Universität St.Gallen

Gravity with History: On Incumbency Effects in International Trade

Peter Egger, Reto Foellmi, Ulrich Schetter, David Torun

October 2024 Discussion Paper no. 2024-04

School of Economics and Political Science, Department of Economics

University of St.Gallen

Editor:	Mirela Keuschnigg University of St.Gallen School of Economics and Political Science Department of Economics Müller-Friedberg-Strasse 6/8 CH-9000 St.Gallen Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St.Gallen Müller-Friedberg-Strasse 6/8 CH-9000 St.Gallen
Electronic Publication:	http://www.seps.unisg.ch

Gravity with History:

On Incumbency Effects in International Trade[1]

Peter Egger, Reto Foellmi, Ulrich Schetter, and David Torun

Author's address: Prof. Dr. Reto Foellmi
SIAW-HSG
Rosenbergstrasse 22
9000/St. Gallen
Phone +41 71 224 2269
Email reto.foellmi@unisg.ch
Website https://sites.google.com/site/foellmireto/home

[1] We thank two anonymous referees, the editor, Giuseppe Berlingieri, Emily Blanchard, Ricardo Hausmann, Stefan Legge, Piotr Lukaszuk, Kalina Manova, Marc Melitz, Marc Muendler, Ralph Ossa, Frank Pisch, and seminar participants at ETSG (2019), SSES (2022), VfS (2022), the Villars Workshop on International Trade (2024), CERGE-EI, U St.Gallen, U Pavia, U Zurich, and the Harvard Growth Lab for helpful comments and discussions. The "Online Appendix" can be found at the bottom of this document. Ulrich Schetter gratefully acknowledges funding from the basic research fund of the University of St.Gallen under grant 1031513.

Emails: pegger@ethz.ch; reto.foellmi@unisg.ch; ulrich.schetter@unipv.it; david.torun@unisg.ch

Abstract

Countries trade more if they liberalized their trade relationship earlier. We derive a gravity equation featuring this path dependence due to sunk market-access costs that generate incumbency effects. We provide supporting evidence for the underlying mechanism and derive an augmented ACR formula (Arkolakis et al., 2012) for the gains from trade that accounts for incumbency effects. A quantification suggests our mechanism explains up to 25% of countries' home shares, and the gains from trade are, on average, 10% larger when allowing for incumbency effects. The analysis further reveals novel distributional effects of trade, boosting real wages but reducing profits.

Keywords

incumbency effects · sunk cost of market access · gravity equation · gains from trade · home bias · path dependence

JEL Classification

F12 · F14 · F15 · F17

1 Introduction

Countries open up to trade gradually over time. When firms start serving foreign markets they, therefore, need to compete against domestic and potentially third-country incumbents. These incumbents have already built up a distribution network, got to know market structures, and established their brand. An extensive literature on industrial organization shows how incumbency effects can deter entry (see, e.g., Noh and Moschini, 2006; Arping and Diaw, 2008; Goolsbee and Syverson, 2008; Schivardi and Viviano, 2011), and it is well-known that sunk market access costs can generate hysteresis effects in international trade (e.g., Baldwin, 1988; Baldwin and Krugman, 1989; Dixit, 1989; Alessandria et al., 2021). It is less well understood how incumbency effects matter for today's trade volumes and the gains from trade.

In this paper, we seek to contribute to closing this gap by introducing incumbency effects into a tractable dynamic trade model with heterogeneous firms. Our model nests the canonical version of a Melitz (2003)-Chaney (2008) model as a special case. The key novelty is that a firm's market access costs are (weakly) decreasing with its tenure in the market. When confronted with market entry by foreign competitors, incumbent firms therefore weigh discounted future profits against lower market access costs. This leads to less exit by (low-productive) incumbents and, hence, less entry vis-à-vis the canonical Melitz-Chaney model.[1] We derive a gravity equation that explicitly accounts for these incumbency effects. It implies that, ceteris paribus, countries trade less if they liberalized their trade relationship later. Accordingly, trade flows are shaped not only by trade frictions today—as (implicitly) assumed in standard gravity equations—but also depend on the history of trade liberalizations. In other words, there is strong path dependence in international trade.[2] We provide evidence in support of the main mechanism exploiting the structure of our model, and then discuss implications. We show that incumbency effects can explain up to 25% of countries' home shares and that they give rise to an augmented ACR formula (Arkolakis et al., 2012) for the welfare gains from trade. These gains are, on average, 10% larger when accounting for incumbency effects. Incumbency effects further generate novel distributional consequences of trade liberalizations through increased competition, which yields higher real wages but lower profits.

Our work is motivated by two strands of empirical evidence. First, trade is impacted

[1]This is in line with, e.g., Schivardi and Viviano (2011), who show that entry barriers are empirically associated with a lower productivity of incumbent firms.

[2]The term "path dependence" is often used to refer to situations where random shocks can have long-lasting effects. A prominent example is agglomeration effects, where small random differences in the initial allocation can have long-lasting and large consequences. In this paper, we use the term "path dependence" to refer to a situation where outcomes are not only shaped by where we are today (in terms of trade and market-access costs), but also by which route we took to get to this point.

by shadows of history. This is well known for colonial ties, which are associated with substantially more trade even decades after independence (Head et al., 2010). In Section 2.1, we further document that countries that were separated by the Iron Curtain ceteris paribus trade substantially less even today. We argue that incumbency effects can help explain these long-lasting effects of historical trade barriers. Second, our key assumption of decreasing fixed market access costs and, hence, sunk upfront investments is in line with persistence in firms' export status (Roberts and Tybout, 1997; Eaton et al., 2007; Ruhl and Willis, 2017; Alessandria et al., 2021). It is also consistent with Krishna et al. (2021), who show that the fixed (documentation) costs of using preferential trade agreements decrease with firms' experience exploiting the agreement. To further motivate our analysis, we consider the market exit of Colombian firms in response to a large negative trade shock with Venezuela in Section 2.2. We show that firms are less likely to stop exporting a product in response to the shock the longer their tenure in Venezuela, controlling for firm and product fixed effects and flexibly for a firm's pre-shock sales at the product-destination level. This suggests that fixed market access costs are indeed decreasing with tenure.

We start our theoretical analysis of tenure-dependent market access costs by considering a simple example with two countries and myopic firms in Section 3. We use this example to show that incumbency effects can give rise to selection into exporting and a home bias of international trade flows even with free trade between perfectly symmetric countries. In the remainder of the paper, we generalize this basic insight to a set-up with forward-looking firms, many asymmetric countries, arbitrary trade frictions, and an arbitrary sequence of trade liberalizations. In Section 4, we present our model. Analogous to the static version of the Melitz-Chaney model (see, e.g., Melitz and Redding, 2014a) there is a fixed cost of entry. Upon entry, firms receive a random productivity draw from a Pareto distribution. Serving a market involves a fixed cost of market access and a variable iceberg trade cost. The key difference to the canonical set-up is that our model is dynamic, with fixed market access costs that decrease with tenure in a market.[3] Firms are forward-looking: in their (market) entry decision, they weigh the present value of future profits against that of fixed costs. They do not anticipate any changes to the exogenous trade costs, but have otherwise perfect foresight.

These dynamics notwithstanding, our model is tractable enough to analytically characterize the equilibrium. We show this in Section 5, where we derive a gravity equation for international trade. First, we solve the equilibrium for the case where all trade liberalizations happen simultaneously in the very first period. In this scenario, tenure effects are

[3] The original Melitz (2003)-model is dynamic. Nevertheless, the majority of the literature considers a static (steady-state) version of this model. We refer to this version with a Pareto distribution of firm productivities as the canonical Melitz-Chaney model.

trivially absent, and the gravity equation reduces to the one in the Melitz-Chaney model. Second, we allow for an arbitrary sequence of trade liberalizations. In this case, there is a novel, multiplicatively separable incumbency term in the numerator of the gravity equation, which captures the impact of history on current trade flows. Our theory predicts that, ceteris paribus, countries trade more if they liberalized their trade relationship earlier.

The gravity equation can be estimated following standard steps from the literature. It implies, however, that in addition to proxies for fixed and variable trade costs today, we need to control for the history of trade liberalizations. Our theory can thus help explain the sizable and long-lasting implications of historical trade barriers. In fact, in Section 6.1, we show that our theory maps exactly onto our empirical specification of Section 2.1. Incumbency effects can thus help understand why historical colonial ties and the Iron Curtain impact trade even today.

There may, however, be other channels through which the Iron Curtain and colonial ties impact trade today. Section 6.2 therefore turns to the underlying mechanism. Specifically, if i liberalizes its trade relationship with j, the profit potential of firms from i in j increases, which facilitates survival in the market with or without incumbency benefits. Consequently, in response to a positive trade shock, incumbency effects become less important for firms from i serving j. We derive a regression equation that maps such changes in the importance of tenure to average firm sales (i.e., total exports divided by the number of firms serving a market). Considering relatively large tariff cuts, we find evidence in line with the model's predictions. Interestingly, this implies that incumbency effects partly offset the positive trade effect of a decrease in variable trade costs.

In Section 7, we discuss implications. We begin in Section 7.1 by showing that our theory gives rise to an augmented ACR formula (Arkolakis et al., 2012). Analogous to the standard case, the welfare gains from trade depend only on primal parameters of the model and the share of expenditures on domestic goods (henceforth, the home share). Different from the standard case, however, incumbency effects amplify the costs of moving to autarky, and more so the smaller a country's home share. Intuitively, shutting down trade destroys sunk upfront investments in accessing foreign markets, and this is the costlier, the more important foreign markets are for a country's total sales (i.e., the smaller its home share). In Section 7.2, we explore the quantitative importance of our mechanism. We outline how our model can be solved in relative changes using hat algebra (Dekle et al., 2007) and readily accessible data, while accounting for the endogenous importance of incumbency effects. The quantification reveals that the model explains up to 25% of a country's home share, and implies gains from trade that are—on average—10% higher than predicted by the ACR formula. This difference between our model and the ACR formula is hump-shaped with respect to the home share: Intuitively, for very low home shares, a country's sales are dominated by exports, while for very high home shares, there is little competitive pressure from foreign suppliers. In either case, incumbency effects in the domestic market have limited impact on a country's global sales and, hence, its gains from trade. We finally show that incumbency effects give rise to novel distributional consequences of trade liberalizations. They shield firms from exiting and, hence, increase competition, which benefits real wages but lowers profits.

Relation to the literature. We develop a dynamic variant of the canonical "gravity" version of the Melitz model. Our paper is thus closest related to the extensive gravity literature in international trade. The majority of the literature in this area is static.[4] A series of recent papers introduce dynamics into gravity models of trade and migration (see, e.g., Allen and Donaldson, 2020; Buera and Oberfield, 2020; Atkin et al., 2021; Caliendo et al., 2021). In these papers, history can impact outcomes today through technological diffusion, agglomeration, or the evolution of country capability. These mechanisms have in common that they impact an exporter (or country/region) in the same way across importers. That is, the impact of history is not specific to an exporter-importer pair. As opposed to that, we present a tractable dynamic framework and derive a gravity equation where trade flows are not only shaped by trade frictions today, but also by the pair-specific history of trade liberalizations. In this regard, our paper is closer to Evenett and Venables (2002); Albornoz et al. (2012); Morales et al. (2019), who show that the probability of a firm entering a new market hinges on how similar this new destination is to destinations that the firm already served in past periods. By contrast, in our model, countries that liberalized their trade relationship earlier trade more today.

Our analysis is centered on fixed costs of market access which are decreasing with tenure. It is well-known, at least since Baldwin (1988); Baldwin and Krugman (1989); Dixit (1989), that sunk entry costs can give rise to hysteresis in international trade. Motivated by empirical evidence on firm entry and exit in export markets—see also Section 2— a growing literature introduces sunk costs of market access into theoretical models to analyze exporter dynamics (see, e.g., Costantini and Melitz, 2007; Das et al., 2007; Fitzgerald et al., 2016; Ruhl and Willis, 2017; Eaton et al., 2021, and Alessandria et al.,

[4]See, e.g. the seminal contributions by Eaton and Kortum (2002), Anderson and van Wincoop (2003), Melitz (2003)-Chaney (2008), Arkolakis et al. (2012), and the review in Head and Mayer (2014).

2021 for a survey article).[5] This literature largely focuses on partial equilibrium models and firm-level transitional dynamics. Notable exceptions are, e.g., Burstein and Melitz (2013); Impullitti et al. (2013); Alessandria and Choi (2014); Alessandria et al. (2014), who consider general equilibrium environments, but can only allow for two symmetric countries. Compared to this literature, our dynamics are simpler and, after a shock, our economy immediately jumps to a new steady state. In turn, this allows introducing incumbency effects into a general equilibrium trade model with numerous asymmetric countries, arbitrary bilateral fixed and variable trade costs, and an arbitrary sequence of trade liberalizations while maintaining tractability.

In our gravity framework, incumbency effects yield three important implications compared to a standard structural gravity model. First, incumbency effects provide an additional channel through which iceberg trade-cost shocks impact trade. Consequently, the trade elasticity varies between positive and negative trade shocks and, in general, is not constant. This finding complements a growing body of work that emphasizes heterogeneous effects of trade liberalizations and proposes alternative mechanisms to generate variable trade elasticities such as non-CES demand (Novy, 2013; Carrère et al., 2020; Chen and Novy, 2022), or endogenous firm selection (Melitz and Redding, 2015; Feenstra, 2018; Adão et al., 2020).[6]

Second, in our gravity equation, countries trade ceteris paribus more if they liberalized their trade relationship earlier. With countries opening up to trade gradually over time, our framework provides a novel mechanism to help explain a home bias in international trade, even with free trade and between perfectly symmetric countries. Several explanations for the large empirical home shares have been proposed, including a discontinuity of trade costs at the national border (McCallum, 1995; Anderson and van Wincoop, 2003; Chen, 2004), multi-stage production (Yi, 2010), or non-homothetic preferences (Caron et al., 2014). We add to this literature by showing how tenure effects lead to higher home shares and show quantitatively that our mechanism can explain up to 25% of a country's home share in our preferred specification.

Third, Arkolakis et al. (2012) show that in benchmark structural gravity models, the trade elasticity and a country's home share are sufficient statistics for the welfare gains

[5] An alternative mechanism to create incumbency effects would be to introduce increasing sales over time, which have been found to be important for explaining some of the micro-patterns in the data (see, e.g., Fitzgerald et al., 2016; Ruhl and Willis, 2017). In our model, this would have implications for aggregate trade that are qualitatively similar to those of tenure-dependent fixed costs. We therefore focus on the latter, which are analytically more tractable and are also the predominant assumption in the literature. In Section 2, we provide evidence to corroborate our set-up.

[6] Arkolakis (2010) formulates market-entry costs as being heterogeneous between firms due to heterogeneous market-penetration costs. This fixed-cost heterogeneity does not fundamentally vary with time due to the tenure of firms in a market. However, it leads to a heterogeneous aggregate trade elasticity in a gravity equation when pooling data across consumer markets, as is customary (see Bas et al., 2017).

from trade. We show that this is not the case in our framework and derive an augmented ACR formula that accounts for incumbency effects. Our work thus relates to the literature analyzing various mechanisms through which the gains can differ from the baseline ACR formula. These include multiple industries and input-output linkages (Caliendo and Parro, 2015; Ossa, 2015), variable markups (Edmond et al., 2015; Arkolakis et al., 2019), firm selection (Melitz and Redding, 2015), trade-induced productivity growth (Melitz and Redding, 2014b; Buera and Oberfield, 2020), and non-homothetic preferences (Fajgelbaum and Khandelwal, 2016). We provide an additional channel through which trade impacts welfare—incumbency effects—and show quantitatively that they can lead to gains from trade that are, on average, 10% higher than suggested by the ACR formula.

2 Motivating Evidence

Before turning to the model, we motivate our set-up with two sets of stylized facts. First, regarding the long-lasting implications of historical events for trade. Second, regarding the exiting behavior of Colombian exporters in response to a large negative trade shock.

2.1 History and Trade Today

It is well known that historical events can have long-lasting consequences for trade. It is less well understood why this is the case. In what follows, we will argue that the "shadows of history" in trade can partially be attributed to incumbency effects. We, therefore, begin our analysis by revisiting two historical events and their bearings on trade today: colonial ties and the (fall of the) Iron Curtain. In Section 6.1, we show that our theory maps onto our empirical specification here.

Colonial ties are a common control variable in empirical gravity models. We add to prior work by also considering only sea and airborne trade, which is arguably less impacted by bilateral investments in infrastructure. The focus on the (fall of the) Iron Curtain is—to our knowledge—new. Specifically, we consider trade between pairs of countries that have been on different sides of the Iron Curtain during the Cold War and show that these pairs of countries trade less even today.[7] This provides a relevant set-up for our purposes for two reasons. First, the Iron Curtain was a major impediment to international trade, and its fall was associated with a dramatic shift in the trade environment—cf. Figure O5.2 of the Online Appendix. Second, the original trade barriers

[7]We classify countries as having been behind the Iron Curtain if they were members of the Council for Mutual Economic Assistance (Comecon). The members of this economic association were Albania, Bulgaria, Cuba, Czechoslovakia, East Germany, Hungary, Mongolia, Poland, Romania, the Soviet Union, and Vietnam (cf. Encyclopædia Britannica). We consider countries to have colonial ties if they had a colonial relationship or a common colonizer after 1945 according to the CEPII Gravity database.

and the subsequent liberalization between members and non-members of the Comecon were mainly a matter of geopolitics rather than one of strategic trade policy.

To analyze the effects of these historical events on trade, we focus on two regressors in a standard gravity regression: a dummy indicating whether exporter i and importer j had colonial ties, $\mathbb{1}[\text{Colonial Ties}]_{ij}$, and a dummy that indicates whether (i, j) had been separated by the Iron Curtain, $\mathbb{1}[\text{Sep. by Iron Curtain}]_{ij}$, where $\mathbb{1}[\cdot]$ is the indicator function that is unity if the term in brackets is true and zero otherwise. In summary, we run the following regression

$$\begin{aligned} \log(X_{ij}) &= E_i + M_j + \zeta_{ij}'\boldsymbol{\beta} + \phi_1 \times \mathbb{1}[\text{Sep. by Iron Curtain}]_{ij} \\ &\quad + \phi_2 \times \mathbb{1}[\text{Colonial Ties}]_{ij} + \upsilon_{ij}, \end{aligned} \tag{1}$$

where X_{ij} is aggregate trade from country i to j, E_i and M_j are exporter and importer fixed effects, respectively, υ_{ij} is an error term, and ζ_{ij} is a (column) vector of controls for bilateral trade frictions. Our main coefficients of interest are ϕ_1 and ϕ_2, which measure the semi-elasticity for the Iron Curtain and colonial ties indicator, respectively. We run this regression for 2015 using OLS, and cluster standard errors at both the importer and the exporter level.[8] A detailed description of our data is provided in Online Appendix O2.1. A robustness check using PPML is provided in the Online Appendix, Table O6.3.

Results. The results are summarized in Table 1. The full set of covariates and coefficients is reported in the Online Appendix, Table O6.1. The leftmost column shows the coefficient on our two indicators of interest, controlling only for distance, contiguity, and the fixed effects. The second column reports on the coefficients when including our full set of controls. The main insight from these results is twofold. First, country pairs that were separated by the Iron Curtain trade substantially less even 25 years after the fall of the Iron Curtain, while country pairs connected through colonial ties trade substantially more. The point estimate for the Iron Curtain indicator in column (2), for example, implies that countries that were separated by the Iron Curtain trade about 47 percent ($\approx 1 - \exp(-.63)$) less today, controlling for standard trade frictions. Second, most of the suggested discrepancies are not picked up by standard gravity controls, suggesting that these historical events have effects on trade that are orthogonal to conventional proxies for bilateral frictions.

Colonial ties and the Iron Curtain plausibly impact pair-specific trade flows today through various channels, including trade policy and transportation costs. In columns (3) to (6), we seek to proxy for these, as we explain in detail in Online Appendix O2.2, by

[8] We exclude former Yugoslavian countries, since these countries were (economically) close to both Iron Curtain and non-Iron Curtain countries. The overall share of relations where the separation indicator equals one is around 23%.

Table 1: Gravity Regressions, Year 2015, OLS

	Main Specification		Robustness			
	(1)	(2)	(3)	(4)	(5)	(6)
Sep. by Iron Curtain	-0.572***	-0.628***	-0.572***	-0.553***	-0.930***	-0.418***
	(0.096)	(0.099)	(0.090)	(0.087)	(0.142)	(0.124)
Colonial Ties	0.781***	0.534***	0.507***	0.497***	0.949***	0.935***
	(0.095)	(0.096)	(0.096)	(0.098)	(0.232)	(0.240)
Baseline Controls		✓	✓	✓	✓	✓
Tariffs & NTMs			✓	✓		
CIF/FOB Ratio				✓		
Comext Data					✓	✓
Mean of Dep. Var.	2.469	2.682	2.742	2.918	2.347	2.024
Adj. R-squared	0.737	0.765	0.782	0.796	0.802	0.799
No. of Observations	19,847	17,153	16,195	14,511	6,954	6,849

Notes. This table reports OLS results for regression (1). Standard errors clustered at both the importer and exporter levels are reported in parentheses. All regressions include importer and exporter fixed effects. "Sep. by Iron Curtain" refers to a dummy variable equal to one if the country pair was historically separated by the Iron Curtain and zero otherwise (see footnote 7). "Colonial Ties" refers to a dummy variable equal to one if the country pair either had a colonial relationship or a common colonizer after 1945. The dependent variable is the log of aggregate trade flows. Additional controls have been added as indicated in the respective column, but are not shown for readability. A long table with coefficients on all controls is provided in the Online Appendix, Table O6.1. Column (6) includes only trade via air or sea transport. * significant at 10%-level;** significant at 5%-level;*** significant at 1%-level.
Data source. Atlas of Economic Complexity (trade data in (1)–(4)), CEPII/Head et al. (2010) (covariates), WITS (import tariffs), GTA (NTMs), ITIC (CIF/FOB), Comext (trade data in (5)–(6)).
Results. Authors' computations.

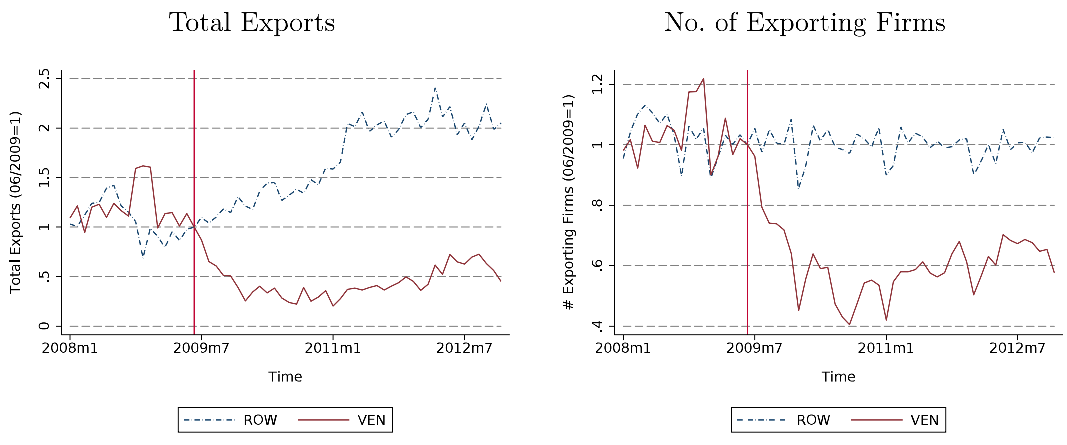
controlling for tariffs and non-tariff barriers (3), CIF/FOB ratios (4), and restricting the mode of transport to air and sea transport only (6). The latter specification relies on the Comext trade data, which has a smaller country coverage. Column (5) therefore repeats our baseline specification from column (2), but considering this smaller set of countries only. Throughout these robustness checks, the results are qualitatively the same.

In summary, these regressions point to sizable, long-lasting, and pair-specific implications of history for trade that are not captured by conventional proxies for trade frictions nor entirely attributable to differences in pair-specific transportation infrastructure. These implications call for a better understanding. We will explore one channel that can help explain such patterns: incumbency effects that may prevent firm entry and exit upon trade liberalizations. To motivate this further, we consider the exit behavior of Colombian exporters in response to a large shock next.

2.2 Tenure and Exit Propensity of Colombian Exporters

In this section, we document that firms are less likely to exit a market the longer their tenure in that market, even when controlling flexibly for their sales. To do so, we use transaction-level data in USD at the 10-digit HS level from ADUANAS for the years

Figure 1: The Venezuela Shock on Colombian Firm Exports



Notes: These figures illustrate monthly total exports from Colombia to Venezuela vs. the rest of the world (left panel), and the number of firms exporting to Venezuela vs. the rest of the world (right panel). Both graphs depict indexed values such that June 2009 is equal to one. See Section 2 for further details.
Data source: ADUANAS (DIAN)
Graph: Authors' representation.

1994 to 2013 provided by the Colombian tax authority (DIAN). We aggregate this data to yearly firm-level exports at the 2-digit HS level and then consider firm exit at the product-destination level. This or similar datasets have previously been used to document exporter dynamics (see, e.g., Roberts and Tybout, 1997; Eaton et al., 2007; Ruhl and Willis, 2017; Alessandria et al., 2021). We add to prior findings (i) by flexibly controlling for log-sales at the product-destination level to emphasize a potential role of tenure-dependent fixed costs, and (ii) by focusing on exit in response to a large negative shock to Colombia's trade with Venezuela rather than on general exporter dynamics.

On July 28, 2009, in response to a dispute over American anti-drug campaigns in Colombia and the supply of weapons to Colombian rebels, the then-president of Venezuela, Hugo Chavez, declared a freeze on diplomatic ties with Colombia and expressed the wish to decrease bilateral trade.[9] While this shock did not entirely close the border, it had a major impact on bilateral trade, as illustrated in Figure 1. From July 2009 to December 2009, Colombian exports to Venezuela fell by over 50%, whereas exports to the rest of the world increased. The number of firms that exported to Venezuela fell by over 40%. By contrast, the number of firms that served the rest of the world remained fairly constant. These are sizable shocks, especially considering the pre-crisis importance of Venezuela as a trading partner. In 2008, the last full year before the crisis, more than 15% of Colombian exports went to Venezuela, and about one-third of all exporters served that market.

[9] See https://www.economist.com/the-americas/2009/09/10/politics-versus-trade, accessed on June 30, 2023.

To analyze the exit response of Colombian firms to this (exogenous) shock, we restrict our sample to all firm-product pairs with positive exports to Venezuela from July 2008 to June 2009. In our baseline specification, we then treat a firm-product pair as an exit if the firm no longer exports that product to Venezuela in the period from January 2010 to December 2010. This allows the shock to manifest gradually over time, as suggested by Figure 1. We run the following linear probability model

$$y_{fp} = \boldsymbol{\lambda}'_{fp}\boldsymbol{\alpha} + \boldsymbol{x}'_{fp}\boldsymbol{\beta} + \boldsymbol{d}'_{fp}\boldsymbol{\gamma} + \epsilon_{fp} \ , \tag{2}$$

where y_{fp} is our exit indicator that takes on value one if firm f stops exporting product p in response to the shock, and zero otherwise. $\boldsymbol{\lambda}$ is a set of tenure dummies $\lambda^1, ..., \lambda^{10}$, where λ^k_{fp} equals one if at the time of the shock firm f exports product p to Venezuela for the k^{th} consecutive period.[10] $\boldsymbol{x}$ is a (set of) size controls and $\boldsymbol{d}$ a set of dummies as specified momentarily. The error term is denoted by ϵ_{fp}, and we apply two-way clustering at the product and firm levels.

In our main—most conservative—specification, we group firms into quintiles using their pre-shock exports of a product to Venezuela. We then control for (i) a full set of product-quintile dummies, and (ii) the firm's pre-shock log-sales of that product with a product-quintile-specific coefficient. This allows to flexibly control for a firm's pre-shock sales of a given product. We add firm dummies to control for general firm-level trends. In Online Appendix O2.2, we show that less restrictive specifications with fewer fixed effects yield the same basic insights.

Figure 2 summarizes the coefficients on the tenure dummies, which measure tenure effects relative to the omitted category of firms with tenure 1 (i.e., of firms who had started exporting to Venezuela in the year before the shock). This figure clearly reveals sizable effects of tenure on exit: the exit probability, ceteris paribus, decreases by more than 40 percentage points when a firm is highly experienced (tenure ≥ 8).

In what follows, we analyze what these tenure effects imply for aggregate trade and how trade is shaped by the history of trade liberalizations.

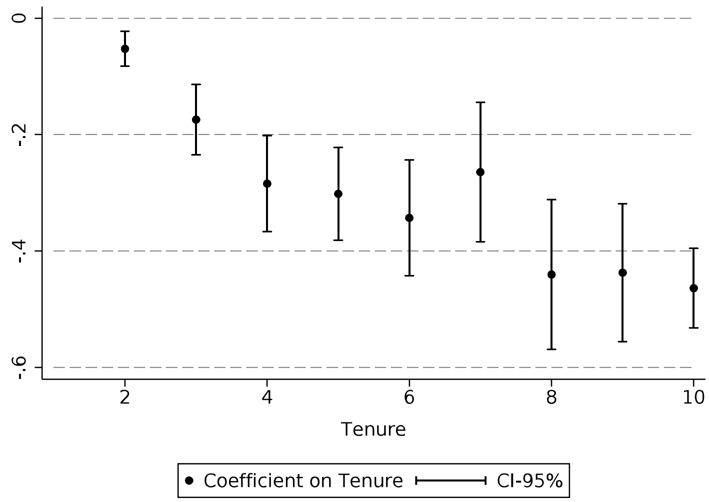
3 A Simple Example

To illustrate our main point of interest, it is instructive to start off with a stylized example, before introducing our model in the next section.

Consider a Melitz model with two perfectly symmetric countries and free trade. Hence, for this example, let us assume that upon trade opening there are no variable trade costs,

[10] To measure tenure in July 2008–June 2009, we aggregate trade flows in preceding years over twelve-month periods from July in a year $t-1$ to June in year t. The basic pattern is the same when aggregating data from August in a year $t-1$ to July in year t and then looking at exit immediately after the shock (i.e., no exports in the period from August 2009 to July 2010). Details are available upon request.

Figure 2: Firm Exit in Response to Venezuela Shock



Notes: This figure summarizes the tenure coefficients from estimating exit probabilities following Equation (2) using OLS. Standard errors are clustered at both the product and the firm level. All tenure levels of 10 years and above have been summarized in one category "tenure 10." Omitted category is tenure=1. Further details are provided in the main text.
Data source: ADUANAS (DIAN)
Graph: Authors' representation.

and market access costs f are the same at home and abroad. In the textbook case of the Melitz model, we then trivially have that the cutoff productivity is the same for every destination market and that

$$\frac{X_{ij}}{X_j} = \frac{1}{2},$$

where X_{ij} denotes trade flows from country i to destination country j, and X_j denotes total expenditure in country j.

Now, suppose that both countries start out under autarky at time 0, and open up to free trade at some time $t > 0$. Economically, this implies that firms must compete against incumbent local firms when they start to export. How is that going to affect international trade? Not at all if fixed costs of market access are constant over time. However, this is no longer true if these costs decrease with tenure in a market. In particular, to keep things simple for now, suppose that firms already serving a market for t periods have to pay a fraction $\alpha < 1$ of fixed costs only, and that they are myopic in their entry decision. Then, in both countries, domestic firms continue serving their home market whenever

$$\pi^v_{ii}(\varphi) \geq \alpha f,$$

where $\pi^v_{ij}(\varphi)$ denotes the variable profits that a firm in country i with productivity φ makes from serving country j. By contrast, firms from country i enter market j whenever

$$\pi^v_{ij}(\varphi) \geq f.$$

It follows that the sequential opening to trade in itself gives rise to selection into exporting and a home bias. In particular, it is straightforward to show that

$$\left(\frac{\varphi_d}{\varphi_x}\right)^{\sigma-1} = \alpha,$$

where φ_d is the cutoff productivity for serving the home market, φ_x the cutoff productivity for serving the foreign market, and $\sigma > 1$ the constant elasticity of substitution in consumption.

In what follows, we develop this argument more carefully and show how it impacts aggregate trade flows between a large set of heterogeneous countries with forward-looking firms, arbitrary variable trade costs, arbitrary fixed market access costs that are weakly decreasing with tenure, and an arbitrary sequence of trade liberalizations.

4 Model

We build on the canonical static Melitz model (as in, e.g., Melitz and Redding, 2014a, Section 6). Analogous to this benchmark, there is free entry, subject to a fixed cost of entry, and serving a market involves a fixed cost of market access. Forward-looking firms decide which markets to serve. The key novelty in our set-up is that fixed cost of market access are decreasing with tenure in that market. That is, serving a market involves higher fixed cost upfront (e.g., in order to establish a brand or a distribution network), and lower fixed cost in the future (e.g., to sustain the brand or the distribution network). In the case where these fixed costs are constant over time, our model reduces to the canonical baseline framework.

Households. There are I countries, indexed by $i, j \in \mathcal{I}$. Country i is populated by a continuum of measure L_i of infinitely lived households. In every period, each household is endowed with one unit of labor that it inelastically supplies to the labor market. Households receive utility from consuming a continuum of horizontally differentiated varieties with preferences given by

$$U_i^0 = \sum_{t=0}^{\infty} (\beta)^t C_i^t \tag{3}$$

$$C_i^t := \left[\int_{\omega \in \Omega_i^t} q_i^t(\omega)^{(\sigma-1)/\sigma} d\omega\right]^{\sigma/(\sigma-1)}, \quad \sigma > 1,$$

where an index t indicates period t, $\beta < 1$ is the discount factor, Ω_i^t denotes the set of varieties available in country i at time t, and $q_i^t(\omega)$ denotes the quantity of variety ω consumed by a representative household in country i at time t. Households can invest in an aggregate asset of domestic firms. The asset yields a rate of return which we denote by r_i^t. The representative household in country i then maximizes (3) subject to the flow budget constraint

$$a_i^{t+1} = (1 + r_i^{t+1})(a_i^t + w_i^t - x_i^t), \tag{4}$$

where a_i^t denotes (pre-determined) asset holdings at the beginning of the period, w_i^t the wage rate, and x_i^t total spending of the representative household in period t.

Following standard steps to solve the household's intra-temporal optimization problem yields its demand for variety ω in country i at time t

$$q_i^t(\omega) p_i^t(\omega) = p_i^t(\omega)^{1-\sigma} (P_i^t)^{\sigma-1} x_i^t,$$

where $p_i^t(\omega)$ denotes the price of variety ω and where

$$P_i^t := \left[\int_{\omega \in \Omega_i^t} p_i^t(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}}$$

is the ideal price index. Intertemporal optimization further requires that in equilibrium we must have

$$1 = \beta(1 + r_i^{t+1}) \frac{P_i^t}{P_i^{t+1}} \tag{5}$$

at all times, i.e., the interest rate adjusts such that households are exactly indifferent between consuming the CES aggregator either today or in the next period.

Firms and production. The competitive environment and production are analogous to the baseline Melitz model, with the only change in assumptions that fixed costs of serving a market vary with tenure in that market. Firms anticipate the evolution of fixed costs when deciding whether to enter a market.

There is free entry into the market for final consumption goods. Entry in country i entails a fixed cost. For simplicity, we assume that this fixed cost is paid in the form of a perpetual annuity of f_{ei} units of domestic labor. Upon entry, a firm is equipped with a new variety ω, and it receives a productivity draw φ from a distribution with CDF $G(\varphi)$, PDF $g(\varphi)$, and positive and unbounded support Φ. To derive a gravity equation, we will later consider the case of a Pareto distribution of productivities, i.e.,

$$G(\varphi) = 1 - \left(\frac{\varphi}{\underline{\varphi}} \right)^{-\theta}, \quad \text{for } \varphi \geq \underline{\varphi}$$
$$\text{and} \quad g(\varphi) = \theta \underline{\varphi}^{\theta} \varphi^{-\theta-1}$$

where $\underline{\varphi} > 0$ is the lowest productivity level, and the shape parameter θ is assumed to satisfy $\theta > \sigma - 1$.[11] A firm with productivity φ needs $1/\varphi$ units of domestic labor in order to produce one unit of its variety.

[11] We assume that $\underline{\varphi}$ is the same across countries. This is for expositional convenience only.

Trade is subject to an iceberg trade cost, that is, $\tau_{ij}^t \geq 1$ units of a variety have to be shipped from country i for one unit to arrive at destination country j. Variable trade costs satisfy the triangle inequality, and $\tau_{ii}^t = 1$ for all i and t. In addition, firms face fixed costs of serving a market: A firm based in country i needs to employ f_{ij} units of domestic labor in every period it is serving market j. These labor units need to be equipped with tenure-dependent units of the CES aggregator.[12] In Section 5, we will show that—in the absence of trade shocks—the CES price index is constant over time. In anticipation of this result, we simplify the exposition by assuming that the resources needed for market access simply scale the market access costs. In particular, suppose that a firm from country i has continuously served market j for $\lambda \geq 0$ periods, where we use λ to denote tenure in a market. Then its fixed costs of serving that market in period t are

$$f_{ij}^t(\lambda) := f_{ij} w_i^t f(\lambda), \tag{6}$$

where $f(\lambda)$ is a tenure-dependent scaling factor satisfying $f(\lambda) \geq 1$ and $f'(\cdot) \leq 0$. Out of these fixed costs, $f_{ij} w_i^t$ are spent on labor in country i, while $[f(\lambda) - 1]\, f_{ij} w_i^t$ are spent on the CES aggregator. Note that Equation (6) nests the common assumption of constant-over-time market access costs in terms of labor as a special case (with $f(\lambda) = 1\ \forall\, \lambda$).

In summary, a firm from country i with productivity φ that started serving households in country j λ periods ago and that wants to sell q units of its variety faces a total cost of

$$\mathbb{C}_{ij}^t(q; \varphi, \lambda) = f_{ij}^t(\lambda) + \frac{\tau_{ij}^t w_i^t}{\varphi} q.$$

Firm behavior. Firms from country i can borrow or lend at the prevailing market interest rate r_i^t as given by Equation (5). Firms are forward-looking. In their (market) entry decision, they weigh discounted future costs against benefits, perfectly anticipating the tenure-dependent evolution of market access cost and any potential endogenous evolution of equilibrium outcomes. They do not, however, anticipate potential changes in the exogenous trade environment. That is, firms expect all τ_{ij} and f_{ij} to remain constant over time, analogous to static gravity equations. We will later examine how the sequence of past shocks to τ_{ij} impact trade flows today. The focus on τ allows disentangling tenure effects from the direct effects of changes in the trade environment.

Given the forward-looking behavior of firms, it will come in handy to introduce the

[12] We introduce this hybrid market access cost for two reasons. First, economically, it is reasonable to assume that accessing a market requires both personnel and financial resources, e.g., for a marketing campaign, and that the marketing budget needed to maintain a brand is smaller than the budget needed to establish a brand. Second, this choice improves the tractability of our model.

following notation

$$\tilde{f}_{ei}^t := \left(f_{ei}\sum_{\varsigma=t}^{\infty} w_i^{\varsigma} R_i^{t,\varsigma}\right)\left(\sum_{\varsigma=t}^{\infty} R_i^{t,\varsigma}\right)^{-1}$$

$$\tilde{f}_{ij}^t(\lambda) := \left(f_{ij}\sum_{\varsigma=t}^{\infty} f(\lambda+(\varsigma-t)) w_i^{\varsigma} R_i^{t,\varsigma}\right)\left(\sum_{\varsigma=t}^{\infty} R_i^{t,\varsigma}\right)^{-1},$$

where

$$R_i^{t,\varsigma} := \prod_{s=t}^{\varsigma} \frac{1+r_i^t}{1+r_i^s}$$

is the discount factor from period $\varsigma \geq t$ to the beginning of period t. In words, $\tilde{f}_{ei}^t$ is an annuity with time t present value equivalent to that of the upcoming stream of fixed entry costs payable by a firm that entered at time t or before. Similarly, $\tilde{f}_{ij}^t(\lambda)$ is the annuity for the stream of market access costs payable by a firm that accessed market j for the first time in period $t-\lambda$, and continuously served that market ever since.

With CES preferences, firms charge the well-known constant mark-up over marginal cost. A firm from country i selling to market j thus offers its product at a price

$$p_{ij}^t(\varphi) = \frac{\tau_{ij}^t w_i^t}{\varphi}\frac{\sigma}{\sigma-1}, \tag{7}$$

and it earns variable profits equal to

$$\pi_{ij}^{v,t}(\varphi) = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}(\tau_{ij}^t w_i^t)^{1-\sigma}(P_j^t)^{\sigma-1}X_j^t\varphi^{\sigma-1}, \tag{8}$$

where X_j^t denotes aggregate expenditure on the CES aggregator in country j as further characterized in Lemma 1 below. Analogous to the above, let $\tilde{\pi}_{ij}^{v,t}(\varphi)$ denote an annuity that has the same time t present value as the infinite stream of variable profits associated with serving market j from period t onward.[13] Then, a firm from country i with productivity φ and tenure λ in market j finds it profitable to continue (or start in case of $\lambda = 0$) serving j whenever

$$\tilde{\pi}_{ij}^{v,t}(\varphi) \geq \tilde{f}_{ij}^t(\lambda). \tag{9}$$

For a given λ, this allows defining a cutoff productivity similar to a canonical Melitz model. In particular, in any period t, $\pi_{ij}^{v,t}(\varphi)$ is increasing in φ and, hence, so is $\tilde{\pi}_{ij}^{v,t}(\varphi)$. By contrast, for a given λ, the right-hand side of Condition (9) is independent of the firm. Accordingly, we can define a cutoff productivity $\varphi_{ij}^{\lambda,t}$ for firms in country i that started serving market j λ periods ago. This cutoff is implicitly defined by

$$\tilde{\pi}_{ij}^{v,t}(\varphi_{ij}^{\lambda,t}) = \tilde{f}_{ij}^t(\lambda), \tag{10}$$

[13] In principle, it is possible for a firm to anticipate its own market exit in the future. However, as we will show in Section 5, this will not be the case in equilibrium. Firms either exit at time $t = 0$ or they expect to be operating forever.

i.e., firms with $\varphi \geq \varphi_{ij}^{\lambda,t}$ (and tenure λ) find it optimal to serve market j. At the time of market entry ($\lambda = 0$), Equation (10) is analogous to the counterpart in the Melitz model, and it holds with equality for the least productive firm serving j from i. With tenure in the market ($\lambda > 0$), however, fixed costs decline, which implies that, in the absence of shocks, Condition (9) is strictly non-binding for *all* firms with positive tenure, as we will explain further in Section 5. This buffer gives rise to tenure effects in aggregate trade.

Free entry implies that in every period the expected profits from entering in country i are non-positive, and they are exactly equal to zero with positive entry. Using the fact that $\left(\tilde{\pi}_{ij}^{v,t}(\varphi)/\tilde{\pi}_{ij}^{v,t}(\varphi_{ij}^{0,t})\right) = \left(\varphi/\varphi_{ij}^{0,t}\right)^{\sigma-1}$ and Equation (10), we obtain the following free entry condition

$$\sum_{j\in\mathcal{I}} \int_{\varphi_{ij}^{0,t}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{ij}^{0,t}}\right)^{\sigma-1} - 1\right] \tilde{f}_{ij}^{t}(0) g(\varphi) d\varphi \begin{cases} \leq \tilde{f}_{ei}^{t} & \text{if } M_{ei}^{t} = 0 \\ = \tilde{f}_{ei}^{t} & \text{if } M_{ei}^{t} > 0 \end{cases}, \tag{11}$$

where M_{ei}^{t} denotes the mass of entrants in period t.

Sequence of events. Before analyzing the equilibrium in Section 5, it is instructive to clarify the sequence of events in our economy. In any period t, the government first announces changes to the (exogenous) trade environment, if there are any. As all such changes are unanticipated, this will result in an instantaneous updating of all asset values in the economy. The values a_i^t, r_i^t (and firm values detailed in the appendix) capture updated values *after* announcement. Finally, firms simultaneously decide on their production and exporting, households on their consumption, prices are determined, and all markets clear.

5 Equilibrium

In this section, we derive the equilibrium in our economy. We start with preliminary considerations on aggregate demand, adjustments to trade-cost shocks, and firm entry. We then derive gravity equations, first for the case of simultaneous trade liberalizations, and then for the general case of sequential liberalizations.

5.1 Preliminary Considerations

In our model, aggregate demand on the CES aggregator comprises household expenditure on consumption plus a share of firm expenditure on fixed market access costs, which are both time-varying. Yet, in the absence of future entry or changes in firms' market access strategies, aggregate demand in country i is a constant-over-time multiple of its wage rate. This follows from noting first that, with CES preferences, total sales of firms in country i are a constant multiple $\sigma/(\sigma-1)$ of total variable costs (i.e., of wage payments to production workers). For a given set of firms and market-access strategies, the latter is a constant multiple of the wage. Second, trade is balanced at all times, i.e., aggregate demand on the CES aggregator in country i is equal to total global sales of firms from i.[14] Now, a triple (i, s, φ) uniquely identifies a firm in our economy, i.e., the firm born with productivity φ in period s in country i. Let $\mathcal{I}^t(i, s, \varphi)$ denote the set of destinations that this firm serves in period $t \geq s$. With this notation, we have:

Lemma 1 *Aggregate—firm plus household—expenditure in country i on the CES aggregator is given by*

$$X_i^t = \frac{\sigma}{\sigma - 1} w_i^t \left[L_i - \sum_{s=0}^{t} \left[M_{ei}^s \left[f_{ei} + \int_{\varphi \in \Phi} \sum_{j \in \mathcal{I}^t(i,s,\varphi)} f_{ij} g(\varphi) d\varphi \right] \right] \right]. \tag{12}$$

Proof Follows from the above. □

Lemma 1 implies that—with a constant set of operating firms—aggregate expenditure in country i is a constant-over-time multiple of the wage rate, analogous to the canonical Melitz model. The basic intuition is that decreasing-over-time firm expenditures on fixed market access costs increase 1-for-1 households' disposable income and, hence, their consumption—see Online Appendix O3.1 for further details. In our model, this implies that in response to a shock, all prices and trade flows adjust in the period of the shock and are constant thereafter.

Lemma 2 *Let there be a shock to trade costs τ at time t. Then, firm sales, prices, wages, and interest rates adjust at time t and remain constant thereafter.*

Proof See Appendix A.1. □

Firms are forward-looking in their entry and market-access decisions. Nevertheless, the equilibrium is highly tractable thanks to Lemmata 1 and 2. These lemmata imply that entry and market-access costs are exogenously given multiples of the wage rate

$$\tilde{f}_{ei}^t = f_{ei} w_i^t$$
$$\tilde{f}_{ij}^t(\lambda) = f_{ij} w_i^t \tilde{f}(\lambda),$$

where

$$\tilde{f}(\lambda) := \left[\sum_{\varsigma=\lambda}^{\infty} f(\varsigma) \left(\frac{1}{1+r} \right)^{\varsigma - \lambda} \right] \frac{r}{1+r}$$

[14] This is because household investment in the domestic aggregate asset just compensates for the time-varying fixed market access cost of the firms. See below and Online Appendix O3.1 for further details.

is an annuity for the multiplier of market access costs for a firm with tenure λ at the beginning of the period, evaluated at the equilibrium interest rate $r := \frac{1}{\beta} - 1$—see Equation (5). In turn, this allows solving for the mass of entering firms.

Lemma 3 *All firms are born at $t = 0$. The mass of entrants in country i is*

$$M_{ei}^{0} = \frac{L_i}{f_{ei}} \frac{\tilde{f}(0)(\sigma - 1)}{(\sigma - 1)\left[(\theta + 1)\tilde{f}(0) - 1\right] + \theta}. \tag{13}$$

Proof See Appendix A.2. □

According to Lemma 3, all firms enter at $t = 0$ (the very first period). Intuitively, this is because the mass of entering firms is independent of the trade environment, analogous to the canonical Melitz model with fixed market access cost in terms of labor and a Pareto distribution of firm productivities. In fact, with $f(\lambda) = 1 \; \forall \; \lambda$, and therefore $\tilde{f}(0) = 1$, Equation (13) reduces to the well-known expression of the canonical Melitz model with free entry (Melitz and Redding, 2014a). With $\tilde{f}(0) > 1$, there are more entrants, reflecting that with higher costs of market access, fewer firms find it profitable to start operating.[15]

5.2 Gravity Equation

Our theory gives rise to a gravity equation for aggregate trade flows, as we now explain.

Simultaneous trade liberalization. To highlight the importance of incumbency effects for aggregate trade, it is instructive to first consider the special case where all trade costs are determined at time $t = 0$ and then held constant forever. In particular, from the previous section, we know that all firms are born at $t = 0$ (Lemma 3) and that the equilibrium is constant over time (Lemma 2). As a consequence, all market entry is simultaneous and incumbency effects—which are our main focus here—trivially have no impact on trade flows. Indeed, the following proposition shows that in this special case, the gravity equation is identical to the one in the canonical Melitz model.

Proposition 1 *Let $\tau_{ij}^{\varsigma} = \tau_{ij}$ for all $0 \leq \varsigma \leq t$. Then time-t bilateral trade shares are given by*

$$\frac{X_{ij}^{t}}{X_{j}^{t}} = \frac{\frac{L_i}{f_{ei}}(\tau_{ij})^{-\theta}(w_i^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}(f_{ij})^{\frac{\sigma-1-\theta}{\sigma-1}}}{\sum_{k\in\mathcal{I}}\frac{L_k}{f_{ek}}(\tau_{kj})^{-\theta}(w_k^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}(f_{kj})^{\frac{\sigma-1-\theta}{\sigma-1}}}. \tag{14}$$

[15] Fixed cost of market access feed back into the mass of entrants in our model because they are partly paid in terms of the CES aggregator. In equilibrium, this implies that the labor demand of the average operating firm is smaller when compared to the labor demand involved in entering, which translates into a larger labor-market-clearing mass of entering firms. See Appendix A.2 for further details.

Proof See Appendix A.3. □

With Proposition 1 as our benchmark, we now turn to characterize the general case of sequential trade liberalizations.

Gravity with history. In an economy with sequential trade liberalizations, the order of trade openings matters for trade flows today. This is because incumbent firms have sunk investments in market access and are, therefore, less prone to exit a market upon negative trade shocks. In aggregate, this implies that—ceteris paribus—countries that started trading earlier trade more today, as we now explain.

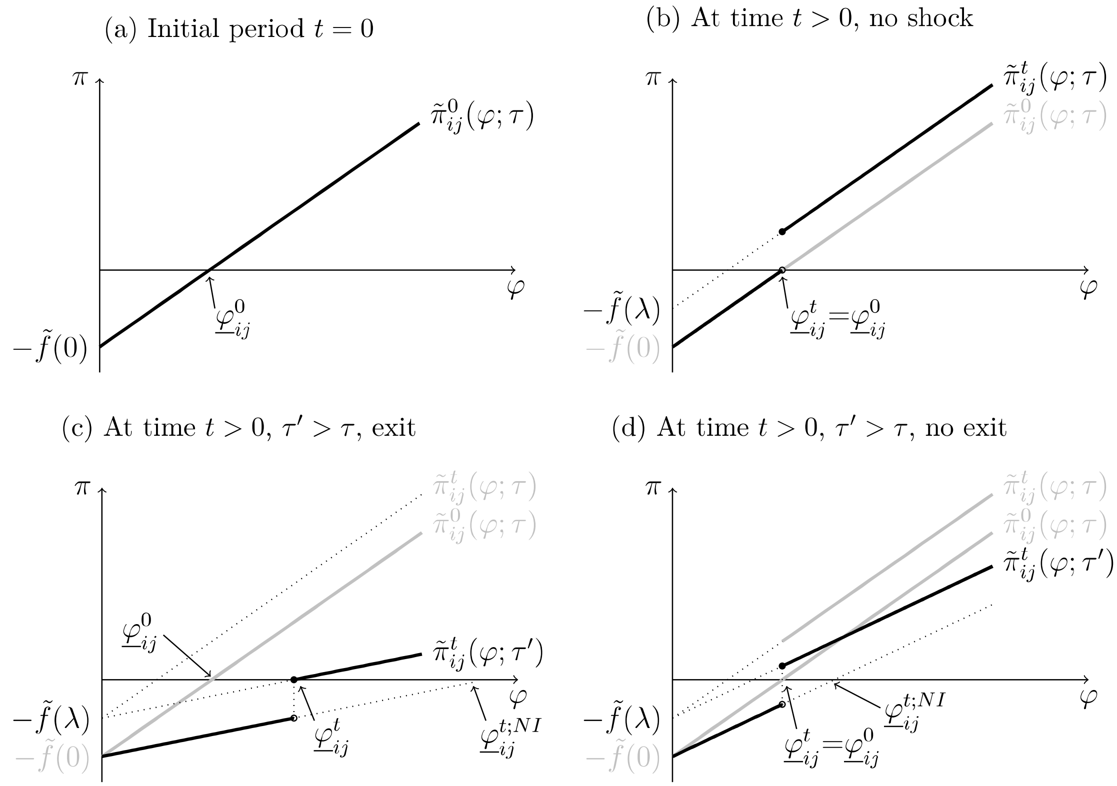
The main mechanism is illustrated in Figure 3, which shows, for various scenarios, firm profits from serving j in annuity terms—$\tilde{\pi}^t_{ij}(\varphi) := \tilde{\pi}^{v,t}_{ij}(\varphi) - \tilde{f}^t_{ij}(\lambda^t_j(i, 0, \varphi))$—as a function of their productivity φ, along with the ensuing productivity of the least productive firm serving i from j at time t. The latter will henceforth be denoted by $\underline{\varphi}^t_{ij}$. For the purpose of illustration, we set $\sigma = 2$ such that $\varphi = \varphi^{(\sigma-1)}$, and normalize $w^0_i f_{ij} = 1$ and neglect general equilibrium effects on the intercepts such that $\tilde{f}^t_{ij}(\lambda) = \tilde{f}(\lambda)$ throughout.

Panel 3a begins by considering the base period upon initial liberalization. By construction, there are no tenure effects, fixed cost of market access are $\tilde{f}(0)$, and the least productive firm serving market j has productivity $\underline{\varphi}^0_{ij} = \varphi^{0,0}_{ij}$, where $\varphi^{0,0}_{ij}$ is the cutoff productivity as defined in Equation (10). That is, the least productive firm serving j just makes zero profits (in annuity terms), analogous to the canonical Melitz model.

This is no longer case as we move forward in time. In particular, suppose there are no trade shocks until at time $t > 0$ trade costs between i and j increase from τ to $\tau' > \tau$. At time t the firm with productivity $\underline{\varphi}^0_{ij}$ has tenure $\lambda = t > 0$ and, hence, benefits from lower market access costs $\tilde{f}(\lambda) < \tilde{f}(0)$. This results in an upward shift of the profit function for all firms $\varphi \geq \underline{\varphi}^0_{ij}$ (i.e., all firms that have served j before the shock), as illustrated in Panel 3b. The upward shift implies that, prior to the shock, the firm with productivity $\underline{\varphi}^0_{ij}$ makes strictly positive profits (in annuity terms) and, hence, introduces a tenure buffer for incumbent firms. It does, however, not result in entry as firms with productivity $\varphi < \underline{\varphi}^0_{ij}$ do not benefit from the lower fixed costs, which implies that $\tilde{\pi}^t_{ij}(\varphi)$ jumps at $\underline{\varphi}^0_{ij}$ in Panel 3b. The least productive firm serving j from i is still $\underline{\varphi}^t_{ij} = \underline{\varphi}^0_{ij}$.

Consider now what happens after the negative trade shock at t, raising trade costs to $\tau' > \tau$. The higher marginal costs of serving the foreign market imply that in response to the shock, $\tilde{\pi}^t_{ij}$ becomes flatter as shown in Panels 3c and 3d. There are two possibilities regarding the qualitative response. If the shock is sufficiently large, the tenure buffer does not suffice to shield all firms from exiting, and some firms with productivity $\varphi > \underline{\varphi}^0_{ij}$ stop serving the market. As a consequence, the least productive firm that finds it profitable to continue serving j after the shock has productivity $\underline{\varphi}^t_{ij} > \underline{\varphi}^0_{ij}$, and Equation (10) is

Figure 3: Profits, Productivity Cutoff, and Exit in Response to Trade Shock



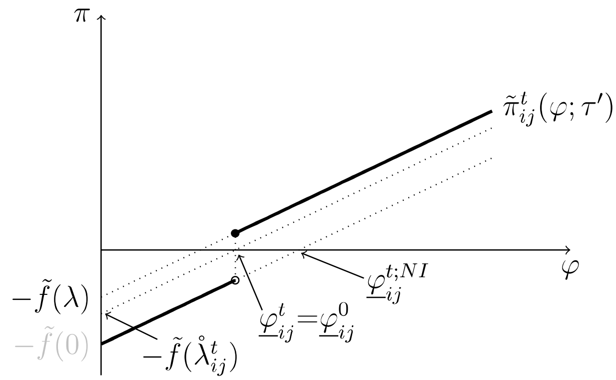
Notes. This figure illustrates how the profit function evolves over time for a given trade pair ij under different trade-cost regimes. For illustrative purposes, we let $\sigma = 2$ (such that $\varphi = \varphi^{(\sigma-1)}$), and we normalize $w^0_i f_{ij} = 1$ and neglect general equilibrium effects on the axes intercepts such that $\tilde{f}^t_{ij}(\lambda) = \tilde{f}(\lambda)$ throughout (in general, $w^t_i f_{ij}$ would be different from one in response to the shock). Note that $\lambda = t > 0$ in Panels 3b–3d. $\underline{\varphi}^{t;NI}_{ij}$ denotes the productivity of the least productive firm serving j from i if there were no incumbency effects.
Graph. Authors' representation.

binding for this firm, as depicted by Panel 3c. Importantly, however, there is still less exit when compared to a model without tenure effects. This is illustrated in the figure by $\underline{\varphi}^t_{ij} < \underline{\varphi}^{t;NI}_{ij}$, where $\underline{\varphi}^{t;NI}_{ij}$ denotes the productivity of the least productive firm that would prevail in the absence of incumbency effects.

If, on the other hand, the shock is small, the tenure buffer is sufficiently large such that even firm $\underline{\varphi}^0_{ij}$ finds it profitable to continue serving the market. There is no exit, and the zero-profit condition is non-binding for the least productive firm that serves j despite the negative trade shock, as illustrated in Panel 3d.

Taking the insights from the graphical representation into account, the equilibrium behavior of firms has two distinct features. First, there is always a unique productivity level $\underline{\varphi}^t_{ij}$ such that all firms in i with productivity $\varphi \geq \underline{\varphi}^t_{ij}$ serve j at t, while all firms

Figure 4: Graphical Illustration of $\tilde{f}(\mathring{\lambda}_{ij}^{t})$ after Trade Shock $\tau' > \tau$



Notes. This figure reproduces the scenario in Figure 3d adding $\tilde{f}(\mathring{\lambda}_{ij}^{t})$, where $\mathring{\lambda}_{ij}^{t}$ is as defined in Equation (15).
Graph. Authors' representation.

with $\varphi < \underline{\varphi}_{ij}^{t}$ do not.[16] Second, incumbency effects may imply that the cutoff firm is *not* indifferent between serving market j or not. To analyze equilibrium trade flows in our economy, it will therefore be convenient to introduce a (hypothetical) tenure level $\mathring{\lambda}_{ij}^{t}$ such that[17]

$$\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^{t}) = w_i^t f_{ij} \tilde{f}(\mathring{\lambda}_{ij}^{t}). \tag{15}$$

In words, $\mathring{\lambda}_{ij}^{t}$ is the tenure level such that the least productive firm that serves j from i would just be indifferent to exit or not—see Figure 4 for a graphical illustration. The tenure level $\mathring{\lambda}_{ij}^{t}$ lies between zero and the actual tenure of the latest entrant. In response to a sufficiently large negative shock, there is exit, and $\mathring{\lambda}_{ij}^{t}$ is exactly equal to the time-t tenure of the least productive firm that continues to serve j, $\lambda_j^t(i, 0, \underline{\varphi}_{ij}^{t})$—see Panel 3c. In response to a sufficiently large positive shock, there is market entry, implying that $\mathring{\lambda}_{ij}^{t} = 0$, analogous to Panel 3a. In general, we have $\mathring{\lambda}_{ij}^{t} \in [0, \lambda_j^t(i, 0, \underline{\varphi}_{ij}^{t})]$. $\mathring{\lambda}_{ij}^{t}$, in sum, indicates the minimum tenure level needed to sustain the level of aggregate trade from i to j after the shock. We will henceforth simply refer to $\mathring{\lambda}_{ij}^{t}$ as the *tenure effect* in trade from i to j (the larger $\mathring{\lambda}_{ij}^{t}$, the more important the incumbency advantage for trade from i to j). With this notation, we can characterize equilibrium trade flows with an arbitrary sequence of trade liberalizations as follows:

Proposition 2 *For any history of trade liberalizations, suppose that at time t there is a*

[16]This is true for an arbitrary sequence of trade liberalizations because at all times firms with higher productivity earn higher variable profits and, hence, they must have a weakly longer tenure in that market. Recall that there is only one firm cohort (Lemma 3).

[17]The expression in Equation (15) assumes that $f(\cdot)$ is continuous. This is for notational convenience only. Alternatively, we could define a (hypothetical) level of $\tilde{f}(\lambda)$.

shock to the trade environment. Then bilateral trade shares are

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}}(\tau_{ij}^t)^{-\theta}(w_i^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}(f_{ij}\tilde{f}(\mathring{\lambda}_{ij}^t))^{\frac{\sigma-1-\theta}{\sigma-1}}}{\sum_{k\in\mathcal{I}}\frac{L_k}{f_{ek}}(\tau_{kj}^t)^{-\theta}(w_k^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}(f_{kj}\tilde{f}(\mathring{\lambda}_{kj}^t))^{\frac{\sigma-1-\theta}{\sigma-1}}} \quad \forall i,j, \tag{16}$$

where $\mathring{\lambda}_{ij}^t$ is as defined in Equation (15) and measures the tenure effect in trade from i to j. In the absence of further shocks, trade shares remain unchanged in future periods.

Proof See Appendix A.4. □

Proposition 2 carries the main message of our paper. It generalizes the basic insight from our simple example in Section 3 to a world with many, asymmetric countries, arbitrary trade frictions, and forward-looking firms. Specifically, it shows how aggregate trade flows are not only influenced by trade costs today, as captured by τ_{ij}^t and f_{ij}, but also by the history of trade liberalizations that resulted in today's trade costs. The latter effect is summarized by the terms $\tilde{f}(\mathring{\lambda}_{kj}^t)$ in Equation (16). These terms imply that, ceteris paribus, countries trade more with each other if they liberalized earlier, as $\tilde{f}(\mathring{\lambda}_{ij}^t)$ is smaller for these relations. This is because firms with longer tenure face lower market access costs and, hence, are less prone to exit in response to a negative trade shock. It further implies that the trade elasticity is not the same for positive and negative trade-cost shocks, and in general it is not constant, as trade shocks also impact the tenure effect. In the special case where $\mathring{\lambda}_{ij}^t$ is the same across all exporters i—or where $\tilde{f}(\cdot)$ is constant—, Equation (16) reduces to the standard gravity Equation (14).

6 Empirical Analysis of Tenure Effects in Aggregate Trade

In the previous sections, we have shown how tenure effects imply that the history of trade liberalizations can have a lasting impact on international trade. This section links this result to our motivating facts from Section 2.1. We then exploit the structure of our model to provide evidence in support of our main mechanism.

6.1 Tenure Effects in Gravity Regressions

As shown in Proposition 2, our theory gives rise to a gravity equation for bilateral trade. This makes the tenure effects very transparent and allows contrasting our results with the large theoretical and empirical literature that is centered on gravity equations. It also allows our theory to directly speak to our motivating regressions from Section 2.1. In particular, taking logs of Equation (16), rearranging terms, and omitting time superscripts for simplicity, we obtain

$$\log(X_{ij}) = \overbrace{\log\left(\frac{L_i}{f_{ei}}(w_i)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}\right)}^{\equiv E_i} + \overbrace{\log\left(X_j\left[\sum_{k\in\mathcal{I}}\frac{L_k}{f_{ek}}(\tau_{kj})^{-\theta}(w_k)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}(f_{kj}\tilde{f}(\mathring{\lambda}_{kj}))^{\frac{\sigma-1-\theta}{\sigma-1}}\right]^{-1}\right)}^{\equiv M_j}$$
$$-\left[\overbrace{\theta\log(\tau_{ij})}^{TC_{var}} + \overbrace{\frac{\theta-(\sigma-1)}{\sigma-1}\log(f_{ij})}^{TC_{fix}} + \overbrace{\frac{\theta-(\sigma-1)}{\sigma-1}\log\left(\tilde{f}(\mathring{\lambda}_{ij})\right)}^{history}\right]. \tag{17}$$

Equation (17) additively separates bilateral trade into exporter-specific terms, importer-specific terms, bilateral (fixed and variable) trade costs, and a bilateral tenure term. It exactly maps onto our empirical specification from Section 2.1. The key novelty of our paper—and the main focus of our analysis—is the last summand in Equation (17). This term implies that, ceteris paribus, trade relations that liberalize later tend to have relatively lower trade flows. Accordingly, we would expect that countries with colonial ties—i.e., countries that plausibly started trading earlier—should ceteris paribus trade more even today. Conversely, countries that were separated by the Iron Curtain—a major impediment to trade as shown in Figure O5.2 of the Online Appendix—should trade less even today. This is in line with what we find in our gravity regressions of Section 2.1. Our theory may thus help explain these sizable, long-lasting, and pair-specific effects of historical events on trade. These effects do not seem to be captured by conventional proxies for trade frictions or be entirely attributable to contemporaneous differences in pair-specific trade policies or transportation infrastructure, as shown in Table 1. Nevertheless, the Iron Curtain and colonial ties dummies in Equation (1) likely capture other effects as well. We therefore analyze our underlying mechanism next.

6.2 Average Firm Sales

In our model, history matters for trade flows today because market access costs decline with tenure, implying less exit and entry in response to a trade shock. We cannot directly observe the fixed market access costs. In this section, we therefore exploit the structure of our model and use average firm sales to provide indirect evidence supporting these effects.

Theory and estimation. In our model, there is a tight connection between the fixed costs of serving a market and the average firm sales in that market (i.e., total exports divided by the number of firms serving a market), analogous to the canonical Melitz model. In particular, it is straightforward to show—see Online Appendix O3.2—that average firm sales from country i to j in period t, $\bar{r}_{ij}^t$, are equal to

$$\bar{r}_{ij}^t = \frac{\sigma\theta}{\theta - (\sigma - 1)} f_{ij} w_i^t \tilde{f}(\mathring{\lambda}_{ij}^t), \tag{18}$$

where, recall, $\mathring{\lambda}_{ij}^t$ captures the *tenure effect* in trade from i to j. Taking log-differences yields

$$\Delta \log \left[\bar{r}_{ij}^t\right] \quad = \quad \Delta \log \left[w_i^t\right] + \Delta \log \left[\tilde{f}(\mathring{\lambda}_{ij}^t)\right] \qquad \forall\ (i, j) \in \mathcal{I} \times \mathcal{I}, \tag{19}$$

where $\Delta \log(x^t) := \log(x^t) - \log(x^{t-1})$. Now, suppose that at time t there is a drop in τ_{ij}^t. Then, ceteris paribus $\tilde{f}(\mathring{\lambda}_{ij}^t)$ weakly increases. This is because in the wave of a positive trade shock, the profit potential of firms from i in j increases (see also the discussion on Figure 3 in Section 5.2). Consequently, incumbency effects become less important for this trade relationship. In fact, in response to a sufficiently large positive shock, there is market entry by new firms (who do not have any incumbency advantages), and tenure has no direct effect on trade from i to j. In our model, this is reflected in larger average firm sales. Intuitively, tenure is less important because firms sell more and, hence, are able to recover higher fixed costs. By contrast, the canonical Melitz-Chaney model predicts no impact on average firm sales after a shock to τ_{ij} when controlling for general equilibrium effects on wages.

To test whether average firm sales from i to j increase in response to a positive shock, we use data on average firm sales from the World Bank's Exporter Dynamics Database and then consider large tariff cuts at the 2-digit HS sector level—see Online Appendix O2.1 for details on the data. We consider tariff changes, as they are arguably an observed shock to τ_{ij} without an (obvious) effect on f_{ij}.[18] In Online Appendix O2.2, we consider a more general event instead: the EU Eastern enlargement. This analysis confirms our main insights presented here.

To control for general exporter and importer trends—as suggested by our theory—, we consider country pairs with tariff cuts and compare these pairs to those with non-negative tariff changes. To include only relatively large shocks, we consider tariff cuts in the top quartile of all tariff reductions in our baseline specification. This yields a still fairly modest cutoff of -2.3%, with a mean decline of gross tariffs of 6.7%. To mitigate concerns regarding mid-year reporting, we further consider changes in average firm sales

[18] One potential concern could be the introduction of a free trade agreement between two parties, which often requires additional effort for firms to benefit from preferential tariffs. To alleviate this concern, we keep for each exporter only those destinations where the exporter's preferential status did not change within our period of analysis. For instance, when we look at changes in average firm sales from year t to $t + 2$, the exporter either has preferential tariffs throughout t to $t + 2$, or it faces MFN tariffs for this entire time span.

from t to $t+2$ in response to a tariff cut at t.[19] We then estimate the following empirical counterpart of Equation (19)

$$\begin{aligned}\Delta^2 \log\left[\bar{r}_{ij,g}^{t+2}\right] &= \phi \times \mathbb{1}[\text{Tariff Cut in the Top Quartile from } t \text{ to } t+1]_{ij,g} \\ &\quad + E_{i,g}^{t+2} + M_{j,g}^{t+2} + \varepsilon_{ij,g}^{t+2}, \qquad (20)\end{aligned}$$

where a subscript g denotes sector g, $\Delta^2 \log\left[\bar{r}_{ij,g}^{t+2}\right] := \log\left[\bar{r}_{ij,g}^{t+2}\right] - \log\left[\bar{r}_{ij,g}^{t}\right]$ is the log-change in average firm sales from t to $t+2$, and $E_{i,g}^{t+2}$ and $M_{j,g}^{t+2}$ are exporter-sector-year and importer-sector-year fixed effects (FEs), respectively, to capture exporter-sector trends in production costs and importer-sector trends in market size. $\varepsilon_{ij,s}^{t+2}$ is an error term. The coefficient of interest is ϕ. Our theory predicts $\phi > 0$, as discussed. We estimate (20) and cluster the standard errors at the exporter-sector, importer-sector, and exporter-importer levels.

Results. Table 2 presents the results. The first three columns in the top panel depict the coefficients from our main specification with sector-year FEs, exporter-sector-year FEs, and both exporter-sector-year and importer-sector-year FEs, respectively. The coefficient of interest is stable and statistically significantly larger than zero throughout, as our model predicts. The point estimates suggest that for country pairs with (relatively large) tariff cuts, growth in the tenure-dependent component of fixed costs was about eight percentage points higher than for the remaining pairs. In other words, the benefit from the drop in τ_{ij} was partly offset by a relative increase in $\tilde{f}(\mathring{\lambda}_{ij})$.

To conclude, we present several robustness tests (columns (4)–(10) in Table 2). In column (4), we apply 90% winsorizing to our dependent variable to test whether our results are driven by a few large changes in average firm sales. This somewhat reduces our estimate but leaves it within one standard error deviation from our main coefficient in the third column. We also control for exporter-importer-year-specific effects, the results of which are reported in the fifth column of the top panel. The coefficient is still in the same ballpark, although the standard error increases substantially. This is because there is little variation in our tariff indicator within a country pair in a given year across sectors.[20] In the bottom panel, we document the sensitivity of our results concerning

[19] This implies that we need to define how tariffs can evolve from $t+1$ to $t+2$ for the "treatment group" (with gross tariff cut of at least 2.3% at time t) and the "control group" (with non-negative tariff changes at t). We allow for a partial rebound of up to one-third of the tariff reduction between $t+1$ and $t+2$ for the "treatment group," and omit observations where average tariffs recover by more than that. Moreover, we verify that the observations in the "control group" did not experience any tariff cuts between $t-1$ and $t+2$, where $t-1$ is also chosen because of potential mid-year reporting effects. All importer-exporter-sector-year observations that did not meet either of these criteria were excluded from the sample.

[20] Only about 1% of the observations in the regression sample have variation in the treatment dummy within all three sets of fixed effects.

Table 2: Impact of Tariff Changes on Average Firm Sales

	Dep. Var.: Log-Change in Av. Firm Sales from t to t+2 Control Group: No Tariff Cuts from t-1 to t+2 Main Specification			Robustness 90% Wins.	Pair-Year FE
𝟙[Tariff Cut]	0.098***	0.089***	0.075**	0.059**	0.063
	(0.016)	(0.016)	(0.033)	(0.026)	(0.072)
Sector-Year FE	✓				
Exp-Sector-Year FE		✓	✓	✓	✓
Imp-Sector-Year FE			✓	✓	✓
Mean of Dep. Var.	0.094	0.093	0.091	0.090	0.089
Adj. R-squared	0.023	0.081	0.091	0.100	0.106
No. of Observations	191,979	187,429	151,694	151,694	145,762
	Dep. Var.: Log-Change in Av. Firm Sales from t to $t+\bar{t}$ Control Group: No Tariff Cuts from t-$\underline{t}$ to $t+\bar{t}$ Robustness				
	$\bar{t}$=2, $\underline{t}$=0	$\bar{t}$=2, $\underline{t}$=2	$\bar{t}$=3, $\underline{t}$=1	$\bar{t}$=3, $\underline{t}$=2	$\bar{t}$=3, $\underline{t}$=3
𝟙[Tariff Cut]	0.058**	0.101***	0.108**	0.131**	0.141**
	(0.029)	(0.039)	(0.045)	(0.052)	(0.057)
Exp-Sector-Year FE	✓	✓	✓	✓	✓
Imp-Sector-Year FE	✓	✓	✓	✓	✓
Mean of Dep. Var.	0.100	0.082	0.121	0.103	0.093
Adj. R-squared	0.091	0.089	0.099	0.098	0.100
No. of Observations	187,631	124,313	108,518	88,235	73,128

Notes. This table reports OLS results from regressing log-changes in average firm sales on an indicator for relatively large tariff reductions, as described in Section 6.2. The regression is outlined in Equation (20). Standard errors clustered at the exporter-sector, importer-sector, and exporter-importer levels are reported in parentheses. The full sample spans the years 1997–2014. The "treatment group" consists of observations with tariff reductions belonging to the top quartile of overall tariff cuts. The "control group" had non-negative changes in average tariffs between the years $t - \underline{t}$ and $t + \bar{t}$. The analysis is conducted at the sector level (2-digit HS).
* significant at 10%-level;** significant at 5%-level;*** significant at 1%-level.
Data source. EDD (average firm exports), WITS (import tariffs).
Results. Authors' computations.

(i) the time span we choose to measure the growth in average firm sales and (ii) how long the reference group without tariff cuts should maintain at least the same tariff level. The results look similar in all specifications, with the coefficient slightly increasing for longer time horizons.[21] Additional robustness checks with respect to our data cleaning and variable definitions are discussed in Online Appendix O2.2.

In sum, our empirical results suggest a significant role of bilateral tenure in shaping aggregate trade flows. This has important implications for the home share and the gains from trade, as we discuss next.

[21] Boehm et al. (2020) document that tariff changes at the 6-digit HS level have a ramp-up phase of up to seven years. Unfortunately, average tariffs at the sector level fluctuate substantially more often than those at the product level, and thus our aggregation level impedes investigating tariff shocks over such a long time span.

7 Incumbency Effects and the Gains from Trade

In this section, we discuss implications of our theory. We begin with deriving an augmented ACR formula (Arkolakis et al., 2012), before exploring the quantitative importance of our main mechanism.

7.1 An Augmented ACR Formula

In their seminal contribution, Arkolakis et al. (2012, ACR) show that under a large class of gravity trade models, a country's home share, $\gamma_{ii} := X_{ii}/X_i$, and the trade elasticity, θ, are two sufficient statistics for its welfare gains from trade. In the limiting case of our model with no incumbency effects, the ACR formula immediately applies. This, however, is no longer the case with incumbency effects, because the impact of trade shocks depends on firms' sunk investment in market access. Interestingly though, with symmetric tenure effects, we can still derive a sufficient statistic for a country's gains from trade based on its home share and primal parameters of our model, as the following proposition shows.

Proposition 3 *Suppose that at time* 0 *all firms have zero tenure in all markets. Let there be no trade shocks until at time* $t > 0$ *country* i *moves to autarky. Then, the change in real income associated with moving country* i *to autarky is given by*

$$\frac{\widetilde{W}_i^{t\star}}{\widetilde{W}_i^t} = \left(\gamma_{ii}^0\right)^{1/\theta} \frac{1 + \gamma_{ii}^0 \mathcal{A}(\lambda_i^t)}{1 + \mathcal{A}(\lambda_i^t)}, \tag{21}$$

where $\widetilde{W}_i^{t\star}$ *(* $\widetilde{W}_i^t$ *) denotes time-*t *welfare with (without) the move to autarky,* $\lambda_i^t = t$ *is the time-*t *tenure of all firms from* i *in all destinations* j*, and where*

$$\mathcal{A}(\lambda_i^t) := \frac{\theta - (\sigma - 1)}{(\sigma - 1)\left[(\theta + 1)\tilde{f}(0) - 1\right] + \theta} \left[\tilde{f}(0) - \tilde{f}(\lambda_i^t)\right].$$

Proof See Appendix A.5. □

Proposition 3 is the direct analogue of Arkolakis et al. (2012, Corollary 1). In the limiting case where $\tilde{f}(\lambda_i^t) = \tilde{f}(0)$ (i.e., in the case without incumbency effects), the welfare change associated with moving to autarky reduces to $(\gamma_{ii}^0)^{1/\theta}$, the famous ACR formula. With incumbency effects, the costs of moving to autarky are higher, and more so the larger the incumbency effects ($\tilde{f}(\lambda_i^t)$ smaller) and the smaller a country's domestic trade share (γ_{ii}^0 smaller). Intuitively, exporting firms have made sunk upfront investments in accessing foreign markets. When moving to autarky, these investments are lost and these losses are more important the more open an economy is.

Proposition 3 points to important implications of incumbency effects for the gains from trade, and it provides a sufficient statistic given a country's home share and primal parameters of our model. It does, however, not fully account for the importance of incumbency effects for trade and the gains from trade, since incumbency effects also impact the home share in the trade equilibrium. We discuss this next.

7.2 Incumbency Effects, Home Share, and the Gains from Trade

In this section, we perform a quantitative analysis of the importance of incumbency effects for the home share and the gains from trade. We begin with summarizing our quantification strategy. Further details are provided in Online Appendix O4.1.

7.2.1 Summary of Quantification Strategy and Parameter Choices

We follow a large body of literature and analyze trade shocks using hat algebra, i.e., by solving for the equilibrium in changes relative to the baseline equilibrium.[22] This procedure avoids the need to specify a large set of bilateral trade frictions and leaves a small set of parameters that need to be calibrated alongside information on trade flows in the baseline equilibrium, as listed in Table O4.1 of Online Appendix O4.1.

Our set-up differs from standard applications of the hat-algebra approach in that the bilateral tenure terms, $\tilde{f}(\mathring{\lambda}_{ij}^{t})$, respond endogenously. To discipline these, we start from baseline equilibria where $\mathring{\lambda}_{ij}^{t} = 0$ for all i, j at some time $t = 0$—an autarky equilibrium in Section 7.2.2 and the current trade equilibrium in Section 7.2.3. We then consider a trade shock at $t' > 0$ and assume that after t' periods, tenure effects have fully materialized, that is $\tilde{f}(t') = 1$.[23] In turn, this implies that in response to the trade shock, we must have $\tilde{f}(\mathring{\lambda}_{ij}^{t'}) \in [1, \tilde{f}(0)]$. To calibrate $\tilde{f}(0)$, we exploit our empirical results of Sections 6.2 and 2.1, respectively, which provide a lower and upper bound, as further discussed in Online Appendix O4.1. In our baseline specification, we choose the midpoint of this interval, which yields $\tilde{f}(0) = 1.485$. Given $\tilde{f}(0)$ and $\tilde{f}(t')$, we solve for $\tilde{f}(\mathring{\lambda}_{ij}^{t'})$—which captures the impact of tenure effects on the counterfactual trade equilibrium—accounting for whether the trade shock leads to firm entry, exit, or whether it leaves the set of active firms in a market unchanged. See Online Appendix O4.1 for further details.

7.2.2 Incumbency Effects and the Gains from Trade

In this section, we use our quantification strategy to assess the importance of incumbency effects for the home share and the gains from trade. To that end, we consider the case where countries start trading with themselves at $t = 0$ and open up to trade with the rest of the world at $t' > 0$. This implies that incumbency effects matter for a

[22] See, e.g., Dekle et al. (2007); Costinot and Rodríguez-Clare (2014); Caliendo and Parro (2015); Eaton et al. (2016); Caliendo et al. (2021).

[23] This assumption is in line with the standard modeling of firm dynamics in trade, where firms incur an upfront investment once and then pay a constant per-period fixed cost of market access (see Alessandria et al. (2021) for a review).

country's home sales, but not directly for its exports. In turn, it allows us to back out $\tilde{f}(\mathring{\lambda}_{ii}^{t'})$ by first constructing counterfactual "closed" economies, where we quintuple τ_{ij} $\forall i \neq j$ starting from a baseline equilibrium that matches the data.[24] From this autarky equilibrium, we then revert our trade shock, first accounting for incumbency effects—which brings the economy back to the baseline equilibrium that matches the data—and, second, ignoring incumbency effects. Figure 5 summarizes our main insights. Technical details are provided in Appendix O4.

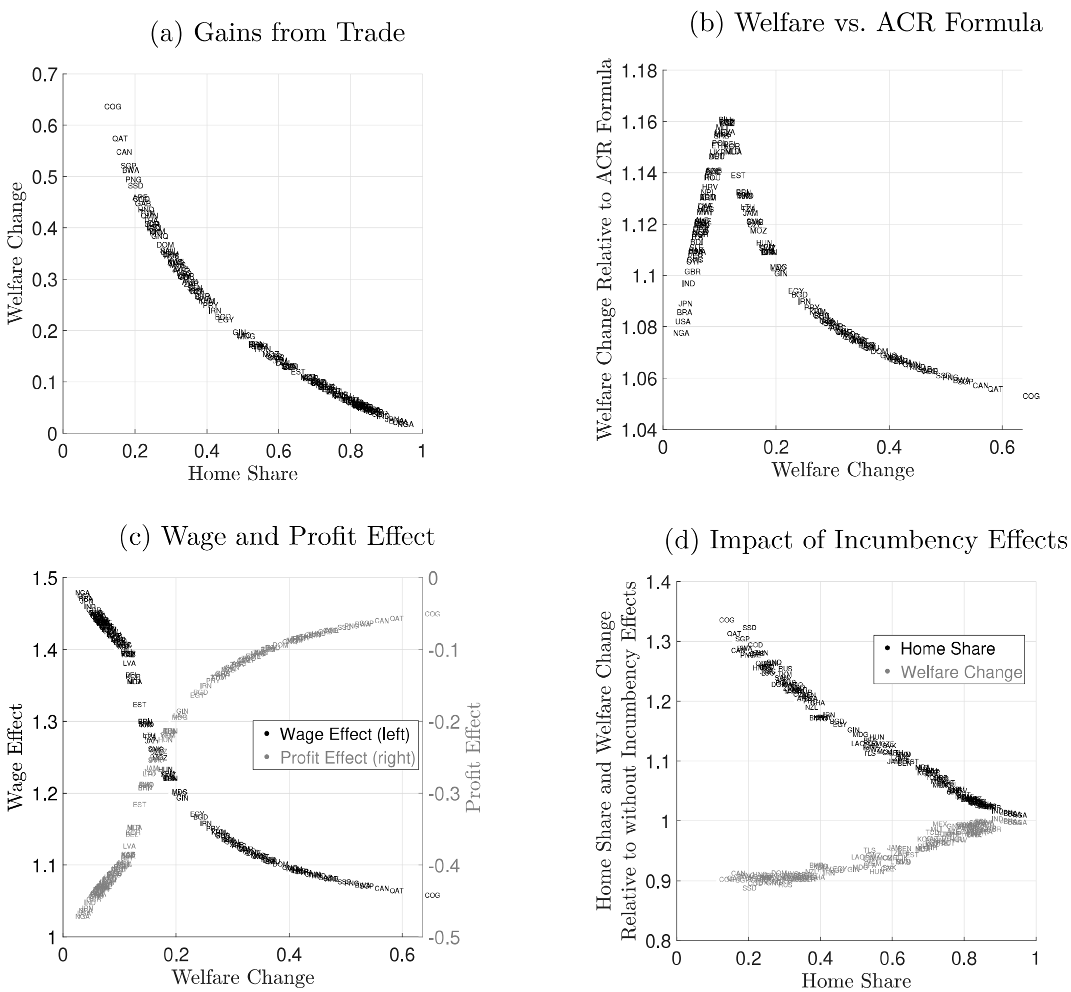
Panel 5a locates each country in a scatter plot with its gains from trade on the vertical axis and its home share in the baseline trade equilibrium on the horizontal axis. As expected, a country's gains from trade are closely connected to its home share. The relationship is, however, somewhat more nuanced when compared to the class of models covered by the ACR formula, as shown in Panel 5b. This figure compares the gains from trade to those obtained by applying the ACR formula to the empirical home share. Two main insights emerge from this figure. First, the gains from trade are on average 10% higher than suggested by the ACR formula. Intuitively, incumbency effects result in less exit by domestic firms. This increases the home share but also competition in the domestic market given the home share and, hence, lowers the aggregate price index. We will get back to this point momentarily.

Second, observe from Panel 5b that the relationship between the gains from trade and the gains suggested by the ACR formula is non-monotonic. For countries with the lowest gains from trade, incumbency effects in the home market are less important because the home share is so large that trade imposes little competitive pressure on domestic firms. As a consequence, the survival of these firms does rely only little on their tenure buffer. As we raise the gains from trade—and, hence, lower the domestic sales share in the trade equilibrium (see Panel 5a)—incumbency effects become increasingly important and the ACR formula increasingly underestimates the gains from trade until we reach a peak at gains of about 10%. At this point, the trade liberalization just induces the least productive domestic firms to stop serving their home market, and tenure effects are maximized.[25] As we keep on increasing the gains from trade, domestic sales and, hence, incumbency effects become again less important for welfare, and the ACR formula

[24]This step is necessary, because for countries with a large home share it need not be the case that in the trade equilibrium the incumbency effects reach their full potential, i.e., that $\tilde{f}(\mathring{\lambda}_{ii}^{t'}) = \tilde{f}(t')$. Indeed, in our baseline quantification this is the case for about one-third of the countries. We note that $\hat{\tau}_{ij} = 5 \; \forall i \neq j$ results in a near autarky equilibrium where the home share is for all countries larger than 99.7%, and 99.99% for the country with the highest home share. Results are almost identical when using $\hat{\tau}_{ij} = 3$ or $\hat{\tau}_{ij} = 7 \, \forall i \neq j$ to construct the counterfactual closed economies.

[25]Figure O5.4 of the Online Appendix reports how the productivity cutoffs of domestic firms and their tenure buffers change when moving from autarky to the trade equilibrium. In line with the discussion above, the difference between our welfare gains and those suggested by the ACR formula in Panel 5b is maximized on the verge of firm exit.

Figure 5: Incumbency Effects and the Gains from Trade



Notes. This figure reports the results of our quantitative investigation of the gains from trade and their relation to incumbency effects, as detailed in Section 7.2.2. First, we create a counterfactual closed economy increasing bilateral trade costs by 400%. Then, we reduce trade costs back to their initial levels once in a version of the model with incumbency effects and once in a version without. Panel 5a relates the welfare gains from trade in our model to the home share in the data. Panel 5b shows the difference between the welfare gains and the ACR formula ($\hat{\gamma}_{ii}^{-1/\theta}$) in the presence of incumbency effects. Panel 5c reports what share of the welfare change is due to changes in the real wage (wage effect) and real profits (profit effect), respectively. That is, the wage and profit effect always sum to 1, and a share above 1 for the wage effect means that the real wage gain is larger than the overall welfare effect. See Online Appendix O4.2 for further details on these effects. Panel 5d divides the home shares and welfare changes by their respective counterparts from a version of the model without incumbency effects.
Data source. See Online Appendix O2.1
Graph. Authors' representation.

underestimates the gains from trade by less.

These effects have important distributional consequences. In the absence of incumbency effects, firms make zero profits in expectation, and welfare depends only on the real wage. With incumbency effects, past sunk investments in market access imply that firms make positive net profits going forward. Trade liberalizations lower these profits as they increase competition, implying that some low-productivity firms survive only thanks to their tenure buffer. At the same time, the lower exit rate and, hence, increased competition in the home market provide an additional positive channel through which trade benefits real wages. As a consequence, workers disproportionately benefit from trade, as may be seen from Panel 5c, which decomposes the overall welfare gain into a wage effect (black, left y-axis) and a profit effect (gray, right y-axis)—see Online Appendix O4.2 for technical details. As the figure shows, the change in real wages is for all countries larger than the overall welfare change, and the difference can be more than 40%.

The previous discussions point to important implications of incumbency effects for trade and the gains from trade. In Panel 5d, we explore these more systematically. This figure compares a country's home share and its gains from trade for the cases with and without incumbency effects. The latter follows from ignoring tenure effects when moving from the closed economies to trade. The figure shows two scatter plots, both using a country's empirical home share on the horizontal axis. The upper (black) scatter plot shows the ratio of a country's home share with relative to without incumbency effects on the vertical axis. The lower (gray) plot depicts the ratio of a country's gains from trade with and without incumbency effects. As expected, the home share is larger with incumbency effects. This effect is sizable and can explain up to about 25% of the home share we observe from the data, and about 11% on average. The effect is stronger for countries with smaller home shares. Intuitively, if the home share is large, incumbency effects are less important as trade imposes less competitive pressure on firms in their domestic market. Shutting off incumbency effects and, consequently, lowering the home share does translate into higher gains from trade (lower, gray plot in Figure 5d). Interestingly though, this welfare difference is smaller than what would be suggested by benchmark models based on the differences in home shares.[26] In other words, our theory provides a novel mechanism that explains up to 25% of the large home shares we observe in the data, but the negative implications for the gains from trade of having large home shares are considerably less severe than suggested by benchmark models.[27]

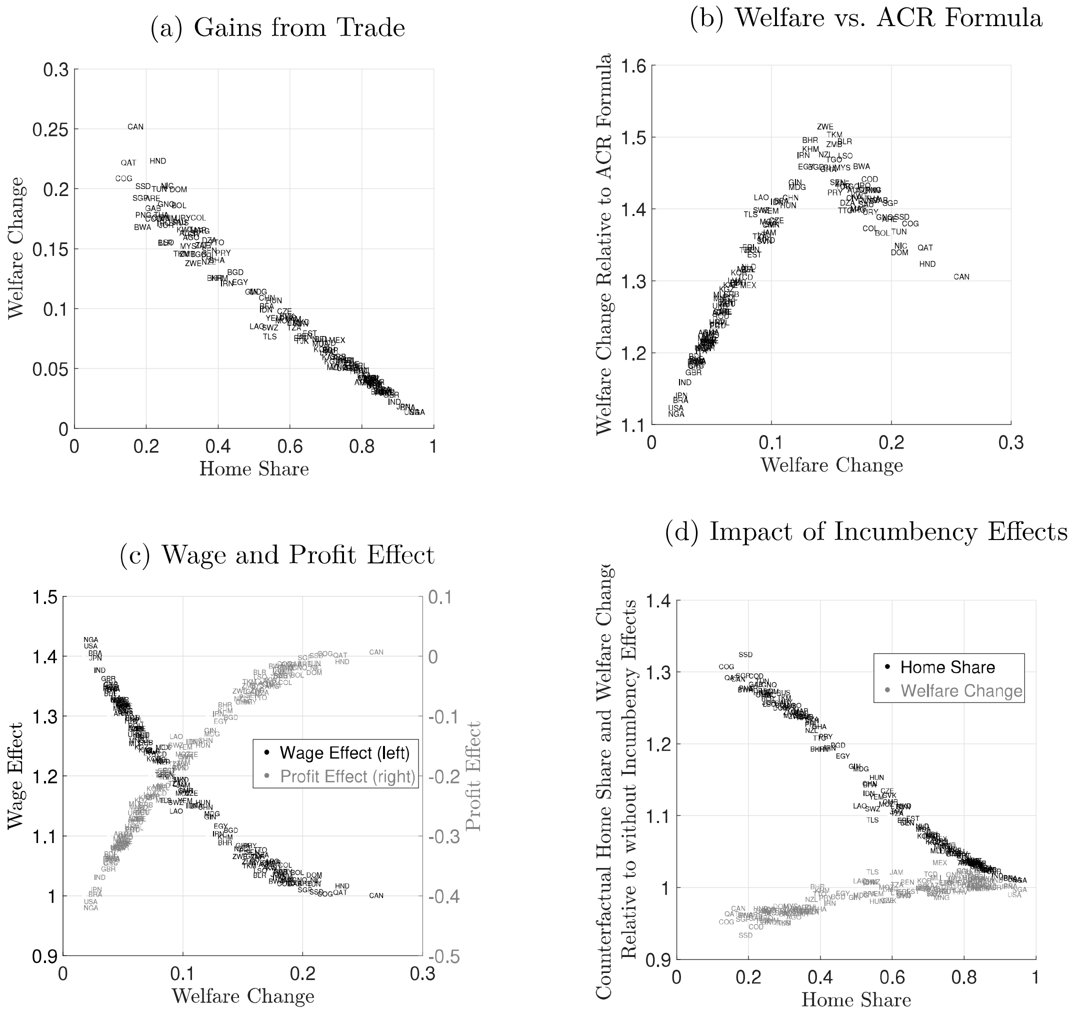
7.2.3 Gains from further Trade Integration

In the previous section, we have considered the gains from trade and focused on moving an economy from autarky to the trade equilibrium we observe from the data—standard

[26]This follows from Panel 5b. In particular, the ratio of applying the ACR formula to the home shares with and without incumbency effects can be derived by multiplying the gray plot in Figure 5d with the inverse of Figure 5b.

[27]Figures O5.5 and O5.6 of the Online Appendix replicate Figure 5 using the lower and upper bounds of $\tilde{f}(0)$, respectively. Qualitatively, all figures yield the same conclusions as those outlined above. A lower value of $\tilde{f}(0)$ is associated with a lower part of the home share that our model can explain. Also, when using the lower bound of $\tilde{f}(0)$, there is exit in every domestic market and, therefore, no hump shape in Figure O5.5b.

Figure 6: Incumbency Effects and the Gains from further Trade Integration



Notes. This figure reports the counterfactual changes that arise after a 15% drop in global variable trade costs, as detailed in Section 7.2.3. Panel 6a relates the welfare gains from trade in our model to the home share in the data. Panel 6b shows the difference between the welfare gains and the ACR formula ($\hat{\gamma}_{ii}^{-1/\theta}$) in the presence of incumbency effects. Panel 6c reports what share of the welfare change is due to changes in the real wage (wage effect) and real profits (profit effect), respectively. That is, the wage and profit effect always sum to 1, and a share above 1 for the wage effect means that the real wage gain is larger than the overall welfare effect. See Online Appendix O4.2 for further details on these effects. Panel 6d divides the counterfactual home shares and welfare changes by their respective counterparts from a version of the model without incumbency effects.
Data source. See Online Appendix O2.1
Graph. Authors' representation.

practice in the literature on international trade. In this section, we use our calibrated model to analyze the implications of further liberalizing by 15%, i.e., of lowering τ_{ij} by 15% for all $i \neq j$. We allow for symmetric incumbency effects in all bilateral trade relationships. That is, we consider the case where at some time $t = 0$ it holds that $\mathring{\lambda}_{ij}^{t} = 0$ for all i, j, and where then the liberalization occurs at $t' > 0$. Figure 6 replicates Figure 5 for this scenario. The importance of incumbency effects is qualitatively the same when considering a move from autarky or further trade integration. The main difference is that in Figure 6d incumbency effects tend to have a smaller impact on the gains from trade when compared to Figure 5d. Intuitively, when further liberalizing starting from a trade equilibrium, also exporters benefit from incumbency effects which benefits consumers via increased competition and, hence, lower prices.[28]

8 Conclusion

This paper introduces incumbency effects into a dynamic general equilibrium trade model that nests the canonical version of the Melitz (2003)-Chaney (2008) model as a special case. We derive a gravity equation that shows how trade is not only shaped by trade frictions today—as typically assumed in the literature—, but also by the history of trade liberalizations. Our theory can explain up to 25% of the domestic trade shares in the data and, more generally, why historical events have long-lasting implications for international trade. Our quantification further implies that the welfare gains from trade are, on average, 10% larger when accounting for incumbency effects.

Our paper provides a first attempt to integrate sunk market access costs into a dynamic general equilibrium model of international trade with many asymmetric countries. In our model, the economy immediately jumps to a new steady state upon a trade liberalization. This allows introducing incumbency effects into the canonical version of the Melitz model while maintaining its tractability. Future work may set out to allow for richer dynamics, thereby providing additional insights into the persistence of incumbency effects in international trade.

[28]To corroborate these results, we have computed several other shocks. Figure O5.7 reports the counterpart of Figure 6 for different shock sizes, focusing on Lithuania (the country with the median home share in the data). Again, the tipping point in Figure O5.7b corresponds to the point where the active domestic firm with the lowest productivity level is on the verge of exit. Figure O5.7d shows that before that point, the welfare changes implied by our model tend to be at least as high as in a version without tenure effects, while the gains are relatively lower for larger shocks. A similar picture arose when focusing on the Côte d'Ivoire and Ecuador, whose domestic trade shares correspond to the first and third quartiles, respectively, of the home shares in the data. In general, the larger the initial home share and the smaller the shock, the more likely that our full model predicts larger gains from trade liberalization than the version without incumbency effects.

Appendix

A Proofs

A.1 Proof of Lemma 2

Let there be a trade shock at time t. To show the desired result, we proceed in three steps. First, we show that equilibrium wages, prices, and aggregate demand are constant over time if (i) any potential firm entry in response to the shock is at time t, and (ii) all firms revise their market access decisions at time t and then continue following this decision. Second, we show that if all potential firm entry in response to the shock is at time t, all firms will indeed perpetually follow their initial market access decision. We finally show that in such case no firm has an incentive to enter at a later stage.

Step 1. With the trade shock at time t and no future changes in the trade environment, we have $\tau_{ij}^{\varsigma} = \tau_{ij}^{t}$ for every i, j, and $\varsigma \geq t$. Similarly, with all firm entry in response to the shock (if any) at time t and all market access decisions fixed at time t, we have $M_i^{\varsigma} = M_i^t := \sum_{s=0}^{t} M_{ei}^s$, and $\mathcal{I}^{\varsigma}(i, s, \varphi) = \mathcal{I}^t(i, s, \varphi)$ $\forall(i, s, \varphi)$ and $\varsigma \geq t$. Hence, aggregate sales at time ς of firms from i are

$$Y_i^{\varsigma} = \sum_{s=0}^{t} \left[M_{ei}^s \int_{\varphi\in\Phi} \sum_{j\in\mathcal{I}^t(i,s,\varphi)} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} (\tau_{ij}^t w_i^{\varsigma})^{1-\sigma} (P_j^{\varsigma})^{\sigma-1} X_j^{\varsigma} \varphi^{\sigma-1} g(\varphi) d\varphi \right], \tag{A.1}$$

where $(P_j^{\varsigma})^{\sigma-1}$ is given by

$$(P_j^{\varsigma})^{\sigma-1} = \left[\sum_{i\in\mathcal{I}} \sum_{s=0}^{t} M_{ei}^s \int_{\varphi\in\Phi} \mathbb{1}\left[j \in \mathcal{I}^t(i,s,\varphi)\right] \left(\frac{\sigma}{\sigma-1} \frac{\tau_{ij}^t w_i^{\varsigma}}{\varphi}\right)^{1-\sigma} g(\varphi) d\varphi \right]^{-1}. \tag{A.2}$$

$\mathbb{1}\left[\cdot\right]$ is an indicator function that takes on value one if the term in brackets is correct, and zero otherwise. Moreover, Lemma 1 implies

$$X_i^{\varsigma} = \frac{\sigma}{\sigma-1} w_i^{\varsigma} \left[L_i - \sum_{s=0}^{t} \left[M_{ei}^s \left[f_{ei} + \int_{\varphi\in\Phi} \sum_{j\in\mathcal{I}^t(i,s,\varphi)} f_{ij} g(\varphi) d\varphi \right]\right]\right]. \tag{A.3}$$

Invoking balanced trade, Equations (A.1) to (A.3) can be reduced to a system of I equations in the I wage rates. This system of equations is the same in every period $\varsigma \geq t$, i.e., indeed wages and, hence, prices, interest rates, and aggregate demand in each country are constant over time.

Step 2. With all prices and aggregate demand constant over time, so are variable profits of all firms serving j from i. Moreover, all wages and interest rates constant implies that

$\tilde{f}_{ij}(\cdot)$ is weakly decreasing over time for all firms serving a market (i.e., for whom λ is increasing), while it is constant over time for any given λ. The former implies that all firms who found it optimal to start serving a market at time t also find it optimal to continue doing so in future periods. The latter implies that for all firms for whom it was optimal *not* to serve market j at t, it is also optimal not to do so in future periods. Hence, indeed, all firms perpetually follow their market access decision.

Step 3. Finally, free entry implies that

$$\sum_{j\in\mathcal{I}} \int_{\varphi_{ij}^{0,t}}^{\infty} \left(\tilde{\pi}_{ij}^{v,t}(\varphi) - \tilde{f}_{ij}^{t}(0)\right) g(\varphi)d\varphi \leq \tilde{f}_{ei}^{t}, \tag{A.4}$$

where in case of strictly positive entry in response to the shock the condition holds with equality. Condition (A.4) also remains unchanged in all periods $\varsigma \geq t$. Therefore, the fact that additional entry is not profitable at the time of the shock implies that it is also not profitable in all subsequent periods.

□

A.2 Proof of Lemma 3

To show the result, we first derive the mass of entrants in the first period, Equation (13), and then proceed by contradiction.

Step 1. Labor market clearing at time $t = 0$ requires

$$M_{ei}^{0}\left[f_{ei} + \sum_{j\in\mathcal{I}} \int_{\varphi_{ij}^{0,0}}^{\infty} \left(\frac{y_{ij}^{0}(\varphi)\tau_{ij}^{0}}{\varphi} + f_{ij}\right) g(\varphi)d\varphi\right] = L_i, \tag{A.5}$$

where $y_{ij}^{0}(\varphi)$ are sales in physical units. Using Equations (7) and (8), and the fact that variable profits are a constant fraction of revenues, Equation (A.5) can be rewritten as

$$M_{ei}^{0}\left[f_{ei} + \sum_{j\in\mathcal{I}} \int_{\varphi_{ij}^{0,0}}^{\infty} \left(\frac{\pi_{ij}^{v,0}(\varphi_{ij}^{0,0})}{w_i^{0}} \left(\frac{\varphi}{\varphi_{ij}^{0,0}}\right)^{\sigma-1} (\sigma - 1) + f_{ij}\right) g(\varphi)d\varphi\right] = L_i. \tag{A.6}$$

By Lemma 2 we know that—in the absence of shocks—the equilibrium is constant. Hence, Equation (10) implies

$$\begin{aligned} \pi_{ij}^{v,0}(\varphi_{ij}^{0,0}) &= \tilde{f}_{ij}^{0}(0) \\ &= f_{ij} w_i^{0} \tilde{f}(0), \end{aligned} \tag{A.7}$$

where the first line uses that $\pi_{ij}^{v,0}(\varphi_{ij}^{0,0})$ is constant and the second line uses

$$\tilde{f}(0) := \left[\sum_{\lambda=0}^{\infty} f(\lambda) \left(\frac{1}{1+r}\right)^{\lambda}\right] \frac{r}{1+r}.$$

$r := \frac{1}{\beta} - 1$ is the equilibrium interest rate, which is the same in all countries. Using Equation (A.7) in (A.6) yields

$$M_{ei}^0 \left[f_{ei} + \sum_{j \in \mathcal{I}} \int_{\varphi_{ij}^{0,0}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{ij}^{0,0}} \right)^{\sigma - 1} (\sigma - 1) \tilde{f}(0) + 1 \right] f_{ij} g(\varphi) d\varphi \right] = L_i. \tag{A.8}$$

Free entry—Equation (11)—implies

$$\sum_{j \in \mathcal{I}} \int_{\varphi_{ij}^{0,0}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{ij}^{0,0}} \right)^{\sigma - 1} - 1 \right] f_{ij} \tilde{f}(0) g(\varphi) d\varphi = f_{ei}. \tag{A.9}$$

Using the Pareto distribution of productivities in Equations (A.8) and (A.9), and rearranging terms yields the expression shown in Equation (13).

Step 2. From Lemma 2, we know that—in the absence of trade shocks—there will be no firm entry after $t = 0$. To show that this is also the case after a trade shock, we proceed by contradiction.[29]

Suppose, by way of contradiction, that in response to a trade shock at time $t > 0$ a mass $M_{ei}^t > 0$ of firms entered in country i. Let $\mathcal{M}_i^t$ denote the set of all firms born in i up to and including time t, and M_i^t the total mass of these firms. Firms in $\mathcal{M}_i^t$ can be uniquely identified by a triple (i, s, φ). Let

$$l^t(i, s, \varphi) := f_{ei} + \sum_{j \in \mathcal{I}^t(i,s,\varphi)} \left(\frac{y_{ij}^t(\varphi) \tau_{ij}^t}{\varphi} + f_{ij} \right)$$

denote the total time-t demand for labor of firm (i, s, φ). Then, for every pair of firms $(i, 0, \varphi), (i, t, \varphi) \in \mathcal{M}_i^t$, i.e., for every pair of firms with equal productivity but born in periods 0 and t, respectively, it holds that $l^t(i, 0, \varphi) \geq l^t(i, t, \varphi)$. In words, the firm born at $t = 0$ demands weakly more labor. This is because (i) conditional on serving a market j, $(i, 0, \varphi)$ and (i, t, φ) demand the same amount of labor to serve j. (ii) For every market j, it holds that $\lambda_j^t(i, 0, \varphi) \geq \lambda_j^t(i, t, \varphi)$ and, hence, $\mathcal{I}^t(i, t, \varphi) \subseteq \mathcal{I}^t(i, 0, \varphi)$. That is, firm (i, t, φ) serves a subset of the markets that firm $(i, 0, \varphi)$ serves, due to (potential) tenure effects.

Now, Equations (A.7) and (A.9) hold for the new entrants with their respective cutoffs, i.e.,

$$\pi_{ij}^{v,t}(\varphi_{ij}^{0,t}) = f_{ij} w_i^t \tilde{f}(0)$$
$$f_{ei} = \sum_{j \in \mathcal{I}} \int_{\varphi_{ij}^{0,t}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{ij}^{0,t}} \right)^{\sigma - 1} - 1 \right] f_{ij} \tilde{f}(0) g(\varphi) d\varphi.$$

[29] From Step 1 we know that the mass of entrants at $t = 0$ is independent of the trade environment in the initial period. We note that in our case this is not sufficient to conclude that there will be no entry upon a future trade shock. This is because—due to the tenure effects—the zero-profit cutoff condition is not necessarily binding for the incumbent firm with the lowest productivity after a trade shock. We will get back to this point when deriving the general gravity equation for sequential trade liberalizations in Appendix A.4.

Derivations along the lines of Step 1 then imply that the labor market in country i would clear if (i) we were in the limiting case where $\mathcal{I}^t(i,t,\varphi) = \mathcal{I}^t(i,0,\varphi)$ and, hence, $l^t(i,t,\varphi) = l^t(i,0,\varphi)$ for every $(i,0,\varphi),(i,t,\varphi) \in \mathcal{M}_i^t$, and (ii) if $M_i^t = M_{ei}^0$. This, however, contradicts $M_{ei}^t > 0$.

□

A.3 Proof of Proposition 1

From Lemma 3 we know that all firms are born at $t = 0$. Lemma 2 implies that they keep their market-access strategies constant over time. Hence, all firms in i with productivity $\varphi \geq \varphi_{ij}^{0,0}$ serve market j at time t, and total exports of i to j are given by

$$X_{ij}^t = M_{ei}^0 \int_{\varphi_{ij}^{0,0}}^{\infty} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} (\tau_{ij} w_i^t)^{1-\sigma} (P_j^t)^{\sigma-1} X_j^t \varphi^{\sigma-1} g(\varphi) d\varphi. \tag{A.10}$$

Solving the integral with the Pareto distribution of firm productivities, using Equation (13) for the mass of entrants, and rearranging terms, we get

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}} (w_i^t \tau_{ij})^{1-\sigma} (\varphi_{ij}^{0,0})^{\sigma-1-\theta}}{\sum_{k\in\mathcal{I}} \frac{L_k}{f_{ek}} (w_k^t \tau_{kj})^{1-\sigma} (\varphi_{kj}^{0,0})^{\sigma-1-\theta}}. \tag{A.11}$$

Equations (8), (10), and the fact that $\tilde{f}_{ij}^t(0) = f_{ij} w_i^t \tilde{f}(0)$ imply

$$(\varphi_{ij}^{0,0})^{\sigma-1-\theta} = \left[(w_i^t)^{\sigma} (\tau_{ij})^{\sigma-1} f_{ij} \tilde{f}(0) \sigma \left(\frac{\sigma}{\sigma-1}\right)^{\sigma-1} (P_j^t)^{1-\sigma} (X_j^t)^{-1} \right]^{\frac{\sigma-1-\theta}{\sigma-1}}.$$

Using this expression in Equation (A.11) and simplifying terms yields Equation (14). By Lemma 2, in the absence of further trade shocks all prices and firms' market access strategies are constant over time and, hence, so are bilateral trade shares.

□

A.4 Proof of Proposition 2

We proceed in two steps. We first characterize firms' market access strategies and then use these to derive the gravity equation.

Step 1. By Lemma 3, all firms are born at time $t = 0$. This implies that for each $(i,j) \in \mathcal{I} \times \mathcal{I}$ there is a unique productivity level $\underline{\varphi}_{ij}^t$ such that all firms in i with productivity $\varphi \geq \underline{\varphi}_{ij}^t$ serve j at t, while all firms with $\varphi < \underline{\varphi}_{ij}^t$ do not.[30] Let $\lambda_{ij}^t := \lambda_j^t(i,0,\underline{\varphi}_{ij}^t) \geq 0$

[30]This is true for an arbitrary sequence of trade liberalizations because at all times firms with a higher productivity earn higher variable profits and, hence, they must have a weakly longer tenure in that market.

denote the time-t tenure of the least productive firm in i serving j. This firm must make non-negative profits from serving j, i.e.,

$$\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^{t}) \geq w_i^t f_{ij} \tilde{f}(\lambda_{ij}^t), \tag{A.12}$$

where we used the fact that by Lemma 2 future profits are constant, and where $\tilde{f}(\lambda) := \left[\sum_{\varsigma=\lambda}^{\infty} f(\varsigma)\left(\frac{1}{1+r}\right)^{\varsigma-\lambda}\right]\frac{r}{1+r}$. Due to the tenure effect, Condition (A.12) may be strictly non-binding. Let us therefore define a hypothetical tenure level, $\mathring{\lambda}_{ij}^t$, at which Condition (A.12) holds with equality, i.e.,[31]

$$\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^{t}) = w_i^t f_{ij} \tilde{f}(\mathring{\lambda}_{ij}^t). \tag{A.13}$$

For all $(i, j, t) \in \mathcal{I} \times \mathcal{I} \times [0, \infty)$, it must hold that $\mathring{\lambda}_{ij}^t \in \left[0, \lambda_{ij}^t\right]$.[32]

Step 2. The remainder of the proof is analogous to the proof of Proposition 1. In particular, using the productivity cutoff $\underline{\varphi}_{ij}^t$, the Pareto distribution of firm productivities, and Equation (13) for the mass of entrants in Equation (A.10), and rearranging terms yields

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}}(w_i^t\tau_{ij}^t)^{1-\sigma}(\underline{\varphi}_{ij}^t)^{\sigma-1-\theta}}{\sum_{k\in\mathcal{I}}\frac{L_k}{f_{ek}}(w_k^t\tau_{kj}^t)^{1-\sigma}(\underline{\varphi}_{kj}^t)^{\sigma-1-\theta}}. \tag{A.14}$$

Equations (8), and (A.13) imply

$$(\underline{\varphi}_{ij}^t)^{\sigma-1-\theta} = \left[(w_i^t)^{\sigma}(\tau_{ij}^t)^{\sigma-1} f_{ij}\tilde{f}(\mathring{\lambda}_{ij}^t)\sigma\left(\frac{\sigma}{\sigma-1}\right)^{\sigma-1}(P_j^t)^{1-\sigma}(X_j^t)^{-1}\right]^{\frac{\sigma-1-\theta}{\sigma-1}}.$$

Using this expression in Equation (A.14) and simplifying terms yields Equation (16). Lastly, Lemma 2 implies, again, that—in the absence of further shocks—bilateral trade shares are constant over time.

□

A.5 Proof of Proposition 3

To show the desired result, we proceed in three steps. We first show that at $t = 0$, welfare is proportionate to $(\gamma_{ii}^0)^{-1/\theta}$, analogous to Arkolakis et al. (2012). We then show how incumbency effects impact welfare over time in the absence of shocks. We finally consider the move to autarky.

[31] If Condition (A.12) holds with equality for more than one λ because $\tilde{f}(\cdot)$ is constant for these λ, we take $\mathring{\lambda}_{ij}^t$ to be the smallest λ for which Condition (A.12) holds with equality.

[32] This follows by contradiction. In particular, (i) if $\mathring{\lambda}_{ij}^t > \lambda_{ij}^t$ it would not be optimal for $(i, \underline{\varphi}_{ij}^t)$ to serve j. And (ii) if Condition (A.13) would be strictly non-binding for $\mathring{\lambda}_{ij}^t = 0$, it would be optimal for some firms with productivity $\varphi < \underline{\varphi}_{ij}^t$ to start serving j, and $(i, \underline{\varphi}_{ij}^t)$ would not be the least productive such firm.

Step 1. Equation (3), Lemma 2, and the fact that the equilibrium interest rate satisfies $r = \frac{1}{\beta} - 1$ imply that welfare is equal to the present value of real income. Let $\widetilde{W}_i^t$ denote the corresponding annuity of real income for country i at time t. We will use this as our measure of welfare.

At time 0, the free-entry condition holds with equality. In combination with Lemma 2 and the fact that $\lambda_j(i, 0, \varphi) = 0$ for all firms and all markets, this implies

$$\widetilde{W}_i^0 = \frac{w_i^0}{P_i^0},$$

i.e., welfare is just equal to the real wage. Equations (8) and (10) imply that

$$P_i^0 = \frac{\sigma}{\sigma - 1} \left(\frac{\sigma w_i^0 f_{ii} \tilde{f}(0)}{X_i^0} \right)^{\frac{1}{\sigma-1}} \frac{w_i^0}{\underline{\varphi}_{ii}^0}. \tag{A.15}$$

To solve for $\underline{\varphi}_{ii}^0$, note first that, at all t, Equation (18) and Lemma 3 imply that

$$X_{ij}^t = \frac{\sigma\theta}{\theta - (\sigma - 1)} f_{ij} w_i^t \tilde{f}(\mathring{\lambda}_{ij}^t) M_{ij}^t, \tag{A.16}$$

where $M_{ij}^t := M_{ei}^0 \left(\frac{\underline{\varphi}}{\underline{\varphi}_{ij}^t} \right)^\theta$ denotes the mass of firms serving j from i at time t. Second, using M_{ij}^t and the fact that all firm birth is at time 0 (Lemma 3) in Equation (12) yields

$$X_i^t = \frac{\sigma}{\sigma - 1} w_i^t \left[L_i - M_{ei}^0 f_{ei} - \sum_{j \in \mathcal{I}} M_{ij}^t f_{ij} \right] \tag{A.17}$$

Solving Equation (A.16) for $f_{ij} M_{ij}^t$ and using it in Equation (A.17) along with Lemma 3 yields

$$X_i^t = \frac{\sigma}{\sigma - 1} w_i^t L_i \left[1 - \frac{\tilde{f}(0)(\sigma - 1)}{(\sigma - 1)\left[(\theta + 1)\tilde{f}(0) - 1\right] + \theta} \right] - \sum_{j \in \mathcal{I}} \frac{\theta - (\sigma - 1)}{(\sigma - 1)\theta} \frac{X_{ij}^t}{\tilde{f}(\mathring{\lambda}_{ij}^t)}. \tag{A.18}$$

Now, at time 0 we have $\mathring{\lambda}_{ij}^0 = 0 \,\forall\, (i, j) \in \mathcal{I} \times \mathcal{I}$. Invoking balanced trade, $\sum_{j \in \mathcal{I}} X_{ij}^t = X_i^t$, and rearranging terms yields

$$X_i^0 = w_i^0 L_i \frac{\sigma\theta \tilde{f}(0)}{(\sigma - 1)\left[(\theta + 1)\tilde{f}(0) - 1\right] + \theta} \tag{A.19a}$$

$$= M_{ei}^0 f_{ei} w_i^0 \frac{\sigma\theta}{\sigma - 1}\ . \tag{A.19b}$$

Dividing Equation (A.16) for the home market (i.e., $i = j$) by Equation (A.19b), using $M_{ii}^t := M_{ei}^0 \left(\frac{\underline{\varphi}}{\underline{\varphi}_{ii}^t} \right)^\theta$, and rearranging terms yields

$$(\underline{\varphi}_{ii}^0)^\theta = \frac{(\sigma - 1)\underline{\varphi}^\theta}{\theta - (\sigma - 1)} \frac{f_{ii} \tilde{f}(0)}{f_{ei} \gamma_{ii}^0}, \tag{A.20}$$

where $\gamma_{ii}^0 := \frac{X_{ii}^0}{X_i^0}$ is country i's home share. Using Equations (A.19a) and (A.20) in Equation (A.15), we get

$$P_i^0 = \frac{\sigma}{\sigma-1}\left(\frac{f_{ii}\left((\sigma-1)\left[(\theta+1)\tilde{f}(0)-1\right]+\theta\right)}{L_i\theta}\right)^{\frac{1}{\sigma-1}}\left[\frac{\theta-(\sigma-1)}{(\sigma-1)\underline{\varphi}^\theta}\frac{f_{ei}}{f_{ii}\tilde{f}(0)}\right]^{\frac{1}{\theta}} w_i^0\left(\gamma_{ii}^0\right)^{1/\theta},$$

which implies

$$\widetilde{W}_i^0 = \frac{w_i^0}{P_i^0} \propto \left(\gamma_{ii}^0\right)^{-1/\theta}.$$

Step 2. In the absence of trade shocks, incumbency effects imply that as of time $t > 0$ welfare is higher than as of time 0. In particular, from Equation (18) we know that

$$M_{ij}^0 f_{ij} w_i^0 \tilde{f}(0) = X_{ij}^0 \frac{\theta-(\sigma-1)}{\sigma\theta}, \tag{A.21}$$

i.e., the annuity of the fixed cost of market access is a constant multiple of aggregate sales. Because this multiple is the same across all destinations, we have

$$\tilde{F}_{mi}^0 := \sum_{j\in\mathcal{I}} M_{ij}^0 f_{ij} w_i^0 \tilde{f}(0) = X_i^0 \frac{\theta-(\sigma-1)}{\sigma\theta}. \tag{A.22}$$

These fixed market access cost decline with tenure and—in the absence of shocks—we have

$$\tilde{\Pi}_i^t := \tilde{F}_{mi}^0 - \tilde{F}_{mi}^t = \tilde{F}_{mi}^0\left[1-\frac{\tilde{f}(\lambda_i^t)}{\tilde{f}(0)}\right],$$

where $\lambda_i^t := t$ is the time-t tenure of all firms from i in all destinations, and where $\tilde{\Pi}_i^t$ is the annuity of aggregate firm profits as of time t. Using Equations (A.19a) and (A.22) we get

$$\tilde{\Pi}_i^t = w_i^0 L_i \underbrace{\frac{\theta-(\sigma-1)}{(\sigma-1)\left[(\theta+1)\tilde{f}(0)-1\right]+\theta}\left[\tilde{f}(0)-\tilde{f}(\lambda_i^t)\right]}_{:=\mathcal{A}(\lambda_i^t)}. \tag{A.23}$$

These profits add to the annuity of households' income as of time t. Hence, in the absence of shocks it holds

$$\widetilde{W}_i^t = \widetilde{W}_i^0\left(1+\mathcal{A}(\lambda_i^t)\right).$$

Step 3. Suppose that at time $t > 0$, the economy moves to autarky. In response to the shock, additional firms start serving the domestic market, and γ_{ii}^t jumps to one. The real wage declines by a factor $(\gamma_{ii}^0)^{1/\theta}$, the standard welfare effect according to Arkolakis et al. (2012). However, in our case, this further impacts the annuity of future profits. The fact that there is entry in response to the shock implies that Equation (A.20) holds both before and after the shock.[33] Using (A.20) in $M_{ii}^t := M_{ei}^0\left(\frac{\underline{\varphi}}{\underline{\varphi}_{ii}^t}\right)^\theta$ therefore implies that

$$\frac{M_{ii}^t}{M_{ii}^{t\star}} = \gamma_{ii}^0,$$

[33] In the initial trade equilibrium, incumbency effects play no role and, hence, a move to autarky induces firm entry in the home market analogous to the canonical Melitz model. Indeed, observe from Equation (A.20) that for $\tilde{f}(0)$ given, a jump of γ_{ii} to one is associated with a decrease in $\underline{\varphi}_{ii}$.

where here and below we use a superscript $\star$ to denote a variable after the shock. In words, only a fraction γ_{ii}^0 of all firms that serve the domestic market under autarky also served it prior to the shock. These are the ones that benefit from the tenure effects and, hence, lower market access costs. Now, the proof of Lemma 3 implies that after the move to autarky the free entry condition is just binding (without further entry) and, hence, the annuity of aggregate profits would be zero in the absence of tenure effects. Derivations analogous to those of Step 2 then imply that

$$\tilde{\Pi}_i^{t\star} = w_i^{t\star} L_i \gamma_{ii}^0 \mathcal{A}(\lambda_i^t),$$

i.e., after the shock, the ratio of the annuity of per household profits over the wage decreases by a factor γ_{ii}^0. This, in combination with Steps 1, 2, and the aforementioned decline in the real wage, implies

$$\frac{\widetilde{W}_i^{t\star}}{\widetilde{W}_i^t} = \left(\gamma_{ii}^0\right)^{1/\theta} \frac{1 + \gamma_{ii}^0 \mathcal{A}(\lambda_i^t)}{1 + \mathcal{A}(\lambda_i^t)}.$$

□

References

Adão, R., Arkolakis, C., and Ganapati, S. (2020). Aggregate Implications of Firm Heterogeneity: A Nonparametric Analysis of Monopolistic Competition Trade Models. Working Paper 28081, NBER Working Paper Series.

Albornoz, F., Pardo, H. F. C., Corcos, G., and Ornelas, E. (2012). Sequential Exporting. *Journal of International Economics*, 88(1):17–31.

Alessandria, G., Arkolakis, C., and Ruhl, K. J. (2021). Firm Dynamics and Trade. *Annual Review of Economics*, 13:253–280.

Alessandria, G. and Choi, H. (2014). Establishment Heterogeneity, Exporter Dynamics, and the Effects of Trade Liberalization. *Journal of International Economics*, 94(2):207–223.

Alessandria, G., Choi, H., and Ruhl, K. (2014). Trade Adjustment Dynamics and the Welfare Gains from Trade. Working Paper 20663, NBER Working Paper Series.

Allen, T. and Donaldson, D. (2020). Persistence and Path Dependence in the Spatial Economy. Working Paper 28059, NBER Working Paper Series.

Alvarez, F. and Lucas, R. E. J. (2007). General Equilibrium Analysis of the Eaton–Kortum Model of International Trade. *Journal of Monetary Economics*, 54(6):1726–1768.

Anderson, J. E. and van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, 93(1):170–192.

Arkolakis, C. (2010). Market Penetration Costs and the New Consumers Margin in International Trade. *Journal of Political Economy*, 118(6):1151–1199.

Arkolakis, C., Costinot, A., Donaldson, D., and Rodríguez-Clare, A. (2019). The Elusive Pro-Competitive Effects of Trade. *The Review of Economic Studies*, 86(1):46–80.

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New Trade Models, Same Old Gains? *American Economic Review*, 102(1):94–130.

Arping, S. and Diaw, K. M. (2008). Sunk Costs, Entry Deterrence, and Financial Constraints. *International Journal of Industrial Organization*, 26(2):490–501.

Atkin, D., Costinot, A., and Fukui, M. (2021). Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom? Working Paper 29500, NBER Working Paper Series.

Baldwin, R. (1988). Hysteresis in Import Prices: The Beachhead Effect. *The American Economic Review*, 78(4):773–785.

Baldwin, R. and Krugman, P. (1989). Persistent Trade Effects of Large Exchange Rate Shocks. *The Quarterly Journal of Economics*, 104(4):635–654.

Balistreri, E. J., Hillberry, R. H., and Rutherford, T. F. (2011). Structural Estimation and Solution of International Trade Models with Heterogeneous Firms. *Journal of International Economics*, 83(2):95–108.

Bas, M., Mayer, T., and Thoenig, M. (2017). From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity. *Journal of International Economics*, 108:1–19.

Boehm, C. E., Levchenko, A. A., and Pandalai-Nayar, N. (2020). The Long and Short (Run) of Trade Elasticities. Working Paper 27064, NBER Working Paper Series.

Borchert, I., Larch, M., Shikher, S., and Yotov, Y. V. (2021). The International Trade and Production Database for Estimation (ITPD-E). *International Economics*, 166:140–166.

Buera, F. J. and Oberfield, E. (2020). The Global Diffusion of Ideas. *Econometrica*, 88(1):83–114.

Burstein, A. and Melitz, M. J. (2013). Trade Liberalization and Firm Dynamics. In *Advances in Economics and Econometrics Tenth World Congress. Applied Economics, Econometric Society Monographs.*, volume 2. Cambridge University Press.

Caliendo, L., Opromolla, L. D., Parro, F., and Sforza, A. (2021). Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement. *Journal of Political Economy*, 129(12):3491–3545.

Caliendo, L. and Parro, F. (2015). Estimates of the Trade and Welfare Effects of NAFTA. *The Review of Economic Studies*, 82(1):1–44.

Caron, J., Fally, T., and Markusen, J. R. (2014). International Trade Puzzles: A Solution Linking Production and Preferences. *The Quarterly Journal of Economics*, 129(3):1501–1552.

Carrère, C., Mrázová, M., and Neary, J. P. (2020). Gravity without Apology: The Science of Elasticities, Distance and Trade. *The Economic Journal*, 130(628):880–910.

Chaney, T. (2008). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review*, 98(4):1707–1721.

Chen, N. (2004). Intra-National versus International Trade in the European Union: Why Do National Borders Matter? *Journal of international Economics*, 63(1):93–118.

Chen, N. and Novy, D. (2022). Gravity and Heterogeneous Trade Cost Elasticities. *The Economic Journal*, 132(644):1349–1377.

Costantini, J. A. and Melitz, M. J. (2007). The Dynamics of Firm-Level Adjustment to Trade Liberalization. In *The Organization of Firms in a Global Economy*. Harvard University Press, Cambridge, Ma.

Costinot, A. and Rodríguez-Clare, A. (2014). Trade Theory with Numbers: Quantifying the Consequences of Globalization. In *Handbook of International Economics*, volume 4, pages 197–261. Elsevier.

Das, S., Roberts, M. J., and Tybout, J. R. (2007). Market Entry Costs, Producer Heterogeneity, and Export Dynamics. *Econometrica*, 75(3):837–873.

Dekle, R., Eaton, J., and Kortum, S. (2007). Unbalanced Trade. *American Economic Review*, 97(2):351–355.

Dixit, A. (1989). Hysteresis, Import Penetration, and Exchange Rate Pass-Through. *The Quarterly Journal of Economics*, 104(2):205–228. Publisher: Oxford University Press.

Eaton, J., Eslava, M., Krizan, C. J., Kugler, M., and Tybout, J. (2021). A Search and Learning Model of Export Dynamics. Working Paper 29100, NBER Working Paper Series.

Eaton, J., Eslava, M., Kugler, M., and Tybout, J. (2007). Export Dynamics in Colombia: Firm-Level Evidence. Working Paper 13531, NBER Working Paper Series.

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

Eaton, J., Kortum, S., Neiman, B., and Romalis, J. (2016). Trade and the Global Recession. *American Economic Review*, 106(11):3401–38.

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, Markups, and the Gains from International Trade. *American Economic Review*, 105(10):3183–3221.

Evenett, S. J. and Venables, A. J. (2002). Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows. Mimeo.

Fajgelbaum, P. D. and Khandelwal, A. K. (2016). Measuring the Unequal Gains from Trade. *The Quarterly Journal of Economics*, 131(3):1113–1180.

Feenstra, R. C. (2018). Restoring the Product Variety and Pro-Competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity. *Journal of International Economics*, 110:16–27.

Felbermayr, G., Teti, F., and Yalcin, E. (2019). Rules of Origin and the Profitability of Trade Deflection. *Journal of International Economics*, 121:103248.

Fitzgerald, D., Haller, S., and Yedid-Levi, Y. (2016). How Exporters Grow. Working Paper 21935, NBER Working Paper Series.

Goolsbee, A. and Syverson, C. (2008). How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines. *The Quarterly Journal of Economics*, 123(4):1611–1633.

Head, K. and Mayer, T. (2014). Gravity Equations: Workhorse, Toolkit, and Cookbook. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, volume 4, chapter 3, pages 131–195. Elsevier.

Head, K., Mayer, T., and Ries, J. (2010). The Erosion of Colonial Trade Linkages after Independence. *Journal of International Economics*, 81(1):1–14.

Impullitti, G., Irarrazabal, A. A., and Opromolla, L. D. (2013). A Theory of Entry into and Exit from Export Markets. *Journal of International Economics*, 90(1):75–90.

Krishna, K., Salamanca, C., Suzuki, Y., and Volpe Martincus, C. (2021). Learning to Use Trade Agreements. Working Paper 29319, NBER Working Paper Series.

McCallum, J. (1995). National Borders Matter: Canada-U.S. Regional Trade Patterns. *The American Economic Review*, 85(3):615–623.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Melitz, M. J. and Redding, S. J. (2014a). Heterogeneous Firms and Trade. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, pages 1–54. Elsevier, 4 edition.

Melitz, M. J. and Redding, S. J. (2014b). Missing Gains from Trade? *The American Economic Review*, 104(5):317–321.

Melitz, M. J. and Redding, S. J. (2015). New Trade Models, new Welfare Implications. *American Economic Review*, 105(3):1105–1146.

Miao, G. and Fortanier, F. (2017). Estimating Transport and Insurance Costs of International Trade. Working Paper 2017/04, OECD Statistics Working Papers.

Morales, E., Sheu, G., and Zahler, A. (2019). Extended Gravity. *The Review of Economic Studies*, 86(6):2668–2712.

Noh, Y.-H. and Moschini, G. (2006). Vertical Product Differentiation, Entry-Deterrence Strategies, and Entry Qualities. *Review of Industrial Organization*, 29(3):227–252.

Novy, D. (2013). International Trade without CES: Estimating Translog Gravity. *Journal of International Economics*, 89(2):271–282.

Ossa, R. (2015). Why Trade Matters After All. *Journal of International Economics*, 97(2):266–277.

Roberts, M. J. and Tybout, J. R. (1997). The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *The American Economic Review*, 87(4):545–564.

Ruhl, K. J. and Willis, J. L. (2017). New Exporter Dynamics. *International Economic Review*, 58(3):703–726.

Schivardi, F. and Viviano, E. (2011). Entry Barriers in Retail Trade. *The Economic Journal*, 121(551):145–170.

Yi, K.-M. (2010). Can Multistage Production Explain the Home Bias in Trade? *American Economic Review*, 100(1):364–93.