%, while the predicted gains in the model that also allows for heterogeneity in fixed costs are closer to the gains suggested by the raw data, at 4.5%. So the full model suggests that CUSFTA "explains" about 90% of the gains reported in our earlier before-after analysis, while the model that only allows for truncation in the Pareto distribution suggests that CUSFTA accounts for about 120% of the observed gains.

The last four rows present other data moments from the 1988-1996 CUSFTA period for comparison with those generated by our calibrated models in response to a the decline in trade costs. The model with only a truncated Pareto distribution in column 2 generates too much domestic exit, domestic exiters that are too small, too many new US exporters, and too much US export growth coming from the extensive margin. The model with heterogeneous fixed costs fits the data moments much better as it leads to domestic exit by fewer but larger firms, and leads to less import growth on the extensive margin.

5.3 Comparing welfare statistics

The last thing we do is compare alternative welfare statistics in our two calibrated models. Remember that the estimated parameters for our two models imply that the import demand is not iso-elastic. The question then becomes, how large is the bias if we were to use welfare statistics that assume iso-elasticity?

The answer depends on what the two models predict in terms of the change in λ, as well as the elasticity θ. In section 4, the numbers we use for θ are 2.9 and 2.65. Table 6 shows the gains from trade liberalization due to CUSFTA in the simulation of the two models. The first row repeats the simulated gains from CUSFTA calculated from the formula in equation (1) shown earlier in Table 5. The second and third rows show ACR's statistic calculated from the simulated data with the two estimates of θ. In both models, the gains from trade calculated ACR's formula are larger than that obtained from equation (1) shown in row 1.

The last two estimates in Table 6 show the gains from trade calculated from formulas that do not rely on iso-elastic import demand functions. The first formula, from Melitz and Redding (2015) and Head et al. (2014), is a local approximation to the gains from trade that

Table 6: Welfare Measurement in Simulated Data

Welfare Statistic	**Melitz-Redding** *Truncated Pareto Only*	**Full model** *Truncated Pareto + Fixed Cost Het.*
Our statistic: $-\frac{1}{\sigma-1}\Delta \ln \lambda^c_{jj}$	5.9%	4.5%
ACR, $\theta = 2.9$ [1]	7.8%	6.2%
ACR, $\theta = 2.65$ [2]	8.5%	6.8%
Melitz and Redding (2015), local elasticity [3]	6.2%	3.8%
Melitz and Redding (2015), Δ lowest prod.[4]	6.0%	0%

[1] ACR welfare statistic with extensive-margin based estimate of trade elasticity $\theta = -\Delta \ln \left(\frac{M_{ij}/M_{jj}}{\tilde{\varphi}_{ij}/\tilde{\varphi}_{jj}} \right)$ or using Zipf's Law calculated as $\theta = \xi \cdot (\sigma - 1)$ where ξ is the elasticity of firm rank with respect to firm employment.
[2] ACR welfare statistic with arc-elasticity proposed by Melitz and Redding (2015) and defined by $\theta_{arc} = \Delta \ln \frac{(1-\lambda_{jj})/\lambda_{jj}}{(1+\tau_{ij})}$
[3] Local welfare statistic (3) with elasticity estimation using $\epsilon_L = (1-\sigma) + \frac{1}{\bar{X}_{jj}/X^{MIN}_{jj}} \cdot \Delta \ln \frac{M_{jj}}{\tau_{ij}}$ from Bas et al. (2017)
[4] Welfare statistic based on change in minimum productivity $\Delta \ln W_j = \Delta \ln \varphi_d$, where φ_d is the productivity of the marginal firm.
Note: Simulated effect of a 23% (13%) reduction in Canadian (US) tariffs in model with a truncated Pareto distribution of productivity and common fixed costs (column 1) and heterogeneous fixed costs (column 2). See text for details.

holds for any productivity distribution:

$$\Delta \ln W_j = -\frac{1}{\epsilon_L} \Delta \ln \frac{\lambda_{jj}}{M^e_j} \tag{3}$$

where M^e is the number of *potential* entrants and ϵ_L is the local trade elasticity. Note that the number of potential entrants M^e is not something that can be observed empirically, but we can use our models' predictions of this variable. As for the local trade elasticity, Bas et al. (2017) estimate it as $\epsilon_L = (1 - \sigma) + \frac{1}{\bar{X}_{jj}/X^{MIN}_{jj}} \cdot \frac{d \ln M_{jj}}{d \ln \tau_{ij}}$, where $\bar{X}_{jj}/X^{MIN}_{jj}$ denotes the ratio of the sales of the average firm to the smallest firm. In our simulated data, this formula gives us $\epsilon_L = 3.75$ and $\epsilon_L = 2.74$ for the models in columns 1 and 2 in Table 6, respectively.

The local approximation of the gains from trade given by equation (3) is shown in row 4 in Table 6. As can be seen, the gains from trade calculated from the local approximation are very similar to those obtained from equation (1) (shown in the first row). The difference of course is that equation (3) is valid only for small changes in trade costs, while we simulate a 23% (13%) tariff cut in Canada (the US). However, note that in our full model with fixed cost heterogeneity, the local approximation is strongly downward biased relative to formula (1), mostly driven by the fact that M_j^e falls in this model.

The last row in Table 6 shows the gains from trade calculated as the change in the minimum productivity of surviving firms. Head et al. (2014) and Melitz and Redding (2015) show that the minimum productivity is a sufficient statistic for the gains from trade for any change in trade costs and for any productivity distribution, but the statistic is only valid if all firms face the same fixed cost of production and exporting. In contrast to M^e which is unobservable, the productivity of the smallest surviving firm can be imputed from the size of the smallest surviving firm. Not surprisingly, the gain from trade calculated from this statistic in the simulated data in the Melitz and Redding (2015) model (column 1) is almost identical to that obtained from our formula in equation (1). This is not the case in the full model that also features heterogeneity in fixed costs. In the simulation of that model, shown in column 2, the productivity of the smallest surviving firm does not change due to CUSFTA so the gains from trade from CUSFTA predicted by that statistic is zero. The reason is of course heterogeneity in fixed costs, where some low productivity firms survive when trade costs fall because some of these firms face low fixed costs.

6 Conclusion

We propose a new sufficient statistic to measure welfare gains in trade models for which the import demand system is not necessarily iso-elastic. This includes a Ricardian model of trade with an arbitrary distribution of productivity and a Melitz model with any pattern of selection into exporting and production as well as any distribution of productivity. The statistic is simple to calculate, as it is just a function of one data moment, the market share of continuing domestic firms, and one parameter, the elasticity of substitution across varieties. When applied to the CUSFTA liberalization period, our statistic indicates an approximately 5% increase in Canadian welfare due to the combination of net domestic exit, net foreign entry and cheaper foreign varieties. These gains are substantially lower than would be implied by statistics that impose iso-elastic import demand on our setting.

There are however some limitations of our proposed statistic. First, it does not measure the potential effect of trade liberalization on productivity growth among incumbent domestic firms (except when applied as described in the appendix (section A2)). Second, the statistic by itself does not tell us what fraction of the implied welfare change is due to changes in trade costs. These limitations are common to other approaches that have been proposed in the literature and highlight the challenge of applying statistics derived from models assuming fixed productivity distributions and well-identified trade cost comparative statics to real world data. Our solution here is to calibrate a generalized Melitz model that does not impose iso-elastic import demand to help answer these questions in the context of Canada's gain from the trade due to CUSFTA.

A third limitation of our statistic is that it can be only be applied to measure the gains from past trade reforms, and cannot tell us the gains from counter-factual changes in trade costs such as a prospective liberalization or a hypothetical return to autarky. Outside of the fairly restrictive class of modeling assumptions that generate iso-elastic import demand, there are two approaches to this issue. One is to take a stand on a specific model (i.e. specifying the parametric distribution of firm productivity and fixed costs), which we do here by leveraging additional micro-data moments. An alternative, proposed by Adao et al. (2020), is to estimate non-parametric functions that capture the contribution of firm heterogeneity to the gains from trade, which requires sufficient data and identification assumptions to allow estimation of the shape of these functions in the parameter space of the counter-factual of interest.

We raise one point for future research. In our calibrated model, we interpret the overlap in size between exporters and non-exporters as due to heterogeneity in fixed exporting costs. Similarly, we model the overlap in size between survivors and exiters as reflecting heterogeneity in fixed production costs. That is obviously a simplification, and perhaps not a good one. Our proposed sufficient statistic does not rely on a specific interpretation of this fact, but we hope that future work will provide richer models for this important stylized fact.

References

R. Adao, C. Arkolakis, and S. Ganapati. Not-parametric gravity: Measurin the aggregate implications of firm heterogeneity. *Working Paper*, 2020.

C. Arkolakis, A. Costinot, and A. Rodriguez-Clare. New trade models, same old gains? *American Economic Review*, 102(1):94–130, 2012.

R. Armenter and M. Koren. Economies of scale and the size of exporters. *Journal of the European Economic Association*, 13(3):482–511, 2015.

M. Bas, T. Mayer, and M. Thoenig. From micro to macro: Demand, supply and heterogeneity in the trade elasticity. *Journal of International Economics*, 108:1–19, 2017.

C. Broda and D. Weinstein. Globalization and the gains from variety. *Quarterly Journal of Economics*, 121(2):541–585, 2006.

B. Chen and D. Jacks. Trade, variety, and immigration. *Economic Letters*, 117:243–246, 2012.

J. Dixon and A.-M. Rollin. Firm dynamics: Employment growth rates of small versus large firms in canada. *Statistics Canada Research Paper*, 11-622-M(25), 2012.

J. Eaton and S. Kortum. Technology, geography and trade. *Econometrica*, 70(5):1741–1779, 2002.

J. Eaton, S. Kortum, and F. Kramarz. An anatomy of international trade: Evidence from french firms. *Econometrica*, 79(5):1453–1498, 2011.

R. Feenstra. New product varieties and the measurement of international prices. *American Economic Review*, 84(1):157–177, 1994.

A. Fernandes, P. Klenow, S. Meleshchuk, M. Pierola, and A. Rodriguez-Clare. The intensive margin of trade: How big and how important? *Working Paper*, 2020.

K. Head and J. Ries. Increasing returns versus national product differentiation as an explanation for the pattern of u.s.-canada trade. *American Economic Review*, 91(4):858–876, 2001.

K. Head, T. Mayer, and M. Thoenig. Welfare and trade without pareto. *American Economic Review Papers and Proceedings*, 104(5):310–316, 2014.

C.-T. Hsieh, N. Li, R. Ossa, and M.-J. Yang. Accounting for the new gains from trade liberalization. *Journal of International Economics*, 127, 2020.

I. Kondo, L. Lewis, and A. Stella. On the u.s. firm and establishment size distributions. *Federal Reserve Board*, working paper, 2019.

P. Krugman. Scale economies, product differentiation and the pattern of trade. *American Economic Review*, 70(5):950–959, 1980.

L. Kueng, M. Yang, and N. Li. The impact of emerging market competition on innovation and business strategy: Evidence from canada. *Journal of Economic Behavior and Organizations*, 2020.

M. Kugler and E. Verhoogen. Prices, plant size, and product quality. *Review of Economic Studies*, 79(1):307–339, 2012.

M. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, 2003.

M. Melitz and S. Redding. New trade models, new welfare implications. *American Economic Review*, 165(3):1105–46, 2015.

M. Melitz and D. Trefler. Gains from trade when firms matter. *Journal of Economic Perspectives*, 26(2):91–118, 2012.

E. Oberfield and D. Raval. Micro data and macro technology. *Working Paper, Princeton University*, 2014.

J. Sutton. Sunk costs and market structure. *MIT Press*, 1991.

D. Trefler. The long and short of the canada-us free trade agreement. *American Economic Review*, 94(4):870–895, 2004.

Appendix

A1: Arkolakis et al. (2012) and Melitz and Redding (2015) as a special case of equation (1)

We now show how other sufficient statistics for the welfare gains from trade proposed in the literature can be derived as special cases of our formula (1). Consider first the case of general productivity distributions but strict sorting into production and export, i.e. the Melitz (2003) model where firms face a common fixed cost of production and exporting and sort based on productivity only. With strict sorting, the term $M_{jj}\tilde{p}_{jj}^{1-\sigma}$ is proportional to the mass of firms that pay the fixed entry cost (M^e) and the average productivity of firms above the domestic productivity cutoff (φ_d), so $M_{jj}\tilde{p}_{jj}^{1-\sigma} \propto M_{jj}\tilde{\varphi}_{jj}^{1-\sigma} = M^e\delta(\varphi_d)$ and $\delta(\varphi_d) \equiv \int_{\varphi_d}^{\varphi_{max}} \varphi^{\sigma-1}dG(\varphi)$. As we discuss in section 5.3, Head et al. (2014) and Melitz and Redding (2015) show how to derive an expression for local welfare changes under strict sorting but without any assumptions on the firm productivity distribution. Here we follow Melitz and Redding (2015) who observe that $\Delta \ln \delta(\varphi_d) = -\gamma(\varphi_d) d\ln \phi_d = -\gamma(\varphi_d) d\ln W$, where $\gamma(\varphi_d)$ is the hazard function for the distribution of log firm size within a market evaluated at the domestic productivity cutoff φ_d. The welfare change (1) can then be expressed as

$$\Delta \ln W_j - \Delta \ln \tilde{\varphi}_{jj}^c = -\frac{1}{(\sigma-1)+\gamma(\varphi_d)}\Delta \ln \frac{\lambda_{jj}}{M_j^e} \tag{4}$$

The term $(\sigma-1)+\gamma(\varphi_d)$ can also be expressed as $\nu + [\gamma(\varphi_d) - \gamma(\varphi_x)]$, where ν is the partial trade elasticity estimated in a gravity equation that holds the domestic productivity cutoff constant. Importantly, Melitz and Redding (2015)'s analysis raises a number of empirical challenges for using equation (4) to measure local changes in welfare in response to trade shocks, which is why we use this approach only in section 5 but not in section 4. In particular, they emphasize that in models with strict sorting into production and export, the partial trade elasticity ν is potentially variable and depends on the shape (hazard function) of the productivity distribution at the export productivity cutoff ($\nu = \sigma - 1 + \gamma(\varphi_x)$). Moreover, even with an estimate of the partial trade elasticity, we need information on the shape of the productivity distribution at both the domestic and export productivity cutoffs (the hazard differential $[\gamma(\varphi_d) - \gamma(\varphi_x)]$) and on changes in the number of (unobserved) firms paying the fixed entry cost in the domestic country M_j^e to approximate local welfare changes. We also note that although the term $\Delta \ln M_j^e$ is quantitatively small in Melitz and Redding (2015), it significantly affects welfare calculations for our full model with heterogeneous fixed costs in Section 5. Additionally, note that (4) only allows the quantification of small trade shocks and therefore local welfare changes. To evaluate global welfare changes implied by large trade shocks, the implementation of (4) would require information the curvature of $\delta(\varphi_d)$ for the equilibria being compared.

These empirical issues make an implementation of Melitz and Redding's approach challenging, which is why empirical researchers often instead use the following approach based on Arkolakis et al. (2012). The ACR welfare statistic can be derived as special case of Melitz and Redding (2015). As they observe, under a Pareto productivity distribution, $\gamma(\varphi_d) = \gamma(\varphi_x) = \theta - (\sigma - 1)$ where θ is the Pareto shape parameter. Pareto productivity also ensures that $d \ln M_j^e = 0$. We can then write the global gains from trade as:

$$\Delta \ln W_j - \Delta \ln \tilde{\varphi}_{jj}^c = -\frac{1}{\theta} \Delta \ln \lambda_{jj} \tag{5}$$

where the Pareto shape parameter θ is also equal to the constant global trade elasticity.[20]

Finally, consider the special case where the set of domestic firms is fixed (e.g. the Armington model). The set of domestic firms is fixed ($X_{jj}^c = X_{jj}$) such that we can substitute the share of continuing domestic firms in total domestic sales (λ_{jj}^c) with the trade share (λ_{jj}). The resulting formula for welfare gains then depends only on the intensive margin elasticity σ and the change in the trade share.[21]

[20] Note that Arkolakis et al (2012) also make the point that their formula applies locally for arbitrary productivity distributions given symmetry and a local estimate of the trade elasticity.

[21] Note that this holds irrespective of whether the foreign country features entry or exit. Balanced trade may require that the terms of trade adjust to reflect changes in exporting status or entry/exit into production for firms in foreign country i, which could change the elasticity of the trade share with respect to a change in variable trade costs, but this does not affect welfare gains for domestic country j conditional on the change in the trade share.

A2: Restricting the set of continuing domestic firms to capture domestic within-firm productivity effects

Note that our welfare formula (equation (1)) estimates changes in the real wage up to changes in the productivity of continuing domestic firms. Although these effects are often absent in theoretical models, including the set of models considered by ACR, there is some empirical evidence for within-firm productivity effects of trade liberalization operating through various channels. Most relevant for our setting, Melitz and Trefler (2012) review the literature for CUSFTA and estimate a 5.4% increase in Canadian manufacturing productivity due to within-firm productivity growth due to new exporters investing in productivity, existing exporters investing in productivity, and improved access to US intermediate inputs (see their Table 2).

While our analysis is focused on the welfare implications of firm selection effects, our formula can potentially capture welfare gains associated with increases in domestic within-firm productivity, provided one is willing to specify a restricted set of continuing domestic firms that do not have productivity changes. By re-calculating equation (1) using the domestic revenue share of domestic continuing firms with no productivity changes, any productivity growth by the other domestic continuing firms is captured by the formula. Intuitively, if we observe a larger fall in λ^c for the restricted set of continuing firms than the full set of continuing firms, it implies a relative increase in domestic revenues for the non-restricted continuing firms that captures the welfare effects of their increased productivity.

In our setting, a natural way to specify this restriction is to use only domestic continuing firms in sectors that had low initial tariffs in 1988 (and hence minimal tariff reductions due to CUSFTA). Table 7 reports results comparing our welfare estimate using the full set of continuing firms with those that restrict continuing firms to those in sectors with initial tariffs below 5% or 3% (recall that 8% is the nominal average tariff prior to CUSFTA). The results are consistent with increased within-firm productivity in sectors with larger tariff reductions. Using a 5% threshold, the increase in welfare is only slightly larger, implying only small gains from within-firm productivity growth by domestic continuing firms. With a 3% threshold, the additional increase in welfare (over our benchmark) is similar in magnitude to the productivity gains estimated in Melitz and Trefler (2012). We stress that these calculations are only meant to be illustrative of the potential for our formula to capture within-firm productivity effects on welfare, as it is challenging to specify which firms are unaffected by a trade liberalization ex-ante and our results based on relative domestic spending shares for different continuing domestic firms are subject to the same caveats about identification as those based on selection effects.

Table 7: Restricting the set of continuing domestic firms to capture domestic within-firm productivity effects

Continuing domestic firms in:	Share 1988 (λ^c)	Share 1996 ($\lambda^{c'}$)	Welfare gain 1988-1996 ($\Delta \ln W$)
All sectors (benchmark)	0.575	0.503	5%
Sectors with initial tariffs below 5%	0.345	0.300	5.2%
Sectors with initial tariffs below 3%	0.208	0.156	10.7%

Notes: Columns 1 and 2 report the continuing domestic firm share of total domestic spending. Column 3 reports our welfare statistic based on equation (1) (the log difference between column 2 and column 1 divided by $(1/(\sigma-1))$). Row 1 reports the results calculating the statistics when treating all domestic continuing firms as continuing, while rows 2 and 3 restrict the set of continuing domestic firms to only those in sectors with initial average tariffs (and hence CUSFTA related tariff reductions) below a specific threshold.

A3: Details on Quantitative Models and Computational Solution Algorithm

In this appendix we offer a detailed discussion of the quantitative trade models of section 5 that deviate from iso-elastic import demand. Before we begin, it is worth re-emphasizing that our full model exhibits both truncated Pareto firm productivities and overlap in selection. As a result, the model with truncated Pareto firm productivities but strict sorting is a special case of our full model. Since the model with truncated Pareto firm productivities and strict sorting has been extensively analyzed by Melitz and Redding (2015), we mostly focus on our full model, but return to the Melitz and Redding (2015) model in Appendix A4.

To fix ideas, let ω index different firms in the data, each of which produces a differentiated CES variety. We assume the firms use a production function $y(\omega) = \varphi(\omega) \cdot L(\omega)$, where $y(\omega)$ are physical units of the differentiated variety ω, $L(\omega)$ is the number of workers firm ω

uses and $\varphi(\omega)$ denotes the firm-specific productivity draw, which we assume is drawn from a Pareto distribution with shape parameter θ and a truncation parameter. Firm productivity will affect revenues in the domestic market $X_{jj}(\omega)$ as well as export revenues $X_{ji}(\omega)$, where j is the home market and i is the foreign market.

Our first extension of the standard model in Melitz (2003) is that the fixed operating cost firms must pay per period is a deterministic function of the firm's initial productivity draw. That is, for a firm with productivity $\varphi(\omega)$,

$$f_{jj}(\omega) = f_d + \beta \cdot \ln \varphi(\omega) \tag{6}$$

This assumption is important to account for the fact that large firms might exhibit increased exit in response to a trade shock, so that marginal firms are not exclusively small. Our formulation nests the usual assumption that $\beta = 0$, such that all firms face a common fixed operating cost f_d as in Melitz (2003). Note that these are still "fixed costs" in the sense that they are invariant to demand shocks, including those related to changes in trade costs. The intuition for this assumption is that high productivity production processes may require higher set-up fixed costs (see Sutton (1991) and Kugler and Verhoogen (2012)). The resulting domestic profit net of fixed costs can therefore be defined as

$$\pi_{jj}(\omega) = X_{jj}(\omega) - w_j L_{jj}(\omega) - w_j \cdot f_{jj}(\omega) \tag{7}$$

We model overlap in the size distribution of exporting and non-exporting firms by allowing firms to face different fixed costs of exporting to the foreign market. We follow Armenter and Koren (2015) in modeling these exporting fixed costs as *i.i.d.* draws, but unlike Armenter and Koren (2015), we build a full general equilibrium two-country model instead of a partial equilibrium model because we our statistic stresses the importance of measuring domestic entry and exit. Intuitively we think of the variation in fixed export costs across firms as capturing heterogeneity in exporting opportunities. We assume that these draws follow a Pareto distribution with scale parameter f_x and shape parameter α and we will calibrate the shape and the scale parameter to match the overlap in the size of exporters and non-exporters in each country. This leads to the following definition of exports profits:

$$\pi_{ji}(\omega) = X_{ji}(\omega) - w_j L_{ji}(\omega) - w_j \cdot f_{ji}(\omega) \tag{8}$$

Note that if the shape parameter of the Pareto distribution for export fixed costs α is sufficiently high, then export fixed costs are not very dispersed and we return to a world with strict sorting into exports based only on productivity (and hence firm size).

To model selection into exit and exporting, we extend the standard Melitz (2003) framework and directly relate the exit and exporting decisions and therefore connect overlap in exit with overlap in exports. To this end, we allow all firms to finance losses in the domestic markets (inclusive of fixed production costs) with profits in export markets. As a result, the selection equation for exiting can be written as

$$\pi_{jj}(\omega) + \pi_{ji}(\omega) < 0 \tag{9}$$

In other words, a firm will only exit if the sum of domestic and export profits net of fixed costs is negative. Note that in a standard Melitz model with strict sorting, it is unnecessary to allow for the possibility of financing domestic losses with export profits, since a firm will only make an export profit if its domestic profits exceed the fixed production costs. In contrast, in our model, export profits can be differently distributed than domestic profits, so that there is the possibility that a firm generates export profits while also generating domestic losses. Additionally, it should be noted that selection equation (9) directly relates overlap along the export margin with overlap along the domestic exit margin.

For the export selection decision, firms must make positive export profits after paying export fixed costs and must have positive profits after paying production and exporting fixed costs.

$$\begin{aligned} &\pi_{ji}(\omega) \geq 0 \\ &\pi_{jj}(\omega) + \pi_{ji}(\omega) \geq 0 \end{aligned} \tag{10}$$

To model entry and exit during the pre-period, we assume that firms face idiosyncratic, proportional productivity shocks drawn from a log-normal distribution with mean 1 and a variance we calibrate to match the pre-period annual exit rate of 9%. The assumption of proportional productivity shocks ensures that firm growth rates are independent of firm size, consistent with Gibrat's law. Firms exit if their profits, net of fixed operating and possibly export fixed costs, become negative as in equation (9). Note that even in the absence of our fixed cost assumptions, random proportional productivity shocks would lead to exit by some firms that are not the smallest in the initial equilibrium. However, the variance of shocks required to match the degree of overlap we observe in the data is extremely high and implies counter-factually high exit rates.

Finally, we follow Melitz and Redding (2015) and model deviations from Pareto productivity by using a truncated Pareto distribution, which adds one additional parameter to the usual two-parameter Pareto distribution. Note that for a given value of the Pareto shape parameter θ, a lower value for the truncation parameter compresses the firm size distribution. This has implications for the relative size of exporters versus non-exporters (used by Melitz and Redding (2015) to calibrate the parameter), the relative size of exiters and continuers, and the overall firm size distribution.[22]

Allowing for overlap in exit or export selection implies that exact analytical solutions become infeasible. To understand why, it is important to remember that firm size is by definition not sufficient to perfectly determine whether firms export or exit in any model with overlap. As a consequence, selection cutoffs will be firm-specific and can imply that very productive and large firms might not decide to export due to high export fixed costs, while very unproductive and small firms might export due to low export fixed costs. Additionally, redefining heterogeneity in terms of a summarizing "net profit" term will not collapse firm heterogneity into a single dimension, as domestic and foreign net profits can be distributed differently and respond differently to shocks.

Because these issues make an exact analytical solution impossible, we use a computational solution based on profit-maximization for a discrete number of firms that have productivity

[22]Note that setting a low value of the truncation parameter is in some ways equivalent to picking a larger value of the θ parameter for our simulations, in the sense that both compress the distribution of firm size and generate larger trade elasticities and domestic net exit terms due to the greater importance of marginal firms compared to the (largest) continuing firms. An additional advantage of using a truncated Pareto distribution instead of a Pareto distribution in that our simulation results are more robust to extreme values.

and export fixed costs drawn at random. We use a quasi-random Sobol sequence of uniform numbers that, with our chosen parameters, approximate a truncated Pareto distribution for productivity and a Pareto distribution of export fixed costs. Our solution algorithm is as follows:

1. We first guess the number of firms that pay the fixed entry cost in each country and receive productivity and export fixed cost draws. We make an initial guess about which of these firms produce and export to the other market and solve for their equilibrium profits and the terms of trade.[23]

   - We determine firm selection, given the current aggregate equilibrium quantities as follows. Given vectors, which track the domestic and export selection decisions, we search for the "most profitable deviations" from the current selection patterns. For example, a firm is currently active in the domestic (or foreign) market might exhibit large profit losses which can be reduced by letting the firm exit the domestic (or foreign) market. Similarly, a firm that is currently not active in the domestic (or foreign) market, might generate much higher profits by participating in the domestic (or foreign) market and therefore enter. We check for the existence of these most profitable deviations from current selection patterns, which can be exit or entry into exporting, production, or both.
   - As additional constraints of the most profitable selection moves, we also impose the selection equations (9) and (10). As previously discussed, these conditions allow firms to finance domestic losses (net of production fixed costs) with export profits.
   - We continue to check for most profitable deviations until there are no profitable deviations from the equilibrium (within some tolerance to prevent cycling).

2. We then check whether the free entry condition is satisfied, i.e. the total profits in equilibrium are equal to the number of entrants multiplied by the fixed entry cost.

   - If not, we increase or decrease the number of entrants, and repeat step 1 to solve for equilibrium given the new set of entrants.

[23] We have experimented with initial guesses that all firms export and produce, or that only the firm with the highest productivity draw exports and produces, obtaining similar results.

- We repeat step 2 until expected profits from entry, ex-ante, converge to zero (within some tolerance).

While our solution algorithm is much faster than considering every potential combination of firm decisions, simulating a very large number of firms is still computationally costly, so in our simulation each "firm" represents approximately 100 plants in the data.

A4: Details on Calibration of Models and Targeted Moments

For our calibration, we take $\sigma = 3.7$ and $\theta = 2.9$ as discussed in section 4, and normalize the Pareto scale parameter to 1 for both countries. We experimented with an untruncated Pareto distribution for the firm productivity distributions. However, the combination of overlap in selection as well as untruncated Pareto, implies too many very large firms for Canada, which in turn leads to strong deviations from Zipf's Law for firm size. To address this issue, we therefore use the simplest deviation from the traditional Pareto distribution, which is a truncated Pareto distribution as in Melitz and Redding (2015). The combination of a high truncation cutoff for the Pareto distribution of firm productivities and our assumptions on the nature of fixed costs together therefore help us simultaneously match the empirically observed overlap in selection as well as Zipf's Law of firm size.

To capture the size differences between Canada and the US, we also take the ratio of manufacturing workers in the US to Canada (similar to the population ratio) as given and equal to 9.1. The standard Melitz model features four additional parameters for each country – the variable trade cost τ, the fixed entry cost f_e, the fixed operating cost f_d and the fixed exporting cost f_x. We will assume that initial variable trade costs are the same in each country.

In the end, our calibration features an additional four parameters for each country – the shape parameter for the iid export fixed cost draws α, the β parameter governing the dependence of fixed operating costs on firm productivity, the Pareto truncation parameter, and the variance of the proportional productivity shocks used to match exit rates during the pre-period.

We pick the model parameters to match several data moments that are standard for cali-

Table 8: Data and calibrated moments

Data moment		Canada Data	Canada Full model	Canada Melitz-Redding	US Data	US Full model	US Melitz-Redding
Firm count (hundreds)		380	430	414	1610	1689	1692
Exporter count (hundreds)		137	154	153	290	319	329
Fraction exporters		36%	36%	37%	18%	19%	19%
Dom. spending share λ_{jj}		79%	79%	79%	98%	99%	98%
Non-exporter size[1] (percentile)	20^{th}	0.37	0.74	0.85	0.33	0.88	0.79
	50^{th}	1.00	1.00	1.00	1.00	1.00	1.00
	80^{th}	2.85	3.40	1.26	3.78	7.46	1.45
Exporter size[1] (percentile)	20^{th}	0.84	1.19	2.51	2.89	6.43	2.76
	50^{th}	2.25	2.22	3.45	9.23	10.77	3.75
	80^{th}	6.41	5.07	6.20	30.45	27.34	6.05
Exiter size[2] (percentile)	20^{th}	0.02	0.81	0.60	0.03	0.20	0.60
	50^{th}	0.31	0.91	0.63	0.37	0.21	0.62
	80^{th}	2.40	2.13	0.68	3.05	1.04	0.67
Continuer size[2] (percentile)	20^{th}	0.16	0.65	0.71	0.20	0.18	0.71
	50^{th}	1.00	1.00	1.00	1.00	1.00	1.00
	80^{th}	6.80	3.55	2.62	5.08	2.26	2.05
Exit rate		0.09	0.09	0.09	0.09	0.09	0.09
Zipf coefficient[3]		0.56-1.06	1.05	1.29	0.4-1.06	0.67	1.55

[1] All exporter vs non-exporter moments are relative to median non-exporter.
[2] All exiter vs continuer moments are relative to median continuer.
[3] See Dixon and Rollin (2012) Table 8 and Kondo et al. (2019) Table 11.
Sources: Calculated from micro-data of Canada's Annual Survey of Manufacturing or US Manufacturing Census. See text for details

bration of Melitz-style models. We exactly match the domestic spending shares λ_{jj} and the fraction of these firms that export and approximately match the number of manufacturing firms in each country (with each simulation firm representing a bit more than 100 establishments). The other standard moment for a model with Pareto productivity is the firm size distribution, which implies a Zipf's law coefficient of $\xi = \frac{\theta}{\sigma-1}$. Note that selection into exporting, even under strict sorting, already leads to some deviation between the Zipf's law coefficient estimated from firm employment and the one implied by a model without exporting. Because of our focus on overlap, we go beyond the single size moment implied by Zipf's law and use multiple novel moments of the firm size distribution to discipline our model. For exporter/non-exporter overlap, we target the 20th, 50th, and 80th percentile plant-level employment for exporters and non-exporters in 1988 for Canada and 1987 for the United States. For continuer/exiter overlap, we target the initial 20th, 50th, and 80th percentile plant-level employment of firms that continue and those that exit over the 1988 to 1996 period (1987 to 1997 period for the United States).[24] We also target an exit rate of 9% based on annual exit rates in the pre-period. Our firm size measure is total employment, and we include fixed operating costs and fixed exporting costs (but not fixed entry cost) in the firm employment measures generated by the model. This implies that exporters employ additional labor compared to non-exporters and more productive firms have even higher employment than is predicted by productivity differences alone.

Although we relax the assumptions about strict sorting and allow for deviations from Pareto productivity, we maintain the same assumptions we made in the derivation of equation (1) in section 2. These also help us close the model and solve for the equilibrium. Assuming balanced trade allows us to solve for terms-of-trade/relative wage effects in general equilibrium, and is a reasonable assumption over this period.[25] Assuming free entry, such that ex-ante expected profits (or total profits net of all fixed costs and the fixed entry costs) are zero, helps solve for the number of firms that pay the fixed entry cost.

A5: Parameters for the US and additional calibration details

In this section we report the parameters for the US in our two country models and add some details on which moments drive the values of the calibrated parameters for Canada and the US.

As mentioned in section 5, the Pareto truncation parameter influences the size of the largest firms, which is especially important for the US data, as the largest US firms are huge.

[24] Note that for Canada we also have data from the 1978-1988 pre-period, which gives very similar size moments.

[25] Canada-US manufacturing trade was close to balanced over 1988-1996, with Canada running a small bilateral trade surplus that changed only slightly as a share of total trade.

Table 9: Parameters and Forcing Variables for US

	Melitz-Redding *Truncated Pareto only*	**Full model** *Truncated Pareto + fixed cost heterogeneity*
Productivity Pareto truncation	4.3	-
Trade friction τ	2.15	2.15
Entry fixed cost	1.3	1.3
Production fixed cost intercept f_d	1.8	0.25
Production fixed cost slope β	-	2.72
Export fixed cost location f_x	0.295	0.01
Export fixed cost shape α	-	0.19
Exit shock standard deviation	0.08	0.028

Notes: Column 1 shows the model parameters for the model with a truncated Pareto distribution and common fixed costs (the Melitz and Redding (2015) model). Column 2 shows the calibrated parameters for the model a truncated Pareto distribution and heterogeneity in fixed costs of production and exporting.

As a result, the Pareto truncation parameter is either higher than the corresponding parameter for CAN in the Melitz-Redding (truncated-Pareto with strict sorting) model or untruncated Pareto fits better for our full model with overlap. The Melitz-Redding model does a slightly better job than the full overlap model at matching the size of the largest exporters, but obviously fails to match the overlap in the data. The relatively small amount of truncation we use does not improve the model's ability to fit the relative size of the median exporter to non-exporter in the data, although it provides a better fit to the average size difference as in Melitz and Redding (2015). The Melitz-Redding model implies a Zipf's law coefficient that is significantly higher than any empirical estimate, reflecting the absence of very large firms. By making the largest firms significantly smaller than in the data, the truncated Pareto distribution does increase the size of domestic exiters relative to continuers, even under strict sorting. However the size of exiting domestic firms relative to continuers is still lower than in the data or our full model. Note that choosing a lower truncation point could increase the size of the domestic net exit term and the average size of exiting firms further. This would provide a better fit to the changes we observe during the CUSFTA period, but at the cost of even greater violation of Zipf's law.

Although modeling overlap in the size distribution of continuers and exiters amplifies differences in firm size for a given productivity distribution, it also increases the average revenue of domestic exiters relative to continuers and hence the domestic net exit term when there is a decline in variable trade costs. This occurs because of the heterogeneity in exporting fixed costs implied by overlap. Some smaller firms can offset the decline in domestic profits with profits from exporting, while some larger firms cannot. Combined with the operating fixed cost assumption that makes the most productive firms less profitable, this allows the model to increase the average domestic revenue, and hence importance for welfare, of domestic firm exit.

A6: Simulation of CUSFTA including reductions in fixed exporting costs

Table 10 presents results similar to those reported in our main Table 5 but also allowing for a reduction in export fixed costs that is proportionate to the reduction in import tariffs. Specifically this means the fixed cost of exporting to Canada falls by 23% and the fixed cost of exporting to the US falls by 13.3%, similar to the reduction in iceberg trade costs. The additional welfare gains from this change are small, particularly for the full model with fixed cost heterogeneity. This reflects the fact that fixed exporting cost heterogeneity is very large in our calibration and the density of firms near the threshold is low, such that average reductions in fixed exporting costs induce only small amounts of entry. In the model with strict sorting, the reduction in average fixed exporting costs leads to somewhat larger effects as marginal exporters are quite large due to the Truncated Pareto productivity distribution.

Table 10: Simulated effects of CUSFTA including reduction in fixed exporting costs

	Melitz-Redding *Truncated Pareto Only*	**Full model** *Truncated Pareto + Fixed Cost Het.*
Market Shares:		
$\Delta \ln X^c_{jj}/X_{jj}$ [1]	7.50%	6.38%
$\Delta \ln \lambda$ [2]	-27.05%	-19.52%
$\Delta \ln \lambda^c$ [3]	-19.55%	-13.15%
Total Welfare Change	7.24%	4.87%

[1] Percentage change in domestic revenues of *continuing* Canadian firms/all Canadian firms.

[2] Percentage change in total domestic sales of *all* Canadian firms/total sales in Canadian market.

[3] Percentage change in total domestic revenues of *continuing* Canadian firms/total sales in Canadian market.

Note: Simulated effect of a 23% reduction in Canadian tariffs and fixed cost of exporting to Canada, combined with a 13% reduction in US tariffs and fixed cost of exporting to US. Table reports models with a truncated Pareto distribution of productivity and common fixed costs (column 1) and heterogeneous fixed costs (column 2). See text for details.