

Crawford School of Public Policy



Centre for Applied Macroeconomic Analysis

What Explains Excess Trade Persistence? A Theory of Habits in the Supply Chains

CAMA Working Paper 11/2021 January 2021

Mariarosaria Comunale

Bank of Lithuania Vilnius University Centre for Applied Macroeconomic Analysis, ANU

Justas Dainauskas London School of Economics and Political Science

Povilas Lastauskas

Bank of Lithuania Vilnius University

Abstract

International trade flows are volatile, imbalanced, and fragmented across off-shored supply chains. Yet, not much is known about the mechanism through which trade flows adjust in response to shocks over time. This paper derives a dynamic gravity equation from a theory of habits in the supply chains that generates autocorrelated bilateral trade flows that are heterogeneous across different country pairs. We estimate our version of the dynamic gravity equation for 39 countries over the period of 1950-2014 and find that the transmission of local and global trade shocks is fundamentally different. We show that the trade persistence coefficient falls from 0.91 to 0.35 when we depart from the existing empirical gravity models that draw inference from the pooled coefficient estimates without controlling for the variation in the unobservable global factors. Thus, our approach escapes the excess trade persistence puzzle and adds to the explanation of the sharp decline and the rapid recovery of the global trade flows during the "Great Trade Collapse" of 2008-09. In addition to the traditional variables in the gravity equation, we also show that a cross-country habit asymmetry creates bilateral and multilateral trade imbalances, which are an important determinant of bilateral trade flows both theoretically and empirically.

Keywords

Dynamic Gravity Equation, Habits, Trade Persistence, Trade Imbalance, Global Shocks, Parameter Heterogeneity

JEL Classification

C23, F14, F41, F62

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

The Centre for Applied Macroeconomic Analysis in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

What Explains Excess Trade Persistence? A Theory of Habits in the Supply Chains^{*}

Mariarosaria Comunale[†]

Bank of Lithuania, Vilnius University, and Australian National University

Justas Dainauskas[‡] London School of Economics and Political Science

Povilas Lastauskas[§] Bank of Lithuania and Vilnius University

Abstract

International trade flows are volatile, imbalanced, and fragmented across off-shored supply chains. Yet, not much is known about the mechanism through which trade flows adjust in response to shocks over time. This paper derives a dynamic gravity equation from a theory of habits in the supply chains that generates autocorrelated bilateral trade flows that are heterogeneous across different country pairs. We estimate our version of the dynamic gravity equation for 39 countries over the period of 1950-2014 and find that the transmission of local and global trade shocks is fundamentally different. We show that the trade persistence coefficient falls from 0.91 to 0.35 when we depart from the existing empirical gravity models that draw inference from the pooled coefficient estimates without controlling for the variation in the unobservable global factors. Thus, our approach escapes the excess trade persistence puzzle and adds to the explanation of the sharp decline and the rapid recovery of the global trade flows during the "Great Trade Collapse" of 2008-09. In addition to the traditional variables in the gravity equation, we also show that a cross-country habit asymmetry creates bilateral and multilateral trade imbalances, which are an important determinant of bilateral trade flows both theoretically and empirically.

JEL Classification: C23, F14, F41, F62.

Keywords: Dynamic Gravity Equation; Habits; Trade Persistence; Trade Imbalance; Global Shocks; Parameter Heterogeneity.

^{*}For useful comments and suggestions, we thank Dennis Novy, Vanessa Gunnella, Milda Norkutė, Ziran (Josh) Ding, and other participants at the European Central Bank's Trade Expert Network Workshop as well as the Bank of Lithuania seminar series. For their kind help with the data and/or estimation, we thank Jan Ditzen, Michele Mancini, and Ingrida Šimkutė. The views expressed in this paper are those of the authors and do not necessarily represent the official views of the Bank of Lithuania or the Europystem.

[†]Address at: mcomunale@lb.lt, Applied Macroeconomic Research Unit (TMTS), Bank of Lithuania, Totorių 4, LT-01121, Vilnius, Lithuania; Faculty of Economics and Business Administration, Vilnius University, Saulėtekio al. 9, 2nd Building, LT-10222 Vilnius, Lithuania; Center for Applied Macroeconomic Analysis (CAMA), Australian National University, Canberra, Australia

[‡]Address at: j.dainauskas@lse.ac.uk, Department of Economics, London School of Economics and Political Science, 32 Lincoln's Inn Fields, London WC2A 3PH, United Kingdom.

[§]Address at: p.lastauskas@trinity.cantab.net, CEFER, Bank of Lithuania, Totorių 4, LT-01121, Vilnius, Lithuania; Faculty of Economics and Business Administration, Vilnius University, Saulėtekio al. 9, 2nd Building, LT-10222 Vilnius, Lithuania

1 Introduction

It takes time and resources to implement trade liberalization policies. That is why the patterns of "who trades with whom" are regionally-biased and slow to adjust (Eichengreen and Irwin (1998)). But at the same time, the value of "how much is traded" among those who partner up is surprisingly volatile, especially when countries are hit by global shocks. One such example is the "Great Trade Collapse" (GTC) of 2008-09 during which the world GDP shrank by 1%, while the value of global trade flows slumped by some 10% in a remarkably synchronized fashion across the world (Alessandria et al. (2010)). Another more recent example is the worldwide disruption to trade flows caused by the outbreak of the COVID-19 pandemic. As things stand, the bulk of the modern trade literature relies on the ubiquitous gravity equation to predict the value of trade flows across countries. And it is notoriously successful at predicting both "who trades with whom" as well as "how much is traded" when trade shocks are local or country-specific. But when trade shocks are global, the observed value of trade flows adjusts by more and more rapidly than predicted by the standard gravity equation. We call this discrepancy the "excess trade persistence" puzzle.

This paper derives a dynamic gravity equation from a theory of habits in supply chains. Habits offer a simple framework to capture the dynamics of the global trade network in reduced form, where the production of the final goods requires intermediate imports dispersed across space. It appeals to the inter-temporal frictions on the globalized production belt line, such as assembling, disbanding, or swapping foreign suppliers in response to shocks. Our theory offers several advantages. First, habits predict autocorrelated trade flows, where the trade persistence coefficient is heterogeneous across different country pairs. Second, cross-country habit asymmetry creates differences in homebias. This causes trade imbalance to drive the value of bilateral trade flows in addition to standard measures, such as aggregate income and geographic distance. Third, habits enhance the geographic distance component of trade costs, because distance applies not only to goods that are "made here, sold there", but also to intermediate inputs that are "bought, sold, and bought again". Fourth, habits create "inward" and "outward" multilateral trade resistance that is not only time-varying, but also enters the dynamic gravity equation in contemporaneous and lagged form. This leads to a fundamentally different transmission of local and global trade shocks, since multilateral trade resistance terms are strongly correlated with foreign demand, foreign supply, as well as trade imbalance, and capture the variation in the unobservable global factors.

Motivated by our theoretical model, we further establish two distinct empirical causes of the excess trade persistence puzzle. First, the prevalent methods of estimating (dynamic) gravity equations do not appropriately account for the global trade shocks. For instance, the standard "country" fixed effects are time-invariant, while the "time" fixed effects are homogeneous for all country pairs. This implies that shocks originating from third countries are not fully reflected in either the source or the destination economies. Second, the inference is commonly drawn from the pooled gravity equation coefficient estimates, which ignores the fact that trade flows between some country pairs are significantly more persistent than others. Contrary to the antecedents, we exploit the relatively large temporal dimension of our panel and retain the cross-country parameter heterogeneity. We also explicitly account for the variation in the unobservable global factors by modeling the multilateral trade resistance terms empirically as the cross-sectional averages of all country-specific regressors. Our results show that absent of the unobservable global factors, the value of the pooled

trade persistence coefficient is 0.91, which is comparable to the existing estimates in the literature. However, this estimate is biased upwards almost three-fold relative to our benchmark model specification that retains the cross-country parameter heterogeneity and introduces the unobservable global factors (i.e., the average trade persistence is 0.35). If we expend the unobservable global factors, but retain parameter heterogeneity, the cross-country average trade persistence coefficient nonetheless shrinks to 0.55. This provides strong evidence in favor of a modern trade theory that predicts heterogeneous trade persistence across different country pairs, such as our proposed framework of habits in the off-shored supply chains.

Despite considerable efforts, not much is still known about the mechanism through which the value of trade flows adjusts in response to either local or global trade shocks over time. The standard gravity equation due to Anderson (1979), Anderson and van Wincoop (2003), and Feenstra (2016) remains the workhorse framework for trade policy analysis in the context of permanent, unilateral, and exogenous trade shocks. But the standard gravity equation is static and it is silent about the transitional dynamics. Several others extend the gravity equation into a dynamic setting using the neo-classical theory of capital accumulation (e.g., Yotov and Olivero (2012); Alvarez (2017); and Anderson et al. (2020)). The neo-classical theory suggests that the trade persistence coefficient corresponds to the annual share of undepreciated capital stock. Yet the empirical estimates of the capital depreciation rate suggest that it is mostly homogeneous across countries and equals around 10% (see IMF (2015)). The neo-classical theory therefore predicts high and homogeneous trade persistence, which is consistent with our pooled estimates absent of global factors, but inconsistent with the sharp decline and the rapid recovery of the global trade flows observed during the GTC.

While habits are not rooted in the first principles as strongly as the process of capital accumulation, they are a widely-established tool of characterizing dynamic properties of fundamentals in the macro-finance literature (e.g., Abel (1990); Campbell and Cochrane (1999); Ravn et al. (2006, 2007); and Herbst and Schorfheide (2016)). Admittedly, capital accumulation plays a role in the persistence of virtually all macroeconomic fundamentals. But our theory of habits in the supply chains admits a much more flexible domain for the trade persistence coefficient that is unrestricted by the capital depreciation rate. The habit framework also nests the static gravity equation à la Anderson and van Wincoop (2003) as a special case when all bilateral habits are infinitesimally weak. And when habits are strong for any given country pair, they cause less volatile and more persistent bilateral trade flows that are consistent with Anderson et al. (2020). Consequently, the theory of habits in the supply chains delivers an intermediate degree of trade persistence relative to the static and the neo-classical gravity equations, such that consistent with the data, it predicts sharp and heterogeneous international trade flow adjustments in response to global shocks.

This paper is related to several other strands of the international macroeconomics and the modern trade literature. First, numerous contributions examine the causes and consequences of the GTC, such as Alessandria et al. (2010); Bems et al. (2010); Altomonte et al. (2012); Antonakakis (2012); Levchenko et al. (2010); Eaton et al. (2016); Novy and Taylor (2020), and others. Second, the analysis of the bilateral trade persistence goes back to Eichengreen and Irwin (1998). But we develop an economic theory to support the dynamic nature of the gravity equation. Third, a number of studies explore the persistence of trade costs during the period of hyper-globalisation (e.g., Anderson and van Wincoop (2004); Disdier and Head (2008); Zwinkels and Beugelsdijk (2010);

Head and Mayer (2014)). Trade cost persistence is related to the dynamic properties of the multilateral trade resistance portrayed in our model and aligns with the stylized facts established by Baldwin and Taglioni (2006). Fourth, Serlenga and Shin (2007) were the first to explore the role of contemporaneous unobservable global factors in the context of the gravity equation. But our empirical modeling of both the contemporaneous and lagged unobservable global factors is motivated by the theory of habits in the supply chains. Fourth, a number of studies examine the welfare consequences of mitigating exogenously pre-existing trade imbalances (e.g., Davis and Weinstein (2002); Dekle et al. (2007, 2008)). While we do not discuss the consequences for the welfare gains from trade in this paper, we uncover a theoretical and empirical significance of country-specific trade imbalance as a structural determinant of bilateral trade flows. The importance of trade imbalance on the implications of globalization shocks has recently been emphasized by Dix-Carneiro et al. (2020) whose focus on labor reallocation and unemployment dynamics is distinct from ours.

The rest of this paper is organized as follows. We set out by describing a select-few stylized empirical facts about the global trade flows in Section 2. The first part of Section 3 presents the general equilibrium model of the world economy with habits in the supply chains. The second part of Section 3 derives the dynamic gravity equation. Section 4 describes the data and the setup of several panel regression techniques applied in this paper. We also discuss the different choices related to the empirical modeling of the unobservable global factors. We then present the estimates of the dynamic gravity equation coefficients and contrast them with the results in the existing literature. Section 5 analyzes the inferred extent of the cross-country parameter heterogeneity, its source, and its relationship to the proposed theoretical model. Finally, Section 6 summarizes and concludes.

2 Motivation

The focus of our analysis is on the interaction between international trade flows and the global business cycle. To that end, figure 1 visualizes the patterns of international trade among several major country groups before, during, and after the Great Recession. Specifically, it depicts the export value indices over the period of 2000-2014 for the global economy, the US, the EU, and other selected groups of countries, such as Brazil, Russia, India and China (abbreviated as BRICS), the group of seven (G7), and a cohort of other emerging and developing countries. There are three stylized facts that stand out the most. First, independently of the country group, the figure demonstrates a sharp and synchronized global decline in the value of international trade in response to the global financial crisis of the 2008-2009. This time period is famously coined as the "Great Trade Collapse" by Alessandria et al. (2010) (see the shaded area of figure 1). Second, the recovery from the GTC is remarkably heterogeneous. In particular, the BRICS recovered most rapidly followed by other emerging and developing countries, leaving the EU and G7 well behind. Third, the value of international trade is substantially more volatile than aggregate income, which has declined by only around 1% globally during this time (not displayed). This indicates relatively low persistence in the value of international trade, particularly in response to global shocks, thereby illustrating the essence of what we call the "excess trade persistence" puzzle.

Despite their prominence, these empirical stylized facts are not yet assimilated into the modern





Notes: The figure depicts the export value indices over the period of 2000-2014 for the global economy, the US, the EU, and other selected groups of countries, such as Brazil, Russia, India and China (abbreviated as BRICS), the group of seven (G7), and a cohort of other emerging and developing countries. There are three stylized facts that stand out the most. The reference year is 2005 when the index value is equal to 100. The shaded area represents a time period known as the "Great Trade Collapse" (GTC).

trade theories that give rise to the so-called "gravity equation". On the one hand, the standard gravity equation due to Anderson (1979), Anderson and van Wincoop (2003), and Feenstra (2016) is static and silent about the transitional dynamics. On the other hand, the existing dynamic extensions of the gravity equation based on the neo-classical theory of capital accumulation, such as Yotov and Olivero (2012), Alvarez (2017), or Anderson et al. (2020)), struggle to explain large, sharp, and heterogeneous adjustments of international trade flows observed in the data. Moreover, the existing estimates of (dynamic) gravity equations predominantly rely on methodologies that do not appropriately account for the global trade shocks and neglect the fact that trade flows between some country pairs are significantly more persistent than others (e.g., Egger (2000); Micco et al. (2003); Helpman et al. (2008); and Feenstra (2016)). Specifically, the application of "country" and "time" fixed effects and drawing inference from the pooled gravity coefficient estimates is a remarkably common practice. But this "run-of-the-mill" approach implicitly assumes homogeneous cross-country transitional dynamics despite their inherent structural and institutional differences. And it also implies that despite their influence on the value of bilateral trade flows, shocks originating from third countries are not fully reflected in either the source or the destination economies. Another feature of the gravity equation that is often overlooked both in theory and in practice is the prevalence of the global trade imbalances. Though country-specific trade imbalances shrunk in the aftermath of the GTC, they remain far from being perfectly balanced and constitute an important conduit of local and global trade tensions (IMF (2019); Beirne et al. (2020)).

We proceed with a detailed description of a theory of habits in the supply chains. The theory provides a justification for our proposed augmentations of the empirical gravity model. Namely, (i) retaining parameter heterogeneity when drawing inference; (ii) incorporating unobservable global factors; (iii) and adding country-specific trade imbalance as a determinant of bilateral trade flows in addition to the standard measures, such as aggregate income and geographic distance.

3 Theoretical Model

Consider a world economy evolving over discrete time t = 0, 1, 2, ... that comprises of a finite number of countries indexed by $i, j \in n = \{1, 2, ..., N\}$. Each country is populated by two types of interacting agents: consumers and producers. The producers in each country operate in two different sectors: wholesale and distribution. The wholesale sector is populated by a unit mass of firms indexed by $\omega \in [0, 1]$. All economies are open to trade wholesale varieties with one another, but bilateral trade flows are dampened at the intensive margin by Samuelson's 'iceberg costs'.¹ Once the tradable goods arrive at the docks of each destination, the distributor merges the imported and domestically-produced wholesale varieties into a composite good. The consumers can only purchase the composite good and supply an inelastic fraction of their time endowment as labor to the wholesale firms. There is no entry or exit of the wholesale firms, but the production technology of the distributor is subject to habits based on the country-specific volume of exports. In equilibrium, multilateral trade imbalance arises when different country pairs are subject to asymmetrical homebias in consumption measured by the import penetration ratios. In turn, home-bias asymmetry

¹Iceberg costs is the catch-all bilateral trade resistance term, which subsumes both tariff and non-tariff barriers to trade, including the exogenously determined geographic distance.

arises when different country pairs are subject to heterogeneous habits of the distributor and/or iceberg costs.

3.1 Supply Side

Wholesale varieties are imperfectly substitutable and produced using linear labor-intensive technology: $m_{ij,t}(\omega) = z_{i,t}h_{ij,t}(\omega)$, where $i, j \in n$. Aggregate labor productivity in country *i* denoted as $z_{i,t}$ is covariance-stationary and exogenously given. The hours of labor spent by workers domiciled in source country *i* to produce varieties that are sold in destination *j* are denoted as $h_{ij,t}(\omega)$. Delivering one unit of the wholesale variety from the source country *i* to the destination *j* costs $d_{ij} - 1 > 0$ relative to the unit costs (i.e., iceberg cost).²

Once the wholesale varieties arrive at the destination country, the local distributor aggregates them into an infinitely-divisible composite good according to the constant elasticity of substitution (CES) production technology augmented by multiplicative habits:

$$x_{ij,t} = \left[\int_{0}^{1} \left(m_{ij,t}(\omega)x_{ij,t-1}^{\chi_{ij}}\right)^{1-1/\eta} d\omega\right]^{1/(1-1/\eta)},$$
(3.1)

where $\eta > 1$ is the elasticity of substitution, $\chi_{ij} > 0$ denotes the habit intensity, and $x_{ij,t-1}$ is the stock of habit.³ Habits offer a simple framework to capture the dynamics of the global trade network in reduced form, where the production of the final goods requires intermediate imports $m_{ij,t}(\omega)$ dispersed across space. It appeals to the inter-temporal frictions on the globalized production belt line, such as assembling, disbanding, or swapping foreign suppliers in response to shocks. Because $\partial \ln x_{ij,t}/\partial \ln x_{ij,t-1} = \chi_{ij}$, habits cause trade flow adjustments to be gradually decaying, permanent, or explosive in response to shocks when $\chi_{ij} \in (0, 1)$, $\chi_{ij} = 1$, or $\chi_{ij} > 1$, respectively. And when habits are infinitesimally weak, such that $\chi_{ij} \to 0$, trade flows are static as per usual.

Let $P_{ij,t}(\omega)$ denote the price of variety ω that is produced in economy *i* and sold in destination j at time *t*. The distributor chooses the amount of wholesale varieties to purchase $m_{ij,t}$ by minimizing the total expenditure on intermediate imports $\tilde{P}_{ij,t}x_{ij,t} - \int_0^1 P_{ij,t}(\omega)m_{ij,t}(\omega)d\omega$ subject to the augmented CES preferences in equation (3.1). The first-order condition with respect to $m_{ij,t}(\omega)$ gives rise to the following optimal demand schedule for wholesale varieties:

$$m_{ij,t}(\omega) = x_{ij,t} x_{ij,t-1}^{\chi_{ij}(\eta-1)} \left[\frac{P_{ij,t}(\omega)}{\tilde{P}_{ij,t}} \right]^{-\eta}.$$
(3.2)

²By assumption, the model maintains the triangular equation at all times, namely, $d_{ij} \leq d_{i\iota}d_{\iota j}$ for all $i, j, \iota \in n$, such that direct shipment of merchandise is always the least expensive route.

³The existing literature provides several ways of modeling habits, sometimes referred to as "catching up with the Joneses". In macro-finance, the stock of habit enters the lifetime utility of the consumer as a function of past consumption, which introduces richer autocorrelation structure and improves the model-implied fit of the observed data (Abel (1990); Campbell and Cochrane (1999); Herbst and Schorfheide (2016)). In closed and open economy macroeconomics, the stock of habit enters the CES preferences as a function of past consumption of individual varieties (i.e., "deep habits"), which generates counter-cyclical mark-up adjustments (Ravn et al. (2006, 2007)). In this paper, the stock of habit enters the CES production technology, which is dual to the CES preferences, but the stock of habit is aggregate and independent of individual varieties of intermediate imports. An interesting extension that is not considered in this paper is to incorporate "deep habits" when firms are subject to idiosyncratic productivity shocks à la Melitz (2003). But the gravity equation is an aggregate relationship and the aggregate stock of habit is sufficient to generate autocorrelated bilateral trade flows specific to each country pair that is of key interest in this paper.

The demand for intermediate imports is increasing in the contemporaneous stock of the composite good $x_{ij,t}$ and, since $\chi_{ij} > 0$, it is increasing in the stock purchased in the previous period $x_{ij,t-1}$, but decreasing in the relative price of that variety $P_{ij,t}(\omega)/\tilde{P}_{ij,t}$.

The wholesalers are monopolistically-competitive. They recognize the demand schedule of the distributors in each destination when setting the optimal price for their variety consistent with the strategy of "pricing-to-habits" (Ravn et al. (2007)).⁴ The optimal wholesale price is derived by maximizing the nominal profits $P_{ij,t}(\omega)m_{ij,t}(\omega) - d_{ij}MC_{i,t}m_{ij,t}(\omega)$, subject to the demand schedule in equation (3.2), which gives rise to the standard expression for a fixed price-cost margin:

$$P_{ij,t}(\omega) = \left(\frac{\eta}{\eta - 1}\right) d_{ij} M C_{i,t}, \qquad (3.3)$$

where $MC_{i,t}$ are the unit costs of producing the wholesale variety. In the absence of idiosyncratic labor productivity shocks in the wholesale production technology, the unit costs for all wholesalers are homogeneous and equivalent to the nominal hourly wage rate in the source country $W_{i,t}$, normalized by the aggregate labor productivity $z_{i,t}$, such that $MC_{i,t} = W_{i,t}/z_{i,t}$.

3.2 Demand Side

Each destination $j \in n$ is populated by a representative consumer characterized by CES preferences over consumption of composite goods originating from each source country:

$$c_{j,t} = \left[\sum_{i=1}^{N} x_{ij,t}^{1-1/\eta}\right]^{1/(1-1/\eta)}.$$
(3.4)

The representative consumer chooses the amount of goods to purchase from any trade partner by minimizing the total expenditure $P_{j,t}c_{j,t} - \sum_{i=1}^{N} \tilde{P}_{ij,t}x_{ij,t}$, subject to the CES preferences, which gives rise to the optimal demand schedule for composite goods:

$$x_{ij,t} = c_{j,t} \left(\frac{\tilde{P}_{ij,t}}{P_{j,t}}\right)^{-\eta}, \qquad (3.5)$$

where $\tilde{P}_{ij,t}$ is the price level at which the distributor in destination j breaks-even, whereas $P_{j,t}$ denotes the aggregate consumer price index. The demand for composite goods is increasing in the aggregate consumption of the destination country $c_{j,t}$, but decreasing in the relative price of composite goods originating from each source country $\tilde{P}_{ij,t}/P_{j,t}$.

Suppose the representative consumer derives utility from the consumption of an infinitelydivisible basket of composite goods $c_{j,t}$, while the aggregate hours of labor are supplied inelastically. The lifetime utility of the representative consumer is therefore given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \ln\left(c_{j,t}\right),\tag{3.6}$$

where \mathbb{E}_t is the rational expectations operator and $\beta \in (0, 1)$ stands for the time preference para-

⁴When the stock of habit $x_{ij,t-1}$ is independent of individual varieties, as is our proposed setting, "pricing-tohabits" implies local currency pricing (LCP). But the choice of invoicing currency does not play an important role in our model, since all exchange rate movements are subsumed in the multilateral resistance terms of the gravity equation.

meter. The representative consumer is subject to an indefinite sequence of budget constraints:

$$c_{j,t} + \mathbb{E}_t[\zeta_{j,t,t+1}b_{j,t+1}] \le b_{j,t} + w_{j,t}h_j + \varpi_{j,t},$$
(3.7)

where $b_{j,t}$ is the net real stock of internationally-traded one-period bonds, $\zeta_{j,t,t+1}$ is the real price of the one-period bond, $w_{j,t} = W_{j,t}/P_{j,t}$ is the real hourly wage rate, h_j are the inelastically supplied hours of labor relative to the total endowment of time, and $\varpi_{j,t}$ is the real aggregate profit dividend. The profit dividend forms part of the representative household income, since they own the wholesale firms that are domiciled in their country. Any supernormal profits or losses that are accrued by the domestic wholesale firms are therefore transferred to the representative consumer.

When the national stock of bonds is in zero net supply (i.e. $b_{i,t} = 0$), the aggregate consumption in each source country equals the wage bill and the profit dividends. However, if the aggregate stock of bonds is allowed to be positive (negative), then aggregate consumption would increase (decrease) in the contemporaneous stock of bonds, but decrease (increase) in the stock of bonds held until maturity in the next period. As a consequence, international borrowing and lending through such implicitly complete financial market structure generates consumption smoothing behavior and ultimately allows for short-run and long-run multilateral trade imbalance to arise. Whether or not the multilateral trade imbalance is explosive over time depends entirely on the time preference parameter β . If β is bounded between zero and unity, as is usually the case, then it can be shown that the transversality condition holds, since the long-run real rate of interest is strictly non-negative. And if so, then the long-run multilateral trade imbalance is equivalent to the present discounted value of the trade balance, which is constant and finite in the steady state (see Appendix A.5). Any country $j \in n$ can therefore sustain a current account deficit perpetually without violating the transversality condition, but only if it is a sufficiently large net creditor to begin with.

The representative consumer chooses the aggregate consumption $c_{j,t}$ and the aggregate stock of bonds $b_{j,t}$ by maximizing their lifetime utility (3.6), subject to an indefinite sequence of budget constraints (3.7), taking the aggregate profit dividend $\varpi_{j,t}$, the real price of one-period bonds $\zeta_{j,t,t+1}$, and the real wage bill $w_{j,t}h_j$ as given. The first-order conditions give rise to the standard Euler equation and the perfect consumption risk sharing relationship, respectively:

$$1 = \beta \mathbb{E}_t \left[\frac{c_{j,t}}{\zeta_{j,t,t+1} c_{j,t+1}} \right], \tag{3.8}$$

$$q_{ij,t} = \frac{Q_{ij,t}P_{i,t}}{P_{j,t}} = \frac{c_{j,t}}{c_{i,t}}.$$
(3.9)

The Euler equation (3.8) captures the tendency for consumers to smooth consumption expenditure over time relative to the contemporaneous stream of income by saving or borrowing against the expected future income. The complete financial market structure and additively separable preferences give rise to the standard consumption risk-sharing relationship à la Backus and Smith (1993) as shown in equation (3.9). It states that absent of arbitrage opportunities, a rise in the marginal utility of consumption in the source country *i* relative to the destination *j* is followed by an identical depreciation in their real bilateral exchange rate in percentage terms. The steady state of the real exchange rate is not equal to unity when the stock of internationally traded bonds are not in zero net supply. To elaborate, the real exchange rate in this model is covariance-stationary, such that $q_{ij,t} = q_{ij} = c_j/c_i$ in the long-run. As a consequence, if the source country *i* was subject to a long-run trade deficit against the destination country *j*, such that $c_i > c_j$, then *ceteris paribus* the value of the real exchange rate in the long-run steady state must be less than unity and vice versa.

3.3 General Equilibrium

General equilibrium is a set of dynamic processes characterizing a unique set of state and control variables given by $\{b_{i,t+1}, \zeta_{i,t,t+1}, x_{ij,t}, m_{ij,t}, p_{ij,t}, \tilde{p}_{ij,t}, w_{i,t}, \varpi_{i,t}, z_{i,t}, c_{i,t}, q_{ij,t}, \pi_{ij,t}, \xi_{i,t}, \Xi_{i,t}\}_{t=0}^{\infty}$ for all $i, j \in n$ that are consistent with the utility-maximizing behavior of the representative household and the profit-maximizing behavior of the representative firm. The dynamic processes are conditional on pre-determined variables $\{b_{i,t}, z_{i,t-1}, x_{ij,t-1}\}_{t=0}^{\infty}$, labor productivity shocks $\{\epsilon_{i,t}\}_{t=0}^{\infty}$, and fixed parameters $d_{ij}, \chi_{ij}, \rho_i, \sigma_i, \mu_i, h_i, \eta, \beta$. Due to the standard price level indeterminacy, general equilibrium is defined in terms of relative prices $p_{ij,t} = P_{ij,t}/P_{j,t}$ and $\tilde{p}_{ij,t} = \tilde{P}_{ij,t}/P_{j,t}$.

The relative prices are independent of ω in equilibrium, because all wholesale firms are subject to country-specific uncertainty associated with aggregate labor productivity over time, but they do not face any idiosyncratic risk. As a consequence, the general equilibrium is symmetric, such that wholesale prices are homogeneous across firms for any given source country and they are set as a constant mark-up over unit costs (see equation (3.3)), which implies not only $P_{ij,t}(\omega) = P_{ij,t}$ and $m_{ij,t}(\omega) = m_{ij,t}$ for all $\omega \in [0, 1]$, but also $h_i = \sum_{j=1}^{N} h_{ij,t}$. The equilibrium supply of intermediate goods for each wholesale variety is thus perfectly price inelastic in each source country. But as long as $\chi_{ij} > 0$, the break-even price index of the distributor $\tilde{P}_{ij,t}$ does not correspond to the aggregate wholesale price index $P_{ij,t}$, because it is influenced by the stock of habit:

$$\tilde{P}_{ij,t} = x_{ij,t-1}^{-\chi_{ij}} P_{ij,t}.$$
(3.10)

Similarly, the duality problem gives rise to the aggregate consumer price index as a function of the break-even price indices from each source country:

$$P_{j,t} = \left[\sum_{i=1}^{N} \tilde{P}_{ij,t}^{1-\eta}\right]^{1/(1-\eta)}.$$
(3.11)

Observe that $\lim_{\chi_{ij}\to 0} \tilde{P}_{ij,t} = P_{ij,t}$, whereas more generally $P_{ij,t}/\tilde{P}_{ij,t}$ is not equal to unity, such that the aggregate stock of intermediate goods in equilibrium is increasing (decreasing) in the current (past) stock of the composite good:

$$m_{ij,t} = x_{ij,t} x_{ij,t-1}^{-\chi_{ij}}.$$
(3.12)

The multilateral trade flows can therefore be expressed as a geometrically decaying function of the demand for intermediate goods in the past:

$$x_{ij,t} = m_{ij,t} x_{ij,t-1}^{\chi_{ij}} = \prod_{s=0}^{\infty} m_{ij,t-s}^{\chi_{ij}^s},$$
(3.13)

thereby introducing the central mechanism through which dynamic gravity effects are sustained even in the symmetric general equilibrium. Let $X_{ij,t} = P_{ij,t}x_{ij,t}$, $C_{j,t} = P_{j,t}c_{j,t}$, and $Y_{j,t} = \bar{P}_{j,t}y_{j,t}$ denote the aggregate nominal value of trade flows, consumption, and output, respectively. Notice that absent of government expenditure and capital formation, the deflator pertaining to the gross domestic product, namely $\bar{P}_{j,t}$, is identical to the consumer price index derived from the duality problem only if trade is balanced at all times. The aggregate income of each country therefore amounts to its global sales of composite goods to each destination in each time period measured in domestic currency units:

$$Y_{j,t} = \sum_{i=1}^{N} Q_{ij,t} X_{ji,t}.$$
(3.14)

By contrast, the aggregate consumption in each economy is equal to its global expenditure on composite goods at any given time period:

$$C_{j,t} = \sum_{i=1}^{N} X_{ij,t}.$$
(3.15)

In this model, the difference between aggregate income and aggregate consumption defines the multilateral trade balance:

$$NX_{j,t} = Y_{j,t} - C_{j,t} = \sum_{i=1}^{N} Q_{ij,t} X_{ji,t} - \sum_{i=1}^{N} X_{ij,t}, \qquad (3.16)$$

such that aggregate consumption is proportional to aggregate income:

$$C_{j,t} = Y_{j,t} \Xi_{j,t}, \tag{3.17}$$

where $\Xi_{j,t} = 1/(1 + \sum_{i=1}^{N} q_{ij,t}\pi_{ji,t} - \sum_{i=1}^{N} \pi_{ij,t}) = 1/(1 + \xi_{j,t})$ captures multilateral trade imbalance relative to the total consumption expenditure and $\pi_{ij,t} = X_{ij,t}/C_{j,t}$ is defined as the import penetration ratio. The trade imbalance term is strictly non-negative $\Xi_{j,t} > 0$ at each time period, since $\xi_{j,t} = NX_{j,t}/C_{j,t} \in (-1,\infty)$ is the ratio of net exports to aggregate consumption expenditure.

In the related gravity literature, $\Xi_{j,t}$ is traditionally assumed to be equal to unity or simply exogenously given. But instead of implicitly ruling out the possibility of multilateral trade imbalance or assuming that they prevail *ad hoc*, our model provides a theoretical justification for their existence in both the short-run and the long-run. The model remains consistent with the accounting identity in which the import penetration ratios associated with each source country sum up to unity in the destination country, such that $\sum_{i=1}^{N} \pi_{ij,t} = 1$. Hence, an increase in the consumption of foreign goods in relative terms (i.e., a rise in $\pi_{ij,t}$ for any $i \in n$) implies a decline in the consumption of domestic goods in relative terms (i.e., a fall in $\pi_{jj,t}$). Consistent with Obstfeld and Rogoff (1996), international trade in this model is one of the forces through which local or country-specific labor productivity shocks are deflected in the short-run by saving or borrowing against the permanent income. While a negative labor productivity shock at home leads to a decline in aggregate consumption in the special case of autarky, some of the loss in productivity in open economies can be substituted with foreign production by running a transitory trade deficit.

When the influence of labor productivity shocks fades away in the long-run, and the economy reverts back to the steady state, the term $\Xi_{j,t}$ would only be equal to unity if there were no struc-

tural heterogeneities across countries (i.e., symmetric steady state). And since consumers supply labor to the wholesale firms inelastically, there can only be two dimensions along which structural heterogeneities distort the import penetration ratios across countries: (i) tariff and non-tariff barriers to trade subsumed within the iceberg costs d_{ij} ; and (ii) technological import dependence encapsulated by the stock of habits in the production technology $x_{ij,t-1}^{\chi_{ij}}$, the differences of which are driven by the parameter characterizing the habits of the distributor χ_{ij} . If economies differ in either of these two dimensions, then trade is not balanced in the long-run and the term $\Xi_{j,t}$ is endogenously and indefinitely shifted away from unity.

Formally, when all exporters adopt LCP strategies, import and export prices are proportional, such that $P_{ij,t} = (\eta/(\eta - 1))d_{ij}MC_{i,t} = d_{ij}P_{ii,t}$. In turn, the break-even price index is $\tilde{P}_{ij,t} = d_{ij}P_{ii,t}x_{ij,t-1}^{\chi_{ij}}$, such that the import penetration ratio is given by $\pi_{ij,t} = (\tilde{P}_{ij,t}/P_{j,t})^{1-\eta} = (d_{ij}P_{ii,t}x_{ij,t-1}^{\chi_{ij}}/P_{j,t})^{1-\eta}$. Although the numerator $P_{j,t}$ itself is also influenced by both d_{ij} and $x_{ij,t-1}^{\chi_{ij}}$, the consumer price index is a function of trade costs and habits across all trade partner countries, while the frictions pertaining to the break-even price index in the denominator $\tilde{P}_{ij,t}$ are bilateral. Consequently, if either habits or trade costs are asymmetric across countries, such that $\chi_{ij} \neq \chi_{ji}$ and/or $d_{ij} \neq d_{ji}$, then $\Xi_{j,t} \neq 1$ in the long-run. And as discussed above, long-run trade imbalances are sustained by a corresponding imbalance in the capital account, which in this model corresponds to a permanent and sustainable inflow or outflow of bonds.⁵

3.4 Dynamic Gravity Equation

We have thus far established a well-defined and internally-consistent general equilibrium in a model of the world economy characterized by habits in the wholesale distribution network. But the general equilibrium itself does not provide an identifiable link between the empirical estimates of international trade flow persistence and the intensity of habits in our model. In the absence of habits in the supply chains, Anderson and van Wincoop (2003) establish the ubiquitous "gravity equation" approach of taking static trade models with homothetic preference aggregators to the data.⁶ The static gravity equation links the bilateral trade flows to aggregate income in the source and destination countries as well as the unobservable bilateral and multilateral trade resistance. This section of the paper shows that habits extend the static gravity equation into a dynamic counterpart in which bilateral trade flows depend on: (i) the lagged bilateral trade flows capturing

⁵When consumers share the risk arising from country-specific shocks, the real exchange rate fluctuations are endogenous to labor productivity shocks as shown in equation (3.9). However, the terms of trade are largely insulated from exchange rate movements, since all exporters are "pricing-to-habits". That said, the export revenue generated in each destination measured in the domestic currency units is given by $Q_{ji}P_{ij,t}x_{ij,t} = d_{ij}Q_{ji}P_{ii,t}x_{ij,t}$ at all times. As a consequence, the endogenous exchange rate fluctuations are absorbed by the aggregate profit dividend $\varpi_{j,t}$ when converting the offshore profits back to the currency units of the source country. When exchange rate volatility is exceedingly high, the aggregate profit dividend under pricing-to-habits may become negative (i.e., losses). In order to ensure strictly positive real profit dividends, the domain of parameter η can be restricted to the close vicinity of unity. In other words, pricing-to-habits would in principle be non-viable if wholesale firms operated in a perfectly competitive market structure characterized by $\eta \to \infty$. An interesting extension that goes beyond the scope of this paper is to consider variable price mark-ups.

⁶In principle, equation (3.13) alone could naively be used to determine the level of bilateral trade flow persistence and to identify parameter χ_{ij} for any country pair. But in a general equilibrium environment, the naive approach does not account for the fact that the stock of intermediate trade flows $m_{ij,t}$ and the stock of composite goods $x_{ij,t}$ are determined simultaneously, which causes the estimates of trade persistence to be biased. We circumvent the problem of simultaneity by replacing the intermediate trade flows entering equation (3.13) with a function of the multilateral trade resistance, which extends the static Anderson and van Wincoop (2003) approach into a dynamic framework.

their heterogeneous persistence across different country pairs; (ii) multilateral trade imbalance in the destination economy; (iii) trade resistance in the form of bilateral geographic distance as well as overall propensity to trade in the source and destination countries; and (iv) the aggregate income in the source and destination countries relative to the world economy as a whole.

Proposition 1. Let $\theta_{i,t} = Y_{i,t}/Y_t$, where $Y_t = \sum_{j=1}^N Y_{j,t}$. Then the share of the source country aggregate income relative to the world income is a function of its export prices and the outward multilateral resistance:

$$\theta_{i,t} = (\Phi_{i,t} P_{ii,t})^{1-\eta}, \tag{3.18}$$

where

$$\Phi_{i,t} = \left[\sum_{j=1}^{N} Q_{ji,t} \theta_{j,t} \Xi_{j,t} \left(\frac{d_{ij} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}}\right)^{1-\eta}\right]^{1/(1-\eta)}$$
(3.19)

Proof. Consider the demand for composite goods $X_{ij,t} = C_{j,t}[\tilde{P}_{ij,t}/P_{j,t}]^{1-\eta}$, the break-even price index $\tilde{P}_{ij,t} = P_{ij,t}x_{ij,t-1}^{-\chi_{ij}}$, and the aggregate consumption identity $C_{j,t} = Y_{j,t}\Xi_{j,t}$. Substitute each of these schedules into the aggregate income identity of the source country $i \in n$ to obtain $Y_{i,t} = \sum_{j=1}^{N} Q_{ji,t}Y_{j,t}\Xi_{j,t}(d_{ij}P_{ii,t}x_{ij,t-1}^{-\chi_{ij}}/P_{j,t})^{1-\eta}$. Solve the above for $P_{ii,t}^{1-\eta}$, which gives $P_{ii,t}^{1-\eta} =$ $Y_{i,t}/[\sum_{j=1}^{N} Q_{ji,t}Y_{j,t}\Xi_{j,t}(d_{ij}x_{ij,t-1}^{-\chi_{ij}}/P_{j,t})^{1-\eta}] = (Y_{i,t}/Y_t)/[\sum_{j=1}^{N} \theta_{j,t}Q_{ji,t}\Xi_{j,t}(d_{ij}x_{ij,t-1}^{-\chi_{ij}}/P_{j,t})^{1-\eta}]$ or simply $P_{ii,t}^{1-\eta} = \theta_{i,t}\Phi_{i,t}^{\eta-1}$, where $\Phi_{i,t}$ measures the outward multilateral resistance and $\theta_{i,t}$ is the share of the source country income in the world economy.

Proposition 2. If the price of imports $P_{ij,t}$ from each source country $i \in n$ are set in local currency terms at each destination $j \in n$, and the production technology of final exported goods exhibits constant returns to scale, then import prices are proportional to the outward multilateral resistance of each source country:

$$P_{ij,t} = \frac{d_{ij}\theta_{i,t}^{1/(1-\eta)}}{\Phi_{i,t}}.$$
(3.20)

Proof. When the production technology of final exports exhibits constant returns to scale, the unit costs of production $MC_{i,t}$ are independent of the trade flows. And if exporters set their prices abroad in local currency terms, then they are proportional to the unit costs of production, namely $P_{ij,t} = (\eta/(\eta-1))d_{ij}MC_{i,t}$ for all $j \in n \setminus i$, since $d_{ii} = 1$, such that $P_{ii,t} = (\eta/(\eta-1))MC_{i,t}$. It follows that import and export prices are proportional to one another: $P_{ij,t} = d_{ij}P_{ii,t}$. And if so, then using Proposition 1 to substitute out the export price gives rise to an expression for import prices as a function of outward multilateral resistance in the source country: $P_{ij,t} = d_{ij}\theta_{i,t}^{1/(1-\eta)}/\Phi_{i,t}$.

Lemma 1. The gravity equation is dynamic when habits are non-zero, such that $\chi_{ij} > 0$ for all $i \in n \setminus j$. And when habits are asymmetric across countries, such that $\chi_{ij} \neq \chi_{ji}$ for all $i \in n \setminus j$, and/or the inward and outward the bilateral iceberg costs are non-identical, such that $d_{ij} \neq d_{ji} > 1$ for all $i \in n \setminus j$, the gravity equation is subject to the multilateral trade imbalance:

$$A_{ij,t} = X_{ij,t} \times \frac{Y_t}{Y_{i,t}Y_{j,t}} = \Xi_{j,t} \left[\frac{d_{ij}^{1+\chi_{ij}}}{\Phi_{i,t}\Phi_{i,t-1}^{\chi_{ij}}P_{j,t}} \right]^{1-\eta} \left[\frac{\theta_{i,t-1}^{-\eta}}{A_{ij,t-1}^{1-\eta}Y_{j,t-1}^{1-\eta}} \right]^{\chi_{ij}}.$$
 (3.21)

Proof. When consumers adopt homothetic preferences with constant elasticity of substitution, the demand for imports is given by $X_{ij,t} = C_{j,t} [\tilde{P}_{ij,t}/P_{j,t}]^{1-\eta}$. The aggregate consumption is proportional to aggregate income, such that $C_{j,t} = Y_{j,t} \Xi_{j,t}$, where $\Xi_{j,t} = 1/(1+\xi_{j,t})$ and $\xi_{j,t} = NX_{j,t}/C_{j,t}$, thus $X_{ij,t} = Y_{j,t} \Xi_{j,t} [\tilde{P}_{ij,t}/P_{j,t}]^{1-\eta}$. Using Proposition 1 to substitute out the break-even price index $\tilde{P}_{ij,t} = P_{ij,t}x_{ij,t-1}^{-\chi_{ij}}$ gives $X_{ij,t} = \Xi_{j,t} \left[d_{ij}x_{ij,t-1}^{-\chi_{ij}}/(\Phi_{i,t}P_{j,t}) \right]^{1-\eta} Y_{i,t}Y_{j,t}/Y_t$. Next, Proposition 2 is used to substitute out the real stock of habits $x_{ij,t-1}$ with the nominal value trade flows given by $X_{ij,t-1} = P_{ij,t-1}x_{ij,t-1}$, such that $x_{ij,t-1} = X_{ij,t-1}\Phi_{i,t}/(d_{ij}\theta_{i,t}^{1/(1-\eta)})$, which implies that $X_{ij,t} = \Xi_{j,t} \left[d_{ij}^{1+\chi_{ij}}/(\Phi_{i,t}\Phi_{i,t-1}^{\chi_{ij}}P_{j,t}) \right]^{1-\eta} \left[\theta_{i,t-1}/X_{ij,t-1}^{1-\eta} \right]^{\chi_{ij}} Y_{i,t}Y_{j,t}/Y_t$. Finally, let $A_{ij,t}$ denote the size-adjusted bilateral trade flows, such that $A_{ij,t} = X_{ij,t}Y_t/(Y_{i,t}Y_{j,t})$. Then it follows that $A_{ij,t} = \Xi_{j,t} \left[d_{ij}^{1+\chi_{ij}}/(\Phi_{i,t}\Phi_{i,t-1}^{\chi_{ij}}P_{j,t}) \right]^{1-\eta} \left[\theta_{i,t-1}/(A_{ij,t-1}^{1-\eta}P_{i,t-1}) \right]^{\chi_{ij}}$ (see Appendix B). \Box

Taking the natural logs on both sides of (3.21) and imposing the identity, $\theta_{i,t} = Y_{i,t}/Y_t$, gives the theoretically-grounded regression model for the dynamic gravity equation:

$$\ln A_{ij,t} = \underbrace{\chi_{ij}(\eta - 1) \ln A_{ij,t-1}}_{\text{size-adjusted bilateral trade flow persistence}} + \underbrace{\ln(\Xi_{j,t})}_{\text{multilateral trade imbalance}} - \underbrace{(1 + \chi_{ij})(\eta - 1) \ln d_{ij} + (\eta - 1) \ln P_{j,t} + (\eta - 1) \ln \Phi_{i,t} + \chi_{ij}(\eta - 1) \ln \Phi_{i,t-1}}_{\text{bilateral and multilateral trade resistance}} + \underbrace{\chi_{ij}\eta \ln Y_{t-1} - \chi_{ij}\eta \ln Y_{i,t-1} + \chi_{ij}(\eta - 1) \ln Y_{j,t-1}}_{\text{aggregate income}}.$$
(3.22)

According to the above dynamic gravity equation, the persistence of the size-adjusted bilateral trade flows $A_{ij,t}$ is increasing in the intensity of the habits in the wholesale distribution network $\chi_{ij} > 0$. Moreover, the size-adjusted bilateral trade flows $A_{ij,t}$ are increasing in the multilateral trade imbalance of the destination economy $\Xi_{j,t}$, decreasing in the iceberg costs d_{ij} , and decreasing in the lagged output of the home economy $Y_{i,t-1}$, but increasing in the multilateral trade resistance (consumer price index) of the destination country $P_{j,t}$, increasing in the current and lagged multilateral trade resistance of the home country $\Phi_{i,t}$ and $\Phi_{i,t-1}$, increasing in the lagged global output Y_{t-1} , and increasing in the lagged output of the destination economy $Y_{j,t-1}$.

Proposition 3. When habits are non-zero, but iceberg costs and habits are symmetrical across countries, such that $d_{ij} = d_{ji}$ and $\chi_{ij} \rightarrow \chi > 0$ for all $i \in n \setminus j$, the gravity equation is dynamic, but all global trade flows are balanced, such that:

$$\lim_{\chi_{ij} \to \chi \forall i \in n \setminus j} \ln A_{ij,t} = \chi(\eta - 1) \ln A_{ij,t-1} - (1 + \chi)(\eta - 1) \ln d_{ij} + (\eta - 1) \ln P_{j,t} + (\eta - 1) \ln \Phi_{i,t} + \chi(\eta - 1) \ln \Phi_{i,t-1} + \chi \eta \ln Y_{t-1} - \chi \eta \ln Y_{i,t-1} + \chi(\eta - 1) \ln Y_{j,t-1},$$
(3.23)

since $\lim_{\chi_{ij}\to\chi\,\forall\,i\in n\setminus j}\Xi_{j,t}=1$ under the assumption that $d_{ij}=d_{ji}$.

Proposition 4. When habits are infinitesimally weak, such that $\chi_{ij} \to 0$ for all $i \in n \setminus j$, the

gravity equation is static à la Anderson and van Wincoop (2003):

$$\lim_{\chi_{ij} \to 0 \,\forall \, i \in n \setminus j} \ln A_{ij,t} = (1 - \eta) \left[\ln d_{ij} - \ln \Phi_{i,t} - \ln P_{j,t} \right], \tag{3.24}$$

since $\lim_{\chi_{ij}\to 0 \forall i \in n \setminus j} \Xi_{j,t} = 1$ assuming that iceberg costs are symmetrical, such that $d_{ij} = d_{ji}$, which implies that $\lim_{\chi_{ij}\to 0 \forall i \in n \setminus j} \Phi_{i,t} = P_{i,t}$.

Observe that the multilateral trade resistance terms are theoretically equivalent to the consumer price index in the static, but not the dynamic, gravity model. This occurs even if the trade costs are symmetric both ways, because trade imbalance introduces a wedge between aggregate consumption and income, which transpires into the outward multilateral resistance, such that the aggregate export revenue need not equal the aggregate import expenditure at any point in time. That said, the inward and outward multilateral resistance terms, namely $P_{j,t}$ and $\Phi_{i,t}$, are nonetheless dual to one another (see Proposition 1). This means that $P_{j,t}$ and $\Phi_{i,t}$ are defined up to a single normalization in the dynamic gravity equation, just as they are defined in the static gravity equation.

There are several notable differences between the standard static gravity equation and the dynamic gravity equation augmented with habits in the supply chains. First, habits predict positively autocorrelated trade flows, where the trade persistence coefficient is heterogeneous across different country pairs, since $\chi_{ij} > 0$. This implies that even if the world economy is in a recession due to a fall in Y_t , country pairs subject to habits bounded between zero and unity $\chi_{ij} \in (0,1)$ experience a less pronounced decline in their bilateral trade flows, compared to a static specification in which habits are absent (i.e., $\chi_{ij} \to 0$). However, as habit intensity approaches unity, such that $\chi_{ij} \to 1$, the more resilient trade flows become to trade shocks. Second, bilateral trade flows from source country i to destination j are increasing in the multilateral trade imbalance in the destination economy j. As such, a rise in the trade surplus of the destination economy j, ceteris paribus, increases trade inflows from all source countries i. This novel feature of the dynamic gravity equation is consistent with the consumption smoothing mechanism in our model, since the multilateral trade imbalance arises when the destination country leverages its permanent income against all local labor productivity shocks. Third, habits enhance the geographic distance component of trade costs, because distance applies not only to goods that are "made here, sold there", but also to intermediate inputs that are "bought, sold, and bought again". This is reflected in a coefficient next to the iceberg costs that is increasing in the habits $\chi_{ij} > 0$ and hints at disproportionate welfare gains from trade in disguise. Fourth, habits create "inward" and "outward" multilateral trade resistance that is not only time-varying, but also enters the dynamic gravity equation in contemporaneous and lagged form. This leads to a fundamentally different transmission of local and global trade shocks, the details of which are discussed in Section 4.2 of this paper.

4 Empirical Analysis

In this section of the paper, we describe the data and the methodology used to estimate the dynamic gravity equation (3.22). Due to the richness of the regression model derived directly from the theoretical model, which admits a lagged dependent variable, time-invariant heterogeneity, as well as time-varying unobservable factors, we recognize several panel data techniques that are

Table 1: Data Description

Variable	Data	Description	Measurement Units
$ \begin{array}{c} \ln(A_{ij,t}) \\ \ln(\Xi_{j,t}) \\ \ln(Y_{i,t}) \\ \ln(Y_{j,t}) \\ \ln(Y_{t}) \\ \ln(Y_{t}) \end{array} $	$\begin{array}{c} \mathrm{FLOW}_{ij,t} \\ \mathrm{TB}_{j,t} \\ \mathrm{GDP}_{i,t} \\ \mathrm{GDP}_{j,t} \\ \mathrm{GDP}_{t} \end{array}$	Size-Adjusted Bilateral Trade Flows Multilateral Trade Imbalance Source Country Aggregate Income Destination Country Aggregate Income World Aggregate Income	U.S. dollars, Millions Gross Share, Percent U.S. dollars, Millions U.S. dollars, Millions U.S. dollars, Millions

Data Sources: Penn World Tables 9.1 by Feenstra et al. (2015), IMF Direction of Trade Statistics (DOTS) Database, World Bank Database.

Data Coverage: All variables cover the period of 1950-2014 and there are 39 countries that include both advanced and emerging markets, namely Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Colombia, Cyprus, Denmark, Egypt, Finland, France, Germany, Greece, Iceland, India, Ireland, Israel, Italy, Japan, Luxembourg, Mexico, Morocco, Netherlands, New Zealand, Norway, Peru, Philippines, Portugal, South Africa, Spain, Sweden, Switzerland, Turkey, Great Britain, the United States, and Venezuela.

available for our estimation purposes. We discuss the advantages and disadvantages of different econometric techniques in the existing panel data literature and present the case for our preferred baseline model specification. For completeness and robustness, our empirical results illustrate a broad range of model specifications and alternative estimation techniques. Once we establish the baseline coefficient estimates of the dynamic gravity equation, we cross-validate the habit theory of trade persistence with reference to other competing theories in the existing trade literature.

4.1 Data

The data used to estimate the dynamic gravity equation is displayed in Table 1. All time series are mapped directly to the variables defined in the theoretical model. The value of size-adjusted bilateral trade flows (FLOW_{*ij,t*}) represents the dependent variable, explicitly defined in equation (3.21). Consistent with equation (3.17), multilateral trade imbalance (TB_{*j,t*}) is measured as the reciprocal of the gross net export share in private consumption expenditure. As is usual in the trade literature, aggregate income in the source country (GDP_{*i,t*}), destination country (GDP_{*j,t*}), and world economy (GDP_{*t*}) is measured by the nominal gross domestic product in each location.

4.2 Unobservable Trade Resistance

There exist several alternative techniques of modeling the unobservable bilateral and multilateral trade resistance empirically. In the conventional static gravity equation (3.24), bilateral trade resistance (d_{ij}) is time-invariant, while multilateral trade resistance $(P_{j,t} \text{ and } P_{i,t})$ is static. Starting with Feenstra (2016), a large stream of the trade literature adopted a panel regression model with: (i) unobservable time-invariant heterogeneity (i.e., country fixed effects); and (ii) an unobservable homogeneous trend (i.e., time fixed effects). Both *country* and *time* fixed effects are expected to simultaneously capture the unobservable inward and outward trade resistance for each country pair and for each time period. We call the conventional strategy as the "Fixed Effects" (FE) approach. The antithesis of the conventional FE approach is to ignore all unobservable bilateral and multilateral trade resistance altogether. Specifically, following Pesaran and Smith (1995), given a relatively large time dimension T in our sample, we can estimate (N - 1)N number of country-

specific regressions, one for each unique country pair, and average all of the coefficient estimates across all of the country pairs. We call this restrictive strategy the "Mean Group" (MG) approach.

Unlike the FE estimator, which provides pooled coefficient estimates that are homogeneous for all country pairs, the key advantage of the MG estimator is that it reflects the observed crosssectional heterogeneity of the panel by generating country-specific coefficient estimates. In the context of the dynamic gravity equation, this means that the MG estimator distinguishes between country pairs for which trade flows are persistent and unbalanced and for those that are not, while the FE estimator "paints with a broad brush". Heterogeneous trade persistence is a property we wish to retain in our empirical estimates, given that the theoretical model of the dynamic gravity equation (3.24) predicts a trade persistence coefficient $\chi_{ij}(\eta-1)$ to be heterogeneous across different country pairs. However, the main disadvantage of the MG approach is that it accounts for neither time-varying nor time-invariant unobservable heterogeneity. Specifically, if we take the inference drawn about the coefficient of trade persistence based on the rudimentary MG estimator at face value, then it is as if geographic distance between countries or their overall propensity to trade have no differential impact on the degree of trade persistence across any country pairs. As a consequence, we also consider a "Hybrid Fixed Effects" approach (HFE), which reflects both the observed and the unobserved heterogeneity of the panel.

There are two important reasons why, despite their popularity, none of the aforementioned approaches are chosen as the preferred technique in this paper. First, the homogeneous time fixed effects do not appropriately reflect the fact that the unobservable time-varying multilateral resistance can be strongly correlated with observable regressors in the dynamic gravity equation. In practice, we have every reason to believe that it is indeed the case. Our theoretical model endogenously links the multilateral resistance to trade flows, trade imbalance, and aggregate income (see equation (3.19)). Our empirical results provide additional support for this hypothesis, which we discuss below in Section 4.4. In fact, Anderson and Yotov (2010) and Anderson (2011) argue that the unobservable inward and the outward multilateral resistance may be heterogeneous across different country pairs. And if so, then the correlation between the observable regressors and the unobservable time-varying inward and the outward multilateral resistance may also be heterogeneous, which the time fixed effect approach is unable to tackle (see Kapetanios et al. (2017)). Second, the gravity equation in this paper is dynamic, not static as is traditionally the case. And while this may seem rather innocuous, Pesaran and Smith (1995) show that neglected parameter heterogeneity associated with the FE approach generates biased and inconsistent coefficient estimates when the panel regression model is indeed dynamic. This observation is particularly alarming, since the existing trade literature tends to ignore parameter heterogeneity in spite of the three-dimensional data structure, which comprises of the source country, the destination country, and time.

We recognize two further MG-based techniques that are able to not only reconcile parameter heterogeneity, but also proxy the unobservable time-varying multilateral trade resistance specific to each country pair. The first technique is known as the "Augmented Mean Group" (AMG) estimator. Following Eberhardt and Teal (2013), AMG involves estimating the standard FE regression model with individual and time fixed effects, extracting the pooled time fixed effect coefficients, and then using their vector as an additional regressor (i.e., 'unobservable common factor') in an otherwise standard MG regression model. Consequently, the AMG coefficient estimates are heterogeneous for each country pair, analogous to the regular MG approach. But unlike the regular MG estimator, the AMG coefficient estimates also reflect the fact that the unobservable common factor exerts a heterogeneous influence on bilateral trade flows for each country pair at each point in time. The second technique is referred to as the "Common Correlated Effects Mean Group" (CCEMG) estimator. Following Pesaran (2006), Kapetanios et al. (2011), and Chudik and Pesaran (2015), CCEMG replaces the homogeneous time fixed effects with unobservable common components, which then enter the panel regression model as additional regressors. The main difference between AMG and CCEMG techniques is that CCEMG usually measures the unobservable common components as the cross-sectional averages of all variables entering the regression model (i.e., "global factors") rather than pre-estimating a set of pooled coefficients next to the homogeneous time trend.

Despite the flexibility of the AMG estimator, our preferred approach of estimating the dynamic gravity equation is the CCEMG estimator, because it is more general and subject to fewer assumptions. Specifically, if the pre-estimated pooled regression in the AMG approach generates inconsistent coefficient estimates, due to the fact that our panel regression model is dynamic, then the inference drawn from the subsequent country pair-specific regressions is inaccurate because it inherits those inconsistencies. Furthermore, the dynamic gravity equation depicted in equation (3.22) incorporates four types of unobservable trade resistance (i.e., d_{ij} , $P_{j,t}$, $\Phi_{i,t}$, and $\Phi_{i,t-1}$). Unlike the aforementioned techniques, CCEMG accounts for the fact that the unobservable timevarying multilateral resistance is dynamic, not static (i.e., both $\Phi_{i,t}$, and $\Phi_{i,t-1}$ must be controlled for). It does so by explicitly incorporating proxies for the contemporaneous and lagged unobservable time-varying multilateral trade resistance. Those proxies are the unobservable global factors entering the gravity equation with country-pair-specific factor loadings. And those unobservable global factors are measured as the cross-sectional averages of all variables entering the dynamic gravity equation, including the contemporaneous as well as lagged trade flows.

In fact, there are several arguments why the vector of unobservable common factors should consist solely of the cross-sectional average trade flows. Theoretically, if N is sufficiently large, the cross-sectional average of the trade imbalance variable tends to zero, because the net trade flows of the world economy as a whole are always balanced. Similarly, the cross-sectional averages of aggregate income are strongly related to the world aggregate income, which enters the dynamic gravity equation by default (see equation (3.22)). For this reason, we also consider one more approach, which we call the "Restricted Common Correlated Effects Mean Group" (CCEMGR), in which the vector of unobservable common factors is based solely on the cross-sectional averages of the contemporaneous and lagged trade flows. But in order to gauge the relative importance of retaining parameter heterogeneity or incorporating unobservable global factors on the inference drawn about the trade persistence coefficient in our empirical model, we consider the Pooled Common Correlated Effects (CCEP) specification. Contrary to the MG approach, which retains parameter heterogeneity, but omits the unobservable global factors, the CCEP approach ignores the intrinsic parameter heterogeneity, but incorporates the unobservablez global factors.

For completeness and robustness, we also consider other variations of the FE approach commonly adopted in the literature (e.g., Piermartini and Yotov (2016); Anderson and Yotov (2020)). While the aforementioned FE approach incorporates both country and time fixed effects, it assumes that all countries share a homogeneous trend component and it does not account for the time-invariant heterogeneity specific to each country pair. For this reason, the FE2 approach replaces the country and time fixed effects with the so-called "time-varying" fixed effects, which allows for a heterogeneous time trend component specific to each country. The FE3 approach applies the standard country and time fixed effects as does the conventional FE approach, but it also controls for the time-invariant heterogeneity specific to each country pair. And finally, the FE4 approach controls for both the heterogeneous time trend as well as time-invariant heterogeneity specific to each country pair, which replaces the country-specific time-invariant heterogeneity.

All approaches described above draw inference about the regression coefficients from a loglinear specification of the dynamic gravity equation. But we admit that all log-linear applications of the bilateral trade flow data entail one simple caveat, which is commonly referred to as the "zero trade problem" due to Santos-Silva and Tenreyro (2007). Specifically, given that our dataset comprises of N = 39 and T = 65, around 10% of total observations TN(N-1) in our sample contain zero entries. This finding documents the fact that the bilateral trade flows between a subset of country pairs during a subset of consecutive time periods were either unrecorded or non-existent. And if so, then the cross-sectional heteroscedasticity caused by the zero entries leads to at least somewhat biased and inconsistent coefficient estimates. A common approach to address the zero trade problem is to use the Poisson Pseudo-Maximum-Likelihood (PPML) approach, which estimates the regression model in a multiplicative form (e.g., Santos-Silva and Tenreyro (2007); Westerlund and Wilhelmsson (2009)). While we incorporate the results from the PPML specification as yet another tentative approach in Appendix C, we recognize several reasons why the results from the CCEMG approach are generally preferred to those of the PPML approach in the context of our empirical application. First, the CCEMG estimator in principle allows the error structure to exhibit unknown heteroscedasticity over time, so long as it is subject to a finite order of integration (see Westerlund (2018)). Second, given the large N and large T nature of the panel, the observed cross-sectional heteroscedasticity in our analysis is dominated by the timevarying component captured by the multi-factor error structure. Third, and most importantly, the existing PPML applications are confined exclusively to the static gravity equations, such as Weidner and Zylkin (2019). A formal extension of the static PPML framework into a dynamic counterpart with a three-dimensional data structure goes beyond the scope of this paper, since it involves nontrivial practical hurdles. Specifically, the zero trade problem in the (lagged) dependent variable, introduction of a multi-factor error structure, as well as retention of parameter heterogeneity.

4.3 Methodology

The empirical adaptation of the theoretical dynamic gravity equation (3.21) is a large N and large T panel regression model. Our panel regression model extends the interactive fixed effects representation of Bai (2009) into a three-dimensional data structure. Specifically, our model captures temporal variation over t = 1, 2, ..., T, but also spatial variation across the source country i = 1, 2, ..., N and the destination country j = 1, 2, ..., N - 1, such that $j \neq i$. Formally,

$$\mathbf{y}_{ij,t} = \beta_{0ij} + \mathbf{x}'_{ij,t} \boldsymbol{\beta}_{ij} + u_{ij,t}, \tag{4.1}$$

$$u_{ij,t} = \lambda'_{ij}\phi_t + \varepsilon_{ij,t}, \tag{4.2}$$

$$\mathbf{x}_{ij,t} = \boldsymbol{\gamma}'_{ij}\boldsymbol{\phi}_t + \boldsymbol{\nu}_{ij,t},\tag{4.3}$$

where $y_{ij,t} := \text{FLOW}_{ij,t}$ are the trade flows, $\beta_{ij} = [\beta_{1ij}, \beta_{2ij}, ..., \beta_{5ij}]'$ is a 5 × 1 vector of coefficients, $\mathbf{x}_{ij,t} = [\text{FLOW}_{ij,t-1}, \text{TB}_{j,t}, \text{GDP}_{i,t-1}, \text{GDP}_{j,t-1}, \text{GDP}_{t-1}]'$ is a 5 × 1 vector of all common and country-specific observable factors, while ϕ_t and λ_{ij} , γ_{ij} represent some configuration of the unobservable vector of common factors and country-pair-specific vectors of factor loadings, respectively. The error terms $\varepsilon_{ij,t}$ and $\boldsymbol{\nu}_{ij,t}$ are assumed to be independently distributed of each other, uncorrelated with the unobservable common factors, and uncorrelated across country pairs.

Each estimation strategy discussed in Section 4.2 is nested as a special case of equations (4.1), (4.2), and (4.3) by choosing an estimator of β_{ij} and imposing restrictions on the inner product of $\lambda'_{ij}\phi_t$. For instance, the configuration of $\phi_t = [1, 1, \tau_t]'$ and $\lambda_{ij} = [\alpha_i, \alpha_j, 1]'$ gives rise to the traditional FE error structure of Feenstra (2016), namely $u_{ij,t} = \alpha_i + \alpha_j + \tau_t + \varepsilon_{ij,t}$, where α_i and α_j are the country fixed effects and τ_t are the time fixed effects. However, FE pools the regressions coefficients by averaging $\mathbf{x}_{ij,t}$ across all source and destination countries, such that $\boldsymbol{\beta} = (\bar{\mathbf{x}}'_t \bar{\mathbf{x}}_t)^{-1} \bar{\mathbf{x}}'_t \bar{\mathbf{y}}_t$ is homogeneous for all $i, j \in n$, where $\bar{\mathbf{x}}_t = 1/\bar{N} \sum_{ij=1}^{\bar{N}} \mathbf{x}_{ij,t}$ and $\bar{\mathbf{y}}_t = 1/\bar{N} \sum_{ij=1}^{\bar{N}} \mathbf{y}_{ij,t}$ denote crosssectional averages, while $\bar{N} = (N-1)N$ measures the number of unique trade pairs. By contrast, the MG approach nullifies the inner product $\lambda'_{ij}\phi_t = 0$, but preserves parameter heterogeneity between different country pairs, such that $\beta_{ij} = (\mathbf{x}'_{ij,t}\mathbf{x}_{ij,t})^{-1}\mathbf{x}'_{ij,t}\mathbf{y}_{ij,t}$ and the inference is drawn from the cross-sectional average of the MG coefficient estimates $\boldsymbol{\beta} = 1/\bar{N} \sum_{ij=1}^{\bar{N}} \beta_{ij}$.

The HFE approach combines the FE and the MG approaches, by imposing $\phi_t = [1,1]'$ and $\lambda_{ij} = [\alpha_i, \alpha_j]'$. Consequently, the only difference between the MG and the HFE approaches is that the multifactor error structure is now given by $u_{ij,t} = \alpha_i + \alpha_j + \varepsilon_{ij,t}$, while the regression coefficients β_{ij} are estimated using the standard MG approach. By contrast, the other common variations of the FE approach adopted in the literature rely exclusively on the pooled coefficient estimator as does the conventional FE approach, but their differences arise from the specification of the multifactor error structure. Specifically, FE2 imposes $\lambda'_{ij}\phi_t = \alpha_{i,t} + \alpha_{j,t}$, FE3 imposes $\lambda'_{ij}\phi_t = \alpha_i + \alpha_j + \alpha_{ij} + \tau_t$, and FE4 imposes $\lambda'_{ij}\phi_t = \alpha_{i,t} + \alpha_{j,t}$. The AMG approach sets $\phi_t = [1, 1, \hat{\tau}_t]'$ and $\lambda_{ij} = [\alpha_i, \alpha_j, \alpha_{ij}]'$, where $\hat{\tau}_t$ are the pre-estimated time fixed effects from the standard FE regression model. Notice that α_{ij} is restricted to equal unity in the FE approach, such that the error structure is given by $u_{ij,t} = \alpha_i + \alpha_j + \alpha_{ij}\hat{\tau}_t + \epsilon_{ij,t}$ and

⁷Analogous to Pesaran (2006), equation (4.3) justifies the use of cross-sectional averages of $\mathbf{x}_{ij,t}$ to proxy the unknown factors in $u_{ij,t}$. This is because ordinary least squares applied to equations (4.1) and (4.2) generally delivers biased and inconsistent coefficient estimates whenever the unobservable common factors ϕ_t are correlated with the regressors $\mathbf{x}_{ij,t}$. And this constraint generally binds, since the dynamic gravity equation (3.22) predicts that bilateral trade flows depend on the inward and outward multilateral resistance, which in turn are functions of bilateral trade flows to and from all trade partners, respectively (see equation (3.19)). But as long as we accommodate the endogeneity of the unobservable common factors are projected onto their cross-sectionally weighted averages, which renders consistency of the coefficient estimates under quite general assumptions set out by Pesaran (2006) and Chudik and Pesaran (2015).

the time fixed effects exert a heterogeneous response for each country pair. The AMG estimates the regression coefficients β_{ij} in a similar way as a regular MG approach as opposed to pooling β_{ij} across all country pairs.

The CCEMG approach imposes $\phi_t = [1, 1, \mathbf{z}'_t]'$ and $\lambda_{ij} = [\alpha_i, \alpha_j, \alpha'_{ij}]'$, where $\mathbf{z}_t = [\bar{\mathbf{y}}_t, \bar{\mathbf{x}}'_t]'$ is the vector of unobserved common factors proxied by the cross-sectional averages of the dependent and independent variables.⁷ In order to see what role, if any, is played by retaining the intrinsic parameter heterogeneity, the CCEP approach adopts an identical error structure as the CCEMG approach, but applies the pooled regression coefficient estimator similar to the FE approach. And finally, the CCEMGR approach simply removes $\bar{\mathbf{x}}_t$ from the proxied unobserved common factors \mathbf{z}_t and replaces it with $\bar{\mathbf{y}}_{t-1}$. This means that the error structure of the CCEMG, CCEMGR, and CCEP approaches is given by $u_{ij,t} = \alpha_i + \alpha_j + \mathbf{z}'_t \alpha_{ij} + \varepsilon_{ij,t}$, which distinguishes between unobservable country-specific time-invariant heterogeneity captured by α_i and α_j as well as unobservable timevarying heterogeneity specific to each country pair captured by the inner product of $\mathbf{z}'_t \alpha_{ij}$.

4.4 Coefficient Estimates

Consider the coefficient estimates of the dynamic gravity equation presented in Table 2. Each column displays the values of the coefficient estimates that are obtained using one of the seven different techniques described in Sections 4.2 and 4.3.⁸ Our preferred baseline model specification titled CCEMG is presented in column (1), which incorporates the time-invariant country- and country-pair-specific heterogeneity, controls for the unobservable global factors, and also estimates the regression coefficients specific to each country pair, such that the intrinsic parameter heterogeneity is retained. We compare and contrast our preferred baseline model estimates to other techniques commonly applied in the existing literature in order to emphasize the importance of incorporating flexible time trends and retaining parameter heterogeneity in the empirical adaptation of the dynamic gravity equation. Specifically, instead of estimating pooled regression coefficients that are homogeneous for all country pairs as is traditionally the case, we examine whether the sequence of first estimating all coefficients specific to each country pair and only then averaging across country pairs makes a significant difference when drawing inference. For direct comparability reasons, Table 2 displays only the pooled or the averaged coefficient estimates, but we demonstrate and discuss the extent of parameter heterogeneity of our preferred baseline model specification in Section 5.2.

The first line of Table 2 presents the (pooled or average) coefficient estimates associated with the lagged dependent variable (FLOW_{*ij,t-1*}), which we define as the "trade persistence coefficient" (i.e., pooled/average β_{1ij} in equation (4.1)). We draw particular attention to the value of the trade persistence coefficient because it summarizes the rate at which trade flows adjust to shocks and simultaneously elicits the heterogeneity in the magnitude of the habits parameter specific to each country-pair (χ_{ij}). Specifically, β_1 is mapped directly to $1/\bar{N} \sum_{ij=1}^{\bar{N}} \chi_{ij}(\eta - 1)$ in equation (3.22), where $\eta > 0$ and $\bar{N} = N(N - 1)$ measures the total number of unique country pairs in our sample. First, the trade persistence coefficient is significantly different from zero and unity for all seven different techniques. This implies that following a random shock, trade flows generally

 $^{^{8}}$ For the sake of clarity and space, the additional results generated using other variations of the FE and PPML approaches are relegated to Table 5 in Appendix C.

		\$		\$			
	CCEMG	ΕĒ	MG	CCEP	HFE	AMG	CCEMGR
VARIABLES	(1) FLOW $_{ij,t}$	$\mathop{\mathrm{FLOW}}_{ij,t}$	(3) FLOW _{ij,t}	(4) FLOW $_{ij,t}$	(5) FLOW $_{ij,t}$	(6) FLOW _{ij,t}	(7) FLOW $_{ij,t}$
$\mathrm{FLOW}_{ij,t-1}$	0.347^{***}	0.907^{***}	0.548^{***}	0.374^{***}	0.488^{***}	0.433^{***}	0.460^{***}
TB	(0.00825)0.075***	(0.00451) 0.219***	(0.00643) 0 803***	(0.0161) 0.612^{***}	(0.00711) 0.865***	(0.00720) 0 839^{***}	(0.00730) 0 770***
LU _{j,t}	(0.126)	(0.0279)	(0.0714)	(0.0801)	(0.103)	(0.0980)	(0.0989)
${ m GDP}_{i,t-1}$	-0.312^{***}	-0.00174	-0.183^{***}	-0.296^{***}	-0.232***	-0.197***	-0.210^{***}
	(0.0778)	(0.00749)	(0.0149)	(0.0338)	(0.0283)	(0.0298)	(0.0422)
${ m GDP}_{j,t-1}$	-0.117	-0.0239^{***}	-0.132^{***}	-0.195^{***}	-0.134^{***}	-0.0234	-0.0634
c 2	(0.0954)	(0.00714)	(0.0150)	0.0271	(0.0306)	(0.0325)	(0.0454)
GDP_{t-1}	0.228	-0.0419	0.322^{***}		0.456^{***}	0.536	0.594^{***}
	(0.201)	(0.0258)	(0.0397)		(0.0996)	(0.114)	(0.160)
Time Fixed Effects	Z	γ	Z	Ζ	Z	Z	Z
Country Fixed Effects	Υ	Υ	Z	Z	Υ	Υ	Υ
2							
Unobservable Global Factors	Υ	Z	Ζ	Υ	Z	Υ	Υ
Constant	-0.628	2.035^{***}	3.711^{***}			8.443***	-4.471**
	(3.552)	(0.353)	(0.522)			(1.631)	(2.033)
Observations	70,579	70,604	61,551	70,526	70,579	70,579	70,579
Number of pairs	1,473	1,480	1,152	1,468	1,473	1,473	1,473
Adj. R-squared	0.79	0.90	0.75	0.84	0.74		0.77
Note: Robust standard errors in pa	arentheses.						
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$	1.						

Table 2: Dynamic Gravity Model

revert back to the trend gradually rather than instantaneously as is implied by the static gravity equation due to Anderson and van Wincoop (2003). Second, and more importantly, our estimates demonstrate a remarkable difference between the standard FE approach, which generates a pooled trade persistence coefficient estimate of 0.91 that is homogeneous for all country pairs (see column (2) in Table 2), and all other techniques that retain estimated parameter heterogeneity and/or incorporate some measure of the unobservable global factors. This implies that following a random shock, trade flows revert back to the trend at a considerably faster rate than suggested by the neo-classical gravity equation pioneered by Yotov and Olivero (2012).

In particular, in the absence of unobservable global factors, the pooled coefficient estimator generates economically and statistically significant excess trade persistence equivalent to almost three-fold the magnitude of our baseline coefficient estimate of 0.35 (see columns (1) and (2) in Table 2). While the excess trade persistence discrepancy related to the FE approach is robust across the board, it is not solely attributable to parameter heterogeneity. Notice that the MG approach generates an average trade persistence coefficient estimate of 0.55 (see column (3) in Table 2), in spite of retaining parameter heterogeneity, because the MG approach does not incorporate flexible time trends and therefore does not account for the role that global shocks play in driving the bilateral trade flows. By contrast, the CCEP approach ignores the parameter heterogeneity and generates a pooled trade persistence coefficient of 0.37 (see column (4) in Table 2), because it incorporates a comprehensive set of unobservable global factors. If anything, the presence of the unobservable global factors in the CCEP approach tends to generate a significantly lower trade persistence coefficient compared to the MG approach, which retains the parameter heterogeneity, but expends the unobservable global factors. However, controlling for the unobservable global factors and retaining parameter heterogeneity both lead to a significantly lower trade persistence coefficient than predicted by the conventional FE approach. This conclusion is further reinforced by the trade persistence coefficient estimate of 0.43 generated by the AMG approach (see column (6) in Table 2), which retains parameter heterogeneity and incorporates one single unobserved common factor. Specifically, the AMG estimate of the trade persistence coefficient is significantly lower than the MG estimate, but significantly higher than the CCEP or the CCEMG estimates. Hence, if the panel regression model controls for the unobservable global factors and retains parameter heterogeneity, then the excess trade persistence disappears.

The second line of Table 2 presents the (pooled or average) coefficient estimates associated with the multilateral trade imbalance in the destination country $(TB_{j,t})$, which we define as the "trade imbalance coefficient" (i.e., pooled/average β_{2ij} in equation (4.1)). Consistent with the theory of habits in the supply chains described in Section 3, the multilateral trade imbalance $TB_{j,t}$ is measured as the reciprocal of the gross share of net exports in total consumption expenditure. As shown in equations (3.21) and (3.22), multilateral trade imbalance enters the dynamic gravity equation with unitary elasticity, such that and increase in the multilateral trade deficit in the destination country $j \in n \setminus i$ should *ceteris paribus* attract more trade flows from each source country i = 1, 2, ..., N. And our empirical estimates indeed correspond to the theoretical predictions of the model. Specifically, we obtain positive and statistically significant trade imbalance coefficients using all seven estimation techniques presented in Table 2 (and Table 5 of Appendix C). Notice that the lowest trade imbalance coefficient estimate of 0.22 is obtained using the FE approach (see column (2) of Table 2), which excludes the unobservable global factors and generates a pooled trade imbalance coefficient. The largest trade imbalance coefficient estimate of 0.98 is obtained using the CCEMG approach (see column (1) of Table 2), our preferred baseline model specification, which incorporates unobservable global factors with pair-specific factor loadings and retains the parameter heterogeneity. The remaining estimation techniques generate the trade imbalance coefficient estimates that are generally closer to the CCEMG approach compared to the FE approach.

The remaining lines of Table 2 present the (pooled or average) coefficient estimates associated with the lagged source country aggregate income $(\text{GDP}_{i,t-1})$, lagged destination country aggregate income (GDP_{j,t-1}), and lagged world aggregate income (GDP_{t-1}). According to the habit model of the dynamic gravity equation depicted in equation (3.22), the size-adjusted bilateral trade flows (FLOW_{ij,t}) are ceteris paribus increasing in $GDP_{j,t-1}$ and GDP_{t-1} , but decreasing in $GDP_{i,t-1}$. Consistent with the theory, we do find evidence of a small, negative, and statistically significant coefficient estimate for $\text{GDP}_{i,t-1}$ across the board. We also find that the coefficient estimate for GDP_{t-1} is positive and statistically significant, but only for some of the estimation techniques that retain parameter heterogeneity. However, contrary to the theoretical predictions, the coefficient estimates for $\text{GDP}_{i,t-1}$ are generally negative, but statistically significant only in the absence of the unobservable global factors and the CCEP approach. As a consequence, our preferred benchmark model specification CCEMG delivers coefficient estimates that are the most theoretically consistent. And when a subset of the unobserved global factors are removed (see CCEMGR in column (6) of Table 2), the coefficient estimate for our single observable global factor GDP_{t-1} becomes relatively large, positive, and statistically significant. It follows that both the observable and the unobservable global factors generally play an important role when drawing inference.

The fact that our results resolve the excess trade persistence puzzle should come as no surprise. It is a well-known empirical stylized fact that imports and exports are the most volatile components of aggregate demand. At the same time, shocks to the unobservable global factors have a tendency to cause sharp, widespread, and synchronized trade flow adjustments. For instance, the so-called "Great Trade Collapse" of the 2008-09 famously coined by Alessandria et al. (2010) and more recently the COVID-19 pandemic. The fact that highly volatile and strongly co-moving bilateral trade flows are widely-observed is hardly consistent with the traditional panel regression models that exclude the unobservable global factors. Once the global factors are appropriately taken into the account, we successfully assimilate low trade persistence and high trade flow volatility relative to output as observed in the data. Our results therefore emphasize an important distinction that the choice of trade partners is indeed a sluggish and persistent process, but the actual value of trade flows changes rapidly, especially in response to the global trade shocks.

4.5 Empirical Implications

Our opening results presented in Section 4.4 establish an important empirical stylized fact. Specifically, in general, traditional panel regression models, proposed by Feenstra (2016), Piermartini and Yotov (2016), Anderson and Yotov (2020) and others that are based solely on country-specific fixed effects, time fixed effects, and a pooled coefficient estimator, provide misleading inference when extended from a static to a dynamic gravity equation setting. In particular, traditional panel regression models generate an exceedingly upwardly-biased estimate of the trade persistence coefficient. There are two distinct sources of this upward bias, which we refer to as "excess trade persistence". First, in keeping with Chudik and Pesaran (2015), omitted unobservable global factors cause the trade persistence coefficient estimates to be biased and inconsistent, because they ignore strong cross-sectional dependence of bilateral trade flows across different country pairs, which stems from the country-specific and time-varying multilateral trade resistance in equation (3.22). Second, in accordance with Pesaran and Smith (1995), the estimates of the trade persistence coefficient in a dynamic gravity equation are biased and inconsistent if the coefficient estimator neglects parameter heterogeneity. Based on the premise that the estimate of the trade persistence coefficient in our baseline model specification incorporates the unobservable global shocks and retains parameter heterogeneity, this section of the paper compares and contrasts the magnitude and the robustness of our coefficient estimates to those in the existing literature.

The most well-known existing theory of bilateral trade flow persistence due to Yotov and Olivero (2012) and Anderson et al. (2020) is based on the standard neo-classical capital accumulation equation. As per usual, the neo-classical theory introduces an infinitely-divisible measure of aggregate capital stock, which depreciates at a deterministic rate $\delta \in [0, 1]$ per every time period and requires investment into new capital stock in order to preserve the balanced growth path. The dynamics of the aggregate capital stock are then linked to the bilateral trade flows through a Cobb-Douglas production function and standard homothetic preferences across the domestic and foreign varieties from which a dynamic gravity equation is derived. The main advantage of the neo-classical theory of trade persistence is that it hinges on capital accumulation and exploits one of the most fundamental sources of dynamics in the real business cycle literature. However, the main disadvantage of the neo-classical theory is that it predicts a highly restrictive domain for the trade persistence coefficient that is at odds with the empirical evidence. In particular, Yotov and Olivero (2012) show that the neo-classical theory predicts a trade persistence coefficient equivalent to $1-\delta$ and estimate δ , measuring the annual rate of capital depreciation, to be anywhere from 0.06 to 0.14. Conversely, IMF (2015) estimates that the value of δ lies in the interval of 0.04 and 0.1, depending on the time period and whether the country is advanced or developing. If we take the neo-classical theory at face value, it follows that an empirically plausible lower bound for the annual trade persistence coefficient is around 0.86. But the lower bound of 0.86 merely corresponds to some of our exceedingly upwardly-biased and inconsistent estimates that neglect parameter heterogeneity and exclude unobservable global factors (see column (2) in Table 2). Once we incorporate the pair-specific fixed effects and flexible time effects and refrain from the pooled coefficient estimator, the magnitude of the trade persistence coefficient shrinks by around 2-3 times. Specifically, our baseline model specification predicts a trade persistence coefficient of 0.35 (see column (1) in Table 2), which in the light of the neo-classical theory implies that 65% of global capital stock depreciates every single year (i.e., up to 16 times more than the IMF (2015) estimates).

Despite how simple and elegant the neo-classical framework of trade persistence is, the striking discrepancy between theory and evidence suggests a more pragmatic view that capital accumulation forms only a subset, but perhaps not the core, of the trade persistence mechanism. And in support of this view, Section 3 of this paper develops a competing theory of trade persistence that extends the relative habits framework of Ravn et al. (2006) to capture a reduced form mechanism of inertia in the globalized wholesale distribution network. The main advantages of the theory of habits in

the supply chains is that it presents not only a much more flexible identification for the domain of the trade persistence coefficient, but also allows for a heterogeneous magnitude across different country pairs. Specifically, the habits framework predicts a trade persistence coefficient measured as $\chi_{ij}(\eta - 1)$. In this theoretical identity, parameter $|1 - \eta|$ stands for trade elasticity, where $\eta > 0$ is the elasticity of substitution, the value of which generally ranges between 5 and 10 in the related literature (see Anderson and van Wincoop (2004) for evidence and Arkolakis et al. (2012) for the application). Conversely, parameter $\chi_{ij} > 0$ measures the intensity of habits specific to any given country pair. If we take the values of η from the literature and combine them with our CCEMG estimate for the trade persistence coefficient of 0.35, then our theoretical model predicts a lower (upper) bound for the habits parameter to be 0.035 (0.07). This value is even more conservative than 0.1, which was originally assumed in the seminal contribution of Ravn et al. (2006).

However, contrary to the traditional predictions in the gravity literature, there exists some evidence that the trade elasticity $|1 - \eta|$ is in fact time-varying as opposed to constant and equal to the elasticity of substitution in CES preferences as is traditionally considered to be the case (e.g., Anderson and van Wincoop (2003)). Specifically, Boehm et al. (2020) measure the short-run and the long-run trade elasticities by exploiting recurring exogenous tariff changes for identification purposes. The authors find substantially smaller values of trade elasticities equal to around 0.7 in the short-run and 1.75 in the long run in absolute value terms. Looking at this new evidence through the lenses of our empirical model, indicates that the habits-induced persistence can be quite large in the short-run (i.e., 0.35/0.7 = 0.5), but declines by around 2.5 times in the long-run (i.e., 0.35/1.75 = 0.2). And since we analyze more than 60 years worth of data across advanced and developing economies with remarkably different industrial structures, the time-variation and heterogeneity of trade elasticities across countries is well-expected (see Imbs and Mejean (2017) for the cross-country evidence). That said, the implied average long-run persistence parameter of 0.2 is nonetheless compatible with the relatively sharp and synchronized international trade flow adjustments in response to large shocks as described in Section 2.

We also conduct a robustness check, where the dynamic gravity equation is re-estimated for all seven different specifications presented in Table 2 by excluding the multilateral trade imbalance from the vector of regressors (see Table 6 in Appendix C). However, even if the multilateral trade imbalance is removed from the empirical model, which contradicts our theoretical model, the coefficient estimates remain broadly unchanged. If anything, the trade persistence coefficient is significantly lower, not higher, when the multilateral trade imbalance is controlled for. We therefore concur with the long-standing trade literature that incorporates multilateral trade imbalance as a weakly exogenous regressor in a static gravity equation (e.g., Davis and Weinstein (2002), Dekle et al. (2007), and Dekle et al. (2008)). Our dynamic extension provides a simple alternative theory in which quantitatively small habits can resolve the excess trade persistence puzzle, and also rationalizes why trade persistence is heterogeneous across different country pairs. However, the disadvantage of the theory of habits in the supply chains is that it is conceptually more difficult to derive from the first principles. It is also more challenging to verify their external validity compared to the rate of capital depreciation. For this reason, Section 5.3 explores the empirical link between our country-specific estimates of trade persistence coefficients and the corresponding country-specific indicators of participation in global value chains.

5 Cross-Validation

5.1 Prediction Performance "Horse Race"

We have thus far established that controlling for the unobservable global factors and retaining the country-pair-specific parameter heterogeneity when estimating the coefficients of the dynamic gravity equation leads to a significantly lower trade persistence coefficient than predicted using the conventional estimation strategies documented in the literature. The benefits of adopting our empirical approach are two-fold. First, our preferred empirical strategy is consistent with the theory of habits in the supply chains, which predicts heterogeneous trade persistence coefficients across different country pairs and generates inward and outward multilateral resistance with lags that strongly correlate with foreign demand and foreign supply shocks. Second, unlike the static gravity equation due to Anderson (1979), Anderson and van Wincoop (2003), and Feenstra (2016), which predicts zero trade persistence, or the neo-classical gravity equation due to Yotov and Olivero (2012), Alvarez (2017), and Anderson et al. (2020), which predicts a trade persistence coefficient of around 0.8-0.9, our preferred estimation strategy delivers a cross-country average trade persistence coefficient equal to around 0.35, which is able to rationalize the sharp and synchronized international trade flow adjustments in response to global trade shocks observed in the data.

In order to illustrate that our preferred estimation strategy, titled CCEMG, outperforms the leading rival empirical strategies in terms of the data fit, especially in response to global trade shocks, such as the "Global Trade Collapse" of 2008-2009, we conduct a so-called "horse race" for the predictive performance of different empirical estimation strategies of our dynamic gravity model presented in equations (4.1)-(4.3). Specifically, Table 3 compares the Root Mean Square Errors (RMSE) calculated using the CCEMG, MG, CCEP, and FE methodologies (see Sections 4.2 and 4.3 for more details). The in-sample RMSEs are presented for the full data sample, the observed "good times", and the observed "bad times", in order to compare different model performance inside and outside of time periods characterized by global trade shocks. Consistent with Kose et al. (2020), the "bad times" represent the global recession years, namely 1975, 1982, 1991, and 2009, while the "good times" are all of the remaining years in our data sample that spans 1950-2014. The term $w = \{0, 1, 2, 3\}$ further indicates the length of the windows surrounding the recession years (i.e., number of years before and after global trade shocks that are included in the "bad times" sample in addition to the outlined recession years).

According to our RMSE calculations presented in Table 3, the CCEMG approach delivers the most accurate data fit not only throughout the entire data sample, but also during solely "good times" or "bad times". Recall that the CCEMG approach controls for the unobservable global factors and retains the country-pair-specific parameter heterogeneity. The runner-up methodologies are the MG approach, which retains the country-pair-specific parameter heterogeneity, but expends the unobservable global factors, and the CCEP approach, which controls for the unobservable global factors, but ignores the country-pair-specific parameter heterogeneity. The conventional FE approach, which expends the unobservable global factors and ignores the country-pair-specific parameter heterogeneity delivers the largest RMSE value and predicts the least accurate data fit. The reason why the performance of the FE approach is inferior to the MG, CCEP, and CCEMG approaches is because the latter all deliver a lower trade persistence coefficient than the FE approach

	Full Sample		"Bad 7	Γimes"			"Good	Times"	
Method		w = 0	w = 1	w = 2	w = 3	w = 0	w = 1	w = 2	w = 3
CCEMG MG CCEP FE	0.38 0.44 0.47 0.55	0.41 0.50 0.47 0.65	0.41 0.49 0.47 0.63	0.40 0.48 0.46 0.60	0.41 0.48 0.47 0.59	0.38 0.44 0.43 0.54	0.37 0.43 0.42 0.52	0.37 0.42 0.42 0.51	0.35 0.40 0.39 0.49

 Table 3: Root Mean Square Error

Note: This figure presents the Root Mean Square Errors (RMSE) calculated using different methods of estimating the coefficients in a dynamic gravity equation. The in-sample RMSEs are presented for the full data sample, the observed "good times", and the observed "bad times" in order to compare different model performance inside and outside of time periods characterized by global trade shocks. Consistent with Kose et al. (2020), the "bad times" represent the global recession years, namely 1975, 1982, 1991, and 2009, while the "good times" are all of the remaining years in our data sample that spans 1950-2014. The term $w = \{0, 1, 2, 3\}$ further indicates the length of the windows surrounding the recession years (i.e., number of years before and after global trade shocks that are included in "bad times" in addition to the recession years). The values in bold indicate the smallest RMSE.

(see Table 2). That said, all of the methodologies we consider perform marginally better during the "good times" rather than the "bad times", because "bad times" occur less frequently.

For the sake of robustness, we calculate the RMSEs for numerous other methodologies considered in this paper and generally reach the same outcome (see Table 7 in Appendix C). While Tables 3 and 7 calculate RMSEs based on the "good times" and "bad times" sub-samples, they nonetheless rely on the dynamic gravity equation coefficient estimates from the entire data sample. In order to ensure that our findings are robust, we also present Table 8 in Appendix C, which calculates both the RMSEs as well as the coefficient estimates based solely on the "good times" and "bad times" sub-samples. Due to the limited number of time periods in the "bad times" subsample, not all methodologies can be successfully implemented, since the "mean group" techniques that retain parameter heterogeneity rely on a sufficiently large temporal dimension of the panel. However, the outcome regarding the superiority of the CCEMG approach generally holds (the only viable rival during "normal times" is PPML4 though CCEMG strictly dominates during global recessions). Consequently, we conclude that even allowing for more structurally-relevant variables, instead of absorbing them into time-varying fixed effects and resorting to reduced-form regressions, our model outperforms the standard empirical approaches documented in the literature and successfully resolves the excess trade persistence puzzle.

5.2 Parameter Heterogeneity

Our results have thus far established the importance of retaining parameter heterogeneity across all country pairs, since it is one of the drivers of the excess trade persistence puzzle and a source of bias and inconsistency of parameters in dynamic gravity models (see Pesaran and Smith (1995) and our discussion in Sections 4.4 and 4.5). We now present Figure 2, which demonstrates the extent of the cross-country parameter heterogeneity in our sample as well as the associated country-specific uncertainty surrounding the coefficient estimates. Our focus is limited to the trade persistence and the trade imbalance coefficients calculated using the CCEMG approach described in Sections 4.2 and 4.3, which retains the country-pair-specific parameter heterogeneity and controls for the unobservable global factors. For the sake of clarity and space, the coefficient estimates specific



Figure 2: Cross-Country Heterogeneity of Trade Persistence and Trade Imbalance Coefficient Estimates (CCEMG)

Notes: Subplots (a) and (b) display the cross-country heterogeneity and the uncertainty surrounding the CCEMG coefficient estimates β_{1i} and β_{2i} , respectively, where *i* stands for the "source" country i = 1, 2, ..., N (i.e., the market from which exports originate). The magnitude of the red dots is measured by the vertical distance and denotes the CCEMG coefficient estimates specific to each source country *i*. The names of the source countries are displayed on the horizontal axis. The country-specific CCEMG coefficient estimates are calculated as an average across all *N* destinations indexed by *j* from which the source country *i* imports. The blue bars surrounding the CCEMG coefficient estimates are the 95% confidence intervals.

to each country pair are averaged across all destinations j for each source country i, resulting in N number of coefficient estimates that we report out of the total of N(N-1) number of unique country pairs for which coefficient estimates exist. In general, we establish pervasive country-specific parameter heterogeneity clustered around the average coefficient estimates presented in Table 2 with few and far between outliers.

Nearly all of the trade persistence coefficients turn out to be positive, statistically significant, and their value is scattered around the interval of -0.2 and 0.5 (see Figure 2a) compared to the cross-sectional average of 0.35 (see column (1) in Table 2). The country-specific estimates of the trade persistence coefficients contain two notable outliers, namely South Africa, where it is small, negative, and statistically significant, and Luxembourg, where it is not significantly different from zero. The trade persistence coefficients in all other countries are significantly different from zero and unity. Conversely, the dispersion of the estimated trade imbalance coefficients is considerably larger (see Figure 2b); namely, it ranges from around -2 in Peru to nearly 2.5 in Greece. The majority of the trade imbalance coefficients are statistically significant (i.e., 26 out of 39) and clustered around the cross-sectional average unitary elasticity. Other countries exhibit statistically insignificant trade imbalance coefficients.

Due to a relatively large number of coefficient estimates (i.e., N = 39), and the fact that the trade persistence coefficient estimate outliers are relatively small, the inference drawn from the cross-sectional average of the trade persistence coefficients is arguably not susceptible to the presence of those outliers. While there exist larger outliers of the trade imbalance coefficients, they are both positive outliers (e.g., Greece and Venezuela) as well as negative outliers (e.g., Cyprus, Peru, and South Africa). As a consequence, the inference drawn from the cross-sectional averages of the trade imbalance coefficients is largely unbiased by the presence of outliers. We also document a largely symmetric and fat-tailed distribution of the country-pair-specific trade imbalance coefficient estimates in Figure 4 in Appendix C. We find that for any given destination country, the bilateral trade imbalance coefficients are remarkably heterogeneous, which are likely to depend on the structural differences between source and destination economies. This implies that bilateral trade reforms may exhibit consequences for international trade flows and the corresponding trade imbalance of countries not directly targeted by the reforms. In fact, using a static gravity equation with homogeneous coefficients, Cunat and Zymek (2019) find that bilateral imbalances depend on aggregate imbalances only if they are explained jointly with the multilateral resistance terms and the structural differences, such as production and spending patterns or trade wedges, which points to the heterogeneous influence of the trade-network-wide factors analyzed in this paper.

5.3 Trade Persistence & Global Value Chains

The pervasive parameter heterogeneity presented in Section 5.2 raises an important question about what drives cross-country differences in the estimates of the trade persistence coefficients. This section of the paper shows that the theoretical model presented in Section 3, in principle, links the estimated trade persistence coefficients to the country pair-specific participation in Global Value Chains (GVCs). We also discuss the empirical obstacles we encounter when identifying the habit parameters from country-pair-specific estimates of the trade persistence coefficients.

The theoretical model presented in Section 3 distinguishes between two different indicators of

participation in GVCs that can be mapped directly to those measured by Casella et al. (2019) for instance. Specifically, the domestic value-added ($\text{DVA}_{ij,t}$) and the foreign value-added ($\text{FVA}_{ij,t}$) in domestic exports expressed as a share of domestic exports, such that $\text{DVA}_{ij,t} + \text{FVA}_{ij,t} = 1$. If the source country *i* is considered as the "domestic" economy and destination *j* is the "foreign" economy, then using equation (3.12), it can easily be shown that

$$FVA_{ij,t} = \frac{X_{ij,t} - M_{ij,t}}{X_{ij,t}} \equiv 1 - x_{ij,t-1}^{-\chi_{ij}} \in [0,1] \bigg|_{\chi_{ij} > 0},$$
(5.1)

where $X_{ij,t}$ $(M_{ij,t})$ are the nominal trade flows of final (intermediate) goods from origin *i* to destination *j*, while $x_{ij,t-1}$ are the analogous lagged real trade flows of final goods. It follows that $\text{FVA}_{ij,t}$ (DVA_{ij,t}) is increasing (decreasing) in the habits parameter χ_{ij} . As a consequence, the unobservable and time-invariant deep habit parameter χ_{ij} could in principle be mapped directly to the time-averages of DVA_{ij,t} and/or FVA_{ij,t}, because they are observable. Formally,

$$\chi_{ij} = -\frac{\ln\left(\frac{1}{T}\sum_{t=1}^{T} \text{DVA}_{ij,t}\right)}{\ln\left(\frac{1}{T}\sum_{t=1}^{T} x_{ij,t-1}\right)},\tag{5.2}$$

such that $\lim_{DVA_{ij}\to 1} \chi_{ij} = 0$ nests the classical "made here, sold there" case of arms length trade. Holding all else constant, the lower is the share of intermediate imports sourced from origin i in destination j (i.e., the closer $DVA_{ij} \equiv 1/T \sum_{t=1}^{T} DVA_{ij,t}$ is to unity), the closer is the deep habit parameter χ_{ij} to zero. And by extension, if destination j does not rely on intermediate imports from origin i, then the nominal value of trade flows from origin i to destination j are expected to be volatile rather than persistent, since the trade persistence coefficient is measured as $\beta_{1ij} = \chi_{ij}(\eta-1)$, where $\eta > 0$ by assumption (see equation (3.22)).

While in theory the mapping between χ_{ij} and DVA_{ij} is relatively straightforward, identifying the habit parameter χ_{ij} empirically is much more difficult. Notice that the trade persistence coefficient $\beta_{1ij} = \chi_{ij}(\eta - 1)$ identifies χ_{ij} and η jointly. One of the most common assumptions in modern trade theory is that the elasticity of substitution η is time-invariant and homogeneous across countries (e.g., Anderson and van Wincoop (2003)). Yet even the cross-sectionally averaged trade persistence coefficients presented in Figure 2 suggest evidence against this assumption. Specifically, if η was truly homogeneous across countries, then as long as $1 < \eta < \infty$, such that final imports from different source countries are considered to be imperfect substitutes, as is usually assumed to be the case, β_{1i} would be strictly non-negative, since $\chi_i > 0$ by definition. And yet in South Africa it is negative and statistically significant (see Figure 2) as is the case for some other bilateral trade persistence coefficients β_{1ij} (not displayed), not all of which can be treated as a sampling error. Consequently, direct mapping between χ_{ij} and DVA_{ij} is ultimately difficult to establish from the trade persistence coefficient estimates, because, contrary to the theory, our estimates suggest that parameter $\eta > 0$ is also likely to be country-specific rather than symmetric across all countries. Not least because our sample considers 39 developed and developing countries in which market structures are not only radically different at any given point in time, but also transformed at a heterogeneous pace over time.

To shed some light on the source of the country-specific heterogeneity in our trade persistence



Notes: The figure depicts habits derived from averaged domestic value added and trade flows data, as suggested in equation (5.2). We have used data from 39 countries over the period of 1990-2014.

	All	$t_{\beta_{1ij}} > 1.64$	$t_{\beta_{1ij}} > 1.96$	$t_{\beta_{1ij}} > 2.575$
	(1)	(2)	(3)	(4)
VARIABLES	$\ln \beta_{1ij}$	$\ln \beta_{1ij}$	$\ln \beta_{1ij}$	$\ln \beta_{1ij}$
$\ln \chi_{ij}$	-0.0536	0.0242	0.0304	0.0407^{**}
	(0.0463)	(0.0202)	(0.0194)	(0.0202)
Colony	0.568^{**}	0.198	0.122	0.214^{*}
	(0.225)	(0.180)	(0.191)	(0.122)
Common language	0.115^{*}	0.0205	0.0421	0.0134
	(0.0688)	(0.0388)	(0.0357)	(0.0335)
$\ln(\text{Distance})$	-0.101***	-0.0707***	-0.0489***	-0.0490***
	(0.0323)	(0.0174)	(0.0163)	(0.0147)
Constant	-0.404	0.208	0.0557	0.376
	(0.684)	(0.321)	(0.306)	(0.290)
Observations	1,302	923	864	725
R-squared	0.174	0.220	0.234	0.253

 Table 4: Persistence Parameters and Global Value Chains

Notes: Robust standard errors associated with the Huber/White/sandwich coefficient estimates are displayed in parentheses. All regression models incorporate source- and destination-country-specific fixed effects. *** p < 0.01, ** p < 0.05, * p < 0.1.

coefficient estimates, we extract the β_{1ij} estimates in equation (4.1), construct the theory-consistent χ_{ij} from equation (5.2), and then investigate the link between the two empirically. Figure 3 presents a distribution of the model-implied habit parameters. As expected, the habit parameters are indeed small (i.e., between 0 and 0.28) and the cross-country distribution appears to be skewed and bimodal, suggesting two distinct clusters: one close to zero and the other equal to around 0.12. In order to validate the theory of habits in the supply chains, it is important to understand whether such small habit parameter values can successfully explain the variation in our trade persistence coefficient estimates. If so, then the habit framework serves not only as a theoretical motivation for an empirical model of the dynamic gravity equation, but also as a theoretical tool that helps explain the excess trade persistence puzzle.

Our empirical strategy is motivated by the theoretical identity $\beta_{1ij} = \chi_{ij}(\eta - 1)$ that comes from the habit framework, where η is not directly observable, but expected to be country-specific as argued above. Moreover, the value of trade persistence coefficients β_{1ij} are estimates and are subject to estimation errors. Consequently, we conduct a three-step selection exercise using 10%, 5%, and 0.5% critical values from the one-tail Normal distribution. This amounts to running separate regression models, which select only those trade persistence coefficient estimates that pass the respective critical value thresholds. Our empirical model therefore projects the estimates of β_{1ij} not only on the model-implied χ_{ij} constructed from the GVC database, but also the country fixed effects, and other standard determinants of the trade elasticity in the gravity literature (from CEPII dataset Mayer and Zignago (2011)) that are likely to capture the cross-country variation in β_{1ij} . In particular, we control for "Colony", which assumes the value of unity if at least 9% of the population speak the same language in both countries; and "Distance", which captures the geographic distance between the capital cities in both countries.

Table 4 presents the results from the log-linearized version of $\beta_{1ij} = \chi_{ij}(\eta - 1)$. According to our calculations, the habit parameter χ_{ij} exhibits positive conditional correlation with the identified β_{1ij} if statistically significant estimates are used (see the line marked with $\ln(\chi_{ij})$ in Table 4). Trade flows appear to be more persistent if countries share colonial ties and/or a common language (see the lines marked with "Colony" and "Common language" in Table 4). By contrast, greater geographic distance is associated with a significantly lower trade persistence between countries (see the line marked with "ln (Distance)" in Table 4). This implies that greater geographic proximity leads to more persistent value of international trade flows over time. We also implement a non-linear version of the auxiliary trade persistence model in multiplicative form using the Poisson regressions (see Table 9 in Appendix C). The Poisson and the log-linearized model results are closely aligned, thereby further reinforcing the robustness of the outcomes we report. While we demonstrate that heterogeneity of the structural trade persistence coefficient plays a prominent role in explaining the excess trade persistence puzzle, which may be related to GVCs through habits in the supply chains, more elaborate evidence is needed to fully characterize the structural differences in trade elasticities across different country pairs (see Boehm et al. (2020) for recent evidence on the time-variation of trade elasticities).

6 Concluding Remarks

International trade flows are volatile, imbalanced, and fragmented across off-shored supply chains. Yet, not much is known about the mechanism through which trade flows adjust in response to shocks over time. As things stand, the bulk of the modern trade literature relies on the ubiquitous gravity equation to predict the value of trade flows across countries. And it is notoriously successful at predicting both "who trades with whom" as well as "how much is traded" when trade shocks are local or country-specific. But when trade shocks are global, the observed value of trade flows adjusts by more and more rapidly than predicted by the standard gravity equations presented in the literature. While the static gravity equation remains the workhorse framework for trade policy analysis in the context of permanent, one-off, and exogenous trade shocks, it is silent about the transitional dynamics. By contrast, the neo-classical gravity equation that relies on the theory of capital accumulation predicts excessively persistent international trade flows that are difficult to square with the sharp and synchronized trade adjustments in response to global shocks.

This paper derives a dynamic gravity equation from a theory of habits in supply chains. Our theory offers several advantages. First, habits predict autocorrelated trade flows, where the trade persistence coefficient is heterogeneous across different country pairs. Second, cross-country habit asymmetry creates differences in home-bias. This causes trade imbalance to drive the value of bilateral trade flows in addition to standard measures, such as aggregate income and geographic distance. Third, habits enhance the geographic distance component of trade costs, because distance applies not only to goods that are "made here, sold there", but also to intermediate inputs that are "bought, sold, and bought again". Fourth, habits create "inward" and "outward" multilateral trade resistance that is not only time-varying, but also enters the dynamic gravity equation in contemporaneous and lagged form. This leads to a fundamentally different transmission of local and global trade shocks, since multilateral trade resistance terms are strongly correlated with foreign demand, foreign supply, as well as trade imbalance, and capture the variation in the unobservable global factors. But despite these new channels, our model conveniently nests the leading rival models of the gravity equation.

We estimate the dynamic gravity equation for 39 countries over the period of 1950-2014 using several dynamic panel regression techniques that retain country-pair-specific parameter heterogeneity and control for the unobservable global factors. We show that in addition to the standard variables in the gravity equation, multilateral trade imbalance is an important determinant of bilateral trade flows both theoretically and empirically. We establish two root causes of the excess trade persistence puzzle. First, the standard fixed effects regression model does not appropriately account for the global shocks transmitted through the time-varying multilateral trade resistance specific to each country pair. This is because the "country" fixed effects are time-invariant, while the "time" fixed effects are homogeneous for all country pairs, such that trade flows for each country pair are counterfactually disconnected from trade shocks originating from third countries. Second, the pooled regression coefficients ignore the fact that some country pairs exhibit stronger habits than others, which is important because trade flows between some country pairs are significantly more persistent than others. Our results show that absent of the unobservable global factors, the value of the pooled trade persistence coefficient is 0.91, which is biased upwards almost three-fold relative to our generalized regression model specification that incorporates proxies for the unobservable global factors and retains parameter heterogeneity. We also establish that our empirical strategy, preserving more structurally-relevant variables relative to the main alternatives, is also capable to provide a better fit of the data, particularly so during the time periods characterized by global recessions.

Our empirical estimates document pervasive heterogeneity of the trade persistence coefficients across countries. We demonstrate that our theory predicts a direct mapping between habits and the trade persistence coefficients, which are related to the time-averaged participation in global value chains. Despite some success, the question of what drives the cross-country differences in the empirical estimates of the trade persistence coefficients remains an open discussion. While the habits model makes valuable progress in terms of resolving the excess trade persistence puzzle and offers an alternative framework to the neo-classical theory of trade persistence, in the end we call for a more structural approach to tackle the dynamics of the global trade network and heterogeneity in trade elasticities. In particular, we encourage more research aimed at separating the short- and the long-run run effects in trade elasticities, which may portray substantial structural heterogeneity as is recently illustrated by Boehm et al. (2020)). Another area that we forfeit to future research is dynamic non-linear panel regression models, which would be able to appropriately account for the "zero trade problem", but simultaneously retain parameter heterogeneity and enrich the model specification with unobservable global shocks.

References

- Abel, A. B. (1990). Asset Prices under Habit Formation and Catching Up with the Joneses. American Economic Review 80(2), 38–42.
- Alessandria, G., J. P. Kaboski, and V. Midrigan (2010). The Great Trade Collapse of 2008–09: An Inventory Adjustment? IMF Economic Review 58(2), 254–294.
- Altomonte, C., F. D. Mauro, G. Ottaviano, A. Rungi, and V. Vicard (2012). Global Value Chains During the Great Trade Collapse: A Bullwhip Effect? Working Paper Series 1412, European Central Bank.
- Alvarez, F. (2017). Capital Accumulation and International Trade. <u>Journal of Monetary</u> Economics 91(C), 1–18.
- Anderson, J. E. (1979). A Theoretical Foundation for the Gravity Equation. <u>American Economic</u> Review 75(1), 106–116.
- Anderson, J. E. (2011). The Gravity Model. Annual Review of Economics 3(1), 133–160.
- Anderson, J. E., M. Larch, and Y. V. Yotov (2020). Transitional Growth and Trade with Frictions: A Structural Estimation Framework. The Economic Journal 130(630), 1583–1607.
- Anderson, J. E. and E. van Wincoop (2003). Gravity with Gravitas: A Solution to the Border Puzzle. American Economic Review 93(1), 170–192.
- Anderson, J. E. and E. van Wincoop (2004). Trade Costs. <u>Journal of Economic Literature</u> <u>42</u>(3), 691–751.
- Anderson, J. E. and Y. V. Yotov (2010). The Changing Incidence of Geography. <u>American Economic</u> Review 100(5), 2157–86.
- Anderson, J. E. and Y. V. Yotov (2020). Short run gravity. <u>Journal of International Economics</u> <u>126</u>, 103341.
- Antonakakis, N. (2012). The Great Synchronization of International Trade Collapse. <u>Economics</u> <u>Letters</u> $\underline{117}(3)$, 608–614.
- Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New Trade Models, Same Old Gains? The American Economic Review 102(1), 94–130.
- Backus, D. K. and G. W. Smith (1993). Consumption and Real Exchange Rates in Dynamic Economies with Non-Traded Goods. Journal of International Economics 35(3), 297 316.
- Bai, J. (2009). Panel Data Models With Interactive Fixed Effects. Econometrica 77(4), 1229–1279.
- Baldwin, R. and D. Taglioni (2006). Gravity for Dummies and Dummies for Gravity Equations. Working Paper 12516, National Bureau of Economic Research.
- Beirne, J., N. Renzhi, and U. Volz (2020). Persistent Current Account Imbalances: Are they Good

or Bad for Regional and Global Growth? ADBI Working Papers 1094, Asian Development Bank Institute.

- Bems, R., R. C. Johnson, and K.-M. Yi (2010). Demand Spillovers and the Collapse of Trade in the Global Recession. IMF Economic Review 58(2), 295–326.
- Boehm, C. E., A. A. Levchenko, and N. Pandalai-Nayar (2020). The Long and Short (Run) of Trade Elasticities. Working Paper 27064, National Bureau of Economic Research.
- Campbell, J. Y. and J. H. Cochrane (1999). By Force of Habit: A Consumption?Based Explanation of Aggregate Stock Market Behavior. Journal of Political Economy 107(2), 205–251.
- Casella, B., R. Bolwijn, D. Moran, and K. Kanemoto (2019). Improving the Analysis of Global Value Chains: the UNCTAD-Eora Database. Transnational Corporations 26(3), 115–142.
- Chudik, A. and M. H. Pesaran (2015). Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors. <u>Journal of Econometrics 188</u>(2), 393–420.
- Cunat, A. and R. Zymek (2019). Bilateral Trade Imbalances. CESifo Working Paper Series 7823, CESifo.
- Davis, D. R. and D. E. Weinstein (2002). The Mystery of the Excess Trade (Balances). <u>American</u> Economic Review 92(2), 170–174.
- Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced Trade. <u>American Economic Review</u> <u>97(2)</u>, 351–355.
- Dekle, R., J. Eaton, and S. Kortum (2008). Global Rebalancing with Gravity: Measuring the Burden of Adjustment. IMF Staff Papers 55(3), 511–540.
- Disdier, A. C. and K. Head (2008). The Puzzling Persistence of the Distance Effect on Bilateral Trade. Review of Economics and Statistics 90(1), 37–48.
- Dix-Carneiro, R., J. P. Pessoa, R. Reyes-Heroles, and S. Traiberman (2020). Globalization, Trade Imbalances, and Labor Market Adjustment. 2018 Meeting Papers 890, NBER Meeting 2020.
- Eaton, J., S. Kortum, B. Neiman, and J. Romalis (2016). Trade and the Global Recession. <u>American</u> Economic Review 106(11), 3401–38.
- Eberhardt, M. and F. Teal (2013). No Mangoes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis. Oxford Bulletin of Economics and Statistics 75(6), 914–939.
- Egger, P. (2000). A note on the proper econometric specification of the gravity equation. Economics Letters 66(1), 25 31.
- Eichengreen, B. and D. Irwin (1998). The Role of History in Bilateral Trade Flows. In <u>The</u> Regionalization of the World Economy, pp. 33–62. National Bureau of Economic Research, Inc.

Feenstra, R. (2016). Advanced International Trade: Theory and Evidence Second Edition (2 ed.).

Princeton University Press.

- Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015). The Next Generation of the Penn World Table. American Economic Review 105(10), 3150–3182.
- Head, K. and T. Mayer (2014). Chapter 3 Gravity Equations: Workhorse, Toolkit, and Cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), <u>Handbook of International Economics</u>, Volume 4 of Handbook of International Economics, pp. 131 195. Elsevier.
- Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. The Quarterly Journal of Economics 123(2), 441–487.
- Herbst, E. P. and F. Schorfheide (2016). <u>Bayesian Estimation of DSGE Models</u> (1 ed.). The Econometric and Tinbergen Institutes Lectures. Princeton University Press.
- Imbs, J. and I. Mejean (2017). Trade Elasticities. <u>Review of International Economics</u> <u>25</u>(2), 383–402.
- IMF (2015). Making Public Investment More Efficient. IMF Policy Papers, International Monetary Fund, Washington, D.C.
- IMF (2019). Global Imbalances. IMF Staff Note, International Monetary Fund, Washington, D.C.
- Kapetanios, G., C. Mastromarco, L. Serlenga, and Y. Shin (2017). <u>Modelling in the Presence of Cross-Sectional Error Dependence</u>, pp. 291–322. Cham: Springer International Publishing.
- Kapetanios, G., M. H. Pesaran, and T. Yamagata (2011). Panels with Non-Stationary Multifactor Error Structures. Journal of Econometrics 160(2), 326–348.
- Kose, M. A., N. Sugawara, and M. E. Terrones (2020, February). Global Recessions. CEPR Discussion Papers 14397, C.E.P.R. Discussion Papers.
- Levchenko, A. A., L. T. Lewis, and L. L. Tesar (2010). The Collapse of International Trade During the 2008-09 Crisis: In Search of the Smoking Gun. IMF Economic Review 58(2), 214–253.
- Mayer, T. and S. Zignago (2011, December). Notes on CEPII s distances measures The GeoDist database. Working Papers 2011-25, CEPII research center.
- Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71(6), 1695–1725.
- Micco, A., E. Stein, and G. Ordonez (2003, October). The currency union effect on trade: early evidence from EMU. Economic Policy 18(37), 315–356.
- Novy, D. and A. M. Taylor (2020, October). Trade and Uncertainty. <u>The Review of Economics</u> and Statistics 102(4), 749–765.
- Obstfeld, M. and K. S. Rogoff (1996). <u>Foundations of International Macroeconomics</u>, Volume 1 of MIT Press Books. The MIT Press.

- Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. Econometrica 74(4), 967–1012.
- Pesaran, M. H. and R. Smith (1995). Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. Journal of Econometrics 68(1), 79–113.
- Piermartini, R. and Y. Yotov (2016, July). Estimating Trade Policy Effects with Structural Gravity. School of Economics Working Paper Series 2016-10, LeBow College of Business, Drexel University.
- Ravn, M., S. Schmitt-Grohé, and M. Uribe (2006). Deep Habits. <u>The Review of Economic</u> Studies 73(1), 195–218.
- Ravn, M. O., S. Schmitt-Grohe, and M. Uribe (2007). Pricing to Habits and the Law of One Price. American Economic Review 97(2), 232–238.
- Santos-Silva, J. and S. Tenreyro (2007). The Log of Gravity. <u>Review of Economics and Statistics</u> 88, 641–658.
- Serlenga, L. and Y. Shin (2007). Gravity Models of Intra-EU Trade: Application of the CCEP-HT Estimation in Heterogeneous Panels with Unobserved Common Time-Specific Factors. <u>Journal</u> of Applied Econometrics 22(2), 361–381.
- Weidner, M. and T. Zylkin (2019). Bias and Consistency in Three-Way Gravity Models. Papers 1909.01327, arXiv.org.
- Westerlund, J. (2018). CCE in Panels with General Unknown Factors. <u>The Econometrics</u> Journal 21(3), 264–276.
- Westerlund, J. and F. Wilhelmsson (2009). Estimating the gravity model without gravity using panel data. Applied Economics 43(6), 641–649.
- Yotov, Y. V. and M. Olivero (2012). Dynamic Gravity: Endogenous Country Size and Asset Accumulation. Canadian Journal of Economics 45(1), 64–92.
- Zwinkels, R. C. and S. Beugelsdijk (2010). Gravity Equations: Workhorse or Trojan Horse in Explaining Trade and FDI Patterns Across Time and Space? <u>International Business Review</u> <u>19</u>(1), 102 – 115.

A Theoretical Model: Technical Details

A.1 Distributor

The production technology of a distributor adopts the following functional form:

$$x_{ij,t} = \left[\int_{0}^{1} \left(m_{ij,t}(\omega) x_{ij,t-1}^{\chi_{ij}}\right)^{1-1/\eta} d\omega\right]^{1/(1-1/\eta)},$$
(A.1)

The distributor operates in a perfectly competitive market structure, such that they minimize production costs by choosing the amount of commodities to import form each sector subject to the above augmented CES production technology

$$\min_{\{m_{ij,t}(\omega)\}} \tilde{P}_{ij,t} x_{ij,t} - \int_{0}^{1} P_{ij,t}(\omega) m_{ij,t}(\omega) d\omega$$
s.t. $x_{ij,t} = \left[\int_{0}^{1} \left(m_{ij,t}(\omega) x_{ij,t-1}^{\chi_{ij}} \right)^{1-1/\eta} d\omega \right]^{1/(1-1/\eta)}$. (A.2)

The first order condition is given by:

$$\tilde{P}_{ij,t}x_{ij,t}^{1/\eta}(m_{ij,t}(\omega)x_{ij,t-1}^{\chi_{ij}})^{-1/\eta}x_{ij,t-1}^{\chi_{ij}} - P_{ij,t}(\omega) = 0,$$
(A.3)

$$\Rightarrow m_{ij,t}(\omega) = \left[\frac{P_{ij,t}(\omega)}{\tilde{P}_{ij,t}}\right]^{-\eta} x_{ij,t} x_{ij,t-1}^{\chi_{ij}(\eta-1)}, \tag{A.4}$$

Distributors break-even when the total revenue is equal to the total costs:

$$\tilde{P}_{ij,t}x_{ij,t} = \int_{0}^{1} P_{ij,t}(\omega)m_{ij,t}(\omega)d\omega$$
(A.5)

The break-even price index of the distributors is then derived by substituting the demand for intermediate imports into the 'zero-profit' condition:

$$\tilde{P}_{ij,t} = \left[\int_{0}^{1} (P_{ij,t}(\omega)x_{ij,t-1}^{-\chi_{ij}})^{1-\eta}d\omega\right]^{1/(1-\eta)}.$$
(A.6)

The aggregate demand for intermediate imports is therefore derived by integrating across all varieties of intermediate imports:

$$m_{ij,t} = \int_{0}^{1} m_{ij,t}(\omega) d\omega,$$

$$= x_{ij,t} x_{ij,t-1}^{\chi_{ij}(\eta-1)} \int_{0}^{1} \left[\frac{P_{ij,t}(\omega)}{\tilde{P}_{ij,t}} \right]^{-\eta} d\omega,$$

$$= x_{ij,t} x_{ij,t-1}^{-\chi_{ij}}, \qquad (A.7)$$

such that the dynamic demand for aggregate imports is given by

$$x_{ij,t} = m_{ij,t} x_{ij,t-1}^{\chi_{ij}}.$$
(A.8)

A.2 Duality Problem

The representative consumer minimizes the consumption expenditure on composite goods from each source country subject to CES preferences:

$$\min_{\{x_{ij,t}\}} P_{j,t} c_{j,t} - \sum_{i=1}^{N} \tilde{P}_{ij,t} x_{ij,t}$$

s.t. $c_{j,t} = \left[\sum_{i=1}^{N} x_{ij,t}^{1-1/\eta}\right]^{1/(1-1/\eta)}$

The first-order condition with respect to the demand for a composite good $x_{ij,t}$ from any source country i = 1, ..., N is given by

$$P_{j,t}c_{j,t}^{1/\eta}x_{ij,t}^{-1/\eta} - \tilde{P}_{ij,t} = 0.$$
(A.9)

Rearranging the above gives the demand schedule for each composite tradable good:

$$x_{ij,t} = c_{j,t} \left(\frac{\tilde{P}_{ij,t}}{P_{j,t}}\right)^{-\eta},\tag{A.10}$$

The consumer price index is derived by substituting the above demand schedule into the CES preferences displayed above and solving for $P_{j,t}$, which gives rise to the following expression:

$$P_{j,t} = \left[\sum_{i=1}^{N} \tilde{P}_{ij,t}^{1-\eta}\right]^{1/(1-\eta)}.$$
(A.11)

A.3 Consumption Smoothing

The consumer maximizes the lifetime utility subject to an indefinite sequence of budget constraints by choosing the aggregate consumption and the aggregate stock of bonds:

$$\max_{\{c_{j,t},b_{j,t+1}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \log\left(c_{j,t}\right) \quad \text{s.t.} \ c_{j,t} + \mathbb{E}_t [\zeta_{j,t,t+1} b_{j,t+1}] = b_{j,t} + w_{j,t} h_j + \varpi_{j,t},$$

which is re-written in the form of a Current Value Lagrangian:

$$\max_{\{c_{j,t},b_{j,t+1}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\log \left(c_{j,t} \right) + \Lambda_{j,t} \left(b_{j,t} + w_{j,t} h_j + \varpi_{j,t} - c_{j,t} - \mathbb{E}_t [\zeta_{j,t,t+1} b_{j,t+1}] \right) \right].$$

The first order conditions define the shadow price of consumption and the stochastic discount factor:

$$\frac{1}{c_{j,t}} - \Lambda_{j,t} = 0 \quad \Leftrightarrow \quad \Lambda_{j,t} = \frac{1}{c_{j,t}},\tag{A.12}$$

$$\Lambda_{j,t}\zeta_{j,t,t+1} - \beta \mathbb{E}_t \left[\Lambda_{j,t+1}\right] = 0 \quad \Leftrightarrow \quad 1 = \beta \mathbb{E}_t \left[\frac{c_{j,t}}{\zeta_{j,t,t+1}c_{j,t+1}}\right]. \tag{A.13}$$

A.4 Wholesalers

The optimal nominal flexible price of the intermediate good exporters $P_{ij,t}(\omega)$, who adopt the "pricing-to-habits" strategy, is one that maximizes the current monopolistically-competitive profit dividends denominated in local currency units:

$$\max_{\{P_{ij,t}(\omega)\}} [P_{ij,t}(\omega) - d_{ij}MC_{i,t}] m_{ij,t}(\omega)$$

s.t. $m_{ij,t}(\omega) = \left[\frac{P_{ij,t}(\omega)}{\tilde{P}_{ij,t}}\right]^{-\eta} x_{ij,t} x_{ij,t-1}^{\chi_{ij}(\eta-1)}.$

The first-order conditions with respect to the nominal price $P_{ij,t}(\omega)$ is given by

$$(1-\eta)m_{ij,t}(\omega) + \eta \left(\frac{d_{ij}MC_{i,t}}{P_{ij,t}(\omega)}\right)m_{ij,t}(\omega) = 0,$$
(A.14)

or alternatively

$$P_{ij,t}(\omega) = \left(\frac{\eta}{\eta - 1}\right) d_{ij} M C_{i,t}.$$
(A.15)

A.5 Transversality Condition

Consider iterating the household budget constraint forwards in the symmetric equilibrium:

$$b_{j,t} = \zeta_{j,t,t+1} \mathbb{E}_t [b_{j,t+1} + \underbrace{c_{j,t} - \varpi_{j,t} - w_{j,t}h_j}_{nx_{j,t}}],$$

$$= \zeta_{j,t,t+1} \mathbb{E}_t [\zeta_{j,t+1,t+2}(b_{j,t+2} - nx_{j,t+1}) - nx_{j,t}],$$

$$= \zeta_{j,t,t+1} \zeta_{j,t+1,t+2} b_{j,t+2} - \zeta_{j,t,t+1}(nx_{j,t} + \zeta_{j,t+1,t+2}nx_{j,t+1}),$$
...
$$= \zeta_{j,t,t+S} b_{j,t+S} - \sum_{s=0}^{S} \zeta_{j,t,t+s+1}nx_{j,t+s}.$$
(A.16)

Next, note that the stochastic discount factor $\zeta_{j,t,t+S} \in (0,1)$ for all s = 1, 2, ..., S as long as the real rate of interest is strictly non-negative. Assuming that foreign economies would only be willing to lend to the domestic economy at a positive rate of interest, it follows that

$$\lim_{S \to \infty} \zeta_{j,t,t+S} = \zeta_{j,t,t+1} \times \zeta_{j,t+1,t+2} \times \zeta_{j,t+2,t+3} \times \dots \times \zeta_{j,S-1,S} = 0.$$
(A.17)

As a result, the stock of debt is clearly non-explosive. To fully convince yourself, consider evaluating the iterated form of the budget constraint along the balanced growth path:

$$b_j = \beta^S b_j - nx_j \sum_{s=0}^S \beta^{1+s} \quad \Rightarrow \quad \lim_{S \to \infty} b_j = -nx_j \sum_{s=0}^\infty \beta^{1+s}, \tag{A.18}$$

$$= -nx_j\left(\frac{\beta}{1-\beta}\right) > -\infty.\Big|_{\beta \in (0,1)}$$
(A.19)

B Dynamic Gravity Equation

Consider the optimal demand for imports, the aggregate consumption identity, and the break-even price index, respectively:

$$X_{ij,t} = P_{ij,t} x_{ij,t} = C_{j,t} \left[\frac{\tilde{P}_{ij,t}}{P_{j,t}} \right]^{1-\eta},$$
(B.1)

$$C_{j,t} = Y_{j,t} \Xi_{j,t}, \tag{B.2}$$

$$\tilde{P}_{ij,t} = P_{ij,t} x_{ij,t-1}^{-\chi_{ij}}.$$
(B.3)

Now substitute (B.3) and (B.2) into (B.1) to obtain

$$X_{ij,t} = Y_{j,t} \Xi_{j,t} \left[\frac{P_{ij,t} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}} \right]^{1-\eta}.$$
 (B.4)

Next, consider the aggregate income identity:

$$Y_{i,t} = \sum_{j=1}^{N} Q_{ji,t} X_{ij,t},$$
(B.5)

Substituting (B.4) into (B.5) gives

$$Y_{i,t} = \sum_{j=1}^{N} Q_{ji,t} Y_{j,t} \Xi_{j,t} \left[\frac{P_{ij,t} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}} \right]^{1-\eta},$$
(B.6)

Note that the import price $P_{ij,t}$ is proportional to export price $P_{ii,t}$, where the proportionality corresponds to the iceberg costs:

$$P_{ij,t} = d_{ij}P_{ii,t}.\tag{B.7}$$

Substituting (B.7) into (B.6) gives

$$Y_{i,t} = P_{ii,t}^{1-\eta} \sum_{j=1}^{N} Q_{ji,t} Y_{j,t} \Xi_{j,t} \left[\frac{d_{ij} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}} \right]^{1-\eta},$$
(B.8)

Now let $\theta_{i,t} = Y_{i,t}/Y_t$, where $Y_t = \sum_{j=1}^N Y_{j,t}$. Then solving (B.8) for the export price scaled by the trade elasticity $P_{i,t}^{1-\eta}$ gives

$$P_{ii,t}^{1-\eta} = \frac{Y_{i,t}}{\sum_{j=1}^{N} Q_{ji,t} Y_{j,t} \Xi_{j,t} \left[\frac{d_{ij} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}}\right]^{1-\eta}},$$

$$= \frac{\theta_{i,t}}{\sum_{j=1}^{N} Q_{ji,t} \theta_{j,t} \Xi_{j,t} \left[\frac{d_{ij} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}}\right]^{1-\eta}},$$

$$= \theta_{i,t} \Phi_{i,t}^{\eta-1},$$
(B.9)

where

$$\Phi_{i,t} = \left[\sum_{j=1}^{N} \theta_{j,t} Q_{ji,t} \Xi_{j,t} \left(\frac{d_{ij} x_{ij,t-1}^{-\chi_{ij}}}{P_{j,t}}\right)^{1-\eta}\right]^{1/(1-\eta)}$$
(B.10)

defines the 'multilateral resistance' of destination i to trade flows from all trade partner countries $j \in n \setminus i$. Next, substitute (B.9) out of (B.4) using the proportionality condition (B.7) to obtain

$$X_{ij,t} = \Xi_{j,t} \left[\frac{d_{ij} x_{ij,t-1}^{-\chi_{ij}}}{\Phi_{i,t} P_{j,t}} \right]^{1-\eta} \frac{Y_{i,t} Y_{j,t}}{Y_t}.$$
 (B.11)

Finally, note that

$$X_{ij,t-1} = P_{ij,t-1}x_{ij,t-1} = d_{ij}P_{ii,t-1}x_{ij,t-1},$$

such that the stock of habits can be replaced by

$$x_{ij,t-1}^{-\chi_{ij}} = \left(\frac{X_{ij,t-1}}{d_{ij}P_{ii,t-1}}\right)^{-\chi_{ij}},$$

= $\left(\frac{X_{ij,t-1}\Phi_{i,t-1}}{d_{ij}\theta_{i,t-1}^{1/(1-\eta)}}\right)^{-\chi_{ij}}.$ (B.12)

Substituting (B.12) into (B.11) therefore gives a dynamic gravity equation:

$$X_{ij,t} = \Xi_{j,t} \left[\frac{d_{ij}^{1+\chi_{ij}}}{\Phi_{i,t} \Phi_{i,t-1}^{\chi_{ij}} P_{j,t}} \right]^{1-\eta} \left[\frac{\theta_{i,t-1}}{X_{ij,t-1}^{1-\eta}} \right]^{\chi_{ij}} \frac{Y_{i,t} Y_{j,t}}{Y_t},$$
(B.13)

$$\frac{X_{ij,t}Y_t}{Y_{i,t}Y_{j,t}} = A_{ij,t} = \Xi_{j,t} \left[\frac{d_{ij}^{1+\chi_{ij}}}{\Phi_{i,t}\Phi_{i,t-1}^{\chi_{ij}}P_{j,t}} \right]^{1-\eta} \left[\frac{\theta_{i,t-1}}{X_{ij,t-1}^{1-\eta}} \right]^{\chi_{ij}},$$
(B.14)

$$=\Xi_{j,t} \left[\frac{d_{ij}^{1+\chi_{ij}}}{\Phi_{i,t}\Phi_{i,t-1}^{\chi_{ij}}P_{j,t}} \right]^{1-\eta} \left[\frac{\theta_{i,t-1}^{\eta}}{A_{ij,t-1}^{1-\eta}Y_{j,t-1}^{1-\eta}} \right]^{\chi_{ij}},$$
(B.15)

since $A_{ij,t}^{1-\eta} = [X_{ij,t}Y_t/(Y_{i,t}Y_{j,t})]^{1-\eta} = [X_{ij,t}/(\theta_{i,t}Y_{j,t})]^{1-\eta}$, thus $X_{ij,t}^{1-\eta}/\theta_{i,t} = A_{ij,t}^{1-\eta}Y_{j,t}^{1-\eta}/\theta_{i,t}^{\eta}$. Taking natural logs on both sides of (B.13) thus gives rise to the dynamic gravity equation regression function specification:

$$\ln A_{ij,t} = \underbrace{\chi_{ij}(\eta - 1) \ln A_{ij,t-1}}_{\text{size-adjusted bilateral trade flow persistence}} + \underbrace{\ln(\Xi_{j,t})}_{\text{multilateral trade imbalance}} - \underbrace{(1 + \chi_{ij})(\eta - 1) \ln d_{ij} + (\eta - 1)P_{j,t} + (\eta - 1) \ln \Phi_{i,t} + \chi_{ij}(\eta - 1) \ln \Phi_{i,t-1}}_{\text{bilateral and multilateral trade resistance}} + \underbrace{\chi_{ij}\eta \ln Y_{t-1} - \chi_{ij}\eta \ln Y_{i,t-1} + \chi_{ij}(\eta - 1) \ln Y_{j,t-1}}_{\text{aggregate income}},$$
(B.16)

since $\ln \theta_{i,t-1} = \ln Y_{i,t-1} - \ln Y_{t-1}$.

Methodologies)
Common
(Other
Model
Gravity
Dynamic
Table 5:

VARIABLES	$\begin{array}{c} {\rm FE2} \\ (1) \\ {\rm FLOW}_{ij,t} \end{array}$	$\begin{array}{c} {\rm FE3} \\ (2) \\ {\rm FLOW}_{ij,t} \end{array}$	$\begin{array}{c} {\rm FE4} \\ (3) \\ {\rm FLOW}_{ij,t} \end{array}$	$\begin{array}{c} \operatorname{PPML} \\ (4) \\ \operatorname{FLOW}_{ij,t} \end{array}$	$\begin{array}{c} \operatorname{PPML2} \\ (5) \\ \operatorname{FLOW}_{ij,t} \end{array}$	$\begin{array}{c} \operatorname{PPML3} \\ (6) \\ \operatorname{FLOW}_{ij,t} \end{array}$	$\begin{array}{c} \text{PPML4} \\ (7) \\ \text{FLOW}_{ij,t} \end{array}$
${ m FLOW}_{ij,t-1}$ ${ m TB}_{j,t}$ ${ m GDP}_{i,t-1}$ ${ m GDP}_{j,t-1}$ ${ m GDP}_{t-1}$	0.908^{***} (0.00322)	$\begin{array}{c} 0.743^{***} \\ (0.00632) \\ 0.418^{***} \\ (0.0334) \\ -0.0198^{**} \\ (0.00963) \\ -0.0529^{***} \\ (0.00935) \end{array}$	0.682^{***} (0.00732)	$\begin{array}{c} 0.956^{***}\\ (0.00327)\\ 0.166^{***}\\ (0.0404)\\ -0.00798\\ (0.0123)\\ -0.00602\\ (0.0158)\end{array}$	0.965^{***} (0.00264)	$\begin{array}{c} 0.770^{***} \\ (0.00939) \\ 0.320^{***} \\ (0.0390) \\ -0.0348^{***} \\ (0.0114) \\ -0.0759^{***} \\ (0.0153) \end{array}$	0.743^{***} (0.00971)
Time fixed effects	Ν	Υ	Ν	Υ	Z	Υ	N
Country fixed effects	Ν	Υ	Ν	Y	Z	Υ	N
Time-varying country fixed effects	Υ	Ν	Υ	Z	Υ	Z	Υ
Pair fixed effects	N	Υ	Υ	Z	Z	Υ	Υ
Unobservable Global Factors	Ν	Ν	Ν	N	Z	N	N
Constant	1.110^{**} (0.0406)	3.987^{***} (0.194)	3.900^{***} (0.0906)	0.764^{**} (0.218)	0.497^{***} (0.0392)	4.536^{**} (0.289)	3.790^{***} (0.144)
Observations Number of pairs R-squared <i>Note:</i> Robust standard errors in parent PPML specifications.	70,604 1,480 0.92 theses; $FLOV$	$\begin{array}{c} 70,602\\ 1,480\\ 0.91\\ V_{ij,t} \ \mathrm{measured} \end{array}$	70,602 1,480 0.93 I in levels and	71,365 1,487 0.96 I only non-zei	71,312 1,485 0.97 ro lagged trac	71,364 1,487 0.96 le flows are re	71,311 1,486 0.97 :tained in all

C Annex

Taura	• بالاللقالية •	diavity in	IT AA) TAMATA			(an	
	CCEMG (1)	FE	MG	CCEP (4)	$\begin{array}{c} \text{HFE} \\ (5) \end{array}$	AMG (6)	CCEMGR (7)
VARIABLES	$FLOW_{ij,t}$	$FLOW_{ij,t}$	$\mathrm{FLOW}_{ij,t}$	$FLOW_{ij,t}$	$FLOW_{ij,t}$	$FLOW_{ij,t}$	$\mathrm{FLOW}_{ij,t}$
$\mathrm{FLOW}_{ij,t-1}$	0.435^{***}	0.908***	0.584^{***}	0.443^{***}	0.527^{***}	0.475^{***}	0.496^{***}
$\mathrm{TB}_{j,t}$	(0.00789)	(0.00448)	(0.00609)	(0.0160)	(0.00684)	(0.00704)	(0.00736)
${ m GDP}_{i,t-1}$	-0.269^{***}	-0.00182	-0.181^{***}	-0.269^{***}	-0.231^{***}	-0.194^{***}	-0.208***
- 2 c - 1	(0.0612)	(0.00754)	(0.0143)	(0.0325)	(0.0262)	(0.0277)	(0.0399)
${ m GDP}_{j,t-1}$	-0.102	-0.0322***	-0.0812^{***}	-0.142^{***}	-0.0705***	0.0405	-0.0150
	(0.0725)	(0.00712)	(0.0135)	(0.0241)	(0.0265)	(0.0294)	(0.0434)
GDP_{t-1}	0.354^{**}	-0.0270	0.251^{***}		0.338^{***}	-0.0513	0.487^{***}
	(0.181)	(0.0258)	(0.0376)		(0.0884)	(0.102)	(0.154)
Time Fixed Effects	Z	Υ	Z	Z	N	Z	Z
Country Fixed Effects	Υ	Y	N	Ζ	Υ	Υ	Υ
Unobservable Global Factors	Y	Z	Z	Y	Z	Y	Υ
Constant	-3.434	1.860^{***}	3.887^{***}			8.950^{***}	-3.309^{*}
	(2.669)	(0.352)	(0.498)			(1.459)	(1.964)
Observations	70,591	70,604	62,211	70,560	70,579	70,579	70,591
Number of pairs	1,475	1,480	1,167	1,471	1,473	1,473	1,475
Adj. R-squared	0.77	0.90	0.74	0.79	0.73		0.76
<i>Note</i> : Robust standard errors in	parentheses.						
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.05$	0.1						

Imbalance)
Trade
(Without
Model
Gravity
Dynamic
Table 6:

	Full Sample		"Bad 7	Γimes"			"Good	Times"	
Method		w = 0	w = 1	w = 2	w = 3	w = 0	w = 1	w = 2	w = 3
CCEMG MC	0.38	0.41	0.41	0.40	0.41	0.38	0.37	0.37	0.35
CCEP	0.44	$0.30 \\ 0.47 \\ 0.45$	0.49	0.48	$0.48 \\ 0.47 \\ 0.50$	0.44	0.43 0.42	0.42	0.40
${ m FE}$ ${ m FE2}$	$\begin{array}{c} 0.55 \\ 0.52 \end{array}$	$\begin{array}{c} 0.65 \\ 0.61 \end{array}$	$\begin{array}{c} 0.63 \\ 0.58 \end{array}$	$\begin{array}{c} 0.60\\ 0.56\end{array}$	$\begin{array}{c} 0.59 \\ 0.55 \end{array}$	$\begin{array}{c} 0.54 \\ 0.50 \end{array}$	$\begin{array}{c} 0.52 \\ 0.49 \end{array}$	$\begin{array}{c} 0.51 \\ 0.47 \end{array}$	$\begin{array}{c} 0.49 \\ 0.46 \end{array}$
FE3 FE4	$\begin{array}{c} 0.53 \\ 0.49 \end{array}$	$\begin{array}{c} 0.61 \\ 0.56 \end{array}$	$\begin{array}{c} 0.59 \\ 0.54 \end{array}$	$\begin{array}{c} 0.57 \\ 0.52 \end{array}$	$\begin{array}{c} 0.56 \\ 0.51 \end{array}$	$\begin{array}{c} 0.51 \\ 0.47 \end{array}$	$\begin{array}{c} 0.50 \\ 0.46 \end{array}$	$\begin{array}{c} 0.49 \\ 0.45 \end{array}$	$\begin{array}{c} 0.47 \\ 0.43 \end{array}$
PPML PPML2	$\begin{array}{c} 0.54 \\ 0.44 \end{array}$	$0.67 \\ 0.65$	$\begin{array}{c} 0.64 \\ 0.61 \end{array}$	$\begin{array}{c} 0.61 \\ 0.60 \end{array}$	$0.61 \\ 0.59$	$\begin{array}{c} 0.55 \\ 0.53 \end{array}$	$0.53 \\ 0.52$	$\begin{array}{c} 0.52 \\ 0.50 \end{array}$	$\begin{array}{c} 0.50 \\ 0.48 \end{array}$
PPML3 PPML4	0.44 0.34	$\begin{array}{c} 0.65 \\ 0.62 \end{array}$	$\begin{array}{c} 0.62 \\ 0.59 \end{array}$	$\begin{array}{c} 0.59 \\ 0.57 \end{array}$	$0.59 \\ 0.57$	$\begin{array}{c} 0.55 \\ 0.53 \end{array}$	$0.53 \\ 0.52$	$\begin{array}{c} 0.53 \\ 0.52 \end{array}$	$0.52 \\ 0.51$

 Table 7: Root Mean Square Error (Full Sample, Extensive List of Methods)

Note: This figure presents the Root Mean Square Errors (RMSE) calculated using different methods of estimating the coefficients in a dynamic gravity equation. The in-sample RMSEs are presented for the full data sample, the observed "good times", and the observed "bad times" in order to compare different model performance inside and outside of time periods characterized by global trade shocks. Consistent with Kose et al. (2020), the "bad times" represent the global recession years, namely 1975, 1982, 1991, and 2009, while the "good times" are all of the remaining years in our data sample that spans 1950-2014. The term $w = \{0, 1, 2, 3\}$ further indicates the length of the windows surrounding the recession years (i.e., number of years before and after global trade shocks that are included in "bad times" in addition to the recession years). The values in bold indicate the smallest RMSE.

	Full Sample		"Bad 7	Γimes"			"Good	Times"	
Method		w = 0	w = 1	w = 2	w = 3	w = 0	w = 1	w = 2	w = 3
CCEMG	0.38	-	0.09	0.26	0.32	0.37	0.34	0.30	0.25
MG	0.44	-	0.32	0.39	0.42	0.43	0.41	0.39	0.33
CCEP	0.47	-	-	0.44	0.47	0.46	0.44	0.42	0.38
\mathbf{FE}	0.55	0.65	0.62	0.60	0.59	0.54	0.52	0.51	0.49
FE2	0.52	0.62	0.60	0.58	0.57	0.52	0.50	0.49	0.47
FE3	0.53	0.60	0.60	0.57	0.57	0.52	0.50	0.49	0.47
FE4	0.49	0.57	0.56	0.54	0.53	0.49	0.47	0.46	0.44
PPML	0.54	0.53	0.70	0.64	0.60	0.54	0.49	0.49	0.49
PPML2	0.44	0.50	0.60	0.54	0.51	0.44	0.40	0.40	0.39
PPML3	0.44	0.37	0.44	0.44	0.43	0.43	0.40	0.39	0.47
PPML4	0.34	0.35	0.35	0.35	0.34	0.34	0.32	0.31	0.44

Table 8: Root Mean Square Error (Sub-Samples, Extensive List of Methods)

Note: This figure presents the Root Mean Square Errors (RMSE) calculated using different methods of estimating the coefficients in a dynamic gravity equation. The in-sample RMSEs are presented for the full data sample, the observed "good times", and the observed "bad times" in order to compare different model performance inside and outside of time periods characterized by global trade shocks. Consistent with Kose et al. (2020), the "bad times" represent the global recession years, namely 1975, 1982, 1991, and 2009, while the "good times" are all of the remaining years in our data sample that spans 1950-2014. The term $w = \{0, 1, 2, 3\}$ further indicates the length of the windows surrounding the recession years (i.e., number of years before and after global trade shocks that are included in "bad times" in addition to the recession years). The values in bold indicate the smallest RMSE.

Figure 4: Distribution of Trade Imbalance Coefficient Estimates

Note: The figure presents CCEMG estimates of the trade imbalance coefficient (β_{2ij}) in the dynamic gravity model presented in equations (4.1)-(4.3). The trade imbalance coefficient estimates characterize 39 countries (i.e., up to 1482 country pairs) over the period of 1950-2014. Some country pair estimates are highlighted with according abbreviations.

	All	$t_{\beta_{1ij}} > 1.64$	$t_{\beta_{1ij}} > 1.96$	$t_{\beta_{1ij}} > 2.575$
	(1)	(2)	(3)	(4)
VARIABLES	$\beta_{1ij} > 0$	$\beta_{1ij} > 0$	$\beta_{1ij} > 0$	$\beta_{1ij} > 0$
$\ln \chi_{ij}$	0.0532	0.0438	0.0470^{*}	0.0555^{*}
	(0.0485)	(0.0271)	(0.0275)	(0.0286)
Colony	0.365^{*}	0.194	0.130	0.189
	(0.187)	(0.165)	(0.174)	(0.153)
Common language	0.0809^{*}	0.0397	0.0536	0.0241
	(0.0459)	(0.0372)	(0.0355)	(0.0327)
$\ln(\text{Distance})$	-0.104***	-0.0634***	-0.0484***	-0.0474***
	(0.0208)	(0.0169)	(0.0162)	(0.0149)
Constant	0.853	0.421	0.300	0.559
	(0.615)	(0.353)	(0.351)	(0.346)
Observations	1,304	925	866	727
Pseudo R-squared	0.020	0.018	0.017	0.017
Notes: Robust standa	rd errors ass	ociated with the	e Huber/White	/sandwich coef-

Table 9: The Poisson Model of Persistence Parameters and Global Value Chains

Notes: Robust standard errors associated with the Huber/White/sandwich coefficient estimates are displayed in parentheses. All regression models incorporate source- and destination-country-specific fixed effects. *** p < 0.01, ** p < 0.05, * p < 0.1.