The Gravity of Experience*

Pushan Dutt†
INSEAD
and
London Business School

Ana Maria Santacreu
Federal Reserve Bank of St. Louis

Daniel A. Traca
NOVA School of Business and Economics

April 2016

Abstract

In this paper, we establish the importance of experience in international trade for reducing trade costs and facilitating bilateral trade. In an augmented gravity framework, with a very comprehensive set of fixed-effects and trend variables, we find that a 1% increase in experience at the country-pair level reduces trade costs by 0.12% and increases bilateral exports by 0.48%. We utilize multiple identification strategies including matching and instrumental variables, and conduct a host of placebo tests to show that experience facilitates trade by reducing bilateral trade costs. We also provide evidence for spillovers in experience by showing the adjustment of the extensive margin of trade and by directly estimating spillovers across 4-digit sectors. Our findings help address the puzzle that despite falling trade costs, most of the rapid growth of world trade is due to the growth of the trade among countries that traded with each other prior to 1950 rather than due to the expansion of trade among new trade partners. To illustrate our findings, we build a trade model with heterogeneous firms where export-experience reduces trade costs, introducing heterogeneity in firm entry decisions for export markets and dynamics for the extensive and intensive margins.

JEL Classification: F10, F14

Keywords: Gravity model; Trade costs; Experience; Extensive and intensive margin

---

* We thank Thierry Mayer, Marc Melitz, Helene Rey and Andrew Rose for very useful comments and suggestions. We also thank audiences at the Royal Economic Society Conference, the World Bank, Oxford University, LBS, INSEAD, Georgetown University, Paris School of Economics, NOVA, ABFER, and University of California, Irvine.

† Corresponding author: INSEAD, 1 Ayer Rajah Avenue, Singapore 138676; Pushan.Dutt@insead.edu
Currently on leave at London Business School, Regent's Park, London NW1 4SA United Kingdom; pdutt@london.edu

1 Introduction

World trade has exploded over the last few decades, a phenomenon often associated with a decline in trade costs. Surprisingly, as Figure 1 shows, virtually all that growth has occurred between country-pairs that were already trading prior to 1950, rather than due to the emergence of trade between new trading partners. We use the IMF-DOTS data over the period 1948-2006 and decompose the evolution of world trade into trade between country-pairs with strictly positive trade prior to 1950 and the emergence of trade between new trading partners. Over the subsequent 55 years, less than 25% of the increase in world trade is attributed to the emergence of new trading partners, and 75% to the increase in trade between partners that had strictly positive trade prior to 1950 (Helpman, Melitz and Rubinstein, 2008).[1] The puzzle is why the decline in trade costs that contributed to the growth of world trade is overwhelmingly confined to countries with a long history of trading.

At the same time, the underlying determinants of trade costs remain poorly understood (Head and Mayer, 2013a). Multiple studies have identified a persistent and even rising role for gravity variables such as distance, borders, language and colonial ties.[2] Prior work, in order to reconcile these persistent effects of such gravity variables for trade, often alludes to informational costs, cultural differences and the importance of business and social networks in overcoming informal barriers to international trade (Rauch, 1999; Greif, 1994, Chaney 2014). For instance, Grossman (1998) argues that estimated distance effects are too large to be explained by shipping costs and that cultural differences and lack of familiarity account for the persistence of distance, while Anderson and van Wincoop (2004) stress the role of information barriers, contracting costs and insecurity. Head, Mayer and Ries (2010) attribute the decline in trade between countries that shared a colonial link to the depreciation of trade-promoting capital embodied in institutions and networks. Head and Mayer (2013b), drawing on the analogy of dark matter, coin the term 'dark trade cost' and argue that

[1] If we adjust trade flows for origin and destination size and distance, we find that 80% of the increase in world trade is attributable to country-pairs that had positive trade prior to 1950.

[2] In a wide range of gravity estimates, the inhibiting influence of distance on bilateral trade seems to have increased since the 1970s (Disdier and Head, 2008; Head and Mayer 2013a), coinciding with a fall in shipping, air cargo and communication costs and the rise of internet technologies. Borders also matter a lot - Head and Mayer (2013b) examine 71 different estimates of the border effect (based on well-specified structural gravity models) and estimate that a border is equivalent to very high tariffs (about 43.5%) and that actual tariffs explain a very small share of the border effect. Egger and Lassmann (2012) in their meta-analysis find that sharing a common language increases bilateral trade by 44%. As with distance, they find not simply a persistent effect but an increasing language effect on trade over time. Finally, Head, Mayer and Ries (2010) find that while trade declines once a colony becomes independent of the colonizer, the trade for country-pairs that had been in a colonial relationship in the past remains 27% higher than trade of countries that were never in colonial relationships.

these gravity variables capture some unmeasured and unknown sources of resistance.[3]

Some papers attempt to directly incorporate these forces. For instance, Rauch and Trindade (2002) find a substantial trade-enhancing effect of Chinese networks and argue that these networks reduce information barriers to trade.[4] Anderson and Marcouiller (2002) and Berkowitz et al (2006) show that insecurity, associated both with contractual enforcement problems and opacity of regulations, lowers trade substantially. Dutt and Traca (2010) find that corruption impedes trade when formal tariff barriers are low. Allen (2012) and Chaney (2014) develop trade models that explicitly incorporate the search process producers use to acquire information about market conditions and trading partners. Finally, Guiso, Sapienza and Zingales (2009) demonstrate the importance of bilateral trust for international trade and link bilateral trust to cultural similarities between pairs of countries.

In this paper, we take as given that there are some unobserved or 'dark' trade costs, only some of which are captured by traditional gravity variables. Our main hypothesis is that a key factor driving the decline in such trade costs is the cumulative experience in international trade. When a country starts exporting to a new destination, a large component of trade costs is related to the novelty and uncertainty of selling in an unfamiliar environment, identifying customer preferences, engaging with foreign shipping agents, customs officials or consumers, and navigating an uncharted legal and regulatory context (Anderson and van Wincoop, 2004). Kneller and Pisu (2011) survey UK firms and find that identifying contacts and building relationships with decision makers, customers and partners are two most commonly cited export barriers. Experience from repeated local interaction can be effective in gaining familiarity, acquiring information, and building contacts. These in turn, contribute to dampening costs associated with the shipment, border crossing, and distribution in the destination country. Hence the accumulation of experience works to overcome the informational, contractual and cultural barriers, some of which are captured by gravity variables, suggesting that experience reduces trade costs and expands bilateral trade flows.

Our empirical specification augments the bilateral gravity equation, the literature´s workhorse, to account for the role of experience, measured at the level of the country-pair. We base our experience measure on the number of years of positive trade between a pair of countries. At the country-pair level, we have sufficient variability in experience, both across countries and over time, which allows us to measure experience very precisely and identify its importance in lowering trade costs and increasing trade.[5] We allow for both

[3] Head and Mayer (2013b) show that 72%–96% of the trade costs associated with distance and borders are attributable to the dark sources (read unknown) sources of resistance.

[4] Head and Ries (1998) find smaller effects of immigration links for Canadian bilateral trade. Bastos and Silva (2012) use Portuguese data to show that larger stocks of emigrants in a given destination, increases export participation and intensity.

[5] Data on bilateral trade is consistently available from 1948. By contrast, firm-level trade data is not widely available for

depreciation in experience over time and for diminishing returns in experience for trade. In a standard gravity specification, with country-year fixed effects, we estimate that a 1% increase in experience increases bilateral trade by 0.9%. In terms of number of years, for a county-pair with the median level of experience, an extra year of trade raises trade by nearly 4%.

Interpreting the estimates of experience as causal requires that experience at the country-pair level is exogenous to unobserved bilateral trade costs. This is especially challenging in our context since omitted variables affect both the current value of trade and our experience measure based on the number of years of strictly positive trade. We first account for this with a very demanding specification that includes country-pair fixed effects and country-pair specific trends. The former accounts for all time-invariant unobserved variables that effect both experience and bilateral trade while the latter accounts for secular decline in unobserved trade costs at the country-pair level. In this preferred specification, we rely on *breaks* in trade to identify the coefficient. For country-pairs with an unbroken trading relationship, experience will be absorbed by the dummies and pair-specific trends. Here we find that the elasticity of trade with respect to experience declines to 0.45. Equivalently, an extra year of trade between a county-pair raises trade by 2%. This is a very robust finding - it survives a battery of checks where we use different estimation techniques (parametric and non-parametric), deploy alternate samples and data sets, and account for measurement error in our experience variable. We also examine the robustness to misspecification and the sensitivity of our estimates of the point estimate for experience using a recent methodology proposed by Athey and Imbens (2015).

Even in a specification with country-year, country-pair fixed effects, and pair-specific trends, bilateral trade shocks may be correlated over time in an arbitrary fashion, which can again yield biased estimates. To further address potential endogeneity concerns, we employ two alternate strategies. First, we use the matching approach of Abadie and Imbens (2006) and apply it to cross-sectional data, one year at a time. We use very stringent matching criteria, matching on all gravity variables, country size, and restrict matches to the same exporter and subsequently to the same importer. Second, we use two instruments for experience at the country-pair-year level, drawing on the geographic spread of export (e.g., Chaney 2014) and historical ties between countries (Mitchener and Weidenmier, 2008). The first instrument uses experience of the exporter in countries that are contiguous to the destination but not part of a preferential trade arrangement as the destination. The second instrument uses experience in countries that were the same empire or the same administrative entity as the destination for a long period (25-50 years in the twentieth century, 75 years in the nineteenth and 100 years before.) Our main identifying assumption for a causal interpretation is that omitted

large numbers of country-pairs and usually span short time-series. With firm data, there are also censoring and selection (firms die or are acquired) issues.

trade-costs making a given destination differentially more attractive for a particular exporting country for both experience and bilateral exports are orthogonal to the exporter experience in the destination's neighbors and experience in countries that were the same administrative entity in the past. We are able to demonstrate that our instruments are strong and we provide some evidence on the validity of the exclusion restriction. The inclusion of a comprehensive set of dummies allows us to guard against a wide range of threats to our identifying assumption. The IV estimate for experience is in fact marginally larger than the OLS estimate. Finally, we employ a set of placebo tests to assure us that our experience measure facilitates trade by reducing bilateral trade costs. Overall, our results imply that countries with a long bilateral trading relationship are likely to face lower trade costs as compared to country-pairs that have little or no experience in trade, helping explain the dramatic growth of trade among established country-pairs found empirically. In other words, trade costs are dynamic and decline with experience.

Measuring experience at the bilateral level implicitly assumes that there are spillovers in export experience. Such spillovers are partly facilitated by trade associations and export promotion bodies (Lederman, Olarreaga and Payton, 2010), worker mobility (Molina and Muendler, 2013), and partly by simple observation (Segura-Cayuela and Vilarrubia, 2008; Hausman and Rodrik, 2003). For instance, Artopoulos et al (2011) use a detailed case study of firms from four export sectors in Argentina, to show how pioneers' export experience diffuses to other firms who follow the pioneer into exporting. In this case, the experience of an exporter may lower trade costs of other exporters, who in turn may expand exports, and even to non-exporters, who may find it easier to enter that market. While we do not use firm-level data on trade costs, we provide evidence for spillovers in two ways. First, we decompose the impact of experience on the bilateral extensive margin, defined as the number of HS-6 products traded, and the intensive margin, defined as the value of exports per product. The evidence of spillovers is captured by a positive effect of experience on the extensive margin. We find strong evidence of such an effect. Second, we use sectorial data to provide additional evidence for spillovers in experience across 4-digit sectors.

To shed more light on these findings, we build a dynamic version of Chaney's (2008) model of heterogeneous firms, where variable trade costs decline as experience is accumulated over time. We show when firms benefit from the experience of their peers, entry in the first period that trade is similar to Chaney (2008). However, different from other static models of trade with heterogenous firms, the extensive margin increases with experience, as non-exporters see their trade costs declining from the experience of the incumbents. The effect on the intensive margin is more complex: the larger exports of incumbents (who see their variable trade costs decline) conflict with the entry of smaller, new exporters, so the effect on the intensive margin is ambiguous (zero if productivity distribution is Pareto but may be positive for other plausible distributions).

Next we simulate the model by generating an artificial sample of country-pairs that differ in the parameters governing the international trade costs (both variable and fixed trade costs and experience) and choose parameters to generate a distribution of experience, exports and output similar to that in our data. We then run a similar regression with the simulated data of bilateral trade on experience. We show that what matters is export experience that reduces the variable costs of trade. The exercise illustrates the main mechanism of how experience affects trade costs and trade and allows us to compare the results to those in our empirical section.

Our paper contributes to the trade costs and gravity literature in three ways. First, we highlight the role of experience in bridging "unmeasured" trade costs associated with informational barriers and cultural differences, which provides an intuitive solution to the puzzle in Helpman, Melitz and Rubinstein (2008). Our empirical work confirms that these effects are very strong and robust. Second, we provide both direct and indirect evidence that the effects of experience are shared among exporters and non-exporters, which complements the recent literature focusing on the role of networks and search in export decision. (Eaton et al, 2012; Chaney, 2014). Third, we develop a model where trade costs are dynamic and evolve over time, driven by the accumulation of experience. Our model introduces dynamic paths of entry into exports and of export volumes in a simple and intuitive way.

The remainder of the paper is organized as follows. Section 2 augments the traditional gravity specification with experience at the bilateral level; Section 3 presents our empirical findings along with a series of identification strategies and robustness checks; Section 4 documents the importance of spillovers in experience; Section 5 introduces experience in a standard Chaney (2008) model to examine the evolution of the extensive and intensive margins; Section 6 simulates the model and compares the results from the simulated data to those in our empirical section; Section 7 concludes.

2 Experience and the gravity equation

The gravity equation is the current workhorse for estimating the importance of trade costs for bilateral trade. There are several theoretical frameworks supporting the gravity specification, with exports from country o (exporter/origin) to country d (importer/destination) in time t, denoted by $X_{od,t}$, given as

$$\ln X_{od,t} = \alpha_o \mu_{o,t} + \alpha_d \mu_{d,t} - \theta \ln \tau_{od,t} + e_{od,t} \tag{1}$$

$\mu_{o,t}$ and $\mu_{d,t}$ are exporter and importer-year dummies that capture attributes of the exporting- and the importing-country, respectively, including size and their multilateral trade resistance (Anderson and van Wincoop, 2003). $\tau_{od,t}$ measures bilateral trade costs, with $-\theta$ as the elasticity of exports with respect to

trade costs.[6] In the standard equation $\ln \tau_{od,t}$ is specified in terms of bilateral gravity variables, as shown below.

$$\ln \tau_{od,t} = \sum_{m=1}^{M} \gamma_m z_{od,t}^m \tag{2}$$

where $z_{od,t}^m$ are the M gravity variables and γ_m are parameters. Head and Mayer (2013a) perform a meta-analysis and identify as main variables the trade and currency agreements, capturing trade policy, and distance, contiguity, shared language, and colonial links, which measure geographic, cultural, and historical barriers. Substituting (2) in (1) yields an estimable specification

$$\ln X_{od,t} = \alpha_o \mu_{o,t} + \alpha_d \mu_{d,t} - \sum_{m=1}^{M} \theta \gamma_m z_{od,t}^m + e_{od,t} \tag{3}$$

This equation can be estimated using data on bilateral trade flows and the bilateral gravity variables. For bilateral trade flows, two data sources are available. First, International Monetary Fund's *Direction of Trade Statistics* DOTS provides data on aggregate bilateral exports from 208 exporters to 208 importers over the time period 1948-2006. Second, UNCTAD's COMTRADE provides data on bilateral trade between pairs of countries at the Harmonized System 6-digit (HS-6) level of disaggregation. The HS-6 data spans 5017 product categories, for the time period 1988-2006 for 183 importers and 248 exporters. For each year, COMTRADE covers more than 99% of all world trade. The advantage of the DOTS data is the coverage over time, while the COMTRADE data allows us to decompose total exports into an extensive and an intensive margin.[7]

For the gravity variables, we use data from the CEPII gravity database (www.cepii.fr). Geographic distance is measured as the logarithm of the distance (in kilometers) between the two most populous cities. Contiguity is a dummy variable that takes the value 1 if the country-pair shares a common border. Common language is captured by a dummy that equals 1 if the country-pair shares a common official language. Colonial relationship takes the value 1 if a country-pair was ever in a colonial relationship (one country was the colonizer and the other colonized or vice versa). Data on these variables are obtained from the CEPII gravity databases (www.cepii.fr). We also create a dummy variable that captures common law legal origins, from Glaeser and Shleifer (2002) (other finer classifications of civil law, Scandinavian law did not seem to

[6] θ has different interpretations depending on the micro-foundations for the gravity equation. It is the elasticity of substitution (minus one) among varieties in Anderson and van Wincoop (2003), the parameter in the Pareto distribution of firm productivities in Chaney (2008) and the parameter governing the dispersion of labour requirements across goods and countries in Eaton and Kortum (2002).

[7] We use the former to measure experience at the bilateral level and the latter to measure bilateral trade and the extensive and intensive margins.

matter). We include also policy-related gravity variables, from Head and Mayer (2013b). We use a dummy variable that captures membership in a currency union. Data on currency unions are from Head, Ries and Mayer (2010). Multilateral market access is captured a dummy variable that takes the value 1 if both trading partners are members of the GATT/WTO and 0 otherwise. Bilateral preferential trade arrangements are captured by a dummy variable which takes the value 1 if both trading partners are members in a preferential trade arrangement (PTA). Data on WTO membership and PTAs are from the CEPII and updated via the WTO website (www.wto.org). Unilateral preferential access is in terms of the Generalized System of Preferences (GSP) where trade preferences are granted on a non-reciprocal basis by developed countries to developing countries. We code a dummy variable as 1 if the importing county grants a GSP to exporter. GSP data are from Andrew Rose's website and updated from the WTO website.

2.1 Measuring experience

Exporting to a new geographic market entails the discovery of (i) the cheapest, most reliable transport; (ii) the best way to get goods through customs, (iii) the right partner for distributing and promoting the goods locally or (iv) the preferences and dispositions of customers.[8] Although firms may engage in pre-entry research, experience is a vital element of this discovery process. The initial contact with a new market environment unavoidably raises unexpected challenges that push the firm to find quick, imperfect solutions. Experience with the local reality helps the firm gain familiarity and find better, cheaper solutions for future shipments, thereby lowering trade costs. For instance, Artopoulos, Friel and Hallak (2013) use four detailed case studies of export firms in Argentina to show that export success stems from a firm's experience in the business community of the destination country. Similarly, Kneller and Pisu (2011) show that the best predictor of whether a particular firm identifies an export barrier as relevant is explained almost exclusively by experience measured as the number of years the firm has been exporting.

In addition to benefiting the firm, experience is also likely to be shared among networks of firms, who may be organized into industry associations at home, by worker mobility across firms, by the use of common consultants or export promotion agencies, and even by simple observation. Eaton et al (2012) use Colombian data to show that firms that were initially non-exporters within a decade account for nearly a quarter of

[8] Eaton et al. (2012) and Freund and Pierola (2010) emphasize learning in a destination country, where producers need to incur costs to find new buyers or new products. In contrast, in Albornoz et al (2012) uncertainty is not destination-specific and firms learn about export profitability as a whole. Allen (2012) models a search process to acquire market information. For regional agricultural trade in the Philippines, he finds that 93% of the observed gravity relationship is related to informational frictions rather than transport costs.

total exports. Firms in their model learn from both own experience and the experience of rivals. Clerides et al. (1998) find evidence for geographic and sectorial spillovers for the export decision of Colombian plants. Álvarez et al (2007) find evidence that Chilean firms learn both from their experience and experience of others. Iacovone and Javorcik (2010) show that for approximately half of new exported products in Mexico, an additional firm starts exporting the same product in the following year and that this diffusion across firms continues for their entire time period of study (8 years). This implies that experience acquired historically by some exporters contributes to increased familiarity by fellow exporters, and even crosses over to non-exporters.

Much of this research on export experience is restricted to a single country. Available firm-level datasets cover relatively few exporting countries over short time-spans yielding limited dispersion and potential measurement error in experience measures, and, therefore, imprecise estimates for the effects of experience. Our focus is on how experience reduces bilateral trade costs and facilitates trade between pairs of countries. Therefore, we augment the gravity specification in equation (3) with a measure of experience at the country-pair level. At the country-pair level, we have extensive datasets that span all of world trade with deep coverage going back in time allowing us to measure experience relatively accurately. Of course, at the county-pair level, experience must be seen as an aggregation of the experience of all the exporter firms involved in trade from o to d, including benefits that spill over across firms.

We measure experience using the DOTS database. For any pair of countries o and d at time t in our sample, experience is based on the number of years with strictly positive exports from o to d from 1948 or year of independence to year $t-1$. While past cumulated exports from the origin to the destination country environment is a natural measure of (stock of) experience, such a measure is influenced by the unit value of exports. In addition to changes in the price of exports, experience based on past exports would also be influenced by changes in the sectorial composition of country's exports, both in terms of comparisons across countries and its evolution in time, creating unwanted spurious variation.[9]

Figure 2 shows the distribution of number of years of strictly positive trade since 1948 for all country-pairs in the year 2006: 3.1% of the country-pairs have just initiated trade (experience equals 0 years,) while the median and the mean trading relationship are for 14 and 21.7 years. The variable is right-censored for country-pairs that had strictly positive trade on or before 1948: 7.8% of the country-pairs have a trading

[9] We recognize that our measure of experience does not distinguish between small and large shipments and assumes that experience spills over. We check the sensitivity of our results to dropping small shipments. Subsequently, as a further robustness check, we use the cumulated value of exports at the 4-digit industry level as our experience measure. With disaggregate data, changes in composition and unit values are also less of a concern. It also allows us to examine directly, the evidence for spillovers in experience across sectors.

relationship at the maximum of 58 years (from 1948 to 2006). Finally, the peak around years 13 and 14 arises from the breakup of the Soviet Union and of Yugoslavia and Czechoslovakia in 1992.[10] In robustness checks, we will account for zero trade, for the censoring of the experience variable, and for the formation of new countries in Eastern Europe around 1992-1993.

2.2 A non-parametric look at the data

We start with a non-parametric approach to provide evidence on the importance of experience for bilateral trade. We estimate equation (3) with a complete set of 58 dummy variables, one for the number of years of strictly positive trade, so that experience is simply the cumulated number of years of trade. This specification is flexible in that it makes no functional form assumption and allows for diminishing returns in experience. We include all gravity variables as well as county-year fixed effects. With 0 years of trade as the omitted category, each coefficient captures the cumulated impact of experience on bilateral trade for a country-pair relative to a country-pair with zero experience. These coefficient estimates along with the 95% confidence bounds are shown in Figure 3.

The figure illustrates three findings. First, all the experience dummies are significant at 1%. Second, in terms of magnitude, the coefficient estimates imply a very strong role for experience. An additional year of experience, on average, increases bilateral trade by 6%. Alternately, comparing trade between a country-pair with the median level of experience of 22 years, to a country-pair with 0 experience (or half the experience), trade in the former is 165% (66%) higher. Another way to think of the magnitude, is that 5 years of experience is equivalent to both country-pairs joining a preferential trading area or sharing a common language, while 9 years of experience is equivalent to a colonial link or sharing a common currency.[11] Third, the relationship seems approximately log-linear. We do see some evidence for non-linear effects - visually, we see a stronger role for experience from years 1-13 and a decline thereafter.[12]

We next introduce a single flexible measure of the stock of experience in the gravity model, where we allow for depreciation in experience and employ a flexible log-log specification that can capture diminishing returns of trade with respect to experience.

[10] In constructing the experience variable, we coded all countries that were formerly part of the Soviet Union, Czechoslovakia and Yugoslavia as new countries and set experience to zero in their first year of trade after 1992. The exceptions are trade with the Soviet Union which was merged with Russia and with West Germany which was merged with Germany. These choices, while reasonable since exporters plausibly faced a new environment, may also create measurement error in experience.

[11] We base these comparisons on structural gravity estimates in the recent survey paper by Head and Mayer (2013a).

[12] The slight decline at 58 years of experience is an artefact of the COMTRADE data - this decline does not show up if we use the DOTS data.

2.3 Experience-adjusted gravity specification

We follow the learning-by-doing literature and construct an experience measure (Levitt, List and Syverson, 2013) that allows for the depreciation of experience or "forgetting". We code a variable $I_{od,t} = 1$ if there are strictly positive exports from o to d at data t, and define experience at time t as accumulated according to a perpetual-inventory process:

$$E_{od,t} = I_{od,t-1} + \delta E_{od,t-1}$$

The experience for country pair od at time t is the sum of two components: a fraction δ of the previous accumulated experience and the existence of strictly positive trade in the previous period. δ parametrizes the fraction of experience that is retained from one period to the next with $\delta = 1$ indicating complete retention where past interactions count as much as recent interactions and $\delta = 0$ indicating complete forgetting or depreciation. With $\delta = 1$, experience is simply the cumulated number of years of strictly positive trade.

Given that Figure 3 suggests diminishing returns, we use the natural log of experience.[13] Our experience-adjusted specification for trade costs

$$\ln \tau_{od,t} = \sum_{m=1}^{M} \gamma_m z_{od,t}^m - \lambda \ln (E_{od,t}) \tag{4}$$

We expect $\lambda > 0$ with $\lambda < 1$ implying diminishing returns in experience for trade costs. Substituting in (1) yields an estimable specification for the gravity equation that accounts for the effect of experience.

$$\ln X_{od,t} = \alpha_o \mu_{o,t} + \alpha_d \mu_{d,t} - \sum_{m=1}^{M} \theta \gamma_m z_{od,t}^m + \theta \lambda \ln (E_{od,t}) + e_{od,t} \tag{5}$$

The coefficient $\theta\lambda$ is the experience elasticity of trade, with $\theta\lambda < 1$ implying diminishing returns in experience for bilateral trade.

Therefore, we estimate the following non-linear system of equations

$$\ln X_{od,t} = \alpha_o \mu_{o,t} + \alpha_d \mu_{d,t} - \sum_{m=1}^{M} \theta \gamma_m z_{od,t}^m + \theta \lambda \ln (E_{od,t}) + e_{od,t} \tag{6a}$$

$$E_{od,t} = I_{od,t-1} + \delta E_{od,t-1} \tag{6b}$$

[13] We also carried out a series of tests to choose the functional form for experience. First, we tried the standard Box-Cox transformation. Unfortunately, as is common in datasets with a large number of observations, the test rejects the log-specification in favor of the linear and the linear in favor of the log. Other tests such as the Cox-Pesaran test and the Davidson-Mackinnon J-test yield a similar finding. We follow prior practice and pick the specification with the lowest chi-square statistic in the Box-Cox test and the lowest z-statistic in the Cox-Pesaran test and the J-test. This selects the log specification for experience.

3 Empirical estimates of experience-adjusted gravity

We estimate equations (6a and 6b) by non-linear least squares to obtain estimates for δ, the retention parameter, and for $\theta\lambda$, the elasticity of export with respect to experience. We use the aggregate COMTRADE data to measure bilateral trade flows, the DOTS data starting in 1948 to measure experience, and the previously described measures of gravity variables. Given that non-linear least squares is difficult to implement with a large set of dummies, we first obtain a robust estimate of δ and subsequently use this estimate to construct the stock of experience. We then use this experience measure and estimate the elasticity of experience for trade with more comprehensive sets of fixed-effects. In line with Head and Mayer (2013a), our main results will rely on country-year fixed-effects, with standard errors adjusted for clustering on country-pairs. Subsequently, we use multiple identification strategies to establish the effect of experience on trade, including more comprehensive fixed-effects and pair-specific trends, cross-sectional estimates, matching models, and instrumental variable estimates. We also address the potential for biases through a series of robustness checks.

3.1 Retention parameter for experience-adjusted gravity

Column (1) in Table 1 uses a very basic specification without any fixed-effects. We estimate a retention rate of 0.963 implying that slightly less than 4% of experience is lost each year. While we can statistically reject that $\delta = 1$, substantively it is very close to 1. Column (2) adds a time trend. The time trend term is actually negative rather than positive, and the estimated experience elasticity is slightly higher. Therefore it is accumulated experience that matters rather than the passage of time since initiation of trade. Column (3) uses time dummies instead of a time trend which leaves the estimates for retention and experience unaffected. Column (4) uses 188 exporter dummies, 169 importer dummies and 18 year dummies and is the most demanding specification in the non-linear least-squares context. The coefficient on the retention parameter exceeds but is very close to 1, while the elasticity of trade with respect to experience declines marginally to 0.798. Overall our δ parameter remains close to 1 as we add more comprehensive set of fixed-effects.

Country-year fixed-effects, which is the preferred specification for gravity models, entails the use of 5487 dummies which is infeasible with non-linear least squares in terms of computational power. Therefore, we used a simpler way to identify the coefficient for δ. We constructed the experience variable for various values of $\delta \in [0, 1.5]$ in increments of 0.05. We then used this variable in the gravity equation with country-year

dummies and picked the value of δ that yields the best fit.[14] This results in $\delta = 0.995$. In all subsequent specifications, of the gravity equation with country-year dummies we set $\delta = 0.995$. Column (5) in Table 1 shows the elasticity of trade with respect to experience is 0.887. If θ is the Pareto shape parameter from Chaney (2008) or the parameter governing the dispersion of labour requirements across goods and countries in Eaton and Kortum (2002), then a reasonable value is $\theta = 4$. Our estimate in Column (5) of Table 1 implies that $\lambda = 0.221$ so that trade costs decline by 0.221% for every 1% increase in experience. Overall, while there is very little loss in experience over time, we do have diminishing returns in experience for trade costs and overall trade.

Finally, if we estimate a traditional gravity equation without experience as an independent variable we obtain coefficients on the gravity variables that are significantly higher in absolute terms (14% for distance and legal system; 4% for contiguity; 4% for colonial link; and 37% for language). In other words, accounting for experience significantly reduces the magnitude of the coefficients for the traditional gravity variables.

3.2 Identification

Absent a randomized controlled experiment, which is obviously infeasible in the context of countries, establishing causality is a challenge. We employ a variety of identification strategies to provide evidence that experience matters strongly for bilateral trade. At the same time, we recognize that all of these strategies make identification assumptions that are not entirely free of criticism. Instead, we hope that deploying a multitude of identification strategies, robustness checks, and placebo tests, the cumulative evidence is strongly suggestive of the importance of experience.

3.2.1 Pair-specific fixed-effects, pair-specific trends and lagged dependent variable

Our baseline estimate of the elasticity of trade for experience is reproduced in Column 1 of Table 2. This specification uses country-year fixed effects and relies on variation in experience both across country-pairs and within country-pairs over time. The fixed-effects absorb any effects that are particular to changes in variables at the exporter-year level and the importer-year level (e.g., investment in ports, infrastructure, doing business indicators that facilitate or impede trade, changes in the relative importance of sectors at both the exporting or the importing country). While we also use the standard array of pair-specific gravity variables, they remain necessarily incomplete.[15]For instance, any unobserved dyadic effects that affect both

[14]GDP in the importer and exporter country are subsumed within country-year fixed effects.

[15]We also experimented with multiple bilateral gravity variables as additional controls. These include the share of migrants from destination in the origin country (to capture networks,) a common religion dummy from Helpman, Melitz and Rubinstein

the onset of trade (and therefore our experience measure) as well as trade today would lead to an upward bias in our coefficient on experience. Therefore, we account for unobserved time-invariant dyadic effects by including 25,581 dyadic fixed effects. Since these dyadic effects are time-invariant, we also include pair-specific linear time-trends, one for every country-pair. This is the most demanding specification accounting for all exporter and importer-year terms, all time-invariant characteristics at the dyadic level, as well as for any country-pair specific variables that evolve in a linear fashion.[16] For instance, if we believe that trade barriers between country-pairs decline in a gradual fashion (e.g., elimination of tariffs after joining a trading arrangement happens only gradually) these should be accounted to some extent, though not entirely, by the pair-specific trends. In this specification, identification relies on breaks in trade. For country-pairs with a break in trade, the dummy for strictly positive trade $I_{od,t}$ switches from 1 to 0. When trade re-starts for this pair, the experience measure corresponding to this observation is lower given that the retention parameter $\delta < 1$. On the other hand, for country-pairs that have traded continuously since 1948 (e.g., US and Canada,) the experience measure will be absorbed by the country-year dummies or by the pair-specific trends. Column 2 in Table 2 shows that the experience elasticity of trade declines to 0.436 but remains statistically significant and substantive. The estimate implies that for a country-pair with the median level of experience, an additional year of trade increases bilateral exports by 2%.[17]

Next, we attempt to account for slow moving unobserved dyadic influences on trade that may not evolve in a linear fashion by including a lagged dependent variable following Eichengreen and Irwin (1997). This specification also allows us to distinguish between the short vs. long run effect of experience. This estimate is shown in Column 3 of Table 2. [18] The coefficient on experience declines marginally to 0.333. However, this coefficient captures only the short-run experience elasticity of trade. Our estimates imply that the long-run experience elasticity of trade equals $\frac{0.333}{(1-0.720)} = 0.358$. In fact, we are unable to reject that this long-run estimate of experience elasticity equals the estimate in Column 2.[19]

Interestingly, in Columns 2 and 3 we find that in the presence of a comprehensive set of fixed effects

(2008), a dummy that takes the value 1 if both countries are democracies (positive Polity score), political ideology of the ruler in power (coding ideology as left-wing vs. right wing from the Database of Political Institutions), a dummy for country-pairs that were part of the same country in the past (e.g., India and Pakistan,) and further refinements of the colonial links and language dummies (see Dutt and Traca 2010). Our results remain unaffected by these permutations.

[16] Using a pair-specific logarithmic time trend yields nearly identical coefficients.

[17] If we estimate the gravity equation in first-differences, the coefficient on experience is 0.382.

[18] The Least Squares Dummy Variables estimator, is inconsistent in the presence of lagged dependent variables. However when the number of time periods is large, as is the case here, this bias goes to zero.

[19] This test yields a chi-square test statistic of 0.56 and a p-value of 0.45. Adding further lags of the dependent variable increases the coefficient estimate for experience.

and time-trends, neither WTO nor PTA membership has an effect on trade. Only membership in the same currency area has a positive impact on bilateral trade. If we drop the pair-specific time-trends then both WTO and PTA membership matter, indicating that the time-trends capture slow moving changes in formal trade barriers following WTO accession or membership in a PTA (see Baier and Bergstrand, 2007).

3.2.2 Cross-section

While the dyadic specification absorbs all time-invariant dyadic variables, and the pair-specific linear trends captures evolution of unobserved trade costs, it may be argued that there is some arbitrary serial correlation in bilateral trade shocks over time. Such serial correlation in shocks would induce a correlation between past trade (which depends on past shocks) and therefore our experience measure and the current shock. Our second identification strategy therefore uses a pure cross-section relying on differences in experience across country-pairs where serial correlation is not a concern. Column 4 shows the OLS estimate for an arbitrarily chosen year, the year 2004, with country fixed-effects. The estimated coefficient on experience of 1.086 is marginally bigger than our baseline estimate of 0.887 in Column 1. We also estimated the cross-sectional coefficient for each year one at a time (not shown) - these estimates are remarkably stable across years, ranging from 0.75 to 1.2.

3.2.3 Matching estimate

In standard gravity specifications, since we cannot measure trade costs directly, all researchers include only some of their proximate determinants. At the same time, the standard gravity equation implicitly assumes that the relationship between the log of trade and these proxies such as distance, language, colonial ties etc. of true trading costs is linear. This assumption may be questionable. For instance, sharing a common colonial history (and thus similar institutions) and geographic distance may interact with one another so that their effect may not be their summed effects in isolation. Similarly, the existence of border effects implies a non-linear effect of distance. In other words, proxies for trade costs (including experience) may affect bilateral trade in more complex non-linear ways. More importantly, in our context, country-pairs that broadly have a lot of experience in trade may differ systematically from those with little or no experience (lack common support). These country-pairs are not randomly selected, and pooling such different country-pairs may yield an upward bias in our coefficient estimates.

In an idealized setting, for a particular cross-section, we would randomly assign some country-pairs to a treatment group of high experience in exports and others to a control group of low experience. Of course such randomized experiments are not possible in the context of countries. We therefore adopt a

matching approach to ensure that country-pairs with high and low experience in trade are as comparable as possible along key observables. The matching methodology has several advantages. It ensures overlap and comparability between country-pairs with high and low experience so that we can be assured that the results are not are not being driven by outliers and are not sensitive to specific functional form assumptions of the gravity equation (Heckman et al. 1997, Imbens 2004). To the extent that matching is on observables, it also reduces biases due to endogeneity in experience (Dehejia and Wahba 1999). Therefore, our next identification strategy relies on matching models and applies it to the cross-sectional data.[20]

Since these models apply to the average effect of a binary treatment, we first convert an experience measure into a binary measure, by defining the treatment indicator as 1 if E_{od} for country-pair exceeds the median in 2004 across country pairs (this translates into 13 years of positive trade.) As in all matching models, the key goal is to prune observations so that the remaining data is balanced in terms of the covariates between the treated and control groups. Instead of using approximate matching based on Mahalanobis distance or propensity score, we implement very stringent matching criteria. We matched exactly on all binary gravity variables, and on year by default (we present estimates for the year 2004 in the table) and employed coarsened exact matching (CEM) using 10 bins for distance and country size (GDP in origin or destination.)[21] We also restrict matches first to the same exporting country and then to the same importing country. This yields a sample of 6,259 observations when we match on the exporter in Column 5 and 5,823 observations in Column 6 when we match on importer. Note one of the key difference from the OLS estimate in Column 4 of Table 2 is that matching estimate does not rely on country-pairs where experience is always greater (or always lower) than the median experience (example, US-Canada and US-India where experience exceeds the median for both). For each exporting (importing) country, a control and treated destination (origin) show identical values for all discrete gravity variables, and very similar values in terms of distance and GDP, with high and low experience as the only difference.[22] [23]We rely on many-to-many matching to fully utilize available data, and use CEM-generated weights to infer "average treatment effect on the treated

[20] See Baier and Bergstrand (2009) for a matching estimate of the treatment effect of an FTA.

[21] Our distance classification is finer than Eaton and Kortum (2002) who decompose distance effects into only four intervals.

[22] As an example, when we restrict matches to the same exporter, Turkey as exporter and Botswana as importer are a country-pair in the control group with 9 years of positive trade in 2004 while Turkey and Zimbabwe pair are in the treatment group with 24 years of positive trade. If we match on importer, then Haiti has only 5 years of positive exports to Thailand while Honduras has 22 years of experience exporting to Thailand. Honduras-Thailand are in the treatment group while Haiti-Thailand are in the control group.

[23] In the absence of matching, we find that countries classified as treated are more likely to be members of the WTO, in an PTA, receive GSP, share a common currency, are geographically proximate, contiguous, share linguistic, colonial, and legal ties.

(ATET)" (Imbens 2004, Iacus et al. 2011 and 2012.) We examine the average treatment effect on the treated (ATET) since our sample includes only country-pairs with strictly positive trade (with the log of trade as the outcome variable, all zero trade observations are dropped.) After pruning the observations to achieve balance, we also added the gravity variables and country-fixed effects as controls (Iacus et al. 2011).

In both Columns 5 and 6, we find that very similar coefficients on experience which is slightly smaller than the OLS estimates in Column 4. We obtain an experience elasticity close to 1. To show that this is not an artefact of the choice of year, Figures 4a and 4b show the estimate and confidence interval for average treatment effect for experience for each of the years from 1988-2006. The magnitude of the estimated coefficient is very similar over the years and tightly estimated from 1992 onwards. In the initial years, while the coefficient on experience is significant, the confidence intervals are broad. This is due to the use of the very stringent matching criteria so that we obtain few matches in the initial years of the sample.[24]

3.2.4 Instrumental variable estimates

In the matching estimates, we can only account for selection on observables. Any unobserved bilateral trade cost that affects both the onset of trade and trade flows can still result in biased estimates. In fact even if serial correlation is not an issue in the cross-sectional and matching settings, country-pairs with a long history of trade may also be those where unobserved trade costs have simply had more time to decline. Therefore, our final identification strategy relies on instruments for experience $E_{od,t}$. As instruments, we need variables that are correlated with our causal variable of interest, namely experience, but uncorrelated with any other determinants of bilateral exports. In particular, the instrument should matter strongly for experience (strong instrument) but should not affect bilateral exports except through the experience channel (exclusion restriction).

We draw on recent work on the geographic spread of exports to construct our instruments. Chaney (2014) builds a network model of trade and shows that if a French firm exports to a destination, it is then more likely to subsequently enter a new destination that is geographically close to the first destination. Morales et al. (2015) use Chilean data to show that the entry of an exporter in a particular market increases the likelihood of his entry into other similar markets. In particular, they find that firms are more likely to export to countries sharing a border with countries to which they were exporting in the previous period. Similarly, Eaton et al (2008) use data on Colombian exporters to show that these exporters use neighboring markets as stepping stones to other Latin American markets. Going beyond single country exporters, Evenett and

[24] The coefficients remains unchanged if we also match on regional export or regional import experience. We also coded experience as a multi-valued treatment in intervals of 5 years and find similar results. These results are available upon request.

Venables (2002) examine 23 developing countries between 1970 and 1997, and show that a product is more likely to be exported from a certain country if the origin country is supplying the same product to nearby markets. Carrère and Strauss-Kahn (2014) use product level data for non-OECD exporters and find that experience is first acquired in neighboring, easy to access destinations before reaching to more distant, richer partners and ultimately serving the OECD. We exploit this pattern of the geographic spread of exports to construct our first instrument for experience as the experience of the origin o in all countries that are contiguous to destination, and take the simple average across all neighbors. For this to be valid, it instrument should be strongly correlated with experience in the destination, but exogenous to unobserved trade costs faced by exporters from o in the destination d. To ensure that this is a valid instrument, we average only over neighboring countries of d that are not part of a preferential trading arrangement with d. Under a PTA, there may be spatial correlation in unobserved trade costs and excluding these countries may invalidate the exclusion restriction. Call this instrument $E^{nbr}_{od,t}$.

Our second instrument relies on historical links between destination countries based on whether they were part of the same empire and/or administrative entity in the past (Mitchener and Weidenmier, 2008). We average the experience of o over countries that were or are the same state or the same administrative entity for a long period (25-50 years in the twentieth century, 75 years in the nineteenth and 100 years before) as the destination d. This definition covers countries belonged to the same empire (Austro-Hungarian, Persian, Turkish), countries that have been divided (Czechoslovakia, Yugoslavia, India) and countries that belong to the same administrative colonial area (e.g., Philippines and Mexico were subordinated to the New Spain viceroyalty). The data are from CEPII. [25]Call this instrument $E^{same}_{od,t}$.

Our key identifying assumption is that experience of the origin in the neighbors of the destination and experience in countries that were historically part of the same state/empire/administrative entity as the destination is unrelated to bilateral exports from o to d (except through its effect on experience $E_{od,t}$). Overall, we identify the experience effect on bilateral trade by basing it on the systematic component of experience for the neighbors of the destination and experience in countries that were the same as the destination, rather than destination-specific idiosyncrasies. We continue to implement our preferred specification which includes country-year and country-pair dummies as well as pair-specific trends. Inclusion of these comprehensive set of controls allows us to guard against a wide range of threats to our identifying assumption.

Table 2 reports a first-stage partial R^2 of 0.33 and a first-stage F-statistic of 1339.44 which easily clears the first-stage relevance tests, including the Stock-Yogo weak instruments test, indicating that our instruments

[25]Note that 444 country-pairs are neighbors but not part of the same country while 150 country-pairs were part of the same country in the past but are non-contiguous.

are strong so the weak instrument critique does not apply.[26] Column 7 in Table 2 shows that the coefficient on experience increases marginally to 0.481. In fact, the instrumental variable estimates are qualitatively and quantitatively similar to those presented in Column 2 of Table 2. Column 7 of Table 2 also reports the Hansen J-test of overidentification (OID) restrictions, which takes advantage of the overidentification to test instrument exogeneity. It tests the null hypothesis of overall validity of the instruments by analyzing the sample analog of the moment conditions used in the estimation process. Rejection of this null hypothesis casts doubt on the validity of the instruments. As Table 2 shows we obtain a $p-$value of 0.31 showing that our instruments are uncorrelated with the error term in the second-stage gravity equation.

The OID test is not a true test of the exclusion restrictions. It simply checks if all subsets of the instruments asymptotically return the same estimate of the effect of experience. While we use a comprehensive set of dummies and trends, it could be argued that an exporter's experience in neighboring countries simply reflects its region-specific exporting strategy. In this case, experience in neighbors affects bilateral trade not only via experience but also via some unmeasured channel that is exporter-region-year specific. To guard against this possible violation, we also included exporter-region-year dummies in the IV specification.[27] We then obtain a significant coefficient on experience of 0.372. Again, this is similar to a pooled estimate with exporter-region-year dummies where the coefficient is 0.390.

We also evaluated the validity of the instruments by restricting our sample to only oil exporters. It seems reasonable that these countries are likely to export to a destination regardless of whether they have experience in the neighboring countries or in countries that were the same as the destination. We can think of these exporting countries as "always takers" - countries where the instrument does not matter for experience. Any relationship between the instruments and the log of exports for such exporters is indicative of violation of the exclusion restrictions. For these countries, in a reduced form regression of the two instruments (and other gravity variables) on bilateral exports shows that these instruments are statistically insignificant.

3.3 Robustness checks

Next, we perform a further set of robustness checks shown in Table 3. We report only the coefficient of experience with our preferred specification from Column 2 of Table 2 that includes time-varying gravity variables, country-pair dummies and pair-specific trends. First, our experience variable based on the DOTS data is right-censored at 50.46 (equivalent to 58 years of continuous trade since 1948 with a retention parameter

[26] The first-stage yields: $E_{od,t} = \underset{(0.112)}{0.404}\, E^{nbr}_{od,t} + \underset{(0.110)}{0.253}\, E^{same}_{od,t} + controls$

[27] The regional classification is East Asia & Pacific, Europe & Central Asia, Latin America & Caribbean, Middle East & North Africa, North America, South Asia and Sub-Saharan Africa.

of 0.995). To account for the right-censoring, we added a dummy variable for all censored observations. We observe that including this dummy does not change the sign, significance, or magnitude of the estimates.[28] An alternative dataset from the Correlates of War (COW) Project tracks bilateral trade from 1870-2006 (Barbieri, Keshk and Pollins, 2012). Relying on this data to construct experience may mitigate the right-censoring concern. However, the COW data, by going further back in time, requires fairly strong assumptions about shifts in country identities through division, unification, and emergence from colonial rule.[29] Of more concern is the fact that COW provides trade data on former colonies in Asia, Africa and Latin America only when they become independent. In contrast, the DOTS data captures bilateral data for these countries prior to colonization. Therefore, experience constructed on the basis of COW data is also not free of measurement error. For this reason, we use the DOTS-based measure as our main measure of experience, and use the COW-based measure to examine the impact of censoring at 58 years.[30] With experience constructed using COW data, Table 3 shows that the coefficient on experience declines slightly to 0.364.[31]

Our findings for experience may be confounded by shifting political boundaries in Easter Europe following the collapse of communism. Therefore, we sequentially dropped 14 countries that were part of the Soviet Union, the 4 countries that formerly constituted Yugoslavia, and finally Czech Republic and Slovakia. In all cases, in Table 3, we observe a marginal increase in the coefficient of experience.

In the DOTS data, trade below $5,000 is set to zero, given the accuracy levels acknowledged by the IMF. Therefore, very small shipments are excluded in constructing our experience measure. Beyond $5000, we do give the same weight to small and large shipments in constructing the experience measure, based as it is on dummies for positive bilateral trade. This may lead to an over-estimation of the effect of experience. We therefore evaluate the sensitivity of our results to dropping small shipments by dropping the smallest 1 and 5% of bilateral export shipments. Again, there is a decline in the coefficient for experience but it remains strongly significant.

[28] Dropping the censored observations (Austin and Brunner, 2003) entails removing all country-pairs that had strictly positive trade in 1948. Since the increase in the volume of trade since then is mainly attributed to the expansion of trade between these pairs, rather than the emergence of new trading partnerships (Helpman, Melitz and Rubinstein, 2008), we lose 62.5% of our observations, that accounts for a majority of world trade.

[29] For instance, the COW-based experience assumes trade with the Austro-Hungarian as trade with both Austria and Hungary and assigns trade with Zanzibar as trade with Tanzania. More importantly, COW provides trade data on former colonies in Asia, Africa and Latin America only when they become independent. In contrast, the DOTS data captures bilateral data for these countries prior to colonization. Moreover, much of the data is missing prior to 1948.

[30] The correlation between the COW and the DOTS measures exceeds 0.8 for each of the years in our sample (1988-2006), rising to 0.91 in the year 2006.

[31] This is not surprising, since for nearly 90% of the observations, the two measures are very close to one another.

Recent papers by Helpman, Melitz and Rubinstein (2008), Evenett and Venables (2002), and Haveman and Hummels (2004) all highlight the prevalence of zero bilateral trade flows. For the bilateral trade data over the period 1988-2006, 37% of all possible bilateral trade flows show a zero value. Unobserved trade costs can endogenously create zeros and taking logs removes them from the sample, creating selection bias. Helpman, Melitz and Rubinstein (2008), henceforth HMR (2008), argue that the zeros in the trade matrix are not statistical but structural, due to fixed costs of exporting. They show that, in this case, firms self-select (or not) into exporting, which causes not just a selection bias but also a heterogeneity bias, due to changes in the composition of firms that export. We adopt the HMR (2008) methodology to account for these biases. In the first-stage, we estimate a probit that predicts the probability of having positive trade, for each year in the panel, $\varrho_{od,t}$, using the gravity variables and country-fixed effects. For the exclusion restrictions, we follow HMR (2008) and use a common religion index: $\sum_k (\mathrm{R}_{k,o} \times \mathrm{R}_{k,d})$,where $\mathrm{R}_{k,j}$ is the share of religion k in country j ($j = o, d$).[32] HMR (2008) show that including a polynomial in $\hat{\tilde{z}}^*_{od,t} = \Phi^{-1}\left(\varrho_{od,t}\right) + \hat{\tilde{\eta}}^*_{od,t}$ (where Φ is the cdf of the unit-normal distribution and $\hat{\tilde{\eta}}^*_{od}$ is the inverse Mills ratio) controls for firm-level heterogeneity and self-selection. The results for the second-stage regression (which also uses country-year fixed-effects), including the bias-correcting variables, show an elasticity coefficient close to 0.433 in Table 3. The inverse Mills ratio and the polynomial in $\hat{\tilde{z}}^*_{od,t}$ are significant at 1%, with signs similar to ones obtained in HMR (2008), showing the importance of correcting for the biases associated with structural zeros.

Finally, we use the methodology of Santos Silva and Tenreyro (2006) who show that log-linear specification of the gravity model in the presence of heteroskedasticity leads to inconsistent estimates.[33] Their methodology treats bilateral trade as a count variable and uses the Poisson Pseudo-Maximum Likelihood (PPML) to estimate the coefficients. Since the dependent variable is trade, rather than the log of trade, it also accounts for zeros in the trade matrix. One key difference between PPML and OLS is that the former puts a lot more weight on high expected trade dyads, since it tries to minimize the distance between real and expected trade in levels rather than in logs. Since our panel data yields more than 5000 country-year dummies, we apply the PPML methodology for a single year, 2004, with exporter and importer fixed-effects, to obtain convergence.[34]Here the coefficient on experience remains strongly significant - the coefficient of 0.577 lies between the OLS estimate and the within estimate.

[32] The set of religions we use is more comprehensive than that of HMR (2008), including k = Bahais, Buddhist, Chinese Universist, Christianity, Confucian, Ethnoreligionist, Hinduism, Jainism, Judaism, Islam, Shinto, Sikhism, Taoists and Zoroastrian. The data are from the Association of Religion Data Archives.

[33] If the error term in the standard log specification is heteroskedastic, its log is not orthogonal to the log of the regressors, leading to inconsistent estimates of the gravity elasticities.

[34] With this methodology, we also find that experience is significant at 1% each year from 1988-2006.

3.4 Placebo tests

Our contention is that experience allows exporters to learn about trade costs, especially unobserved trade costs (unobserved to the econometrician) in the destination market. To evaluate that this is indeed the case, we implement a series of placebo tests. First, we examined whether the *importing* experience of country o from d matters for *exports* from o to d. To the extent that the importing experience at the country-pair level also cumulates slowly over time, a significant effect of this would lead us to suspect that our exporting experience measure is some proxy for slow moving bilateral ties and not destination-specific trade costs encountered by exporters. This would render the interpretation of our findings questionable. Row 1 of Table 3 under Placebo Tests shows that importing experience of the origin from the destination does not matter for bilateral exports from the origin to the destination.

Second, we randomly varied the dummy for strictly positive trade $I_{od,t}$. For each country pair, we generated a random sequence of 0's and 1's and replaced the dummy $I_{od,t}$ with this random sequence. An experience measure for such a random sequence does not matter for bilateral trade as shown on Row 2. Next, we replaced $I_{od,t}$ only for each pair's first 50% of the observations and subsequently only for the second 50% of the observations. We used these two perturbations to construct additional placebo experience measures. Note that these measures coincide with our experience measure exactly half the time. Row 3 shows that replacing the first 50% of $I_{od,t}$ renders this placebo experience measure insignificant. This indicates that the first 50% of the experience measure has significant informational content, consistent with a retention parameter $\delta < 1$ and diminishing returns in experience. When we replace only the second 50% of $I_{od,t}$,the new placebo measure coincides with our original experience measure for the first 50% of the observations. Not surprisingly, since initial experience counts for more, Row 4 shows that the placebo experience measure that coincides with the original experience measure for the first 50% of observations does matter for bilateral exports. The lower coefficient of 0.13 also seems reasonable - while the first few years of trading experience is critical, additional accumulation of experience matters as well. If we include our original experience measure and either or both of these placebo experience measures, then the placebo measures are insignificant while the original experience measure continues to be strongly significant.

Third, we followed a quasi-randomization procedure and assigned the experience of each exporter in a particular destination to the alphabetical neighbor of the destination country from the same region. This allows us to check if our experience measure is simply picking up omitted variables that are specific to a particular region. Row 5 shows that this placebo measure does not affect bilateral exports. Next, we assign the experience of each exporter in a particular destination to the alphabetical neighbor of the exporter from the same region. Row 6 shows we obtain a negative coefficient on this placebo experience measure. Overall,

these placebo tests demonstrate that it is the experience of the exporter in the destination that really matters for bilateral exports.

Finally, it may be argued that our experience measure is simply a proxy for deeper integration between country-pairs, which may manifest itself as harmonization of worker, product and environmental standards, IP regulations, tax rules, etc. In fact, recent preferential trading arrangements increasingly emphasize such issues over formal trade barriers. As such our experience may be unrelated to unobserved pair-specific trade costs. To examine this, we replace the dependent variable with a measure of bilateral FDI stock from CEPII, available for a single year 2004.[35] We find a strong *negative* role for experience (see Row 7 of Table 3). To the extent that export experience facilitates bilateral trade by reducing trade costs, our finding is consistent with the contention that FDI and exports are substitutes for serving a particular destination.[36]

3.5 Robustness to model misspecification

In arriving at our point estimate for the elasticity of exports with respect to experience, and the standard error of this estimate (call this σ_1), we assumed that the log of experience affects the log of exports in a linear fashion, with all other gravity variables entering independently. Plausibly, different choices for this specification decision may lead to different point estimates of the effect of the experience. For instance, experience may matter more for certain country pairs, say those which are more distant, non-contiguous, and lack colonial, legal, or linguistic links. Athey and Imbens (2015) suggest a new test of robustness to such specification choices by supplementing conventional standard error of the point estimate with a scalar measure of the sensitivity of the estimates to a range of alternative models. We implement their test by splitting the sample into subsamples based on covariate values of every gravity variable, estimating the model separately for each subsample, and then combining the results to form a new weighted estimate of the overall effect.[37] The weights used are the fraction of observations in each subsample. The Athey and Imbens robustness measure is the standard deviation of the point estimate of the effect of the experience over the set of models(call this σ_2). A value of $\frac{\sigma_2}{\sigma_1}$ indicates that point estimate is non-robust and sensitive to model specification.[38]

[35] In this cross-section, we include only exporter and importer fixed-effects.

[36] We also experimented with sales of foreign affiliates of multinationals from origin o, in country d.The coefficient on the export experience measure is again negative but insignificant.

[37] Each of these specifications nests the baseline model as a special case.

[38] Athey and Imbens (2015) compare the ratio $\frac{\sigma_2}{\sigma_1}$ of an experimental setting with randomization to a non-experimental setting. In the experimental setting this ratio is 0.2 while it equals 3.338 for the nonexperimental data. They infer that the results for the experimental data are far more robust than those for the nonexperimental data.

The first row in Table 4 reports our baseline estimates, one with only country-year fixed effects and one with country-year, country-pair fixed effects and pair-specific trends. The subsequent rows split the sample on each of the gravity variables, all of which are binary save distance. For the time-varying binary variables such as WTO and PTA membership, we split the sample based on the values taken by these variables for the first year of data so that the sub-samples include the same country-pairs for all years. For distance, we split the sample into three subsamples based on mean ± 1 standard deviation of distance. The rest of the rows in Table 4 shows the weighted average of the coefficient estimates over these sample splits. We observe a similar coefficient on experience in both columns across sample slits. However, the standard deviation of the coefficient estimates (around the baseline estimate for experience) is 0.043 for the specification with only country-year dummies. This is more than 200% of the standard error of the baseline standard error of 0.018. Contrast this with the specification in the second column of Table 4 that includes all fixed effects and pair-specific trends. Here the standard deviation of estimates across model specifications is 0.03 which is lower than and equals 43% of the baseline standard error of 0.069. Therefore, once we account for time-invariant dyadic effects and pair-specific trends, our results seem very robust to misspecification. This also emphasizes the importance of accounting for unobserved dyadic effects and pair-specific trends.

4 Spillovers in experience

By focusing on trade between country-pairs, we cannot infer whether the decline in trade costs that we attribute to experience are at the firm-level, or whether there are spillovers in experience so that trade costs decline for whole sectors or even for the exporting country as a whole. We adopt two strategies to explore the existence of spillovers. First, we decompose aggregate COMTRADE 6-digit trade data into an extensive and intensive margin. Analyzing how each margin responds to the accumulation of experience, provides insight into spillovers in experience and the extent to which experience reduces fixed vs. variable costs of trade. Second, we analyze 4-digit commodity level trade to directly estimate the impact of spillovers in experience across sectors. We use the bilateral commodity trade data from NBER-UN (Feenstra et al, 2005) available at the 4-digit SITC Rev. 2 level of disaggregation. Even though we lack rich firm-level trade data to accurately measure firm export experience for a large set of destinations, the 4-digit commodity trade data is a reasonably good compromise. It spans years 1962-1999 allowing us to measure experience relatively accurately, and covers 98% of world trade. It also allows us to construct multiple measures of experience - at the industry-country-pair level, as well as destination-specific experience across sectors.

4.1 Inferring spillovers with margins of trade

Following Eaton, Kortum, and Kramarz (2004), Bernard, Jensen, Redding and Schott (2007), and Flam and Nordstrom (2007) we decompose bilateral exports $X_{od,t}$ as the product of an extensive margin ($N_{od,t}$), defined as the number of 6-digit products traded, and an intensive margin ($\overline{x}_{od,t}$), defined as the volume of exports per product so that

$$X_{od,t} = N_{od,t} * \overline{x}_{od,t} \tag{7}$$

Examining the coefficient on experience for the two margins provides insights into whether experience does spill over and whether experience works to reduce the fixed vs. variable costs of trade. In the absence of spillovers and with experience reducing only the fixed costs of trade, neither margin will adjust. If experience reduces variable trade costs but there are no spillovers in experience, we should expect no adjustments in the extensive margin along with an increase in the intensive margin - the number of products exported should remain unaffected as potential entrants do not benefit from experience while incumbent firms increase their exports raising the export per product. The extensive margin will increase with experience only if there are spillovers in experience. If experience spills over and reduces only the fixed costs of trade, the intensive margin should decline (there is no impact on exports of incumbent firms but the new entrants enter at a smaller scale reducing export per product). Finally, if experience spills over and reduces the variable costs of trade, the impact on the intensive margin is ambiguous - exports of incumbent firms increase which raises export per product but entry at a smaller scale reduces export per product.

Table 5 shows the effect of experience on the two margins of trade for our baseline specification and a specification that adds dyadic fixed effects. Columns 1-3 show that both margins increase with experience, with approximately 43% of the increase in overall trade coming via an adjustment of the extensive margin, and 54% coming via the intensive margin. Adding dyadic effects and pair-specific trends in Columns 4-6 does not affect the sign and significance of experience on the two margins, but now 60% of the effect of experience on trade is via the extensive margin and 40% via the intensive margin. Overall, both margins adjust with accumulated experience.[39]

The fact that the extensive margin increases with experience indicates that there are spillovers in experience across 6-digit sectors, with stronger effects in the more demanding specification. The fact that the intensive margin increases with experience allows us to rule out the case that experience reduces only the fixed costs of trade. Overall, these results are consistent with spillovers in experience and experience

[39] These results survive all the robustness checks in the previous section. Results available from authors upon request.

reducing the variable costs of trade.

4.2 Spillovers with disaggregate data

The previous results for the margins allows us to only infer spillovers by seeing that the extensive margin increases with experience. Next, we use data on bilateral trade at the 4-digit SITC level to provide more direct evidence for spillovers. With disaggregate data, we implement another change - measuring experience as cumulated exports rather than cumulated years of positive trade. With aggregate data, since changes in composition of exports over time and in unit values create measurement error in our experience variable, we chose to base the experience measure on number of years of trade. At the disaggregate level, the concerns related to changes in the composition of exports or in unit values over time are mitigated, though not completely eliminated. Therefore, our experience measures are based on the cumulated value of trade, rather than an indicator variable that captures trade vs. no-trade. The measure is more aligned with the learning-by-doing literature and also allows us to capture the intensity of experience since we now distinguish between 'small' and 'large' trade flows.[40][41]

While the NBER-UN 4-digit export data starts in 1962, a significant product reclassification was undertaken in 1983 (from SITC Rev 1 to SITC Rev 2). Given the potential for this re-classification inducing measurement error, for estimation we use data only from 1984 onwards.

We construct own experience at the industry-country-pair level $E^k_{od,t}$ as

$$E^k_{od,t} = X^k_{od,t-1} + \delta_1 E^k_{od,t-1} \tag{8}$$

where $X^j_{od,t-1}$ is the value of exports from o to d in 4-digit industry k and δ_1 is the retention parameter for own-learning. Next, for each origin country, we also measure destination-specific experience as

$$E_{od,t} = \sum_k X^k_{od,t-1} + \delta_2 E_{od,t-1} \tag{9}$$

This measure is based on the sum of exports across sectors to a particular destination and captures spillovers across 4-digit industries. We allow for a distinct retention parameter δ_2. For completeness, we also include a measure of industry-specific experience which is based on cumulating exports across destinations for each sector.

$$E^k_{o,t} = \sum_d X^k_{od,t-1} + \delta_3 E^k_{o,t-1} \tag{10}$$

[40] The NBER-UN data set includes data provided they exceed \$10,000 per year.

[41] Our results also work when we measure experience in terms of years of positive trade.

The industry-specific experience is related to the comparative advantage of a particular sector as well as to experience in production reducing production costs, rather than to experience in exporting to a particular destination that reduces trade costs. We use these three experience measures in a gravity equation for exports at the industry-country-pair level.

$$\ln X^k_{od,t} = \alpha_o \mu_o + \alpha_d \mu_d + \alpha_t \mu_t - \sum_{m=1}^{M} \theta \gamma_m z^m_{od,t} + \theta \lambda_1 \ln\left(E^k_{od,t}\right) + \theta \lambda_2 \ln\left(E_{od,t}\right) + \theta \lambda_3 \ln\left(E^k_{o,t}\right) + e_{od,t} \quad (11)$$

As with the aggregate data, we estimate equations (8)-(11) using non-linear least squares, estimating the three retention parameters indexed by δ and the three experience elasticities indexed by $\theta\lambda$. With more than 5 million observations, we estimate this system first without any dummies (Column 1 in Table 6) and then with origin, destination, and time fixed effects (Column 2 in Table 6). In both specifications we estimate an elasticity of export experience that is industry and destination specific of 0.9 and a retention parameter $\delta_1 = 0.64$. The retention parameter δ_2 for the destination-specific experience is 0.002 in Column 1 and declines to 0 when we add the fixed-effects. This suggests that for destination-specific experience across sectors, there is almost no retention beyond the previous period. At the same time, we do observe spillovers across sectors - a 1% increase in destination-specific experience across sectors, increases bilateral sectorial trade by 0.055%. We can think of this as a lower bound for spillovers since $E^k_{od,t}$ aggregates all trade within each 4-digit sector for a particular. Therefore, the estimated coefficient also encapsulates spillovers across sub-categories within a 4-digit sector.

Using more comprehensive fixed effects is computationally infeasible. Therefore, in Column 3, we set the retention parameters to the estimates in Column 2, and re-estimate the equation (11) with exporter-year and importer-year fixed effects. We observe a slight decline in the coefficient of own experience and a more than halving of the coefficient for destination-specific experience. Despite the decline, there still remains a significant role for both own and destination-specific experience.

Next, we evaluate the robustness of experience to an alternate identification strategy with disaggregate data. We restrict the exporting country to a single origin and add industry-year and destination-year fixed-effects. Since we restrict the data to a single origin country, identification relies on variation in experience by 4-digit industry over time. In this specification, we cannot estimate coefficients on either destination-specific or industry-specific experience since these are subsumed in the destination-year fixed-effects and industry-year fixed-effects, as are all gravity variables including even unobserved pair-specific trade costs. We pick 5 of the largest exporting countries: USA, China, Japan, Germany and India. For each of these we find an experience elasticity of trade ranging from 0.65 to 0.81, which is higher than that of our preferred specification in the aggregate data (with country-pair dummies and pair-specific trends). The higher coefficient is reasonable

since learning opportunities within a particular sector selling in a destination are likely to be stronger.

5 An Illustrative Model

Now, we introduce a model to illustrate the mechanisms behind the results of the empirical section, namely the impact of experience on trade flows and the role of experience spillovers on the dynamics of entry (the extensive margin). The illustrative nature of the exercise has pushed us to choose simplicity, rather than developing a complex model of the learning processes of firms in export markets (see Chaney, 2014).

We build on the Melitz (2003) and Chaney (2008) two-country model of heterogeneous firms with fixed costs of exports, introducing the impact of experience in lowering trade costs. The two countries, denoted o (origin) and d (destination) are symmetric, facing the same structural parameters governing preferences, technology and trade costs. In each country i at time t there is a non-traded final good, Y_{it}, produced competitively using a continuum of differentiated traded intermediate goods from each country, according to

$$Y_{it} = (\int_{k \in O,D} x_{kit}^{\frac{\varepsilon-1}{\varepsilon}} dk)^{\frac{\varepsilon}{\varepsilon-1}} \qquad i = o, d \tag{12}$$

where $\varepsilon \gg 1$ is the elasticity of substitution, and x_{kit} is the intermediate k used in the production of the final good in country i. Intermediate k may be in the continuum O or D, if produced, respectively, in o or d. Two elements should be noted. First, intermediates from o used in d are the exports from o to d, whereas, conversely, the intermediates from d used in o are the exports from d to o. Second, the elasticity of substitution between domestic intermediates is similar to that between domestic and foreign intermediates. Profit maximization by the producers of the final good yield the demand for intermediate k

$$x_{kit} = \left(\frac{p_{kit}}{P_{it}}\right)^{-\varepsilon} Y_{it}, \qquad i = o, d \tag{13}$$

where p_{kit} is the price of intermediate k in i, and P_{it} is the price of Y_{it} with

$$P_{it}^{1-\varepsilon} = \int_{k \in O,D} p_{kit}^{1-\varepsilon} dk, \qquad i = o, d \tag{14}$$

Each intermediate good k is produced by a monopolistic competitive firm according to $x_{kt} = a_k l_{kt}$, where a_k is the firm's productivity and l_{kt} is labor, the only factor of production in our model. Hence its unit cost of production is W_{it}/a_k, where W_{it} is the nominal wage in the country of firm k. The distribution of productivity is constant over time and symmetric across countries, captured by the density $g(a)$ on the support $[1, +\infty]$. In line with Chaney (2008), we simplify and assume that g follows a Pareto distribution, with a scaling parameter $\theta \gg 1$, such that: $g(a) = \theta a^{-(\theta+1)}$ and $P(a > \bar{a}) = \bar{a}^{-\theta}$.

Exports of each intermediate good k is subject to trade costs. There is a fixed export cost, which we assume is prohibitive at the beginning of time, but lowers to $F < \infty$ after a probabilistic event, which allows exports to begin. We will denote by $t = 0$ the time of this event. We assume that the fixed cost is constant, common to all firms and set in terms of the final good. There is also a variable (iceberg) trade cost. To capture the notion that variable trade costs decline as exporters increase familiarity with the local context and discover better and cheaper ways to transport, clear customs and distribute, we set the iceberg trade cost of intermediate k, τ_{kt}, as a function of experience, denoted by $E_{kt} \geq 0$:

$$\tau_{kt} = \tilde{\tau}(g + E_{kt})^{-\lambda} \tag{15}$$

where $\tilde{\tau} > 1$, g is positive but small, and $\lambda \geq 0$ governs the elasticity of the variable trade cost to experience.[42]

5.1 Entry into exports

The inclusion of time introduces a dynamic dimension for exports. For now, we will take the viewpoint of the firm producing intermediate k in country o, and focus on its decision to export to d - hence the subscript kd denotes variables that capture the activity of firm k (from o) in d. After the event that lowers the fixed cost at $t = 0$, the firm must decide whether or not to enter into the export market, and how much to export in each period $t \geq 0$. Firms are forward-looking and we assume that once a firm starts to export, it remains an exporter forever.[43] The present discounted value of all future export profits, contingent on exporting from $t_{kd} \geq 0$ onwards, is

$$\begin{aligned} V(t_{kd}|a_k) &= 0 + \sum_{t=t_{kd}}^{+\infty} \rho^{-t} R_{kdt} dt \qquad \text{where} \\ R_{kdt} &= \left(p_{kdt} - \frac{\tau_{kdt} W_{ot}}{a_k} \right) x_{kdt} - F \end{aligned} \tag{16}$$

The firm chooses when/whether to start exporting, t_{kd}, and then how much to export, x_{kdt}, to maximize $V(.|a_k)$. If $\max V(.|a_k) < 0$, the firm will never export k to d. The pricing decision is static, and the traditional mark-up rule: $p_{kdt} = (\varepsilon/\varepsilon - 1)\, \tau_{kdt} W_{ot}/a_k$ yields the amount of k exported to d as

$$x_{kdt} = \left(\frac{\varepsilon}{\varepsilon - 1} \frac{\tau_{kdt} W_{ot}}{a_k} \right)^{-\varepsilon} \frac{Y_{dt}}{P_{dt}^{-\varepsilon}} \tag{17}$$

[42] If we set $\lambda = 0$, the model reverts to Chaney (2008). Parameter $g \gtrsim 0$ ensures that the iceberg cost is not prohibitive when $E = 0$. It becomes redundant when $E \gg 0$.

[43] This will be true in the equilibrium of the model, since there are no shocks that could make the firm exit, after entry. In the empirical data, there is evidence of breaks in bilateral trade flows - see section 3.2.

and the profits from exporting to d as:

$$\begin{aligned} R_{kdt} &= \tilde{\varepsilon}\left(\frac{a_k}{\tau_{kdt}W_{ot}}\right)^{\varepsilon-1}\frac{Y_{dt}}{P_{dt}^{-\varepsilon}} - F \\ \text{where } \tilde{\varepsilon} &= (\varepsilon-1)^{\varepsilon-1}/\varepsilon^{\varepsilon} \end{aligned} \tag{18}$$

For simplicity, we assume proximity to a steady-state where the aggregate variables W_{ot}, Y_{dt} and P_{dt} can be taken as constant by the firm, in its decision on entry. Later, we will discuss the general equilibrium considerations and confirm the existence of such steady-state. In this case, the trade-off facing the firm relative to the timing of entry is given by

$$\frac{\partial V}{\partial t_{kd}} = \rho^{-t_{kd}}\left[-R_{kdt_{kd}} + \sum_{t=t_{kd}}^{+\infty}\rho^{-(t-t_{kd})}\,\frac{\partial R_{kdt}}{\partial \tau_{kdt}}\,\frac{\partial \tau_{kdt}}{\partial E_{kdt}}\,\frac{\partial E_{kdt}}{\partial t_{kd}}\,dt\right] \tag{19}$$

Delaying entry (increasing t_{kd}) has two potentially conflicting effects. One the one hand, there are the profits of the period, captured by the first term $-R_{kdt_{kd}}$, which can be either positive or negative. On the other, there are the losses on experience associated with delaying entry by one period, captured in the second term, which is negative, provided delaying entry lowers experience ($\partial E_{kdt}/\partial t_{kd} < 0$). When the operational profits are positive ($R_{kdt_{kd}} > 0$), the choice is clearly to enter. When the firm faces negative operational profits at time of entry ($R_{kdt_{kd}} < 0$), the firm faces a trade-off.

5.2 Experience and the dynamics of entry

Firm k, and all other firms, for that matter, benefit from its experience and that of its peers. Following the specification of the empirical model, we assume that: (i) Experience grows every year that there are positive exports to d (but is independent of the volume of exports), and (ii) there is a cumulative depreciation of the effect of past experience. $E_{kdt} = E_{dt} = I_{dt-1} + \delta E_{dt-1} = \sum_{i=1}^{t}\delta^{t-i}I_{di-1}$. If there is no interruption of trade for a country-pair, i.e. $I_{dt} = 1$ for $t > 0$, we obtain $E_{dt} = \left(1-\delta^{t}\right)/(1-\delta)$. Note that, when trade starts, $t = 0$, experience is zero; and when $t = +\infty$, it grows to an upper bound: $(1-\delta)^{-1}$. Our preferred specification - column 5 in Table 1 - estimated $\delta = 0.995$.

With this specification of experience, we obtain $\partial E_{kdt}/\partial t_{kd} = 0$ in (19) and $\partial V/\partial t_{kd}$ is negative (i.e. the firm enters) when R_{kdt} is positive. Firms have no incentive to take short term losses, because their gains from experience do not depend on their own entry decision, but on those of its preceding (more productive) peers. With t_{kd}^{*} denoting the optimal period of entry for firm k, the conditions for optimality imply three types of behavior by firms, in terms of exporting to d: (a) some firms are pioneers that start exporting from period 0 ($\partial V\left(t_{kd}^{*}=0\right)/\partial t_{kd} \leq 0, t_{kd}^{*} = 0$), (b) other firms are laggards that opt to begin exporting at a

later stage ($\partial V\left(t_{kd}^{*}>0\right)/\partial t_{kd}=0, t_{kd}^{*}>0$), and finally, (c) non-exporters opt out of exporting to d for the foreseeable future ($V\left(t_{kd}^{*}\right)<0$). Hence, different from Chaney (2008), we introduce a dynamic path of firm entry, due to the effect of experience on trade costs at the firm-level.

From (18) and (19), the marginal firm at t, i.e. the firm that satisfies the interior solution ($\partial V\left(t\right)/\partial t = -R_{kd}=0$), has productivity $\bar{a}_{dt}$:

$$\bar{a}_{dt}=\left(F\frac{P_{dt}^{-\varepsilon}}{\tilde{\varepsilon}Y_{dt}}\right)^{\frac{1}{\varepsilon-1}} W_{ot} \quad \tilde{\tau}\left[g+E_{dt}\right]^{-\lambda} \tag{20}$$

which declines with E_{dt}. Moreover, since $\partial V\left(t\right)/\partial t=-R_{kd}$ is decreasing, and $V(t|a_k)$ increasing, in a_k, all firms with higher productivity, $a_k>\bar{a}_{dt}$, are also exporting.

The dynamics of entry unfolds as follows. Firms with productivity $a_k \geq \bar{a}_{d0}=\left(F\frac{P_{d0}^{-\varepsilon}}{\tilde{\varepsilon}Y_{d0}}\right)^{\frac{1}{\varepsilon-1}} W_{o0}\tilde{\tau}g^{-\lambda}$ are pioneers. Firms of lower productivity may enter as experience increases: since $E_{dt}=\left(1-\delta^{t}\right)/(1-\delta)$, the firm with productivity $a_k<\bar{a}_{d0}$ enters into exports at t_{kd}^{*} given by:

$$\frac{1-\delta^{t_{kd}^{*}}}{1-\delta}=\left(F\frac{P_{dt}^{-\varepsilon}}{\tilde{\varepsilon}Y_{dt}}\right)^{\frac{1}{\lambda(\varepsilon-1)}}\left(\frac{\tilde{\tau}W_{ot}}{a_k}\right)^{\frac{1}{\lambda}}-g$$

Since E is bounded, $\bar{a}_{dt}$ has a lower bound, and firms with productivity $a_k \leq \left(F\frac{P_{dt}^{-\varepsilon}}{\tilde{\varepsilon}Y_{dt}}\right)^{\frac{1}{\varepsilon-1}} W_{ot}\tilde{\tau}\left[g+(1-\delta)^{-1}\right]^{-\lambda}$ never export.

Finally, we define the extensive margin as the mass of firms from o exporting to d, given by

$$\begin{aligned} N_{dt} &\equiv \int_{k\in O, a_k>\bar{a}_{dt}} g(a_k)da_k=\int_{k\in O, a_k>\bar{a}_{dt}} -\theta a_k^{-(\theta+1)}da_k=\bar{a}_{dt}^{-\theta} \\ &= F^{-\frac{\theta}{\varepsilon-1}}\left(\frac{P_{dt}^{-\varepsilon}}{\tilde{\varepsilon}Y_{dt}}\right)^{-\frac{\theta}{\varepsilon-1}} W_{ot}^{-\theta}\tilde{\tau}^{-\theta}\left[g+E_{dt}\right]^{\theta\lambda} \end{aligned} \tag{21}$$

with total exports given by[44]

$$\begin{aligned} X_{dt} &\equiv \int_{a_k>\bar{a}_{dt}} p_{kt}x_{kt} \quad \theta a_k^{-(\theta+1)}da_k = \\ & \frac{\theta\varepsilon}{\theta-(\varepsilon-1)}F^{1-\frac{\theta}{\varepsilon-1}} \quad \left(\frac{P_{dt}^{-\varepsilon}}{\tilde{\varepsilon}Y_{dt}}\right)^{-\frac{\theta}{\varepsilon-1}} W_{ot}^{-\theta}\tilde{\tau}^{-\theta}\left[g+E_{dt}\right]^{\theta\lambda} \end{aligned} \tag{22}$$

and the intensive margin, i.e. the average volume of exports by exporter, as

$$\frac{X_{dt}}{N_{dt}}=\frac{\theta\varepsilon}{\theta-(\varepsilon-1)}F$$

[44] To obtain the expression for N, we assume $\theta>(\varepsilon-1)$, as usual in the literature (Melitz, 2003)

The presence of laggards, who share the benefits from the experience of incumbents, implies that extensive margin increases with experience. The implications of experience for the intensive margin are more complex. While the exports of incumbents grow due to lower variable costs, the lower exports of laggards (with weaker productivity) that enter into exports reduce the average exports per firm. The implication is that the impact on the intensive margin is ambiguous. Under a Pareto distribution, as assumed here, the two effects cancel out and declines in the variable trade cost leave the intensive margin unaffected - as shown in the previous equation. Dutt et al. (2013) show that, for other plausible distributions, the incumbent effect dominates and the decline in the variable component of trade costs raises the intensive margin which would be in accordance with our empirical results in Section 4.1.

The log-linearization of the expressions for total exports and the extensive margin, eliminating g, yield an expression that is close to our empirical specification

$$\ln X_{odt} = \ln \frac{\theta \varepsilon \tilde{\varepsilon}^{\frac{\theta}{\varepsilon-1}}}{\theta-(\varepsilon-1)} + \frac{\theta}{\varepsilon-1} \ln \frac{Y_{dt}}{P_{dt}^{-\varepsilon}} - \theta \ln W_{ot} - \left[\theta \ln \tilde{\tau}_{odt} + \left(\frac{\theta}{\varepsilon-1}-1\right) \ln F_{od}\right] + \lambda\theta \ln E_{odt} \tag{23}$$

where we have introduced the subscript o for the exporting country. This expression mirrors the empirical gravity equation estimated in (6a), with the first term captured by the constant, the second and third terms captured by country-year dummies, and the terms within the squared brackets proxied by the gravity variables. The accumulation of experience expands bilateral trade with elasticity $\lambda\theta$, a term that was estimated previously in reduced form - column (5) in Table 1 - as $\lambda\theta = 0.884$.

5.3 General equilibrium considerations

Next we introduce general equilibrium conditions to determine W_{ot}, Y_{dt} and P_{dt}. First, the prices of the final goods, given in (14), yield

$$P_{dt}^{1-\varepsilon} = \frac{\varepsilon\theta}{\varepsilon-1}\left(\int_{\bar{a_t}}^{\infty} \frac{\tau_{kdt}^{1-\varepsilon} W_{ot}^{1-\varepsilon}}{a_k} a_k^{-(\theta+1)} da_k + \int_1^{\infty} \frac{W_{dt}^{1-\varepsilon}}{a_k} a_k^{-(\theta+1)} da_k\right) = \frac{\varepsilon}{\varepsilon-1}\frac{\theta}{\theta+1}\left(\tau_{kdt}^{1-\varepsilon} W_{ot}^{1-\varepsilon} \bar{a_t}^{-(\theta+1)} + W_{dt}^{1-\varepsilon}\right) \tag{24}$$

In the labor market, demand includes the production of intermediates for the domestic final good, $L_o \equiv \int_1^{\infty} \frac{x_{kot}}{a_k} \theta a_k^{-(\theta+1)} da_k$, and for the foreign final good, $L_d \equiv \int_{\bar{a_{dt}}}^{\infty} \frac{x_{kdt}}{a_k} \theta a_k^{-(\theta+1)} da_k$. From (17), the equilibrium can be written

$$\bar{L} = \frac{\theta}{\theta-(\varepsilon-1)}\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\varepsilon} \frac{Y_{dt}}{P_{dt}^{-\varepsilon}} W_{ot}^{-\varepsilon}\left[1+\bar{a}_{dt}^{\varepsilon-1-\theta}\right] \tag{25}$$

where $\bar{a}_{dt}$ is given in (20) and $\bar{L}$ is labor supply. Corresponding equations to (24) and (25) exist for the trade partner. Finally, (assuming financial autarky), from (22), the trade balance condition $X_{odt} = X_{dot}$ yields

$$\frac{Y_{dt}}{(W_{ot}/P_{dt})^{\varepsilon-1}} = \frac{Y_{ot}}{W_{dt}^{\varepsilon-1}} \tag{26}$$

The model provides 5 equations, which, taking the final good from o as numeraire: $P_{ot} = 1$, allow us to obtain P_{dt}, W_{ot}, W_{dt}, Y_{ot} and Y_{dt}, in terms of L_o, F, $\tilde{\tau}$ and E_t. We assume that countries are symmetric in that they face the same parameters. This implies that $P_{dt} = 1$, $W_{ot} = W_{dt} = W_t$ and $Y_{ot} = Y_{dt} = Y_t$, rendering (24)-(26) into

$$\begin{aligned} \frac{\varepsilon}{\varepsilon-1}\frac{\theta}{\theta+1}W_t^{1-\varepsilon}\ \left[\tilde{\tau}^{1-\varepsilon}[g+E_t]^{\lambda(\varepsilon-1)}\bar{a_t}^{-(\theta+1)}+1\right] &= 1 \\ \frac{\theta}{\theta-(\varepsilon-1)}\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\varepsilon}Y_tW_t^{-\varepsilon}\ \left[\bar{a}_t^{\varepsilon-1-\theta}+1\right] &= \bar{L} \end{aligned}$$

Bilateral trade and the extensive margin can be obtained from W_t and Y_t, as follows

$$X_t = \frac{\theta\varepsilon}{\theta-(\varepsilon-1)}F^{1-\frac{\theta}{\varepsilon-1}}\quad(\tilde{\varepsilon}Y_t)^{\frac{\theta}{\varepsilon-1}}\,W_t^{-\theta}\tilde{\tau}^{-\theta}\,[g+E_t]^{\theta\lambda}\qquad N_t = F^{-\frac{\theta}{\varepsilon-1}}\,(\tilde{\varepsilon}Y_t)^{\frac{\theta}{\varepsilon-1}}\,W_t^{-\theta}\tilde{\tau}^{-\theta}\,[g+E_t]^{\theta\lambda} \quad (27)$$

Experience, given above as $E_t = \frac{1-\delta^t}{1-\delta}$, determines the influence of the accumulation of experience over time on the system. The accumulation of experience yield a decline in trade costs which, in addition to expanding bilateral trade and the extensive margin, supports in increase in output and real wages through the enhanced gains from trade. This process of economic growth stops the long-run (steady-state) equilibrium, when experience reaches an upper bound given by $Lim_{t\to\infty}E_t = (1-\delta)^{-1}$. In the transition of experience accumulation, the expansion of output further contributes to the rise in exports generated directly by experience, while the rise in wages works to slow down trade growth.

5.4 A numerical exercise

Now, we perform a simple numerical exercise that estimates the effect of experience on bilateral trade. We simulate data for a sample of 201 symmetric country-pairs and estimate the effect of experience using a regression approach. The purpose of this exercise is to understand how a simple model of dynamic trade costs, including its general equilibrium effects, can produce some of the results found in Section 3. It is important to stress that this is not a full-fledge calibration exercise.

The structural parameters of the model are: $\theta, \delta, \lambda, \varepsilon$ and g. Our preferred empirical estimates, in column (2) of Table 2, matched against the structural coefficients in (23), suggest $\lambda\theta = 0.438$ and $\delta = 0.995$. With the Pareto parameter (θ) set to 3.8 (see Ghironi and Melitz, 2005), we obtain $\lambda = 0.232$. We choose $g = 0.01$ and set the elasticity of substitution (ε) to 3.4.[45]

Our data-generating sample of 201 country-pairs varies along three dimensions: (1) the iceberg transport cost parameter $\tilde{\tau}$; (ii) the fixed cost of exports F; and (iii) the number of years that each country-pair has

[45] The literature has found values for ε ranging from 1.5 to 10. We follow Ghironi and Melitz (2005), setting $\varepsilon = 3.4$.

been trading, t. Following Ghironi and Melitz (2005), the constellation of parameters for these 201 symmetric country-pairs is set as follows. The iceberg transport cost, $\tilde{\tau}$, is set randomly from the interval $[1.2, 1.7]$, based on estimates from the literature using distance and geography variables. The number of years of experience is also set randomly from the distribution in our dataset, with a maximum at 56. Then, the fixed cost is calibrated so that (1) the share of exporting firms lies between $[0.1, 0.7]$ and (2) the distribution of total exports is not too far from that in the data.

Based on this data on trade costs and experience, and the ensuing total exports obtained from (27), for each of the 201 country-pairs, we run a regression that mimics the gravity specification, with the gravity variables replaced by our measures of trade costs:

$$log(X_{ij}) = \beta_0 + \beta_1 logE_{ij} + \beta_2 log(\tilde{\tau}_{ij}) + \beta_2 log(F_{ij}) + \beta_4 log(Y) + u_t$$

where $E_{ij} = \frac{1-\delta^{t_{ij}}}{1-\delta}$. Compared with the standard gravity specification, the effects of the domestic incomes in the two countries, captured in (6a) with the two country-time dummies and in (23) with the terms on Y_d and W_o, are replaced here with a single parameter β_4, because domestic income and wages in both countries are the same due to symmetry. Moreover, given the data generation procedure, Y is endogenously determined by trade costs and experience, raising the potential for collinearity and reduced significance.

The regression results with and without Y are reported in Table 7. In the regression without Y, trade costs and experience have the expected effects and are both statistically and economically significant. The inclusion of Y lowers the significance of the estimates, namely of the parameter on variable trade costs, due to collinearity. Moreover, it lowers the magnitude of the effects of fixed costs and experience, because Y captures the indirect effect of these variables on trade flows, through the rise in income (gains from trade). This indirect effect explains also why the coefficient of experience is slightly higher than $\lambda\theta = 0.438$ when income is not included. Hence our simple dynamic model where trade costs decrease with experience has the elements to help us understand the results from Section 2.

As a final robustness check, we simulated a version of the model in which experience does not matter, but where iceberg transport costs vary over time. To be consistent with the previous quantitative exercise, we calibrate this process so that the effective trade costs evolve over time according to $\tau = \tilde{\tau} * \frac{1-\delta^{t_{ij}}}{1-\delta}$. That is, the variable trade costs decrease over time for reasons that do not necessarily have to do with experience. When we estimate the standard gravity equation on the same 201 simulated country pairs, consistent with the model, we obtain a negative and statistically coefficient on both the iceberg and fixed trade costs and a positive coefficient on output. Adding our measure of experience, yields a coefficient of 0 for experience, while both iceberg transport costs and fixed costs continue to have negative and statistically significant

coefficients. Our findings suggest that in a gravity equation derived from a model that does not explicitly account for experience, adding experience as an additional regressor results in an insignificant effect of this variable.

6 Conclusion

In this paper we show that experience matters a lot for bilateral trade flows, and that the effect is strong and persistent. Our non-parametric estimates show that an additional year, on average, increases bilateral trade by approximately 6%, while a parametric specification that controls for country and country-pair fixed effects as well as pair-specific trends and gravity variables, estimates an elasticity of bilateral trade with respect to our measure of experience of 0.48%. We find also that the association of experience with the rise in years of positive trade is subject to a small but significant depreciation, which suggests an upper-bound for experience. We employ multiple identification strategies, and confirm the robustness of our results to traditional econometric concerns of the gravity literature. A series of placebo tests helps shed light on how experience matters for bilateral trade. The role of experience helps explain the dominance of country-pairs that have been trading with each other for a long time in the growth of world trade, first highlighted in Helpman, Melitz and Rubinstein (2008).

We also demonstrate spillovers in experience by showing that the extensive margin responds strongly to experience, that nearly 50% of the impact of experience on bilateral trade is at the extensive margin. We also document the importance of spillovers across 4-digit industry sectors. We recognize that there are interesting dynamics and spillovers at the firm level as an emerging literature has started to document (e.g., Eaton et al, 2012) and that we are unable to shed much light on. However, with the short time-span of existing firm-level data, measuring experience accurately at the firm-level is a non-trivial task. Therefore, our results on spillovers at the level of country-pairs, and at the sectorial level are indicative at best. Addressing the specific effects of experience at the firm-level, and the spillovers across firms remains a challenge for future work, as the time span of firm-level datasets expands.

We also illustrate the impact of experience in a simple dynamic model of international trade with heterogeneous firms, where experience lowers variable trade costs and its benefits are shared across firms. This introduces export pioneers and laggards into the static specification of Melitz-Chaney, a phenomenon in concordance with empirical (Iacovone and Javorcik, 2010) and theoretical work on export emergence (Hausman and Rodrik, 2003). We show that overall exports and the extensive margin increases with experience, which generates rises in wages and output. Subsequently, we rely on the model to generate an artificial sample with

differing trade costs (both variable and fixed trade costs) and experience, chosen to generate a distribution of exports and output consistent with the data. Regressions on the simulated data support a positive effect of experience on bilateral trade.

Given our finding that the benefits from experience tend to be shared among firms and industries, the presence of dynamic effects opens the possibility of external effects and the scope for policy: supporting the entry of early exporters, even temporarily, may lower the trade costs for non-exporters and encourage entry. Reassessing the role of export-promotion in this general, and of export promoting agencies, from the normative and positive perspectives (e.g. in the East-Asian miracle), remains a challenge for future research.

References

[1] Abadie, A., and G. W. Imbens. 2006. "Large sample properties of matching estimators for average treatment effects," Econometrica 74: 235–267.

[2] Aitken, Brian Gordon H. Hanson, Ann E. Harrison. 1997. "Spillovers, foreign investment, and export behavior," Journal of International Economics, 43(1–2), 103-132.

[3] Albornoz, Facundo, Héctor F. Calvo Pardo, Gregory Corcos, and Emanuel Ornelas. 2012. "Sequential exporting." Journal of International Economics, 88(1): 17-31.

[4] Alvarez, R., Faruq, H. and Lopez, R. (2007). "New products in export markets: learning form experience and learning from others," mimeo, Indiana University, Bloomington

[5] Allen, Treb. 2012. "Information Frictions in Trade." Working Paper.

[6] Altonji JG, Elder TE, and Taber CR. 2005. "Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools," Journal of Political Economy 113:151-184.

[7] Anderson, James E., and Douglas Marcouiller. 2002. "Insecurity and the Pattern of Trade: An Empirical Investigation." The Review of Economics and Statistics, 84(2): 342-352.

[8] Anderson, James E., and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." American Economic Review, 93(1): 170-192.

[9] Anderson, James E., and Eric van Wincoop. 2004. "Trade Costs." Journal of Economic Literature, 42(3): 691-751.

[10] Artopoulos, Alejandro, Daniel Friel, and Juan Carlos Hallak. 2013. "Export emergence of differentiated goods from developing countries: Export pioneers and business practices in Argentina," Journal of Development Economics, 105: 19-35.

[11] Athey, Susan and G.W. Imbens (2015). "A measure of robustness to misspecification," American Economic Review, 105 (5), 476–480.

[12] Austin, Peter C. and Lawrence J. Brunner. 2003. "Type I Error Inflation in the Presence of a Ceiling Effect," The American Statistician, Vol. 57, No. 2, pp. 97-104

[13] Baier, Scott L. and Jeffrey H. Bergstrand. 2007 "Do free trade agreements actually increase members' international trade?" Journal of International Economics, 71, 72-95.

[14] Baier, Scott L. and Jeffrey H. Bergstrand. 2009. "Estimating the effects of free trade agreements on international trade flows using matching econometrics," Journal of International Economics, 77 (1), 63-76.

[15] Barbieri, Katherine. 2002. The Liberal Illusion: Does Trade Promote Peace? Ann Arbor: University of Michigan Press

[16] Barbieri, Katherine, and Omar Keshk. 2012. Correlates of War Project Trade Data Set Codebook, Version 3.0.Online: http://correlatesofwar.org.

[17] Bastos, Paulo and Joana Silva. 2012. "Networks, Firms, and Trade," Journal of International Economics, 87 (2), 352-364.

[18] Berkowitz, Daniel, Johannes Moenius, and Katharina Pistor. 2006. "Trade, Law, and Product Complexity," Review of Economics and Statistics, 88(2), 363-373.

[19] Bernard, A. B., Van Beveren, I. and Vandenbussche, H. 2014. "Multi-Product Exporters and the Margins of Trade." Japanese Economic Review, 10.1111/jere.12030.

[20] Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding and Peter K. Schott. 2007. "Firms in International Trade." Journal of Economic Perspectives, 21(3): 105-130.

[21] Carrère, Céline and Strauss-Kahn, Vanessa. 2014. "Developing countries exports survival in the OECD: Does experience matter?" CEPR Discussion Paper No. DP10059.

[22] Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." American Economic Review, 98(4): 1707-1721.

[23] Chaney, Thomas. 2014. “The Network Structure of International Trade,” American Economic Review, 104(11): 3600-34.

[24] Combes, P-P., Lafourcade M. and T. Mayer, 2005. “The Trade Creating Effects of Business and Social Networks: Evidence from France,” Journal of International Economics, 66,1-29.

[25] Clerides S., Lach S., Tybout J., 1998. Is learning-by-exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. Quarterly Journal of Economics 113, 903-47.

[26] Dehejia, RH, and Wahba, S. 1999. “Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs,” Journal of the American Statistical Association 94:1053-1062.

[27] Disdier, Anne-Célia, and Keith Head. 2008. “The Puzzling Persistence of the Distance Effect on Bilateral Trade” The Review of Economics and Statistics, 90(1): 37-48.

[28] Dutt, Pushan, and Daniel Traca. 2010. “Corruption and Bilateral Trade Flows: Extortion or Evasion?” The Review of Economics and Statistics, 92(4): 843-860.

[29] Dutt, Pushan, Ilian Mihov, and Timothy Van Zandt (2013). "The Effect of WTO on the Extensive and Intensive Margins of Trade.” Journal of International Economics, 91(2): 204-219.

[30] Eaton, Jonathan, and Samuel Kortum. 2002. “Technology, Geography, and Trade.” Econometrica, 70(5): 1741-1779.

[31] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2004. “Dissecting Trade: Firms, Industries, and Export Destinations.” American Economic Review, 94(2): 150-154.

[32] Eaton, Jonathan, Samuel Kortum, Brent Neiman, and John Romalis. 2010. “Trade and the Global Recession.” National Bank of Belgium Working Paper Research 196.

[33] Eaton, Jonathan, Marcela Eslava, Maurice Kugler, and James R. Tybout. 2008. “The margins of entry into export markets: Evidence from Colombia," in Elhanan Helpman, Dalia Marin, and Thierry Verdier, eds., The Organization of Firms in a Global Economy, Cambridge: Harvard University Press.

[34] Eaton, J., Eslava, M., Krizan, C., Kugler, M., Tybout, J., 2012. “A search and learning model of export dynamics,” Working paper.

[35] Egger, Peter H. and Andrea Lassmann. 2012. “The language effect in international trade: A meta-analysis,” Economics Letters, 116(2), 221-224.

[36] Evenett, Simon, and Anthony Venables. 2002. “Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows.” www.alexandria.unisg.ch/Publikationen/22177

[37] Flam, Harry, and Hakan Nordström. 2006. “Euro Effects on the Intensive and Extensive Margins of Trade.” CESifo Working Paper Series 1881.

[38] Feenstra, Robert C. and Hiau Looi Kee 2008. “Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity.” Journal of International Economics, 74, 500–518.

[39] Freund, Caroline, and Martha Denisse Pierola. 2010. “Export Entrepreneurs: Evidence from Peru.” The World Bank Policy Research Working Paper Series 5407.

[40] Ghironi, Fabio & Marc J. Melitz, 2005. "International Trade and Macroeconomic Dynamics with Heterogeneous Firms," The Quarterly Journal of Economics, 120(3): 865-915.

[41] Glaeser, Edward L., and Andrei Shleifer. 2002. “Legal Origins.” The Quarterly Journal of Economics, 117(4): 1193-1229.

[42] Greif, Avner, 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies.” Journal of Political Economy, 102(5): 912-950.

[43] Grossman, G. M., 1998. “Comment” (pp. 29–31), in J. A. Frankel (Ed.), The Regionalization of the World Economy, Chicago: University of Chicago Press.

[44] Guiso, Luigi, Paola Sapienza, and Luigi Zingales. “Cultural biases in economic exchange?” The Quarterly Journal of Economics, 124(3), 1095-1131.

[45] Hausmann, Ricardo and Dani Rodrik. 2003. “Economic Development as Self Discovery,” Journal of Development Economics, 72 (2): 603–633.

[46] Haveman, Jon, and David Hummels. 2004. “Alternative hypotheses and the volume of trade: the gravity equation and the extent of specialization.” Canadian Journal of Economics, 37(1): 199-218.

[47] Head, Keith, and John Ries. 1998. “Immigration and Trade Creation: Econometric Evidence from Canada.” Canadian Journal of Economics, 31(1): 47-62.

[48] Head, Keith, Thierry Mayer, and John Ries. 2010. “The Erosion of Colonial Trade Linkages after Independence.” Journal of International Economics, 81(1): 1-14.

[49] Head, Keith, and Thierry Mayer. 2013. "Gravity Equations: Workhorse, Toolkit, and Cookbook." forthcoming in Handbook of International Economics Vol. 4 (ed. Gopinath, Helpman, and Rogoff)

[50] Head, Keith and Thierry Mayer. 2013. "What Separates Us? Sources of Resistance to Globalization." Canadian Journal of Economics, 46(4): 1196-1231.

[51] Heckman JJ, Ichimura H, and PE Todd PE. 1997. "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme," Review of Economic Studies 64: 605-54.

[52] Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." The Quarterly Journal of Economics, 123(2): 441-487.

[53] Iacovone, L. and B. Javorcik. 2010. "Multi-product exporters: product churning, uncertainty and export discoveries," The Economic Journal, 120: 481–499

[54] Iacus SM, King G, and G Porro. 2011. "Multivariate matching methods that are monotonic imbalance bounding," Journal of the American Statistical Association 106:345-361.

[55] Iacus SM, King G, and G. Porro 2012. "Causal inference without balance checking: Coarsened exact matching," Political Analysis 20:1-24.

[56] Imbens G. 2004. "Nonparametric estimation of average treatment effects under exogeneity: A review," Review of Economics and Statistics 86: 4-29.

[57] Kneller, Richard and Mauro Pisu. 2011."Barriers to Exporting: What are They and Who do They Matter to?" The World Economy 34(6), 893-930.

[58] Lederman, Daniel, Marcelo Olarreaga, and Lucy Payton. 2010. "Export promotion agencies: Do they work?" Journal of development economics 91(2): 257-265.

[59] Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." Econometrica, 71(6): 1695-1725.

[60] James Mitchener, K. J. and Marc Weidenmier. 2008. "Trade and Empire," The Economic Journal, 118 (533) 1805-1834.

[61] Molina, Danielken and M. Muendler 2013. "Preparing to export," NBER Working Paper 18962.

[62] Morales, Eduardo, Sheu G., and Zahler, A. 2015. "Extended Gravity," Working paper

[63] Rauch, James E. 1999. "Networks versus Markets in International Trade." Journal of International Economics, 48(1): 7-35.

[64] Rauch, James E., and Vitor Trindade. 2002. "Ethnic Chinese Networks in International Trade." The Review of Economics and Statistics, 84(1): 116-130.

[65] Santos Silva, J. M. C., and Silvana Tenreyro. 2006. "The Log of Gravity." The Review of Economics and Statistics, 88(4): 641-658.

[66] Segura-Cayuela, R., and J. Vilarrubia. 2008. Uncertainty and entry into export markets, Bank of Spain Working Paper

Table 1: Experience and Bilateral Trade

	(1)	(2)	(3)	(4)	(5)
Retention parameter δ	0.963***	0.961***	0.962***	1.019***	0.995
	(0.002)	(0.002)	(0.002)	(0.001)	
Experience	0.753***	0.817***	0.817***	0.798***	0.887***
	(0.011)	(0.011)	(0.011)	(0.009)	(0.018)
Both in GATT/WTO	-0.048***	0.022**	0.020*	0.010	0.193***
	(0.011)	(0.011)	(0.011)	(0.018)	(0.062)
PTA	0.279***	0.258***	0.258***	0.719***	0.724***
	(0.011)	(0.011)	(0.011)	(0.02)	(0.053)
GSP	0.174***	0.021	0.025*	0.157***	0.241***
	(0.015)	(0.015)	(0.015)	(0.018)	(0.048)
Common currency	0.375***	0.414***	0.409***	0.163***	0.278**
	(0.045)	(0.045)	(0.045)	(0.043)	(0.121)
Distance	-1.127***	-1.134***	-1.132***	-1.261***	-1.259***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.018)
Contiguity	0.878***	0.839***	0.842***	0.563***	0.537***
	(0.028)	(0.029)	(0.029)	(0.028)	(0.082)
Colonial relationship	0.869***	0.877***	0.867***	0.913***	0.946***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.088)
Common language	0.362***	0.356***	0.357***	0.423***	0.449***
	(0.015)	(0.015)	(0.015)	(0.016)	(0.039)
Common law	0.83***	0.788***	0.794***	0.406***	0.380***
	(0.018)	(0.018)	(0.018)	(0.021)	(0.053)
GDP exporter	1.018***	1.016***	1.017***	0.443***	
	(0.003)	(0.003)	(0.003)	(0.023)	
GDP importer	0.873***	0.869***	0.873***	0.690***	
	(0.002)	(0.002)	(0.002)	(0.027)	
Time trend		-0.070***			
		(0.001)			
Time dummies	No	No	Yes	Yes	NA
Country dummies	No	No	No	Yes	NA
Country-year dummies	No	No	No	No	Yes
Observations	226223	226223	226223	226223	240873
R-squared	0.69	0.69	0.69	0.76	0.76

Standard errors in parentheses are clustered on country-pair: * significant at 10%; ** significant at 5%; *** significant at 1%

Columns 1-4 show non-linear least squares estimates; Column 5 uses δ with best fit

Table 2: Identification

	(1) Baseline	(2) Pair-fixed effects	(3) Pair-fixed effects + lagged DV	(4) Cross-section	(5) Matching estimate	(6) Matching estimate	(7) IV estimate
Experience	0.887***	0.438***	0.333***	1.086***	1.024***	1.022***	0.481***
	(0.018)	(0.069)	(0.100)	(0.037)	(0.122)	(0.116)	(0.114)
Both in GATT/WTO	0.193***	-0.053	-0.063	0.222	0.025	-0.062	-0.054
	(0.062)	(0.065)	(0.066)	(0.136)	(0.234)	(0.293)	(0.065)
PTA	0.724***	-0.052	-0.010	0.596***	0.196	1.727***	-0.051
	(0.053)	(0.048)	(0.046)	(0.064)	(0.223)	(0.276)	(0.048)
GSP	0.241***	0.122	0.085	0.030	0.329	0.532	0.121
	(0.048)	(0.109)	(0.104)	(0.068)	(0.559)	(0.461)	(0.109)
Common currency	0.278**	0.105**	0.077*	-0.198	-0.121	0.938	0.104**
	(0.121)	(0.049)	(0.044)	(0.150)	(0.539)	(0.915)	(0.049)
Distance	-1.259***			-1.362***	-1.588***	-1.571***	
	(0.018)			(0.025)	(0.085)	(0.085)	
Contiguity	0.537***			0.602***	2.885***	0.972	
	(0.082)			(0.098)	(0.582)	(0.697)	
Colonial relationship	0.946***			0.899***	0.957**	0.685	
	(0.088)			(0.100)	(0.461)	(0.666)	
Common language	0.449***			0.538***	0.603***	0.659***	
	(0.039)			(0.056)	(0.188)	(0.179)	
Common law	0.380***			0.334***	0.303	0.278	
	(0.053)			(0.076)	(0.216)	(0.211)	
Lagged trade			0.068***				
			(0.006)				
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-pair dummies	No	Yes	Yes	No	No	No	Yes
Pair specific trends	No	Yes	Yes	NA	NA	NA	Yes
Observations	240873	240873	203650	19272	6259	5823	240873
R-squared	0.76	0.94	0.94	0.76	0.67	0.61	0.94
First-stage F-statistic							1339.44**
First-stage Partial R^2							0.33
OID test p-value							0.31

Standard errors in parentheses are clustered on country-pair;
* significant at 10%; ** significant at 5%; *** significant at 1%
Retention parameter $\delta = 0.995$; Columns 4-6 uses data only for year 2004; 13 years is the median time trade used to code the experience dummy for the matching estimates; Column 3 matches on all gravity variables, GDP in destination and same exporter; Column 4 matches on all gravity variables, GDP in origin, and same importer

Table 3: Robustness Checks

Alternate Specifications	Coefficient on experience
1. Baseline	0.438***
	(0.069)
2. Right censoring dummy	0.438***
	(0.069)
3. Experience based on Correlates of War data	0.364***
	(0.019)
4. Dropping Former Soviet Union (FSU)	0.465***
	(0.078)
5. Dropping FSU and Yugoslavia	0.477***
	(0.081)
6. Dropping FSU, Yugoslavia and Czechoslovakia	0.481***
	(0.082)
7. Dropping the smallest 1% of observations	0.397***
	(0.070)
8. Dropping the smallest 5% of observations	0.389***
	(0.067)
9. Correcting for selection and heterogeneity bias [HMR, 2008]	0.433***
	(0.070)
10. Correcting for heteroskedasticity [PPML estimates for year 2004]	0.577***
	(0.140)

Placebo Tests		Coefficient on placebo variable
1. Import experience of exporter		0.065
		(0.062)
2. Randomize dummy for all years of positive trade		0.007
		(0.042)
3. Randomize dummy for positive trade (for less than median years of experience)		0.043
		(0.050)
4. Randomize dummy for positive trade (for more than median years of experience)		0.130**
		(0.062)
5. Assign experience to alphabetical neighbor of importer from same region		0.042
		(0.037)
6. Assign experience to alphabetical neighbor of exporter from same region		-0.068*
		(0.035)
7. FDI stock as dependent variable	-0.084**	
	(0.035)	

Standard errors in parentheses are clustered on country-pair: * significant at 10%; ** significant at 5%; *** significant at 1%. The retention parameter δ is set to 0.995 in all rows and columns. All specifications include country-year and country-pair dummies and pair-specific trends.

Table 4: Variation of Coefficient on Experience Across Model Specifications

Variable	Coefficient on experience	Coefficient on experience
Baseline Model	0.887	0.438
(Standard Error σ_1)	(0.018)	(0.069)
Split On Gravity Variable		
Both in GATT/WTO	0.896	0.485
PTA	0.880	0.425
GSP	0.970	0.399
Common currency	0.875	0.420
Distance	0.801	0.420
Contiguity	0.860	0.424
Colonial relationship	0.903	0.382
Common language	0.855	0.425
Common law	0.868	0.434
Standard deviation of coefficient estimates (σ_2)	(0.043)	(0.030)
σ_2/ σ_1	2.378	0.431
Country-year dummies	Yes	Yes
Country-pair dummies	No	Yes
Pair specific trends	No	Yes

Standard errors in parentheses are clustered on country-pair: * significant at 10%; ** significant at 5%;
All splits are based on binary variables except for distance; for distance we split the sample into three parts, 1 std. dev. above mean, mean ± 1 std dev, and 1 std. dev. below mean

Table 5: Spillovers in Experience from Margins of Trade

	(1) Bilateral Exports	(2) Extensive margin	(3) Intensive margin	(4) Bilateral Exports	(5) Extensive margin	(6) Intensive margin
Experience	0.887***	0.378***	0.509***	0.438***	0.155***	0.283***
	(0.018)	(0.010)	(0.014)	(0.069)	(0.027)	(0.061)
Both in GATT/WTO	0.193***	0.259***	-0.066	-0.053	-0.017	-0.036
	(0.062)	(0.036)	(0.046)	(0.065)	(0.029)	(0.057)
PTA	0.724***	0.460***	0.264***	-0.052	-0.041*	-0.011
	(0.053)	(0.033)	(0.038)	(0.048)	(0.023)	(0.041)
GSP	0.241***	0.177***	0.064*	0.122	0.057	0.064
	(0.048)	(0.024)	(0.039)	(0.109)	(0.038)	(0.099)
Common currency	0.278**	0.091	0.187**	0.105**	0.030	0.075*
	(0.121)	(0.078)	(0.079)	(0.049)	(0.026)	(0.044)
Distance	-1.259***	-0.854***	-0.405***			
	(0.018)	(0.012)	(0.012)			
Contiguity	0.537***	0.431***	0.105**			
	(0.082)	(0.065)	(0.052)			
Colonial relationship	0.946***	0.629***	0.317***			
	(0.088)	(0.059)	(0.056)			
Common language	0.449***	0.497***	-0.048			
	(0.039)	(0.024)	(0.030)			
Common law	0.380***	0.150***	0.229***			
	(0.053)	(0.029)	(0.041)			
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country-pair dummies	No	No	No	Yes	Yes	Yes
Pair-specific trends	No	No	No	Yes	Yes	Yes
Observations	240873	240873	240873	240873	240873	240873
R-squared	0.76	0.84	0.54	0.94	0.97	0.85

The retention parameter δ is set to 0.995 in all columns.
Standard errors in parentheses are clustered on country-pair: * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Spillovers in Experience from Disaggregate Trade

	(1) Pooled	(2) Pooled	(3) Pooled	(4) USA	(5) CHN	(6) DEU	(7) JPN	(9) IND
Retention parameter δ_1	0.649***	0.636***	0.636	0.636	0.636	0.636	0.636	0.636
	(0.0005)	(0.0005)						
Retention parameter δ_2	0.002***	0.000001***	0.000001	0.000001	0.000001	0.000001	0.000001	0.0000
	(0.0004)	(0)						
Retention parameter δ_3	1.368***	1.350***	1.350	1.350	1.350	1.350	1.350	1.350
	(0.001)	(0.001)						
Experience (destination and industry-specific)	0.946***	0.878***	0.741***	0.763***	0.715***	0.813***	0.743***	0.653*
	(0.003)	(0.002)	(0.004)	(0.012)	(0.019)	(0.010)	(0.013)	(0.018
Destination-specific experience (across 4-digit industries)	0.054***	0.055***	0.021***					
	(0.001)	(0.001)	(0.003)					
Industry-specific experience (across destinations)	1.107***	1.114***	0.139***					
	(0.001)	(0.001)	(0.004)					
Both in GATT/WTO	0.005***	-0.069***	0.037					
	(0.002)	(0.002)	(0.032)					
PTA	0.090***	0.201***	0.046***					
	(0.002)	(0.002)	(0.011)					
GSP	0.058***	-0.044***	-0.045***					
	(0.002)	(0.006)	(0.012)					
Common currency	0.110***	-0.047***	-0.074**					
	(0.006)	(0.001)	(0.030)					
Distance	-0.208***	-0.257***	-0.046***					
	(0.001)	(0.003)	(0.007)					
Contiguity	0.345***	0.271***	0.096***					
	(0.003)	(0.002)	(0.018)					
Colonial relationship	0.040***	0.179***	-0.017					
	(0.002)	(0.002)	(0.015)					
Common language	0.172***	0.062***	-0.004					
	(0.002)	(0.002)	(0.013)					
Common law	0.097***	0.013***	0.017					
	(0.002)	(0.001)	(0.014)					
GDP exporter	0.070***	0.086***						
	(0.001)	(0.001)						
GDP importer	0.256***	0.293***						
	(0.001)	(0.005)						
Time dummies	No	Yes	NA	NA	NA	NA	NA	NA
Country dummies	No	Yes	NA	NA	NA	NA	NA	NA
Country-year dummies	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year dummies	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	5084671	5084671	5084671	507494	201630	266319	335204	11059
R-squared	0.40	0.43	0.74	0.82	0.82	0.87	0.84	0.76

Standard errors in parentheses are clustered on country-pair: * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Regression results from simulated data

	(1)	(2)
Experience (E)	0.430***	0.440***
	(0.000)	(0.001)
Variable (Iceberg) Trade Cost (τ)	0.026	-0.100***
	(0.004)	(0.114)
Fixed Trade Cost (F)	-3.193***	-3.275***
	(0.003)	(0.009)
Income (Y)	0.583***	
	(0.012)	
Constant	-1.339	-1.034
	(0.008)	(0.013)
Observations	201	201
R-squared	0.99	0.99

Standard errors in parentheses are clustered on country-pair:
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1: Decomposition of World Exports into Trade Between Country-Pairs that Traded Prior to 1950 and Emergence of New Trade

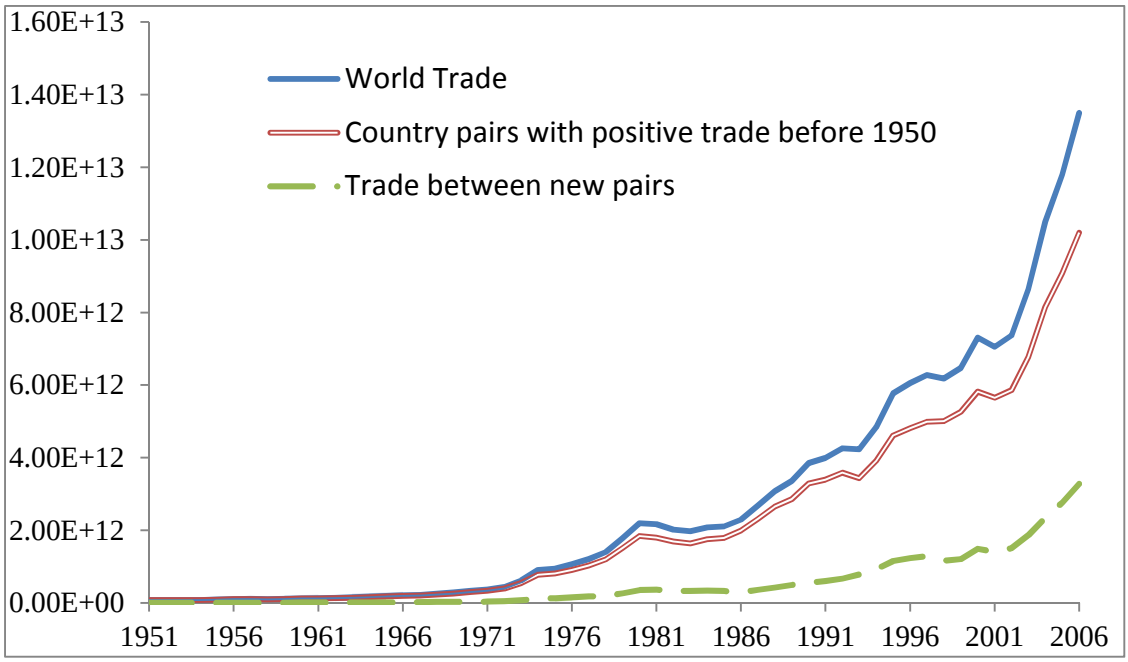


Figure 2: Distribution of Experience in 2006

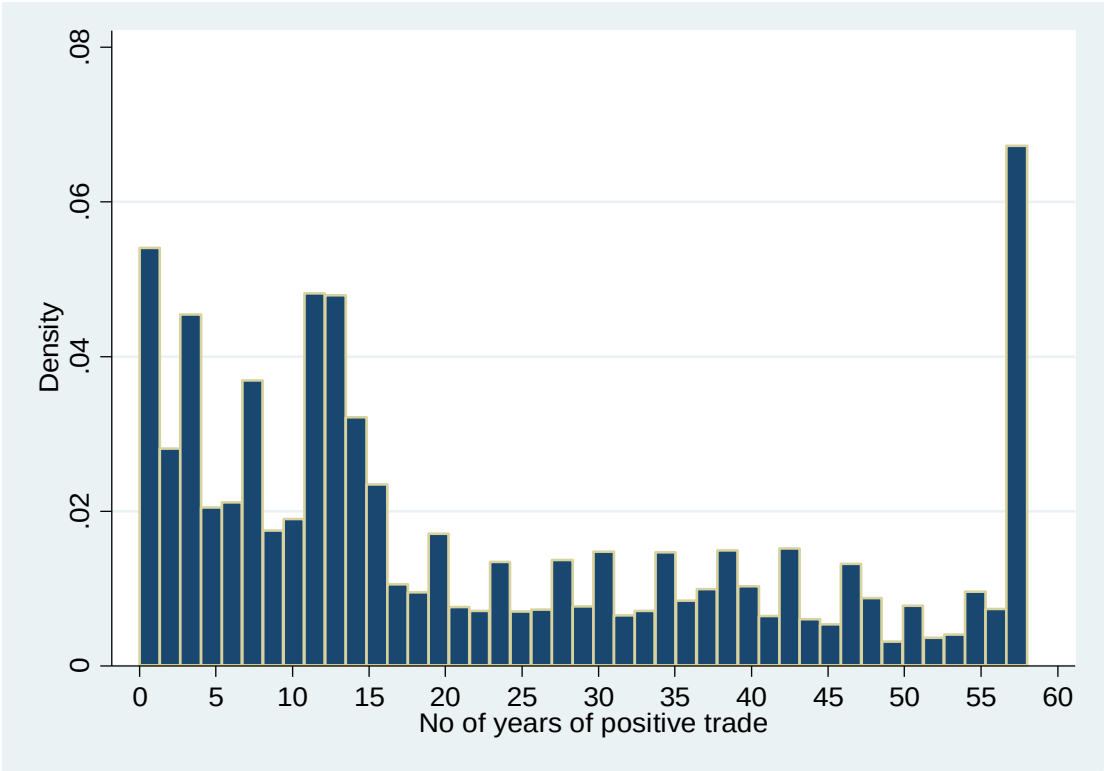
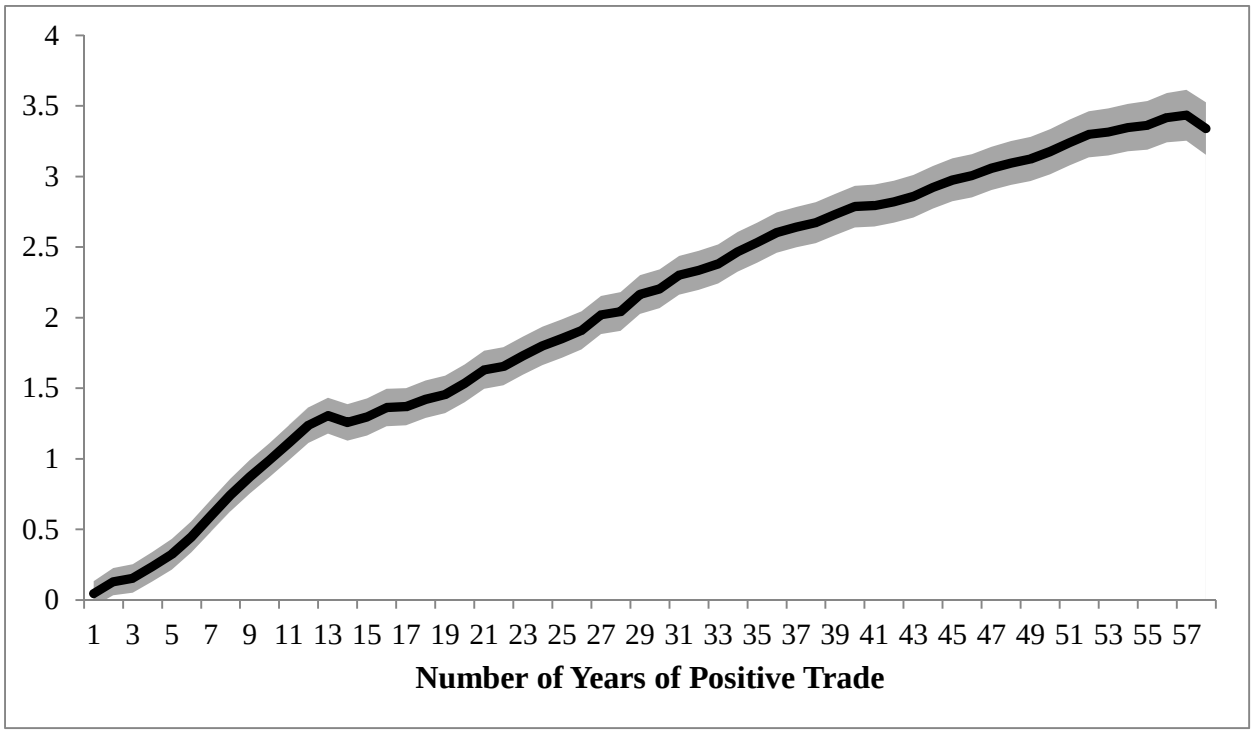
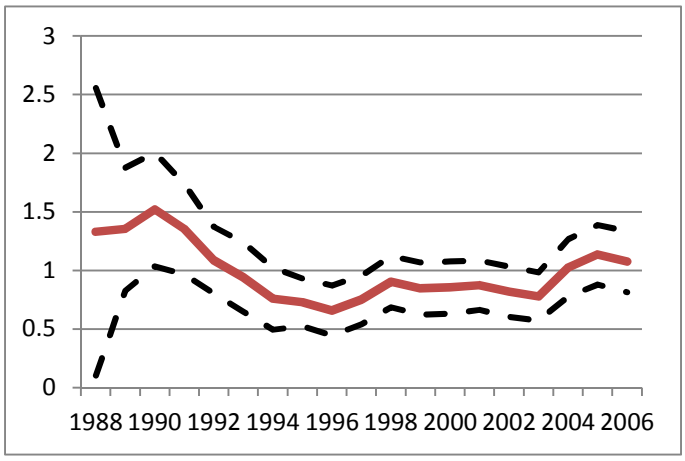


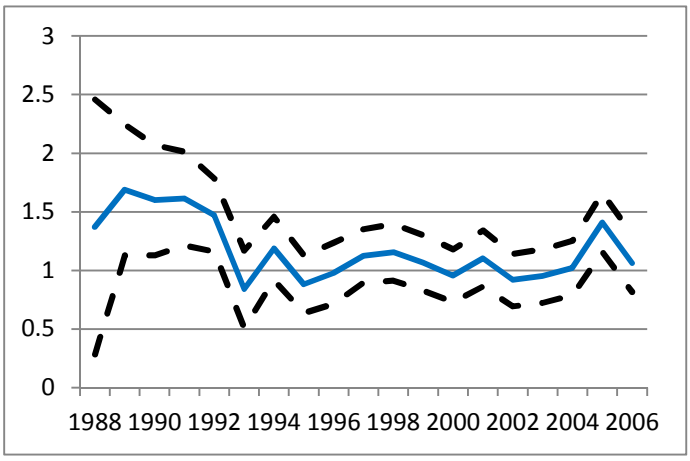
Figure 3: Cumulative Effect of Experience on Logged Bilateral Trade (Coefficient estimate with 95% confidence interval)



Figures 4a and 4b: Matching Estimate for Experience by Year and 95% Confidence Interval



4a: Matched on exporter



4b: Matched on importer