NBER WORKING PAPER SERIES

SHORT RUN GRAVITY

James E. Anderson
Yoto V. Yotov

Working Paper 23458
http://www.nber.org/papers/w23458

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2017

Short Run Gravity
James E. Anderson and Yoto V. Yotov
NBER Working Paper No. 23458
May 2017
JEL No. F10,F14,F15

ABSTRACT

Short run gravity is a geometric weighted average of long run gravity and bilateral capacity. The model features (i) joint trade costs endogenous to bilateral volumes, (ii) long run gravity as a limiting case of efficient investment in bilateral capacities, (iii) a structural ratio of short run to long run trade elasticities equal to a micro-founded buyers' incidence elasticity, and (iv) tractable short and long run models of the extensive margin. Application to manufacturing trade of 52 countries during the globalization period 1988-2006 strongly supports the model. Results solve several time invariance and trade elasticity puzzles in the literature.

James E. Anderson
Department of Economics
Boston College
Chestnut Hill, MA 02467
and NBER
james.anderson.1@bc.edu

Yoto V. Yotov
School of Economics
Drexel University
Philadelphia, PA 19104
and ERI-BAS
yotov@drexel.edu

1 Introduction

Gravity's emergence from the shade [e.g. Head and Mayer (2014)] highlights empirical and conceptual puzzles posed by the model's success. On one hand gravity-fitted bilateral trade flows come very close to the data and estimated coefficients are stable over time. On the other hand, this fit and stability violate intuition and observation about trade behavior over time. Secular invariance of coefficients is a puzzle in light of big technological improvements in transportation and communications over the past 60 years.[1] Higher frequency invariance and good fit is further puzzling in light of pronounced business cycle movement of ocean shipping rates and delivery lags. Imperfectly correlated national expenditures over the business cycle and big secular shifts in the location of economic activity suggest cyclical and trend variation in bilateral trade costs over time and space. Time invariance also appears to conflict with the richly patterned adjustment over time in bilateral trade links emphasized by Chaney (2014), on French firm export market behavior, and the random entry and exit emphasized by Besedes and Prusa (2006), on US 10-digit trade.

Addressing the empirical puzzles poses conceptual problems for gravity theory derived from static equilibrium economics. The short run gravity model of this paper is a conceptual solution. Trade costs vary with volume in the short run due to diminishing returns to a variable factor that is paired with a bilaterally specific fixed factor. Short run gravity nests the standard long run gravity model as a limit approached when bilaterally specific capacities adjust to efficient levels. Application of the model to manufacturing trade data for 52 countries over 1988-2006 yields a good fit with intuitively plausible parameters. The estimated parameters give solutions to the time invariance and related empirical puzzles.

Bilateral capacity is a notion consistent with the network link dynamics modeled by Chaney (2014) and with the link between managers' experience in previous firms and the export performance of their current company described by Mion and Opromolla (2014).

[1] Coe et al. (2002) define the 'Missing Globalization Puzzle' as "the failure of declining trade-related costs to be reflected in estimates of the standard gravity model of bilateral trade" (p.1).

Commonly used assumptions in production theory with fixed factors then yield a log-linear structural gravity model. Short run gravity is a geometric weighted average of the familiar (long run) gravity model and a bilateral capacity variable for shipments from origin i to destination z:

$$[\textbf{Long Run Gravity}(\mathbf{i},\mathbf{z})]^{\rho}\,\lambda(\mathbf{i},\mathbf{z})^{\mathbf{1}-\rho}.$$

[**Long Run Gravity** $(\cdot)$] is the standard gravity model expression for the trade depressing effect of bilateral relative to multilateral resistance.[2] $\lambda(i,z)$ is a bilateral capacity variable that reflects origin-destination-specific investment in bilateral links. ρ is the buyers' short run incidence elasticity, the fraction of trade cost variation borne by buyers.[3] ρ is a combination of the elasticity of substitution in demand and the elasticity of supply, itself micro-founded in the joint supply of output delivered to many destinations at increasing cost due to diminishing marginal product of a variable factor on each link.[4] Our estimation suggests $\rho \in (0.20, 0.37)$, narrowed to $(0.20, 0.24)$ in our favored specifications with a preferred value 0.24. In all robustness checks ρ is significantly far below 1, statistically and economically.

In the short run (i.e. at all observable moments), observed trade flows face rising costs due to diminishing marginal product of variable factors with fixed bilateral capacity on each link. Over time bilateral capacity adjusts, plausibly toward the efficient level. With perfectly efficient origin-destination-specific capacity on all links $\{\lambda^*(i,z)\}$, the model converges to

[2] As demonstrated by Arkolakis et al. (2012), [**Long Run Gravity**$(\cdot)$] derives from a wide array of theoretical microfoundations. Anderson (1979) is the first to offer a theoretical economic foundation for the gravity equation in an Armington-CES setting. Other early contributions to gravity theory include Krugman (1980) and Bergstrand (1985), who obtain gravity from a monopolistic competition and a Heckscher-Ohlin foundations, respectively. Eaton and Kortum (2002), who derive gravity in a Ricardian framework with intermediate goods, and Anderson and van Wincoop (2003), who popularize the Armington-CES model and emphasize the importance of the general equilibrium effects of trade costs, are arguably the most influential gravity theories. More recently, the gravity model has been derived in a heterogeneous firm setting by Chaney (2008) and Helpman et al. (2008). We refer the reader to Anderson (2011), Arkolakis et al. (2012), Head and Mayer (2014), and Costinot and Rodrguez-Clare (2014), and Yotov et al. (2016) for recent surveys of the evolution of the theoretical gravity literature.

[3] The term incidence elasticity is chosen to distinguish it from the passthrough elasticity used to describe incomplete incidence of exchange rate changes on prices, a higher frequency behavior that may reflect other causes than are the focus of this paper.

[4] Short run fixed capacity causes diminishing marginal product of labor in production-*cum*-delivery on any link. Heterogeneous productivity of firms results in a composite supply elasticity with an extensive margin including the dispersion parameter of a Pareto distribution of productivity draws.

the standard long run gravity equation. Over the era of globalization 1988-2006, our results suggest that the growing efficiency of trade capacity increased trade volume by 162%.

Bilateral trade costs are interdependent in the short run, jointly determined across all bilateral volumes. Efficient allocation of the variable factor interacts with size effects in demand at each destination and supply at each origin along with bilateral fixed capacities to determine the system of equilibrium trade costs. Despite this generality, CES-type tractability still obtains. The model encompasses heterogeneous productivity, of industries (as in Eaton and Kortum (2002)) or of firms (as in Chaney (2008)). The property of rising short run marginal cost of trade resembles the model of Arkolakis (2010), but differs in the cause. In Arkolakis (2010) rising trade cost in the long run is due to increasing difficulty of reaching new customers. Long run trade costs are constant in our model.

Short run gravity generates a simple prediction about the relationship between the short run (SR) gravity coefficients and their long run (LR) counterparts:

$$\beta_l^{SR} = \beta_l^{LR}\rho, \tag{1}$$

where l denotes any gravity covariate, e.g. distance, tariffs, etc. Equation (1) is an explanation of the difference between short run and long run trade elasticities based on fixed bilateral capacities.[5] Short run gravity elasticities are around 1/4 the size of long run elasticities based on our parameter estimates. Ruhl (2008) reconciles the difference in a model based on cyclical fluctuations being temporary whereas trade cost changes are permanent. Arkolakis et al. (2011) offer an alternative explanation by introducing a dynamic adjustment process on the part of consumers. Most recently, Crucini and Davis (2016) explain the discrepancy between the short- and long-run elasticities with frictions in distribution, which may lead to the slow adjustment of quantities following a change in relative prices.

[5] The long run elasticity estimates in the trade literature usually vary between 2 and 12. See Eaton and Kortum (2002), Anderson and van Wincoop (2003), and Broda et al. (2006). The corresponding short run elasticity indexes from the IRBC macro literature are usually between 1 and 2. See Backus et al. (1994), Zimmermann (1997), Heathcote and Perri (2002) and Feenstra et al. (2012).

A second implication of equation (1) suggests an explanation of the broad 'missing globalization puzzle' (Coe et al., 2002): the effects of geographic impediments such as distance and other geographic impediments to estimated gravity equations have been stable, or even increasing over time. Time invariance of estimated coefficients of standard cross section gravity regressions is explained by iceberg trade cost proxies that control well for the omitted bilateral capacity variables. Bilateral capacity adjusts over time toward an efficient level that is shown below to depend on the same iceberg trade costs as in the standard cross section model.[6] We estimate highly significant declining border effects associated with rising investment in cross-border capacities relative to domestic capacities. The missing globalization is in the omitted dynamic adjustment of bilateral capacities driven by the same geographic forces that drive spatial variation of bilateral trade flows. The omission is concealed by standard static gravity estimation that is an unbiased estimate of long run gravity. Our tests below based on (1) are consistent with this solution to the puzzle.

The short run gravity model also features a useful theoretical approach to the many zeroes observable in bilateral trade flows, some persistent and some flickering on and off. Changes in persistent zeroes are due to investments in bilateral capacity on the extensive margin as capacity moves toward an efficient long run extensive margin of markets served. Flickering zeroes are due to demand changes combined with period-by-period fixed cost components of variable trade costs such that the origin-destination pair representative firm flips from one side to the other of the breakeven point for serving a particular destination. In the heterogeneous firms version of the model, the breakeven point for not serving a destination applies the least productive active firm on an interior extensive margin.[7]

[6]The puzzles of time-invariant gravity estimates have attracted significant attention in the literature. A series of papers have proposed purely empirical solutions to the puzzles. See Buch et al. (2004), Carrre and Schiff (2005), Brun et al. (2005), Boulhol and de Serres (2010), Lin and Sim (2012), Yotov (2012) Carrre et al. (2013), Bergstrand et al. (2015) and Larch et al. (2016).

[7]Helpman et al. (2008) (HMR) offer the first model of zeroes based on fixed export costs. Firms draw productivities from a Pareto distribution with an upper bound such that zeroes are explained by a sufficiently low upper bound such that no firm at a particular origin receives a draw high enough to cover the fixed cost of entering a particular destination market. Our model for short run zeroes differs only in that fixed capacity implies increasing opportunity cost of serving small markets, hence a U-shaped full average cost curve that can lie everywhere above a small destination market's willingness to pay, even with an unbounded

Section 2 presents the basic theoretical model of joint trade costs of production and delivery to many destinations. Section 3 applies the model to the multi-country setting, yielding the short run gravity model. Section 4 develops and tests two complementary empirical versions of the model. Theoretical Appendix A shows that the basic model encompasses heterogeneous productivities (as in Eaton and Kortum (2002)) and heterogeneous firms (as in Chaney (2008)). Appendix A also draws out the implications of model for a theory of endogenous trade costs such as may be applied to model inference from price comparisons for homogeneous goods across many locations. Appendix B reports on sensitivity experiments with the main empirical models. The main results are robust.

2 A Model of Joint Trade Costs

An origin region produces and ships a product to potentially many destinations. Distribution on each link requires multiple resources in variable proportions, generalizing the usual notion of iceberg-melting trade costs but preserving the essential feature that distribution multiplicatively amplifies production costs. The scalar trade cost factors differ across destinations due to endogenous as well as exogenous geographical features of the world economy. Cost-minimizing resource allocation implies that bilateral trade destinations are imperfect substitutes. Specializing this story to a manageable model, assume two factors are required to produce and distribute to each destination market. One factor is variable in the short run (labor that can be freely allocated across destination markets that have positive capacity) and one is fixed in the short run (capacity in each potential network link).

The variable labor requirement to each destination served includes a fixed labor component, interpretable as a manager. For simplicity the fixed labor requirement is normalized

Pareto productivity distribution. The equilibrium extensive margin thus depends only on technology and taste parameters. Short run entry and exit flickering on the extensive margin are induced by demand or technology shocks that can have any distribution. Other models use demand systems with choke prices to explain zeroes, e.g. quadratic as in Melitz and Ottaviano (2008), or translog as in Novy (2013). While choke prices are realistic, such systems appear to do worse (according to Novy's, 2013, results) at fitting the data than the standard CES gravity equation. We leave empirical validation of the SR gravity model of zeroes to future research.

to one. Labor, including the managerial labor, is mobile in the short run across production for destinations actively served.

The trade activity also requires destination-specific “marketing capital” that is committed (sunk) before the allocation of labor, though it is variable *ex ante*. Marketing capital is thus bilaterally specialized and fixed in the short run. The details of destination specialization, while interesting and important (e.g., Chaney, 2014), are outside the scope of this paper. “Marketing capital” is left vague to encompass both human capital in the form of network connections and physical capital particularized to serve a particular destination. The idea of a retail network in Crucini and Davis (2016) is similar.

Bilateral production and trade of a generic firm in a generic origin and sector to destination z is formally modeled as a Cobb-Douglas function of non-managerial labor $L(z) - 1$, 1 unit of managerial labor, and capital $K(z)$: $x(z) = (1/t(z))K(z)^{1-\alpha}(L(z) - 1)^{\alpha}$. The functional form implies that the manager is required.[8] $x(z)$ is delivered product. $t(z) > 1$ is the technological iceberg-melting parameter from the given origin to destination z, a penalty imposed by ‘nature’ relative to the frictionless benchmark $t = 1$. $t(z)$ also reflects a productivity penalty in the usual sense that would apply to all destinations z uniformly. With all inputs variable (in the long run), the production function exhibits economies of scale. In the short run with $K(z)$ fixed, decreasing returns dominate.[9] For now, all firms are identical in any origin, aggregating to an industry with the generic firm’s characteristics. Appendix Section A.1.1 shows that the short run gravity model extends in all essentials to include the heterogeneous firms model with firm productivities drawn from a Pareto distribution. (Essentially the same short run gravity model extends to incorporate fixed infrastructure at each location and time, adding location-time-specific productivity shifters controlled for

[8]The manager story is just one justification of the formal requirement of a fixed labor cost for each destination served. Choice of the manager is outside the model, while normalization of managerial or fixed cost input to 1 unit of labor is a harmless convention. A useful extension empirically introduces heterogeneity across destinations in the amount of fixed input required. At this level of generality, the production function has the Stone-Geary form.

[9]To see the long run economies of scale, consider a scalar expansion of $K(z), L(z)$. The elasticity of $x(z)$ with respect to a one % rise in $K(z), L(z)$ is equal to $1 - \alpha + \alpha L(z)/[L(z) - 1] > 1$ in the long run. The marginal product of labor is diminishing in the short run, $\partial^2 x(z)/\partial L(z)^2 = (\alpha - 1)\alpha x(z)/[L(z) - 1] < 0$.

econometrically with location-time fixed effects.)

For the generic sector as a whole, labor is drawn from a national labor market in an amount satisfying the value of marginal product condition at the national wage rate. The labor market constraint on short run allocation across destinations is $L = \sum_0^n L(z)$, where L is labor supply to the sector and n is the extensive margin destination. n units of fixed managerial labor are included in L. Labor is efficiently allocated across destination activities with value of marginal product equal to the common wage. Managers are drawn from the common labor pool and are paid the common wage.

The price for delivered output at destination z is $p(z)$. Delivered price $p(z)$ in competitive equilibrium covers costs: $w[L(z) - 1] + r(z)K(z) + w = p(z)x(z)$ where $r(z)$ is the realized (residual) return on the specific capital for delivery to destination z. The first term on the left hand side of the equation is the non-managerial wage bill, the second term the payments to sector specific capital and the third term the payment to the manager.[10]

The extensive margin destination n is reached when the wage required for the marginal manager exceeds the value of sales minus the wage bill for non-managerial labor; i.e., profits are negative. Even with positive profits, destination z may not be served if $K(z) = 0$, no destination-specific capital is allocated. Destination-specific capital $K(z)$ is allocated prior to production and delivery to $n^* \geq n$ potentially active destinations, accounting for $K = \sum_0^{n^*} K(z)$ units of capital. The allocation shares of capital are $\lambda(z) = K(z)/K$. Long run efficient allocation of capital and its extensive margin n^* is analyzed in Sections 3.2-3.3.

The value of production at delivered prices in the generic region and sector is the sum over destinations z of the payments to labor and to managerial capital and the payment to

[10]Managers may also receive a premium as a share of the residual return to an active sector. The share can be rationalized as incentivizing, or as the outcome of bargaining between managers with a disutility of managing and the owners of specific capital. Total cost is then $w[L(z) - 1] + (1 - \pi)r(z)K(z) + [w + \pi r(z)K(z)] = p(z)x(z)$, where π is the manager's premium share. This nod to realism plays no role in the model, hence is suppressed for simplicity.

non-managerial capital given by[11]

$$Y = (L - n)^{\alpha} K^{1-\alpha} C, \tag{2}$$

where

$$C \equiv [\sum_{0}^{n} \lambda(z)(p(z)/t(z))^{1/(1-\alpha)}]^{1-\alpha}. \tag{3}$$

The real activity level

$$R = (L - n)^{\alpha} K^{1-\alpha}$$

in (2) is multiplied by a price index C embedding efficient allocation of the joint activity to delivered output at the many destinations. The setup thus yields a Constant Elasticity of Transformation (CET) joint revenue function for delivered output. (The level of activity R can be taken as exogenous for gravity model purposes, hence it is not necessary to assume that activity R is a Cobb-Douglas function of labor and specific human capital.)

The equilibrium share of sales to each destination z that is served, by Hotelling's Lemma (applied to (2) using (3)), is

$$s(z) = \lambda(z) \left(\frac{p(z)/t(z)}{C} \right)^{1/(1-\alpha)}. \tag{4}$$

It is convenient for later purposes to sort destinations by rank order, beginning with the largest (so in equation (4) $z \in [0, n]$). The local delivery market is $s(0)$ by convention because empirically it is almost universally the largest.

The short run extensive margin n is determined by the value of marginal product of labor in sector n falling below its value on the intensive margin for sectors $z < n$. This implies

[11] The setup here extends the specific factors production model of Anderson (2011) to include a fixed cost. In Anderson (2011) the environment is a GDP function where z denotes a sector in a continuum of fixed size, $t(z)$ is a productivity penalty and there is no fixed cost.

Competitive equilibrium implies that labor is paid a common wage in serving all markets, equal to the value of marginal product of labor in each market. Aggregate labor employed is L, taken as an endowment in solving for Y in (2) below. Price index $C(\cdot)$ results from solving the labor market clearance condition for the wage, then replacing the wage with the resulting reduced form in evaluation aggregate revenue.

that the extensive margin is efficient in the sense of maximizing (2) with respect to n. For simplicity, temporarily think of a continuum of destinations (with 'shares' being densities) and differentiate (2) with respect to n. The first order condition yields

$$\alpha Y/(L-n) = (1-\alpha)s(n)Y.$$

The left hand side is the value of marginal product of non-managerial labor, equal to the wage. The right hand side is the value of marginal product of the extensive margin manager. Manipulating the first order condition, the implication is:

Proposition 1: Extensive Margin. *The smallest market served in short run equilibrium is unique and characterized by:*

$$n = z : s(z) = \frac{\alpha}{1-\alpha}[L-z]^{-1}. \tag{5}$$

Ordering markets by decreasing size, the extensive margin is the smallest market that can be served.

Equation (5) implies that a manager paid w exhausts the entire residual payment in sector n. The wage $= \alpha Y/(L-n)$ is rising in n, while $(1-\alpha)s(n)Y$ decreasing in n; hence the equilibrium extensive margin exists and is unique. (5) readily rationalizes the flickering on and off of bilateral trade that is observable in highly disaggregated data: capacity is in place but insufficient revenue to cover the overhead cost of management indicate temporary shutdown. Section 3.2 analyzes the long run extensive margin of installed capacity.

An implication of extensive margin model (5) is that volume equations for positive trade are not subject to selection bias. That is because firms are identical, in contrast to heterogeneous firms literature. When the model is extended to heterogeneous firms with productivities drawn from an unbounded Pareto distribution following Melitz (2003), the volume of positive trade on any link combines smooth action on both extensive and 'interior' intensive margins in a closed-form structural gravity equation. With proxies for fixed costs that differ from the proxies for iceberg trade costs, the action can be decomposed into the component intensive and extensive margins. (The details are developed below in Appendix Section A.1.1.)

For any market that is served, the equilibrium delivered price $p(z)$ is endogenously determined by the supply side forces described in (2)-(4) interacting with demand forces described by Constant Elasticity of Substitution preferences or technology (in the case of intermediate goods). The intuitive notion of a bilateral trade cost corresponds to $p(z)/p(0)$, a clear idea when $t(0) = 1$, so bilateral trade cost is endogenous. In practice, this is a dangerous simplification because internal delivery costs are not zero, differ across countries and are endogenous just as the bilateral costs are endogenous, cf. Agnosteva et al. (2014) and Ramondo et al. (2016).

Description of the demand side of the market requires an expansion of notation to designate the location of the originating sector. The CES expenditure share for goods from origin i in destination z is given by

$$\frac{X(i,z)}{E(z)} = \left(\frac{\beta(i)p(i,z)}{P(z)}\right)^{1-\sigma}. \tag{6}$$

Here, $X(i,z)$ denotes the bilateral flow at end user prices, $E(z)$ denotes the total expenditure in destination z on goods from all origins serving it, $\beta(i)$ is a distribution parameter of the CES preferences/technology, σ is the elasticity of substitution, and $P(z)$ is the CES price index for destination z.

The market clearing condition for positive bilateral trade from i to z is

$$Y(i)s(i,z) = X(i,z).$$

Using (4) for $s(i,z)$ and (6) for $X(i,z)$ in the market clearing condition to solve for the equilibrium price $p(i,z)$ yields:

$$p(i,z) = \left[\frac{E(z)P(z)^{\sigma-1}\beta(i)^{1-\sigma}[t(i,z)C(i)]^{\eta}}{Y(i)\lambda(i,z)}\right]^{1/(\eta+\sigma-1)}, \tag{7}$$

where $\eta = 1/(1-\alpha) > 1$ is the supply elasticity. The short run equilibrium price in an active origin-destination pair in (7) is an intuitive constant elasticity function of demand shifters $E(z)P(z)^{\sigma-1}$, supply shifters $Y(i)$ and $C(i)$, and the exogenous bilateral friction components in $t(i,z)$ and the contemporaneously exogenous bilateral capacity $\lambda(i,z)$.

Incidence of trade costs to buyers is incomplete: the buyers' incidence elasticity is $d\ln p(i,z)/d\ln t(i,z) = \eta/(\eta+\sigma-1) \equiv \rho$. The incidence elasticity ρ (dropping "buyers" for brevity) plays a key role in the gravity representation of the model. ρ has a deep micro-foundation as a combination of the demand elasticity parameter σ and the supply elasticity parameter η, itself microfounded in the equilibrium of distribution based on the Cobb-Douglas specific factors model. ρ is increasing in η and decreasing in σ.[12]

3 Gravity in Short Run and Long Run

The gravity representation of short run equilibrium trade yields a gravity model with a short run trade cost elasticity equal to the product of the long run trade cost elasticity times the incidence elasticity ρ. The short run elasticity applies everywhere. The short run trade flow and multilateral resistance equations are additionally altered relative to long run gravity by the volume effects of bilateral specific capacity everywhere. Multilateral resistances retain their interpretation as sellers' and buyers' incidence of trade costs. Short run size-adjusted trade flows (the ratio of observed to hypothetical frictionless trade flows) are a geometric weighted average of long run gravity frictions and bilateral specific capacity investment,

[12]Equation (7) can, in principle, account for substantial variation in prices across time and space. Rents to the sector specific factor similarly vary (see Section 3.2). Rents are competitive in the model, so pricing to market behavior in the usual sense is not implied. The model can be extended to monopolistic competition by treating each origin i as a firm. With CES demand, markups are constant when firms shares are small.

where the geometric weight on long run gravity is equal to the buyers' incidence elasticity.

Multilateral resistance for sellers is derived from the global market clearing condition for $Y(i)$. Substitute (7) into (6), then multiply by $E(z)$ and sum over z to obtain the global demand for $Y(i)$. Collect the terms for $Y(i)$ on the left hand side of the market clearance condition and simplify the exponents $1-\rho(\sigma-1)/\eta = \rho$ on $Y(i)$ and $(\sigma-1)/(\eta+\sigma-1) = 1-\rho$ on $\lambda(i,z)$. The result is

$$Y(i)^{\rho} = [\beta(i)C(i)]^{\rho(1-\sigma)} \sum_z [E(z)P(z)^{\sigma-1}]^{\rho} t(i,z)^{\rho(1-\sigma)} \lambda(i,z)^{1-\rho}.$$

Divide both sides by Y^{ρ}. The result is

$$\left[\frac{Y(i)}{Y}\right]^{\rho} = [\beta(i)C(i)\Pi(i)]^{\rho(1-\sigma)} \Rightarrow \frac{Y(i)}{Y} = [\beta(i)C(i)\Pi(i)]^{1-\sigma}, \tag{8}$$

where outward multilateral resistance

$$\Pi(i)^{(1-\sigma)\rho} = \sum_z \left(\frac{E(z)}{Y}\right)^{\rho} \left(\frac{t(i,z)}{P(z)}\right)^{(1-\sigma)\rho} \lambda(i,z)^{1-\rho}. \tag{9}$$

The left hand side of (8) is recognized as a CES share equation for a hypothetical world buyer on the world market, with a world market price index for all goods equal to 1. Short run multilateral resistance in (9) is a CES function of bilateral relative trade costs $t(i,z)/P(z)$, where the elasticity of short run substitution is $(1-\sigma)\rho$. $\Pi(i)$ is homogeneous of degree one in $\{t(i,z)\}$ for given $\{P(z)\}$.

The gravity representation of trade also requires a relationship between the buyers' price index and relative trade costs. Substitute (7) in the CES price definition $P(z)^{1-\sigma} = \sum_i [\beta(i)p(i,z)]^{1-\sigma}$. Then use (8) to substitute for $[\beta(i)C(i)]^{1-\sigma}$ in the resulting equation. After simplification this gives the short run price index as

$$P(z)^{(1-\sigma)\rho} = \left(\frac{E(z)}{Y}\right)^{-(1-\rho)} \sum_i \frac{Y(i)}{Y} \left(\frac{t(i,z)}{\Pi(i)}\right)^{(1-\sigma)\rho} \lambda(i,z)^{1-\rho}.$$

Buyers' price index $P(z)$ is the product of a size effect $[E(z)/Y]^{-(1-\rho)}$ and a CES function of the set of bilateral buyers' incidences:

$$\widetilde{P}(z)^{(1-\sigma)\rho} = \sum_i \frac{Y(i)}{Y} \left(\frac{t(i,z)}{\Pi(i)}\right)^{(1-\sigma)\rho} \lambda(i,z)^{1-\rho}, \ \forall z, \tag{10}$$

with short run elasticity of substitution $(1-\sigma)\rho$. $\widetilde{P}(z)$ is the buyers' short run multilateral resistance, also interpreted as the buyers' short run incidence of trade costs. Simplifying the CES price index, $P(z) = [E(z)/Y]^{(1-\rho)/(\sigma-1)\rho}\widetilde{P}(z)$. Higher relative demand $E(z)/Y$ raises $P(z)$ given $\widetilde{P}(z)$ due to fixed capacities $\{\lambda(i,z)Y(i)\}$. In long run gravity, as effectively $\eta \to \infty$, $\rho \to 1$ and the buyers' market size effect vanishes from the price index $P(z)$.

Use $P(z)^{(1-\sigma)\rho} = [E(z)/Y]^{-(1-\rho)}\widetilde{P}(z)^{(1-\sigma)\rho}$ in sellers' multilateral resistance (9) to yield the more intuitive equivalent form

$$\Pi(i)^{(1-\sigma)\rho} = \sum_z \frac{E(z)}{Y} \left(\frac{t(i,z)}{\widetilde{P}(z)}\right)^{(1-\sigma)\rho} \lambda(i,z)^{1-\rho}, \ \forall i. \tag{11}$$

The final step in deriving short run gravity is to substitute the right hand side of (7) for $p(i,z)$ in (6) and use (8) to substitute for $[\beta(i)C(i)]^{1-\sigma}$ in the resulting expression. After simplification using incidence elasticity $\rho = \eta/(\eta+\sigma-1)$, this gives:[13]

Proposition 2: Short Run Gravity. *Short run gravity is a geometric weighted average of long run gravity and a bilateral capacity variable $\lambda(i,z)$. Short run gravity trade flows are given by:*

$$X(i,z) = \frac{Y(i)E(z)}{Y} \left[\frac{t(i,z)}{\Pi(i)\widetilde{P}(z)}\right]^{(1-\sigma)\rho} \lambda(i,z)^{1-\rho}. \tag{12}$$

where the multilateral resistances $\Pi(i)$, $\widetilde{P}(z)$ are given by (10)-(11)

The first term on the right hand side of (12) is the frictionless benchmark flow at given

[13] $X(i,z) = E(z)(\beta(i)p(i,z))^{1-\sigma} = E(z)^\rho[t(i,z)/P(z)]^{(1-\sigma)\rho}\lambda(i,z)^{1-\rho}H(i)$ where $H(i) = [\beta(i)C(i)]^{(1-\sigma)\rho}Y(i)^{1-\rho}$. Substitute $[Y(i)/Y\Pi(i)^{1-\sigma}]^\rho$ for $[\beta(i)C(i)]^{(1-\sigma)\rho}$ in $H(i)$ and replace $P(z)^{(1-\sigma)\rho}$ with $[E(z)/Y]^{-(1-\rho)}\widetilde{P}(z)^{(1-\sigma)\rho}$. Rearranging the result yields equation (12).

sales $\{Y(i)\}$ and expenditure $\{E(z)\}$. The middle term is the familiar effect of gravity frictions, the ratio of bilateral to the product of buyers' and sellers' multilateral resistance. The difference is that the short run trade elasticity is reduced in absolute value to $(1-\sigma)\rho$. The last term $\lambda(i,z)^{1-\rho}$ is the 'friction' due to inefficient investment in capacity on link i,z. Dividing both sides of (12) by the frictionless benchmark, size adjusted trade is

$$\frac{X(i,z)}{Y(i)E(z)/Y} = \left[\frac{t(i,z)}{\Pi(i)\widetilde{P}(z)}\right]^{(1-\sigma)\rho} \lambda(i,z)^{1-\rho},$$

a geometric weighted average of long run gravity frictional displacement and inefficient link capacity allocation.

Intuition about short run gravity system (10)-(12) is aided by considering an equiproportionate change in all bilateral trade costs $t(i,z)$: $t^1(i,z) = \mu t^0(i,z)$. Intuitively, bilateral trade flows should be unchanged because no relative price changes. Checking the system (10)-(11), $\{\widetilde{P}(z), \Pi(i)\}$ are homogeneous of degree 1/2 in $\{t(i,z)\}$, hence buyers and sellers multilateral resistances change by $\mu^{1/2}$ so indeed bilateral trade flows are constant. As with long run gravity, system (10)-(11) solves for multilateral resistances up to a normalization.

Over time the allocation of destination specific capital $\{\lambda(i,z)\}$ presumably moves toward the efficient level analyzed below in Section 3.2. The efficient allocation matches the long run demand pattern, so that the short run gravity equation (12) approaches the long run gravity equation, intuitively equivalent to $\rho \to 1$.

The preceding derivation of (12) combined with (9) and (10) uses for simplicity the Armington CES/endowments setup of Anderson and van Wincoop (2003), but the same short run gravity structure derives from two alternative structures with endogenous production and heterogeneous productivities. Appendix Section A.1.1 shows that the form of short run gravity in equations (9)-(12) holds exactly for the alternative interpretation of gravity based on heterogeneous productivity draws in a Ricardian model due to Eaton and Kortum (2002).

It also extends to the heterogeneous firms gravity model of Chaney (2008),[14] understanding that the composite supply elasticity combines intensive margin elasticity η above with an extensive margin elasticity based on the dispersion parameter θ of the Pareto productivity distribution. The composite supply elasticity is $\tilde{\eta} = \eta(1 + \theta - \eta) > \eta$ for the intuitive case $\theta > \eta$, the necessary and sufficient condition for the extensive margin to be rising in price. The incidence elasticity becomes $\tilde{\rho} = \tilde{\eta}/(\tilde{\eta} + \sigma - 1) > \rho$ for $\theta > \eta$. In the special case $\theta = \eta$ the intensive and extensive supply responses are perfect substitutes and $\tilde{\rho} = \rho$. Thus, the short run gravity adjustment developed here applies to the wide class of models that have been described in Arkolakis et al. (2012).

3.1 Short vs. Long Run Gravity Elasticities

Standard cross-section gravity inference of the effect of bilateral frictions is interpreted as long run gravity. The short run gravity model (12) nests long run gravity and thus can simply explain three prominent empirical puzzles. To aid intuition, let the vector of bilateral trade costs $t(i, z)$ from the short run gravity specification (12) be an exogenous function of standard gravity variables:

$$t(i, z) = \prod_{l=1}^{L} (1 + tar(i, z)) d(i, z)_l^{\psi_l}, \tag{13}$$

where $tar(i, z)$ denotes ad-valorem tariffs and the vector $d(i, z)$ includes a set of binary and continuous determinants of bilateral trade. Proposition 2 delivers a straightforward relationship between the short run (SR) gravity coefficients and the corresponding long run (LR) counterparts:

Corollary 2.1: Short Run vs. Long Run Gravity Estimates.

$$\beta_l^{SR} = (1 - \sigma)\psi_l \rho = \beta_l^{LR} \rho. \tag{14}$$

[14] The extension is exact for competitive firms and a close approximation for monopolistically competitive firms.

Equation (14) gives structure to address two elasticity puzzles in the trade literature. First, the theoretical constraint $\rho \in (0,1)$ applied to the estimates of direct price shifters such as tariffs at least partially resolves the empirical puzzle posed by estimates of the trade elasticity in cross-section gravity estimations that are much larger than short run estimates used in the macro literature. (See Footnote 5.) Using (14) with estimates of ρ, inference of short run elasticities from the two literatures can be compared, with remaining differences isolated from that due to the missing incidence elasticity. Some portion of remaining difference may be due to incomplete passthrough to high frequency price changes. Equation (7) is the short run gravity model counterpart, as developed further in Appendix A.

Second, equation (1) and its relation to equation (12) helps resolve the time invariance 'distance puzzle' and the more broad 'missing globalization puzzle'. A possible rationale for the puzzling time invariance of estimated trade cost elasticities starts from the plausible idea that iceberg trade cost proxies control for the omitted capacity variables in the standard cross section gravity regression. Then the effect of distance on trade, including its effect on capacity could be falling in absolute value over time. Specifically, divide both sides of (12) by $Y(i)E(z)/Y$ at a point in time, take logs and totally differentiate with respect to distance $D(i,z)$ holding multilateral resistances constant:

$$\left(\frac{d\ln[X(i,z)Y/Y(i)E(z)]}{d\ln D(i,z)}\right)^{SR} = \rho\left(\frac{\partial\ln[X(i,z)Y/Y(i)E(z)]}{\partial\ln D(i,z)}\right)^{LR} + (1-\rho)\frac{\partial\ln\lambda(i,z)}{\partial\ln D(i,z)}.$$

The first term on the right hand side is the special case of (14) for distance. Long run distance elasticity estimates are nearly constant in long run gravity regressions. The intuition that distance elasticities should be declining in the era of globalization is interpreted here as a positive second term. Over time $\lambda(i,z)$ moves toward its efficient level, which level will be a loglinear function of distance as shown below in Section 3.2. Our results below are consistent with a positive second term.[15] The missing globalization puzzle is the broader version of the

[15]In contrast, intuition that the long run partial distance elasticity in the first term should fall over time requires that technological change reduces long distance trade cost more than short distance trade cost. In contrast, *uniform* technological change reduces trade cost at all distances equiproportionately, an inference

distance puzzle, applying to all the long run gravity covariates.

Structural estimation can potentially identify the structural parameter ρ. With ρ estimates in hand, equation (14) offers several opportunities to test the SR gravity theory. (14) implies that the SR gravity estimates should be smaller than the corresponding LR numbers by the factor ρ, suggesting the test $\beta^{SR} - \rho\beta^{LR} = 0$, one for each of the gravity regressors in vector $d(i, z)$.

3.2 Long Run Efficient Allocation

The allocation of capital to destinations z is outside the static model developed above under the realistic assumption that investment is predetermined and generally inefficient relative to current realizations of random variables. It is nevertheless useful to construct a benchmark efficient allocation as an aid to intuition and to learning something about the inefficiency of allocation. A key, though in some sense obvious result is that the general gravity model under inefficient investment nests the standard iceberg trade cost model as a special case of efficient allocation. The difference between actual and hypothetical efficient allocation is presumably due to un-modeled frictions hampering investment in the face of various risks and imperfect information about realizations of natural bilateral resistance and other components of the realized equilibrium. The development of the benchmark allocation provides structure to the econometric application that generates inferred differences between actual and benchmark allocations as residuals.

Economic intuition suggests that standard iceberg trade costs should emerge as a reduced form of all efficient equilibrium production and distribution models because it is consistent with the envelope theorem in the allocation of all relevant resources in distribution. Effectively, geography dictates the allocation of capital as well as the distribution of goods given that efficient allocation of capital. Development of the benchmark special case demonstrates how this works.

that is consistent with the data.

Efficient capital allocation achieves equal returns on investment in each destination served. The average return on investment for the generic sector and economy of Section 2 is given by $\bar{r} = Y_K = (1-\alpha)Y/K$. The return on capital relative to the average for destination z is given by $s(i,z)/\lambda(i,z)$ (Anderson, 2011). If investors perfectly foresee bilateral natural trade costs and the extensive margin, then $\lambda(i,z) = s(i,z)$ in the capital allocation equilibrium actually realized. Then $\lambda^*(i,z) = s(i,z) = \lambda^*(i,z)[p(i,z)/t(i,z)C(i)]^\eta \Rightarrow p(i,z) = t(i,z)C(i)$. Combine this restriction with equation (7) for the market clearing price to solve for the efficient allocation

$$\lambda^*(i,z) = \frac{E^*(z)}{Y^*(i)}\left(\frac{\beta(i)C^*(i)t(i,z)}{P^*(z)}\right)^{1-\sigma} = \frac{E^*(z)}{Y^*}\left(\frac{t(i,z)}{\Pi^*(i)P^*(z)}\right)^{1-\sigma}. \tag{15}$$

Note that (15) is a general equilibrium concept: the multilateral resistances imply that all origins solve for efficient allocation simultaneously. Efficient allocation share $\lambda^*(i,z)$ is decreasing in the cross section of trade pairs in natural trade friction $t(i,z)$, increasing in destination market potential $E(z)P(z)^{\sigma-1}$ and origin utility weight $\beta(i)^{1-\sigma}$. Each of these effects is intuitive. It is also increasing in the 'average economic distance' of the origin from its markets, $\Pi^*(i)$, implying that for markets actually served, relationship-specific investments must be larger to overcome the average resistance.

Notice that $\eta = 1/(1-\alpha)$ plays no role in the efficient allocation equilibrium. This arises because no short run reallocation of labor is needed; the trade flows have converged to the standard gravity model pattern. The 'long run' trade cost elasticity is $1-\sigma$, which exceeds in absolute value the 'short run' elasticity $(1-\sigma)\eta/(\eta+\sigma-1) = (1-\sigma)\rho$, a familiar implication of the envelope theorem. A convenient approximation of the convergence to efficient allocation is the parametric result

$$\lim_{\eta\to\infty}(1-\sigma)\eta/(\eta+\sigma-1) = 1-\sigma.$$

Thus the short run gravity model of Section 2 is effectively related to the long run model as if the elasticity of transformation became infinite through the mechanism of efficient reallocation of specific investment. (Opening the black box of convergence here requires developing the dynamics of $\lambda(i,z)$, discussed in Section 4.)

3.3 Efficient Extensive Margin

In efficient equilibrium, the extensive margin is determined by equation (5) with $s(i,n) = \lambda^*(i,n)$. Order the destinations with ordering $Z(i)$ such that the efficient destination investment shares defined by (15) are decreasing in z: $Z(i) = \mathbf{P}(\{z\}) : d\lambda^*(i,z)/dz < 0$ where $\mathbf{P}(\{z\})$ denotes a permutation of the ordering of $\{z\}$. Then $n(i)^*$ is defined by

$$n^*(i) = z : L(i) - z - \frac{\alpha}{(1-\alpha)\lambda^*(i,z)} = 0; z \in Z(i).$$

Existence and uniqueness of the fixed point $n^*(i)$ is guaranteed because

$$L(i) - z - \frac{\alpha}{(1-\alpha)\lambda^*(i,z)}$$

is decreasing in z. More intuitively,

$$\underline{s}^*(i,\tau) = \frac{\alpha}{1-\alpha}[L(i,\tau) - n^*(i,\tau)]^{-1}. \tag{16}$$

Evidently the efficient equilibrium extensive margin $n(i)^*$ is increasing in origin size $L(i)$ and increasing in origin average economic distance $\Pi^*(i)$, the latter effect because it reduces the relative difficulty of entering the marginal market $t(i,z)/\Pi^*(i)$. A more steeply rising profile of bilateral trade costs for an origin country reduces its extensive margin. Destination size distribution (market potential) $E(z)P(z)^{\sigma-1}$ affects all exporters equally; as markets are more equal, more are served by every origin. The intuitive results on origin and destination size accord with observed characteristics of the extensive margin of trade. The more subtle implications of (16) remain to be explored in applications. Notice that this theory of the extensive margin imposes no structure on the distribution of productivity that contributes implicitly to the variation of $t(i, z)$.

4 Empirical Analysis

Short run gravity theory generates a series of testable predictions. Taking the predictions to data faces two challenges. First, there is no direct data on bilateral capacity. Proxies must be found to substitute for missing data. Second, there is no fully developed theory of the investment trajectory.[16] To maintain focus on short run forces, the application abstracts from details about the progress of formal globalization aside from treatment of FTA implementation and tariff changes. Results here should be viewed as 'proof of concept' rather than an accounting exercise. Knowledge gained will guide further empirical tests and extensions of SR gravity, discussed in the conclusion and the Appendices.

The application focuses on the intensive margin, for two reasons. First, equation (12) for positive trade flows is central as a theory nesting long run gravity with forces departing from long run gravity. Second, the key incidence elasticity parameter ρ is inferred from equation (12), a parameter that helps resolve well-known elasticity puzzles in the empirical gravity literature. The empirical investigation of SR gravity on the intensive margin is a natural first step before taking its implications to the more difficult issue of capturing behavior on the extensive margin.

A guiding intuition for the empirical analysis is that an evolutionary process is gradually approaching efficient capacity allocation in the era of globalization. Uncoordinated buyers' and sellers' agents grope forward experimentally to form links, expanding high rent links and

[16] Full rational expectations development is a chimera. The investments that plausibly form parts of $\lambda(i, z)$ are the result of uncoordinated decisions of many actors on N^2 links where N is the number of countries. See Chaney (2014) for evidence on export dynamics of French firms. Information network theories of links suggest search and random matching conditioned on anticipated probabilities of a match. Simple structures are thus applied here to model an investment trajectory.

contracting low rent links. Extend short run gravity model (12) to add a time dimension τ:

$$X(i,z,\tau) = \frac{Y(i,\tau)E(z,\tau)}{Y(\tau)} \left(\left[\frac{t(i,z,\tau)}{\Pi(i,\tau)P(z,\tau)} \right]^{1-\sigma} \right)^{\rho} \lambda(i,z,\tau)^{1-\rho}. \tag{17}$$

Established econometric techniques and standard proxies control for bilateral trade costs ($t(i,z,\tau)$), the multilateral resistances ($\Pi(i,\tau)$ and $P(z,\tau)$), and the size effects ($Y(i,\tau)$, $E(z,\tau)$, and $Y(\tau)$). $\lambda(i,z,\tau)$ remains.

The challenge posed by estimation of (17) is how to account for evolution of inefficient capacity investment $\lambda(i,z,\tau)$. We use two complementary approaches. The reduced-form approach uses (cross-)border-time fixed effects to pick up the effect of evolving bilateral capacity. Bergstrand et al. (2015) also use time-varying border dummies, designed to capture the impact of globalization on bilateral trade. A similar border-time fixed effects strategy is used by Head et al. (2010) to estimate the decline in 'marketing capital' in the trade of ex-colonies with their former masters relative to continuously non-colonial international trade. In contrast to the current approach, these earlier reduced-form specifications have no theoretical foundation pointing to structural interpretation of the estimated fixed effects. The alternative structural approach is a dynamic extension of theory suggesting an appropriate proxy variable for $\lambda(i,z,\tau)$. The merits and disadvantages of each of the two approaches are discussed with the findings in sections 4.2 and 4.3, respectively. Before that we briefly describe the data.

4.1 Data: Description and Sources

In order to obtain the main empirical results we use the dataset of Baier et al. (2016), which covers total manufacturing bilateral trade among 52 countries over the period 1988-2006.[17]

[17]The 52 countries/regions in the sample include: Argentina, Australia, Austria, Bulgaria, Belgium-Luxembourg, Bolivia, Brazil, Canada, Switzerland, Chile, China, Colombia, Costa Rica, Cyprus, Germany, Denmark, Ecuador, Egypt, Spain, Finland, France, United Kingdom, Greece, Hungary, Indonesia, Ireland, Iceland, Israel, Italy, Jordan, Japan, South Korea, Kuwait, Morocco, Mexico, Malta, Myanmar, Malaysia, Netherlands, Norway, Philippines, Poland, Portugal, Qatar, Romania, Singapore, Sweden, Thailand, Tunisia, Turkey, Uruguay, United States

In addition to spanning over a relatively long time period, the dataset of Baier et al. (2016) has two advantages. First, it includes data on intra-national trade flows.[18] As will become clear shortly, availability of intra-national trade flows data is crucial for the implementation of the reduced form approach to test SR gravity theory. Second, the dataset includes data on applied tariffs, which will enable us to identify the estimate of the trade elasticity of substitution directly from the empirical gravity specification, and to compare it between the short run vs. long run gravity specifications. The original source of the tariff data is the United Nation's TRAINS database. In addition to trade and tariff data and a rich set of fixed effects, we also employ data on standard gravity variables (bilateral distance, colonial ties, etc.), which come from the CEPII distances database (see Mayer and Zignago (2011)), and data on free trade agreements (FTAs), which come from the NSF-Kellogg Database on Economic Integration Agreements of Jeff Bergstrand. For further description of the main dataset we refer the reader to Baier et al. (2016).

In the sensitivity analysis we also experiment with an alternative database. Specifically, we employ the latest edition of the WIOD database, which covers 43 countries over the period 2000-2014. Similar to the dataset from Baier et al. (2016), the WIOD data includes consistently constructed international and intra-national trade flows.[19] The disadvantages of WIOD dataset, as compared to the data used to obtain the main results, are that the WIOD data cover a shorter time period and a smaller number of countries. However, the WIOD dataset also has two main advantages. First, it offers a complete sectoral coverage for each of the countries in the sample. Thus, summing across all sectors will enable us to cover total trade (including international and intra-national trade flows) for each country in the sample. Second, on a related note, we will use the WIOD data to test our theory at the sectoral level. We do take advantage of these features of the WIOD data in the sensitivity analysis

[18] Intra-national trade flows are constructed as apparent consumption, i.e. as the difference between the gross value of total production and total exports. The original trade data come from the UN COMTRADE database, accessed via WITS. The data on total gross production come from the CEPII TradeProd database and UNIDO IndStat database.

[19] The intra-national trade flows in the WIOD database are constructed using input-output linkages. See Timmer et al. (2015) and Timmer et al. (2016) for further details.

that we present in Appendix B, where we demonstrate that our findings are confirmed with the WIOD dataset.

4.2 A Reduced Form Approach to Short Run Gravity

Start with the intuition that in an era of globalization (1988-2006 or 2000-2014 depending on the dataset), cross-border trade capacities $\lambda(i,z,\tau)$ are inefficiently small (network links are inefficiently sparse) while domestic capacities $\lambda(i,i,\tau)$ are inefficiently large (trade links are relatively too dense). Over time capacity investment presumably evolves toward efficiency, so cross-border investment in trade links rises relative to domestic investment in trade links. The reduced from approach looks for evidence in cross-border time fixed effects without imposing any time path of adjustment.

Cross-border-time fixed effects averaged over the cross-destination variation for an origin are structurally interpreted as:

$$[\lambda(i,z,\tau)/\lambda(i,i,\tau)]^{1-\rho} = [\lambda(i,z,0)/\lambda(i,i,0)]^{1-\rho}\mu(i,\tau),\ \ z \neq i$$

where $\mu(i,\tau) > 1$ is the growth factor of $[\lambda(i,z,\tau)/\lambda(i,i,\tau)]^{1-\rho}$. No structure is imposed over time, so $\mu(i,\tau)/\mu(i,\tau-1)$ is not restricted over τ. Initially we average across all cross-border pairs at each point in time, $\mu(i,\tau) = \mu(\tau)$. Subsequently we allow for differential rates of adjustment for origin countries specialized to a difference between developed and developing countries. Other uniformity restrictions are plausible as well in order to supply degrees of freedom to identify the average time variation of primary concern. Idiosyncratic variation of growth of $\lambda(i,z,\tau)/\lambda(i,i,\tau)$ that deviates from the uniform growth assumption is associated with effects that go into the error term of the gravity equation. Averaging the $\mu(\tau)$s to a smooth exponential increase in efficiency as we do for some purposes below, $\tilde{\mu}(\tau) = e^{b_\tau \tau}$ where b_τ equals $1-\rho$ times the growth rate of $\lambda(i,z)/\lambda(i,i)$.

A useful benchmark is a long run estimated panel gravity model that does not control for the efficiency improvements to be looked for based on the short run gravity theory. Thus $\rho = 1$ and the econometric model run in logs with OLS is:[20]

$$\begin{aligned} LN\_X_{ij,\tau} &= \beta_1 LN\_DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 CLNY_{ij} + \beta_4 LANG_{ij} + \beta_5 FTA_{ij,\tau} + \\ & \quad \beta_6 LN\_TARIFF_{ij,\tau} + \beta_7 INTL\_BRDR_{ij} + \pi_{i,\tau} + \chi_{j,\tau} + \epsilon_{ij,\tau}, \end{aligned} \tag{18}$$

The dependent variable in (18) is the logarithm of nominal bilateral trade flows. The regressors are the standard gravity variables including the logarithm of bilateral distance between countries i and j (LN_DIST_{ij}), and indicator variables for contiguous borders ($CNTG_{ij}$), common language ($LANG_{ij}$), colonial ties ($CLNY_{ij}$), and free trade agreements ($FTA_{ij,\tau}$). In addition, we include the log of applied tariffs ($LN\_TARIFF_{ij,\tau}$).[21] $INTL\_BRDR_{ij}$ is an indicator variable that takes a value of one for international trade and it is equal to zero for internal trade. Specification (18) includes exporter-time and importer-time fixed effects to control for the multilateral resistances of Anderson and van Wincoop (2003) and will absorb any other observable and unobservable country-specific characteristics on the exporter and on the importer side, respectively. Finally, following the recommendation of Cheng and Wall (2005), we use 3-year intervals instead of consecutive years in order to obtain our main estimates.[22]

Estimates obtained from specification (18) are reported in column (1) of Table 1. With an $R^2 = 0.88$, gravity delivers its usual strong fit. In addition and as expected, distance and tariffs[23] are strong impediments to trade, while the presence of colonial ties, sharing a common official language, and having FTAs in force all promote bilateral trade. The

[20]Specification (18) is consistent with the traditional empirical approach to estimate gravity with the OLS estimator. We show in the sensitivity experiments that our results are robust to using the PPML estimator instead, which, as demonstrated by Santos Silva and Tenreyro (2006, 2011), simultaneously accounts for heteroskedasticity and takes into account the information contained in zero bilateral trade flows.

[21]Since tariffs are a direct price shifter, the estimate of the tariff elasticity equals the trade elasticity σ.

[22]We demonstrate in the sensitivity experiments below that our main results and conclusions are robust to using all years in the sample. We refer the reader to Head and Mayer (2014) and Yotov et al. (2016) for further discussion of the challenges with the estimation of empirical gravity models and their solutions.

[23]The estimated tariff elasticity suggests a value of σ that is low relative to previous estimates. We suspect this is due to aggregation bias for such a heterogeneous sector as manufacturing.

negative, large, and significant estimate on $INTL\_BRDR$ is also expected, and it confirms the strong presence of international borders, even after controlling for the impact of distance and tariffs. Apart from the insignificant estimate on $CNTG$, all other estimates from column (1) of Table 1 are readily comparable to corresponding estimates in the literature (see the meta analysis estimates from Head and Mayer (2014)). In sum, the results from column (1) of Table 1 are long run benchmarks that confirm the representativeness of our sample and provide long run responses to compare to the short run responses that follow.

The estimates from column (2) of Table 1 are obtained with time-varying cross-border trade fixed effects, motivated by our theory. A series of year-specific dummies for international borders $\sum_{\tau=1991}^{2006} \beta_\tau INTL\_BRDR\_\tau_{ij}$ are introduced to (18):

$$\begin{aligned} LN\_X_{ij,\tau} \;=\; & \beta_1 LN\_DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 CLNY_{ij} + \beta_4 LANG_{ij} + \beta_5 FTA_{ij,\tau} + \beta_6 LN\_TARIFF_{ij,\tau} + \\ & \beta_7 INTL\_BRDR_{ij} + \sum_{\tau=1991}^{2006} \beta_\tau INTL\_BRDR\_\tau_{ij} + \pi_{i,\tau} + \chi_{j,\tau} + \epsilon_{ij,\tau}, \end{aligned} \tag{19}$$

The theoretical interpretation of each of the time-varying border estimates is $\beta_\tau = (1-\rho)\Delta \ln[\lambda(i,z,\tau)/\lambda(i,i,\tau)] = (1-\rho)\Delta \ln \mu(\tau)$ and, by construction, these estimates should be interpreted as deviations from the estimate of β_7. Due to perfect collinearity, we omit the border for the first year of the sample, 1988.

Two findings stand out from the estimates in column (2) of Table 1. First, we see that the estimates of the standard gravity variables are statistically unchanged from column (1), implying that the uniform growth assumption imposed in (19) is approximately valid (non-uniform growth would likely be correlated with cross-section variation in bilateral distance, etc.). Second, the estimates on $INTL\_BRDR_{ij}$ are all positive, statistically significant, and gradually increasing over time. SR gravity theory in an era of globalization is consistent with this finding, hence interpretation with the model provides a structural explanation of the "missing globalization puzzle".

The estimates from column (3) of Table 1 are obtained after allowing for the effects of distance to vary over time. Specifically, we interact the distance variable with dummy variables for each year on our sample.[24] In order to ease the interpretation of our results, we keep the original distance variable and do not include the distance variable for 1988. Thus, the estimate on LN_DIST should be interpreted as the effect of distance in 1988, and all other time-varying distance estimates should be interpreted as deviations from the estimate on LN_DIST. Column (3) implies that distance elasticities are stable over time. The "distance puzzle" of non-declining distance elasticities in international trade (Disdier and Head (2008)) remains.

As demonstrated in Section 3.1, short run gravity theory suggests that an intuitive interpretation of the effects of distance on trade over time is to combine the direct impact that is captured by the distance estimates in column (1) with the indirect effect of distance on efficiency improvements associated with movement toward efficient capacity $\lambda^*(i, z)$. Thus the estimates of the (insignificant) time-varying distance effects for each year should be multiplied by the corresponding portion of efficiency improvement change for that year that is attributable to distance. The implication is that the "distance puzzle" is resolved; while the distance effects will remain strong and negative in 2006, they have steadily decreased over time. (The structural approach to the dynamics of capacity investment that follows provides more evidence buttressing this intuitive implication.)

The estimates from columns (4) and (5) of Table 1 employ pair fixed effects. The motivation for the use of the time-invariant bilateral dummies is twofold. First, these variables will completely account for the impact of all observable and unobservable determinants of bilateral trade, cf. Agnosteva et al. (2014) and Egger and Nigai (2015). Second, on a related note and as demonstrated by Baier and Bergstrand (2007), the use of the pair fixed effects will help mitigate endogeneity concerns related to the trade policy variables in our specification. The main findings from columns (4) and (5) are that (i) the estimates of the trade policy variables are unchanged across the two specifications; and (ii) we observe the gradual

[24] In principle, we could allow for the effects of all gravity variables to vary over time. We chose to focus on distance only for expositional simplicity and because the distance variable has been used most prominently to capture trade costs and their changes.

and economically and statistically significant increase in the estimates on the international border dummies. Exponentiating the border-time cumulative efficiency change estimate in column (5) of Table 1 (0.965) yields an overall efficiency of trade gain from 1988-2006 of 162%: world cross-border trade is 162% larger in 2006 than it would have been with the more inefficient bilateral capacities of 1988.

Column (6) relaxes the uniform efficiency growth restriction to allow for efficiency growth to vary across 'rich' vs. 'poor' exporters, as classified by the World Bank.[25] The estimates from column (6) of Table 1 reveal significant efficiency improvements for each group of countries captured by the positive, significant, and gradually increasing over time estimates on $INTL\_BRDR\_HIGH$ and $INTL\_BRDR\_LOW$. More importantly, low income countries have converged faster toward more efficient trade during the period of investigation. (The result may not be robust to further country level disaggregation, plausibly being driven by few outliers such as China. The difference between high income and developing countries may also an artifact of composition effects, as the sectoral composition of manufacturing changes differently across countries. Sectoral disaggregation is indicated as part of investigating country differences.)

A series of sensitivity experiments demonstrate the robustness of our main results, summarized here with a full report in Appendix B. Motivated by the work of Santos Silva and Tenreyro (2006, 2011) advocating the PPML estimator as an alternative to OLS with the log of gravity, Table 2 replaces the main estimates from Table 1 with those based on the PPML estimator. PPML estimates confirm the OLS results: all border-time estimates are positive, significant and increasing over time. Another experiment uses size-adjusted trade as the dependent variable. The size-adjusted dependent variable specification is consistent

[25] Note that non-uniform efficiency growth does not violate the structural assumptions used to specify short run gravity. In principle origin-specific estimates of the efficiency-improvement effects $\beta_\tau(i)$'s can be identified from variation across destinations z under a uniformity condition imposed on the external (foreign) destinations vs. internal trade. The original World Bank classification includes five income categories: 'High Income OECD', 'High Income Non-OECD', 'Upper Middle Income', 'Lower Middle Income', and 'Low Income' countries. We used the top three categories to form our sample of 'rich' countries, and the bottom two categories to form the group of 'poor' countries.

with theory and tends to reduce the problem of heteroskedasticity associated with the level of trade specification (19). The OLS and PPML estimates using size-adjusted trade as dependent variable confirm the robustness of our main findings. A third experiment uses data for all years rather than 3 year intervals. The OLS and PPML results are very similar to the main estimates from Table 1. The next experiment employs the newly available 2016 WIOD data.[26] Again, the main results from Tables 1 and 2 are closely matched with OLS and PPML estimators respectively. Finally, exploiting the sectoral dimension of the WIOD data we reproduce our main results for Crop and Animal Production, Forestry and Logging, Fishing and Aquaculture, Mining and Quarrying, Manufacturing, and Services. The results naturally vary by sector, but the main finding for the presence of efficiency improvement is present in each sector.

4.3 A Structural Approach to Short Run Gravity

The complementary structural approach to estimating short run gravity adds a theory of investment in bilateral capacity. The investment structure enables (i) direct inference about the buyers' incidence parameter ρ, (ii) additional tests of the short run model, and (iii) consistency checks between the two complementary approaches.[27]

Suppose that intensive margin investment in pair-specific capital moves the current level from its past level toward the efficient level at some rate of log-linear adjustment that is implied by a Cobb-Douglas function of efficient and past levels.[28] The adjustment process

[26] See Data section for a description and details on the WIOD data.

[27] There are advantages and disadvantages to the treatment of efficiency improvements with border-time dummies. As to advantages, dummy variables are exogenous by construction, hence avoid the endogeneity issues raised by our complementary structural approach to dynamic adjustment. Another advantage of the dummy variable approach relative to the structural approach is a comprehensive accounting for all forces that could have contributed to improved efficiency over time, including those omitted from the structural approach. Thus while (19) imposes a uniform efficiency improvement across destinations $z \neq i$ at each time τ, the specification does not impose any pattern whatever on efficiency variation over time. The disadvantage of the efficiency treatment with border-time dummies is that it does not permit identification of ρ, but only the effect of adjustment of cross-border capacity toward the long run over time.

[28] The approach developed here is a first cut at a difficult problem. It leaves out plausibly important effects. Investment in bilateral trade may be systematically affected in a differential fashion by a number of variables reflecting allocations subject to credit constraints. Exchange rate risk's effects on trade flows can be hedged

specification is inspired by *inter alia* Lucas and Prescott (1971), Hercowitz and Sampson (1991), and Anderson et al. (2015) and Eaton et al. (2016) in the gravity context:

$$\lambda(i,z;\tau) = \lambda^*(i,z)^{\delta}\lambda(i,z;\tau-1)^{1-\delta};\ \delta \in (0,1). \tag{20}$$

The parameter δ reflects both costs of adjustment and depreciation, the higher is δ the faster the movement to the efficient level. In the steady state, $\lambda = \lambda^*$.

Operationalization of the adjustment process requires finding observables to replace the unobservable λs. The agents know that efficient allocation implies that $\lambda^*(i,z) = s^*(i,z)$. This suggests that $\lambda(i.z,\tau) \to s(i,z,\tau)$. Groping ahead toward the eventual efficient allocation suggests a specification of the dynamic process replacing $s^*(i,z)$ with $s(i,z,\tau)$. Moving toward operationality, replace the unobservable $\lambda^*(i,z) = s^*(i,z)$ in (20) with $s(i,z,\tau)$. Similarly, replace the unobservable $\lambda(i,z,\tau-1)$ with $s(i,z,\tau-1)$. (Note that $\lambda(i,z)$ is likely unobservable by individual competitive firms in i exporting to multiple zs, as well as by the econometrician.) Then the costly adjustment specification in the spirit of Lucas and Prescott is:[29]

$$\lambda(i,z,\tau) = s(i,z,\tau)^{\delta}s(i,z,\tau-1)^{1-\delta}.$$

The steady state of this process reaches the efficient allocation of long run gravity, is plausible as an approximation to more sophisticated expectations mechanisms, and its simplicity preserves the simple loglinear features of structural gravity.

Substitute the right hand side of the preceding equation for the (implicit) contempora-

for many sectors with minimal cost, but hedging over longer intervals appropriate for fixed commitments is expensive. This suggests that bilateral exchange rate volatility may significantly affect investment in bilateral trade but not variable trade cost. Similarly, bilateral covariance of business cycles may affect investment but not variable trade cost. Such refinements are beyond the scope of this project.

[29]The specification here is in part an expedient to deal with the problem of unobservable capacity allocations, but on a conceptual level it violates rational expectations. The alternative of super-humanly rational expectations is implausible considering the extremely high dimensionality where each bilateral link has potentially many uncoordinated agents and there are very many links with simultaneous interaction. In a simpler environment of perfect information, Anderson et al. (2015) combine the Lucas-Prescott adjustment mechanism with the consumer's inter-temporal maximization problem to derive the optimal accumulation of country-specific physical capital. In contrast, the present case involves a bilateral capacity adjustment on each link, N^2 adjustment paths compared to N in the simpler case.

neous value of $\lambda(i,z,\tau)$ in gravity equation (12). The result is

$$X(i,z,\tau) = \frac{Y(i,\tau)E(z,\tau)}{Y(\tau)} \left[\frac{t(i,z,\tau)}{\Pi(i,\tau)\widetilde{P}(z,\tau)}\right]^{(1-\sigma)\rho} \left[s(i,z,\tau)^{\delta} s(i,z,\tau-1)^{1-\delta}\right]^{1-\rho}.$$

The presence of contemporaneous trade share $s(i,z,\tau)$ on the right hand side of the preceding equation requires solution for contemporaneous bilateral trade flows to yield:

$$X(i,z,\tau) \;=\; Y(i,\tau)\left(\frac{E(z,\tau)}{Y(\tau)}\right)^{\frac{1}{1-\delta(1-\rho)}} \left[\frac{t(i,z,\tau)}{\Pi(i,t)\widetilde{P}(z,\tau)}\right]^{\frac{(1-\sigma)\rho}{1-\delta(1-\rho)}} \left(\frac{X(i,z,\tau-1)}{Y(i,\tau-1)}\right)^{\frac{(1-\delta)(1-\rho)}{1-\delta(1-\rho)}}. \quad (21)$$

Use the standard gravity proxies for bilateral trade costs introduced in the previous section, log-linearize specification (21), and add an error term in order to obtain the econometric model:[30]

$$\begin{aligned} LN\_X_{ij,\tau} \;=\; & \tilde{\beta}_1 LN\_DIST_{ij} + \tilde{\beta}_2 CNTG_{ij} + \tilde{\beta}_3 CLNY_{ij} + \tilde{\beta}_4 LANG_{ij} + \tilde{\beta}_5 FTA_{ij,\tau} + \tilde{\beta}_6 LN\_TARIFF_{ij,\tau} + \\ & \tilde{\beta}_7 INTL\_BRDR_{ij} + \tilde{\beta}_8 LN\_X_{ij,\tau-1} + \tilde{\pi}_{i,\tau} + \tilde{\chi}_{j,\tau} + \tilde{\epsilon}_{ij,\tau}. \end{aligned} \quad (22)$$

Note that the exporter-time fixed effects in the preceding expression will absorb and control for all contemporaneous as well as lagged exporter-specific characteristics, including the structural contemporaneous and lagged size terms, that may affect bilateral trade.[31]

The structural interpretation of the estimated coefficient on $LN\_X_{ij,\tau-1}$ is:

$$\tilde{\beta}_8 = \frac{(1-\delta)(1-\rho)}{1-\delta(1-\rho)}. \quad (23)$$

(23) is used below to recover the structural efficiency parameter ρ in combination with information on δ.

[30] Similar to the analysis in the previous section, in the sensitivity analysis we also experiment by estimating the model with the PPML estimator.

[31] In contrast, exporter specific δ_i must be treated with exporter-specific $\beta_1 - \beta_8$. An exporter-type δ_i consistent with the evidence that border-time fixed effects differ between high income and developing economies. We impose a common δ here since our use of aggregate manufacturing trade confounds dynamic adjustment cost effects with changing composition effects that vary across exporters. A proper investigation of origin variation in δ should be done with disaggregated data.

Specification (22) for short run gravity estimation differs in two important ways from long run gravity specification (18).[32] The first difference between the two estimating equations is the appearance of the lagged dependent variable as a regressor in the SR gravity specification (22). From an econometric perspective short run gravity implies that the standard estimation of gravity without lagged variables may suffer from omitted variable bias. Hypothesizing that the standard gravity estimator controls also control for the omitted capacity variables, the standard estimator may be an unbiased estimator of long run gravity. The no-bias hypothesis can be tested jointly with the structural implications of short run gravity, the second difference between specifications (18) and (22).

The theoretical relationship between long run and short run gravity coefficients in Corollary 2.1, extended to accommodate the theory of investment developed here is:

$$\tilde{\beta}_{SR} = \beta_{LR} \frac{\rho}{1 - \delta(1 - \rho)}, \tag{24}$$

where, the subscripts SR and LR denote estimates from the short run (SR) and from the long run (LR) empirical gravity specifications, respectively. For any given gravity variable coefficient, the combination of equations (23) and (24) enable a joint test of SR gravity structure and the no-bias hypothesis for estimates of long run gravity by checking whether:

$$\frac{\tilde{\beta}_{SR}}{\beta_{LR}} = 1 - \tilde{\beta}_8. \tag{25}$$

Because the gravity specifications include six gravity covariates, test (25) is performed six times, and reported below.

The natural starting point for estimation is a standard/long run gravity model:

$$\begin{aligned} LN\_X_{ij,\tau} \;=\; & \beta_1 LN\_DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 CLNY_{ij} + \beta_4 LANG_{ij} + \beta_5 FTA_{ij,\tau} + \beta_6 LN\_TARIFF_{ij,\tau} + \\ & \beta_7 INTL\_BRDR_{ij} + \pi_{i,\tau} + \chi_{j,\tau} + \epsilon_{ij,\tau}, \end{aligned} \tag{26}$$

[32] In addition to the two estimation-related differences, the structural interpretation of the fixed effects will also be different in the SR gravity specification since the size variables no longer appear with unitary elasticities in the theoretical SR gravity model. However, this has no consequences for gravity estimations.

The only difference between specifications (22) and (26) is the absence of lagged trade in the latter specification. The standard gravity estimates from column (1) of Table 3 are readily comparable to their counterparts from the previous section. The estimates from column (2) of Table 3 are obtained from specification (22) with a lagged dependent variable.

An important potential issue is the dynamic panel bias from the use of a lagged dependent variable, i.e. the Nickell (1981) bias. As explained in Roodman (2009a), the use of sufficiently long time spans may mitigate and even eliminate the Nickell bias by construction. The analysis in this section employs every year in the dataset[33] motivated by this observation. Importantly, even if the time coverage is not long enough to eliminate the dynamic panel bias completely, the 'naive' OLS results from column (2) are useful for our purposes because they establish an upper bound for the key estimate of the coefficient $\tilde{\beta}_8$ on the structural efficiency term $LN\_X_{ij,\tau-1}$.[34] This, in combination with expression (23), will enable us to draw inference about the bounds for the structural efficiency parameter ρ. In particular, we will use the estimates from column (2) to establish a lower bound for ρ. Finally, note that the OLS results from column (2) will be supported by the findings from the more sophisticated econometric specifications that we employ below.

The estimates in column (2) of Table 3 reveal a positive and significant estimate of $\tilde{\beta}_8 = 0.788$ (std.err. 0.009). While the estimate of $\tilde{\beta}_8$ seems fairly large in magnitude, it is quite low to be explained solely by the capital adjustment cost δ from our theory. Capitalize on the structural restriction (23) in combination with an external estimate $\delta = 0.061$[35] in order to recover the incidence elasticity parameter $\rho = 0.202$ (std.err. 0.008). Three

[33] In the sensitivity experiments we also employ the 3-year interval data from the previous section and we demonstrate that our main results are robust to employing this alternative sample.

[34] Specification (26) will deliver an upward bound for the estimate on $\tilde{\beta}_8$ due to positive correlation between the lagged dependent variable and the pair fixed effects in the error term, which will inflate the estimate on $LN\_X_{ij,\tau-1}$ by attributing additional predictive power to it. See Roodman (2009a) for further details and a very informative discussion.

[35] This value for δ is the average adjustment cost parameter estimate from the structural gravity setting of Anderson et al. (2015), obtained under the same assumptions for investment. Specifically, Anderson et al. (2015) recover δ from a second stage regression that links trade openness to capital accumulation. These authors also obtain a distribution of country-specific capital adjustment cost parameters that vary between 1.6 and 13.6.

implications follow.

First, column (2) of Table 3 implies that gravity estimates that do not control for efficiency improvements and adjustment may suffer significant biases. The estimates from column (1) are very significantly larger than the estimates from column (2). This result reinforces the findings of Olivero and Yotov (2012) and Eichengreen and Irwin (1996) who conclude that they "will never run another gravity equation that excludes lagged trade flows" (p.38). But the possible bias may be one of interpretation under a valid 'no-bias' hypothesis that standard gravity covariates also control for variation of the omitted variable.

Second, SR gravity theory as captured by equation (24) implies at least partial solution to the trade elasticity puzzle – the gap between estimates of the trade elasticity σ from the trade and the macro literatures, cf. Ruhl (2008), Arkolakis et al. (2011) and Crucini and Davis (2016).

Third, column (2) of Table 3 implies support for the quantitative structural predictions of SR gravity theory and the 'no-bias' hypothesis. Expression (25) implies the hypothesis test:

$$1 - \tilde{\beta}_8 - \tilde{\beta}_{SR}/\beta_{LR} = 0 \tag{27}$$

The test statistic in column (3) of Table 3 uses the estimated coefficients of columns (1)-(2) of Table 3 in (27) for each of the gravity variables. The null hypothesis cannot be rejected at high levels of significance for all six tests.[36]

The results from the lagged dependent variable empirical model are complements to the border-time dummy variable model under the strong assumption that the Lucas-Prescott adjustment model (20) applies to both. In that case, the border-time dummy variable coefficient $\beta_7+\beta_\tau$ should in theory be declining at rate $b_\tau = (1-\rho)\delta$.[37] The lagged dependent

[36] The largest estimate that we obtain in column (3) is for $CLNY$. However, as can be seen from columns (1) and (3) of Table 3, the estimate of the effect of sharing a common border is insignificant to start with. Thus, even though the corresponding estimate in column (3) is insignificant and supports our theory, we discount this finding.

[37] Divide both sides of (20) by $\lambda(i, z, \tau - 1)$ to yield $\lambda(i, z, \tau)/\lambda(i, z, \tau - 1) = [\lambda^*(i, z)/\lambda(i, z, \tau - 1)]^\delta$ with growth rate δ.

variable model implies $\beta_8 = (1-\rho)(1-\delta)/[1-\delta(1-\rho)]$. This structure can be used in two ways. First, for a given δ, compare the implied value of ρ from the lagged dependent variable model to the implied value of ρ from the border-time dummy variable model. Second, solve the two equations simultaneously for the unique values of ρ and δ that satisfy the equations: $\hat{\rho} = (1-\hat{\beta}_8)(1-\hat{b}_\tau)$ and $\hat{\delta} = \hat{b}_\tau/[\hat{\beta}_8 + \hat{b}_\tau(1-\hat{\beta}_8)]$.

We first recover $\hat{b}_\tau = (1-\rho)\delta$ as the slope of the fitted line of the estimates of the changes in the effects of international borders, from our preferred specification with paired fixed effects from column column (5) of Table 1. The best-fit line in the β_τ–τ coordinate space is plotted in Figure 1. The regression implies that the slope is $\hat{b}_\tau = 0.046$. In combination with the adopted value of $\delta = 0.061$, the estimate of $\hat{b}_\tau = 0.046$ enables us to recover $\hat{\rho} = 1 - \hat{b}_\tau/\delta = 0.246$. This value is somewhat larger but comparable in magnitude to the preceding estimate of $\rho = 0.202$ from the structural approach above. Finally, use the estimates of $\hat{\beta}_8$ and $\hat{b}_\tau$ in the simultaneous equations $\hat{\rho} = (1-\hat{\beta}_8)(1-\hat{b}_\tau)$ and $\hat{\delta} = \hat{b}_\tau/[\hat{\beta}_8 + \hat{b}_\tau(1-\hat{\beta}_8)]$ in order to simultaneously recover unique values of $\hat{\rho} = 0.202$ (std.err. 0.008) and $\hat{\delta} = 0.058$ (std.err. 0.001). ρ is statistically and quantitatively indistinguishable from the estimate of using the external value of $\delta = 0.061$ while the internally generated estimate of δ differs statistically significantly but quantitatively insignificantly from the external value. The close magnitudes of the various inferences of the value of ρ are compelling. Pushing inference to the limit, we note that the lower value of $\hat{\delta} = 0.058$ obtained from the simultaneous solution corresponds more closely to the interest rate values that are used for calibrations in the macroeconomic literature.

Estimates of $\hat{\rho}$ imply restrictions on the deep structural parameters of supply η and demand σ because

$$1/\rho - 1 = (\sigma - 1)/\eta.$$

Using the estimate $\hat{\rho} = 0.20 \Rightarrow \hat{\eta} = (\sigma-1)/4$. Estimate of σ in the literature[38] range from 4 to 10, implying $\hat{\eta} \in [0.75, 2.25]$. The finite short run supply elasticities implied pass a smell

[38] The estimate of $1-\sigma$ inferred in our regressions is lower but we assume it is biased downward.

test (with heterogeneous firms, $\tilde{\eta} < 1$ is readily possible) and provide an intuitive foundation for thinking about short run gravity.

Columns (4)-(8) of Table 3 report results from alternative estimators. Columns (4) and (5) use pair fixed effects for two reasons. First, the pair fixed effects will absorb and control for any omitted time-invariant determinants of bilateral trade. Second, the use of pair fixed effects has proven to be a useful method to address possible endogeneity of FTAs and trade policy in general. In the context of dynamic panel bias caused by the lagged dependent variable (Roodman (2009a)), an additional advantage of pair fixed effects is mitigation of the dynamic panel bias. The so-called Least Square Dummy Variable (LSDV) treatment of the Nickell bias is especially effective for samples with long time coverage. Finally, as demonstrated in Roodman (2009a), in case the LSDV estimator does not eliminate completely the bias in the estimate of the lagged dependent variable, then the remaining bias would be a downward bias owing to a negative correlation between the lagged dependent variable and the remaining error. This, in combination with expression (23), enables us to establish an upper bound for ρ. Using external value $\delta = 0.061$, the result in column (5) give an upper bound $\bar{\rho} = 0.371$ (std.err. 0.011) to pair with the lower bound from column (2) $\underline{\rho} = 0.202$.

Two further results stand out from columns (4)-(6). First, the SR gravity theory test (27) reported in column (6) offer strong support for the theory. Second, in combination with the estimates from the border-time dummy variable model, the results from column (5) imply $\hat{\rho} = 0.368$ (std.err. 0.011) and $\hat{\delta} = 0.072$ (std.err. 0.001). The incidence elasticity estimate of ρ is almost unchanged by the switch from externally to structurally generated δ.

The relatively long time coverage of our sample attenuates endogeneity bias in the LSDV estimator of SR gravity of column (5), but may not eliminate it. Column (7) of Table 3 thus implements an IV estimation along the lines of Frankel and Romer (1999). We construct an instrument for lagged trade by using a reduced-form gravity specification that only includes the standard gravity variables, which are exogenous by definition, and also exporter and importer population, which also are arguably exogenous. We use the second to fifth lags of the newly constructed trade variable as instruments for the lagged dependent variable in specification (22). The estimation results from column (7) support our theory. First, the instruments are good since, based on the specification tests that we perform in the bottom of Table 3, they pass all IV tests. Second, once again, we obtain a positive and significant estimate on the lagged trade term. In combination with (23) and the external value of $\delta = 0.061$, $\rho = 0.242$ (std.err. 0.120). We calculate $\hat{\rho} = 0.242$ (std.err. 0.118) and $\hat{\delta} = 0.061$ (std.err. 0.009) when we combine the estimates from column (7) with those from the main results from the border-time dummy variable model. The close approximation of the two estimated parameter vectors is remarkable. Pushing the results of the various specifications hard, our preferred range for the key incidence elasticity is $\rho \in (0.20, 0.24)$ and our preferred value is 0.24, while we report a conservative range $\rho \in (0.20, 0.37)$.

Note finally that the SR test for the LSDV IV estimation reported in column (8) weakly supports the theory. We cannot reject the null hypothesis for the FTA variable. In contrast, the test statistic on LN_TARIFF rejects the null. We are suspicious of all estimates of the tariff parameter due to aggregation bias and low variation.

We conclude with a brief description of robustness checks. Detailed results and descriptions for each experiment can be found in Appendix B. The first experiment uses interval data instead of yearly data. The second experiment uses size-adjusted trade as the dependent variable. Use of the trade share instead of the level of trade tends to reduce the problem of heteroskedasticity. The third sensitivity experiment uses international trade data only. All three experiments confirm robustness of the main findings. We also experiment with the PPML estimator, yielding estimates that correspond to the upper and lower bound results from Table 3. The next experiment employs the system-GMM estimator of Arellano and Bond (1991) and Arellano and Bover (1995)/Blundell and Bond (1998). Here the results for ρ are higher, though remaining well inside the unit interval. (We suspect the poorer small sample properties of the system-GMM estimator are to blame.) Finally, the main results are reproduced with the aggregate WIOD data as well as for each of the main sectors that are covered in the WIOD dataset. In each case, the new estimates are consistent with and comparable to the main results from Table 3.

5 Conclusion and Extensions

Short run gravity features (i) joint trade costs endogenous to bilateral volumes, (ii) long run gravity as a limiting case of efficient investment in bilateral capacities, (iii) tractable short and long run models of the extensive margin, and (iv) a structural ratio of short run to long run trade elasticities. Despite the complexity, bilateral trade is a geometric weighted average of long run gravity and a bilateral capacity variable, where the weight on long run gravity is the short run buyers' incidence elasticity. The incidence elasticity is itself micro-founded. The theoretical short run gravity model finds strong support in the data and the empirical analysis resolves several empirical puzzles.

The short run gravity model here suggests several extensions. (i) The model can be applied at the firm level, where indeed its assumption of a common Cobb-Douglas distribution 'production' function is most natural. (ii) Another potential extension is to the explanation of income inequality. The rents to destination-specific managerial labor may be linked to inefficient investment and costs of adjustment, inducing income inequality within firms and across firms in a sector. (iii) The model also extends upward to embedding in a multi-sector general equilibrium setting. The stock of labor and human capital is simply the sectoral allocation, possibly with differentials due to search costs. (iv) Empirical implementation and tests of the short run and long run extensive margin implications of SR gravity is an important challenge.

More challenging, perhaps leading outside the scope of the present model, is opening the black box of bilateral capacity. This paper treats capacity investment as effectively on the exporter side only. Since some investment is on the importer side, this may affect the dynamics and even the short run statics of the model. Firm level data and modeling suggests more carefully modeling the 'manager' input, sources of heterogeneity across markets and also across modes of organization: arms length contracting, joint ventures or horizontal integration in a multinational structure. The Arkolakis (2010) model extended to a dynamic setting may be a useful starting point.

References

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov, "Intra-national Trade Costs: Measurement and Aggregation," *NBER Working Paper No. 19872*, 2014.

Anderson, James E., "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 1979, *69* (1), 106–116.

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

Anderson, J.E., "The Gravity Model," *Annual Review of Economics*, 2011, *3*, 133–160.

— **, Mario Larch, and Yoto V. Yotov**, "Growth and Trade with Frictions: A Structural Estimation Framework," *NBER Working Paper No. 21377*, 2015.

Anderson, T. W. and C. Hsiao, "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 1982, *82* (1), 47–82.

Arellano, Manuel and Olympia Bover, "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 1995, *68* (1), 29–51.

— **and Stephen Bond**, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 1991, *58* (2), 277–297.

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

— **, Arnaud Costinot, and Andrs Rodrguez-Clare**, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

— **, Jonathan Eaton, and Samuel S. Kortum**, "Staggered Adjustment and Trade Dynamics," 2011 Meeting Papers 1322, Society for Economic Dynamics 2011.

Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, "Dynamics of the Trade Balance and the Terms of Trade: the J-curve," *American Economic Review*, 1994, *84* (1), 84–103.

Baier, Scott L. and Jeffrey H. Bergstrand, "Do Free Trade Agreements Actually Increase Members' International Trade?," *Journal of International Economics*, 2007, *71* (1), 72–95.

— **, Yoto V. Yotov, and Thomas Zylkin**, "On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration," *School of Economics Working Paper Series, 2016-15*, 2016.

Bergstrand, Jeffrey H., Mario Larch, and Yoto V. Yotov, "Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation," *European Economic Review*, 2015, *78*, 307–327.

Bergstrand, J.H., "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence," *Review of Economics and Statistics*, 1985, *67* (3), 474–481.

Besedes, Tibor and Thomas J. Prusa, "Product differentiation and duration of US import trade," *Journal of International Economics*, December 2006, *70* (2), 339–358.

Blundell, Richard and Stephen Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 1998, *87* (1), 115–143.

Boulhol, Herv and Alain de Serres, "Have Developed Countries Escaped the Curse of Distance?," *Journal of Economic Geography*, January 2010, *10* (1), 113–139.

Broda, C., J. Greenfield, and D. Weinstein, "From Groundnuts to Globalization: A Structural Estimate of Trade and Growth," *NBER Working Paper No. 12512*, 2006.

Brun, Jean-Franois, Cline Carrre, Patrick Guillaumont, and Jaime de Melo, "Has Distance Died? Evidence from a Panel Gravity Model," *World Bank Economic Review*, 2005, *19* (1), 99–120.

Buch, Claudia M., Jorn Kleinert, and Farid Toubal, "The Distance Puzzle: On the Interpretation of the Distance Coefficient in Gravity Equations," *Economics Letters*, June 2004, *83* (3), 293–298.

Carrre, Cline and Maurice Schiff, "On the Geography of Trade. Distance is Alive and Well," *Revue conomique*, 2005, *56* (6), 1249–1274.

_ , Jaime de Melo, and John Wilson, "The Distance Puzzle and Low-Income Countries: An Update," *Journal of Economic Surveys*, 2013, *27* (4), 717–742.

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

_ , "The Network Structure of International Trade," *American Economic Review*, 2014, *104* (11), 3600–34.

Cheng, I-Hui and Howard J. Wall, "Controlling for Heterogeneity in Gravity Models of Trade and Integration," *Federal Reserve Bank of St. Louis Review*, 2005, *87* (1), 49–63.

Coe, D., A. Subramanian, N. Tamirisa, and R. Bhavnani, "The Missing Globalization Puzzle," *IMF Working Paper No. 171*, 2002.

Costinot, Arnaud and Andrs Rodrguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," Chapter 4 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, Elsevier Ltd., Oxford, 2014.

Crucini, Mario J. and J. Scott Davis, "Distribution capital and the short- and long-run import demand elasticity," *Journal of International Economics*, 2016, *100* (C), 203–219.

Disdier, A.-C. and K. Head, "The Puzzling Persistence of the Distance Effect on Bilateral Trade," *Review of Economics and Statistics*, 2008, *90* (1), 37–48.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

_ , _ , B. Neiman, and J. Romalis, "Trade and the Global Recession," *American Economic Review*, 2016, *106* (11), 3401–38.

Egger, Peter H. and Sergey Nigai, "Structural gravity with dummies only: Constrained ANOVA-type estimation of gravity models," *Journal of International Economics*, 2015, *97* (1), 86–99.

Eichengreen, Barry and Douglas A. Irwin, "The Role of History in Bilateral Trade Flows," NBER Working Papers 5565, National Bureau of Economic Research, Inc May 1996.

Feenstra, R.C., Maurice Obstfeld, and Katheryn N. Russ, "In Search of the Armington Elasticity," available for download at `http://www.econ.ucdavis.edu/faculty /knruss /FOR_6-1-2012.pdf`, 2012.

Frankel, J.A. and D. Romer, "Does Trade Cause Growth?," *American Economic Review*, 1999, *89* (3), 379–399.

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

Head, Keith and Thierry Mayer, "Gravity Equations: Workhorse, Toolkit, and Cookbook," Chapter 3 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, Elsevier Ltd., Oxford, 2014.

Heathcote, J.H. and F. Perri, "Financial Autarky and International Business Cycles," *Journal of Monetary Economics*, 2002, *49* (3), 601–627.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 2008, *123* (2), 441–487.

Hercowitz, Zvi and Michael Sampson, "Output Growth, the Real Wage, and Employment Fluctuations," *American Economic Review*, 1991, *81* (5), 1215–1237.

Krugman, P.R., "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 1980, *70* (5), 950–959.

Larch, Mario, Pehr-Johan Norbck, Steffen Sirries, and Dieter Urban, "Heterogeneous Firms, Globalization and the Distance Puzzle," *2016, The World Economy,* forthcoming, 2016.

Lin, Faqin and Nicholas C.S. Sim, "Death of Distance and the Distance Puzzle," *Economics Letters*, 2012, *116* (2), 225–228.

Lucas, Robert E. and Edward C. Prescott, "Investment Under Uncertainty," *Econometrica*, 1971, *39* (5), 659–681.

Mayer, T. and S. Zignago, "Notes on CEPII's Distances Measures: the GeoDist Database," *CEPII Working Paper No. 2011-25*, 2011.

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

— **and Gianmarco I. P. Ottaviano**, "Market Size, Trade, and Productivity," *Review of Economic Studies*, 2008, *75* (1), 295–316.

Mion, Giordano and Luca David Opromolla, "Managers' mobility, trade performance, and wages," *Journal of International Economics*, 2014, *94* (1), 85–101.

Nickell, Stephen J, "Biases in Dynamic Models with Fixed Effects," *Econometrica*, November 1981, *49* (6), 1417–26.

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

Olivero, Mara Pa and Yoto V. Yotov, "Dynamic Gravity: Endogenous Country Size and Asset Accumulation," *Canadian Journal of Economics*, 2012, *45* (1), 64–92.

Ramondo, Natalia, Andrs Rodrguez-Clare, and Milagro Saboro-Rodrguez, "Trade, Domestic Frictions, and Scale Effects," *American Economic Review*, 2016, *106* (10), 3159–3184.

Roodman, David, "How to do xtabond2: An introduction to difference and system GMM in Stata," *Stata Journal*, 2009, *9* (1), 86–136.

_ , "A note on the theme of too many instruments," *Oxford Bulletin of Economics and Statistics*, 2009, *71* (1), 135–158.

Ruhl, Kim J., "The International Elasticity Puzzle," unpublished manuscript, available for download at `http://www.kimjruhl.com/`, 2008.

Santos Silva, J.M.C. and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

_ **and** _ , "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator," *Economics Letters*, 2011, *112* (2), 220–222.

Timmer, M. P., B. Los, and G. J. de Vries, "An Anatomy of the Global Trade Slowdown based on the WIOD 2016 Release," *GGDC research memorandum number 162*, 2016.

_ **, E. Dietzenbacher, B. Los, R. R. Stehrer, and G. J. de Vries**, "An Illustrated User Guide to the World Input?Output Database: the Case of Global Automotive Production," *Review of International Economics*, 2015, *23*, 575?605.

Yotov, Yoto V., "A Simple Solution to the Distance Puzzle in International Trade," *Economics Letters*, 2012, *117* (3), 794–798.

_ **, Roberta Piermartini, Jos-Antonio Monteiro, and Mario Larch**, *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, Geneva: UNCTAD and WTO, 2016.

Zimmermann, C., "International Real Business Cycles among Heterogeneous Countries," *European Economic Review*, 1997, *41* (2), 319–356.

Table 1: SR Gravity with Efficiency Improvements, OLS Estimates

	(1) Standard	(2) Efficiency	(3) DistPuzzle	(4) PairFEs	(5) Efficiency2	(6) Development
LN_DIST	-1.184	-1.185	-1.183			
	$(0.049)^{**}$	$(0.049)^{**}$	$(0.057)^{**}$			
CNTG	-0.047	-0.048	-0.047			
	(0.189)	(0.189)	(0.189)			
CLNY	0.460	0.459	0.460			
	$(0.159)^{**}$	$(0.159)^{**}$	$(0.159)^{**}$			
LANG	0.782	0.782	0.781			
	$(0.107)^{**}$	$(0.107)^{**}$	$(0.107)^{**}$			
FTA	0.413	0.406	0.419	0.165	0.147	0.139
	$(0.062)^{**}$	$(0.062)^{**}$	$(0.065)^{**}$	$(0.041)^{**}$	$(0.041)^{**}$	$(0.041)^{**}$
LN_TARIFF	-1.723	-1.704	-1.719	-1.311	-1.283	-0.990
	$(0.443)^{**}$	$(0.442)^{**}$	$(0.445)^{**}$	$(0.308)^{**}$	$(0.306)^{**}$	$(0.260)^{**}$
INTL_BRDR	-3.235	-3.733	-3.737			
	$(0.299)^{**}$	$(0.324)^{**}$	$(0.336)^{**}$			
INTL_BRDR_1991		0.215	0.381		0.231	
		$(0.091)^{*}$	$(0.116)^{**}$		$(0.092)^{*}$	
INTL_BRDR_1994		0.462	0.551		0.511	
		$(0.098)^{**}$	$(0.158)^{**}$		$(0.098)^{**}$	
INTL_BRDR_1997		0.622	0.508		0.713	
		$(0.122)^{**}$	$(0.166)^{**}$		$(0.121)^{**}$	
INTL_BRDR_2000		0.685	0.664		0.812	
		$(0.132)^{**}$	$(0.174)^{**}$		$(0.131)^{**}$	
INTL_BRDR_2003		0.722	0.713		0.864	
		$(0.149)^{**}$	$(0.197)^{**}$		$(0.144)^{**}$	
INTL_BRDR_2006		0.799	0.680		0.965	
		$(0.160)^{**}$	$(0.209)^{**}$		$(0.158)^{**}$	
LN_DIST_1991			-0.051			
			$(0.030)^{+}$			
LN_DIST_1994			-0.027			
			(0.040)			
LN_DIST_1997			0.034			
			(0.037)			
LN_DIST_2000			0.006			
			(0.038)			
LN_DIST_2003			0.002			
			(0.041)			
LN_DIST_2006			0.035			
			(0.042)			
INTL_BRDR_HIGH_1991						0.314
						(0.237)
INTL_BRDR_HIGH_1994						0.589
						$(0.187)^{**}$
INTL_BRDR_HIGH_1997						0.800
						$(0.258)^{**}$
INTL_BRDR_HIGH_2000						1.009
						$(0.262)^{**}$
INTL_BRDR_HIGH_2003						1.065
						$(0.267)^{**}$
INTL_BRDR_HIGH_2006						1.287
						$(0.306)^{**}$
INTL_BRDR_LOW_1991						0.438
						$(0.177)^{*}$
INTL_BRDR_LOW_1994						0.939
						$(0.214)^{**}$
INTL_BRDR_LOW_1997						1.235
						$(0.226)^{**}$
INTL_BRDR_LOW_2000						1.237
						$(0.275)^{**}$
INTL_BRDR_LOW_2003						1.351
						$(0.322)^{**}$
INTL_BRDR_LOW_2006						1.520
						$(0.351)^{**}$
N	18345	18345	18345	18344	18344	18336
r2	0.880	0.881	0.881	0.935	0.936	0.955

Notes: This table reports results from the international-borders dummy variable approach to test SR gravity theory. All estimates are obtained with the OLS estimator and with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with time-varying international border dummies, designed to capture relative efficiency improvements. Column (3) allows for time-varying distance effects, in addition to the time-varying border dummies. Columns (4) and (5) reproduce the results from columns (1) and (2) but with pair fixed effects. Finally, column (6) uses the specification from column (5) but allows for heterogeneous efficiency improvements across developed vs. developing countries. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 2: SR Gravity with Efficiency Improvements, PPML Estimates

	(1) Standard	(2) Efficiency	(3) DistPuzzle	(4) PairFEs	(5) Efficiency2	(6) Development
LN_DIST	-0.648	-0.662	-0.640			
	(0.065)**	(0.065)**	(0.074)**			
CNTG	0.488	0.500	0.500			
	(0.148)**	(0.146)**	(0.145)**			
CLNY	-0.049	-0.064	-0.066			
	(0.109)	(0.108)	(0.108)			
LANG	0.325	0.343	0.345			
	(0.139)*	(0.138)*	(0.138)*			
FTA	0.142	0.073	0.073	0.318	0.112	0.214
	(0.123)	(0.118)	(0.117)	(0.077)**	(0.067)$^{+}$	(0.062)**
LN_TARIFF	-7.438	-6.803	-6.859	-4.842	-2.699	-2.122
	(1.167)**	(1.227)**	(1.238)**	(0.588)**	(0.547)**	(0.448)**
INTL_BRDR	-2.635	-3.032	-3.069			
	(0.154)**	(0.152)**	(0.164)**			
INTL_BRDR_1991		0.202	0.182		0.147	
		(0.023)**	(0.057)**		(0.016)**	
INTL_BRDR_1994		0.324	0.297		0.249	
		(0.028)**	(0.055)**		(0.024)**	
INTL_BRDR_1997		0.449	0.394		0.401	
		(0.044)**	(0.086)**		(0.035)**	
INTL_BRDR_2000		0.543	0.595		0.481	
		(0.055)**	(0.106)**		(0.041)**	
INTL_BRDR_2003		0.491	0.561		0.498	
		(0.062)**	(0.106)**		(0.043)**	
INTL_BRDR_2006		0.510	0.621		0.570	
		(0.065)**	(0.090)**		(0.040)**	
LN_DIST_1991			0.011			
			(0.023)			
LN_DIST_1994			0.013			
			(0.023)			
LN_DIST_1997			0.028			
			(0.031)			
LN_DIST_2000			-0.029			
			(0.036)			
LN_DIST_2003			-0.039			
			(0.038)			
LN_DIST_2006			-0.061			
			(0.036)$^{+}$			
INTL_BRDR_HIGH_1991						0.436
						(0.054)**
INTL_BRDR_HIGH_1994						0.750
						(0.078)**
INTL_BRDR_HIGH_1997						0.967
						(0.150)**
INTL_BRDR_HIGH_2000						0.988
						(0.227)**
INTL_BRDR_HIGH_2003						0.972
						(0.208)**
INTL_BRDR_HIGH_2006						1.118
						(0.205)**
INTL_BRDR_LOW_1991						0.512
						(0.084)**
INTL_BRDR_LOW_1994						1.162
						(0.123)**
INTL_BRDR_LOW_1997						1.368
						(0.132)**
INTL_BRDR_LOW_2000						1.520
						(0.172)**
INTL_BRDR_LOW_2003						1.900
						(0.187)**
INTL_BRDR_LOW_2006						2.062
						(0.193)**
N	18928	18928	18928	18928	18928	18928

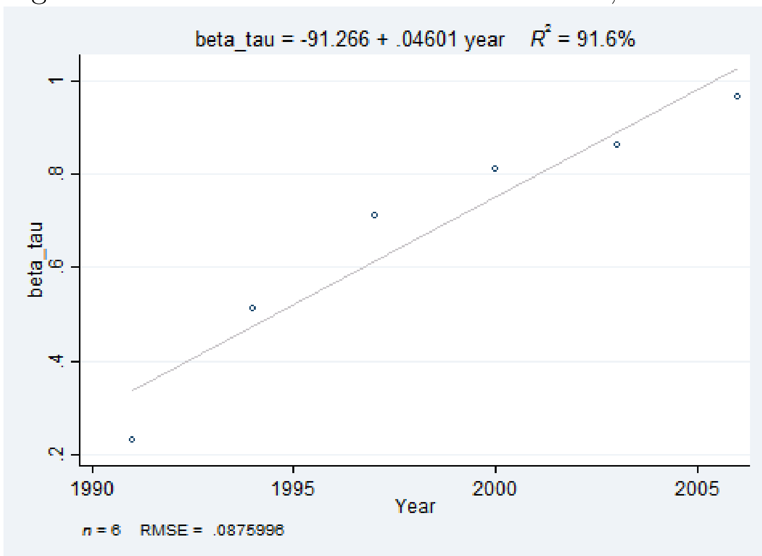
Notes: This table reports results from the international-borders dummy variable approach to test SR gravity theory. All estimates are obtained with the PPML estimator and with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with time-varying international border dummies, designed to capture relative efficiency improvements. Column (3) allows for time-varying distance effects, in addition to the time-varying border dummies. Columns (4) and (5) reproduce the results from columns (1) and (2) but with pair fixed effects. Finally, column (6) uses the specification from column (5) but allows for heterogeneous efficiency improvements across developed vs. developing countries. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 3: SR Gravity & Efficiency

	(1) Standard	(2) SR Gravity	(3) SR Test1	(4) LSDV	(5) SR LSDV	(6) SR Test2	(7) LSDV IV	(8) SR Test3
LN_DIST	-1.186	-0.248	0.003					
	(0.049)**	(0.016)**	(0.008					
CNTG	-0.036	0.005	0.358					
	(0.189)	(0.039)	(1.097)					
CLNY	0.457	0.092	0.112					
	(0.158)**	(0.034)**	(0.074)					
LANG	0.773	0.147	0.022					
	(0.105)**	(0.022)**	(0.028)					
FTA	0.418	0.108	- 0.046	0.170	0.070	- 0.024	0.029	0.081
	(0.062)**	(0.013)**	(0.032)	(0.037)**	(0.016)**	(0.097)	(0.017)$^{+}$	(0.130)
LN_TARIFF	-1.513	-0.348	- 0.018	-1.114	-0.446	0.014	-0.579	-0.265
	(0.378)**	(0.100)**	(0.066)	(0.259)**	(0.118)**	(0.105)	(0.149)**	(0.097)**
INTL_BRDR	-3.229	-0.635	0.015					
	(0.297)**	(0.064)**	(0.019)					
L.LN_TRADE		0.788			0.614		0.746	
		(0.009)**			(0.011)**		(0.124)**	
N	49916	46868		49916	46868		35411	
R^2	0.882	0.959		0.936	0.964		0.964	
ρ		0.202			0.371		0.242	
		(0.008)**			(0.011)**		(0.120)**	
Under Id χ^2							13.595	
p-val							(0.009)	
Weak Id χ^2							3.394	
p-val							(0.009)	
Over Id χ^2							1.350	
p-val							(0.717)	

Notes: This table reports results from the structural investment approach to test SR gravity theory. All estimates are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reports the results from a test of the SR gravity theory according to equation (27). Columns (4)-(6) reproduce the results from columns (1)-(3) but with pair fixed effects, thus implementing the LSDV estimator. Column (7) implements an IV LSDV estimation, and column (8) reports the results from test (27). The "weak identification" (Weak Id) statistics are based on Kleibergen-Paap Wald F values, which are appropriate when the standard error i.i.d. assumption is not met and the usual Cragg-Donald Wald statistic is no longer valid. (The corresponding Cragg-Donald Wald F statistic is 3.781.) Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Figure 1: Evolution of International Borders, 1988-2006.



Note: This figure plots the evolution of the international border estimates from column (5) of Table 1, which are obtained with the OLS estimator and pair fixed effects, over time. The figure also reports the slope of the fitted line, which corresponds to the parameter b_{τ} from the main text.

Appendix A Theoretical Extensions

A.1 Heterogeneous Productivities and Short Run Gravity

Short run gravity can equivalently represent a foundation based on heterogeneous productivities generated by a probability distribution (Eaton and Kortum, 2002). With productivity draws from a Frechet (Type II extreme value) distribution, it is well known that the same long run gravity equation system results with the trade cost elasticity $1 - \sigma$ reinterpreted as the dispersion parameter θ of the Frechet distribution and the taste parameter β_i reinterpreted as the location parameter of the productivity distribution for origin i. The CES demand system is used in the Eaton-Kortum model but σ ends up in a background constant term due to the infinite elasticity of transformation of the Ricardian model.

The simplicity of the Cobb-Douglas production model here allows short run long run gravity to be interpreted in terms of either foundation. For the Eaton-Kortum interpretation, each origin supplies potentially an infinite number of varieties to each destination. Each variety has the same basic Cobb-Douglas production function up to a productivity draw parameter divided by an iceberg trade factor. The proportion of draws favorable enough to permit shipments to destination j relative to shipments from all origins to destination j is a CES-type share function of unit cost of delivered output from origin i relative to an index of unit costs of delivered output to j from all origins. The key link between the two models is that the unit cost of delivered output in equilibrium in the short run model plays the same role in either model. Note also that the long run Cobb-Douglas model with efficient capacities becomes Ricardian, as in the Eaton-Kortum model.

Feenstra (2016) following Chaney (2008) shows that with constant iceberg bilateral trade costs and Melitz (2003) Pareto distributed productivities, a version of the structural gravity model emerges, with the trade cost elasticity having a different interpretation and fixed costs playing a role. The key step in reaching this conclusion is that the combination of Pareto distribution and iceberg bilateral costs leads to a log-linear closed form solution for the productivity of the least productive firm in any origin able to serve each destination from that origin. The Pareto distribution then yields a log-linear expression for the proportion of firms from each origin that serve each destination.

For *purely competitive* firms with Pareto distributed productivities, fixed costs of exporting to a destination, and the increasing variable bilateral trade costs modeled in this paper, Section A.1.1 below shows that the optimized (restricted) profit function has a constant elasticity with respect to productivity. This implies a closed form solution for the minimum productivity firm and thus the proportion of firms that can serve any market. In this case, the heterogeneous purely competitive firms model again yields a standard structural gravity equation. The trade cost elasticity is now a function of the elasticity of substitution σ, the Pareto distribution dispersion parameter θ and the elasticity of transformation $1/(1-\alpha)$.

The monopolistic competition case is examined in Section A.1.3. Unfortunately, the combination of heterogeneous productivities and selection of *monopolistically* competitive firms due to fixed costs does not lead to a tractable short run gravity model. There is generally no closed form solution for the productivity of the least productive firm able to serve a destination. This is because in monopolistic competition the elasticity of optimized (restricted) revenue with respect to productivity is smaller than the elasticity of optimized variable cost with respect to productivity. The implicit relationship between the proportion of firms serving a market on the one hand and market size and origin wage on the other hand retains the qualitative properties of the Melitz-Chaney model, but the implicit function substantially departs from log-linearity. Nevertheless, Section A.1.3 shows that the purely competitive firms model closely approximates the monopolistic competition model as σ is large. Numerical evaluation reveals the two models are close.

A.1.1 Short Run Gravity with Competitive Heterogeneous Firms

Each firm draws a Hicks-neutral productivity scalar $\varrho \geq 1$ from a Pareto distribution $G(\varrho) = 1 - \varrho^{-\theta}$. Index the firms in an (implicit) origin by their productivity draws ρ. Firms have (by assumption) previously committed capital $K(z)$ to each destination z in the same pattern based on expected demand because prior to receiving their productivities, all firms are identical. The profit of firm ϱ is

$$\varrho \frac{p(z)}{t(z)} L(z,\varrho)^{\alpha} K(z)^{1-\alpha} - wL(z,\varrho).$$

Profit maximization by a price taking firm implies the restricted profit function $\varrho^{1/(1-\alpha)}\bar{R}(z)$ where

$$\bar{R}(z) = \left(\frac{p(z)}{t(z)}\right)^{1/(1-\alpha)} K(z)[\alpha^{\alpha/(1-\alpha)} - w^{-\alpha/(1-\alpha)}\alpha^{1/(1-\alpha)}]$$

is the variable profit of the least productive firm (whether it ships to destination z or not).

The zero profit cutoff value of ϱ occurs at

$$\underline{\varrho}(z) = [f(z)/\bar{R}(z)]^{1-\alpha}$$

where $f(z)$ is the fixed cost of exporting to z. The proportion of firms with $\varrho \geq \underline{\varrho}$ is given by $1 - G(\varrho) = \int_{\underline{\varrho}}^{\infty} \theta \varrho^{-\theta-1} d\varrho$. Then the aggregate value shipped to destination z is given by

$$\left(\frac{p(z)}{t(z)}\right)^{1/(1-\alpha)} K(z)\alpha^{\alpha/(1-\alpha)} S(z) \tag{28}$$

where the mass of firms serving destination z

$$S(z) = \int_{\underline{\varrho}}^{\infty} \theta \varrho^{1/(1-\alpha)-\theta-1} d\varrho = \frac{\theta}{\theta + 1 - \eta} \left(\frac{f(z)}{\bar{R}(z)}\right)^{\eta-\theta} \tag{29}$$

using $\eta = 1/(1-\alpha)$. Relative to the text version of short run gravity, $S(z)$ enters everywhere that $K(z) = \lambda(z)K$ appears. $\bar{R}(z)$ is proportional to $[p(z)/t(z)]^{\eta}$. Using this expression in (29), $p(z)/t(z)$ affects the extensive margin of trade through $S(z)$ with elasticity $\eta(\theta - \eta)$. The intensive margin response to a fall in trade cost $t(z)$ acts to crowd out the extensive margin response in (29), completely so at $\theta = \eta$. The mass of active firms falls and average productivity rises when $\eta > \theta$, which is the famous Melitz (2003) case where effectively $\eta \to \infty$.

(29) combines in (28) with the effect of $p(z)/t(z)$ on the intensive margin of trade with elasticity η. The combined effect of $p(z)/t(z)$ on shipments has composite supply elasticity $\tilde{\eta} = \eta(1 + \theta - \eta)$. This composite supply side elasticity combines with the demand side elasticity in the short run equilibrium model of the text to yield a buyers' incidence elasticity equal to

$$\tilde{\rho} = \frac{\eta(1 + \theta - \eta)}{\eta(1 + \theta - \eta) + \sigma - 1} > \rho \tag{30}$$

for $\theta > \eta$, the intuitive case where the extensive margin $S(z)$ is increasing in price $p(z)$. In the special case $\theta = \eta$, the responses of firms on the intensive and extensive margins are perfect substitutes.

In empirical applications, the extensive margin component can be identified if proxies for fixed export costs can be found that are not also proxies for variable trade costs. In the absence of such proxies, estimated trade elasticities have to be considered combinations of intensive and extensive margin responses.

A.1.2 Empirical Implications

The development of the firm level model here points to a rationale for the simple econometric procedure adopted for the empirical work. First, the $\lambda(z)$ allocations are assumed to be the same for any firm, under the assumption that all heterogeneity is in ϱ draws. We can extend the logic to assume that all $z > 0$ external markets are ex post under-supplied to a uniform extent, $\lambda(z)$ is too small by a common fraction and correspondingly $\lambda(0)$ is too large. This permits drawing inferences from multiple observations on the movement of external vs. internal trade over time.

A second aspect of the model with empirical content is that the zeroes in bilateral trade flows can simply be rationalized as the absence, for whatever reason, of a finite positive λ.

A.1.3 Monopolsitic Competition Case

The production function for delivered output by monopolistic competitive firm ϱ to destination z from a representative origin is the same as in the competitive case, $y(z,\varrho) = \varrho L(z,\varrho)^{\alpha} K(z)^{1-\alpha}/t(z)$, where $K(z)$ is assumed to be committed before ϱ is drawn from the Pareto distribution with dispersion parameter θ. The firm faces willingness to pay for its shipments to destination z equal to

$$p(z,\varrho) = y(z,\varrho)^{-1/\sigma} E(z)^{1/\sigma} P(z)^{1-1/\sigma}.$$

Once ϱ is drawn, the firm hires labor (including the manager) to maximize profits (including the premium paid to the manager). The efficient level of labor is such that the wage w is equal to the marginal revenue product of labor:

$$L(z,\varrho) = \varrho^{(1-1/\sigma)/(1-\alpha(1-1/\sigma))} \left(\alpha(1-1/\sigma)E(z)^{1/\sigma}(P(z)/t(z)^{1-1/\sigma} K(z)^{(1-\alpha)(1-1/\sigma)}/w\right)^{1/(1-\alpha(1-1/\sigma))} \tag{31}$$

The wage bill of firm ϱ on shipments to z is w times the right hand side of (31). The value of sales at destination prices follows from substituting the right hand side of (31) into the production function for delivered output and the willingness to pay function for delivered output. The restricted profit function is equal to sales minus the wage bill. It is convenient for present purposes to define the sales $\underline{y}(z)$ and labor $\underline{L}(z)$ of the minimally productive firm with $\varrho = 1$, suppressing the unnecessary details. The restricted profit function of firm ϱ on sales to z is

$$R(z,\varrho) = \underline{y}(z)\varrho^{(1-1/\sigma)(1+\alpha)/(1-\alpha(1-1/\sigma))} - \underline{L}(z)\varrho^{(1-1/\sigma)/(1-\alpha(1-1/\sigma))}. \tag{32}$$

The least productive firm able to serve market z must be productive enough to pay fixed cost $f(z)$. The cutoff firm has productivity

$$\underline{\varrho}(z) = \min \varrho : R(z, \varrho) - f(z) = 0.$$

The definition imposes a plausible restriction that the monopolistic competitive firm will find the smallest root $\underline{\varrho}$ when there are multiple roots solving the equation.

Expression (32) does not yield a closed form solution for ϱ for a given profit level, so it will not yield a convenient structural gravity equation for the mass of firms serving market z from the origin in focus here. In contrast, perfectly competitive firms yield a closed form solution. The monopolisitic competition case comes close to the competitive case as σ grows large, hence $1 - 1/\sigma \to 1$ in the exponents of (32).

Quantitatively, the difference between the monopolistic competition case and the pure competition case can be numerically evaluated by comparing solutions for $\underline{\varrho}(z)$ given reasonable range values of σ with the competitive equivalent solution with $\sigma \to \infty$. Plugging the pair of solutions for into equation (29), the ratio of the mass of sales of monopolistic to competitive firms can be computed, given a value of θ. A representative evaluation case used $f/\underline{Y} = 0.5$, $w\underline{L}/\underline{y} = 0.6$, $\alpha = 0.67$, $\theta = 4$. At $\sigma = 2$, monopolistic firms total sales are 91% of competitive firms sales. At $\sigma = 10$, they are 99% of competitive firms sales. The ratio is monotonic (and equilibrium $\underline{\varrho}$ is unique), convex and 'nearly' linear in $\sigma \in (1, \infty)$.

A.2 Short Run Trade Cost Endogeneity

The short run gravity model implies a theory of short run endogenous trade costs. Equation (7) yields a structure that potentially connects inference about trade costs from price variation to inference from trade flow variation.

Relative price $T(i, z) \equiv p(i, z)/p(i, i)$ is common direct measure of trade cost variation. Use (7) and the result on the CES price index $P(z) = (E(z)/Y)^{-(1-\rho)} \widetilde{P}(z)$ from Section 3 where $\widetilde{P}(z)$ is the buyer's multilateral resistance. Simplify the exponents of the result to give:

$$T(i,z) = \left(\frac{t(i,z)}{t(i,i)}\right)^{\rho} \left(\frac{\lambda(i,z)}{\lambda(i,i)}\right)^{-\rho/\eta} \left(\frac{\widetilde{P}(z)}{\widetilde{P}(i)}\right)^{1-\rho} \left(\frac{E(z)}{E(i)}\right)^{1/\eta}. \tag{33}$$

$T(i,z) \to t(i,z)/t(i,i)$ as $\lambda(i,z)/\lambda(i,i)$ adjusts to its efficient level (see Section 3.2). $t(i,z)/t(i,i)$ is inferred from standard (i.e. long run) gravity equations, while (33) implies that spatial variation of $T(i,z)$ includes the effect of inefficient bilateral capacities and the spatial variation of demand conditions. Demand variation in the last two terms on the right is direct in $E(z)/E(i)$ and indirect through its general equilibrium effects on buyers' relative multilateral resistance $\widetilde{P}(z)/\widetilde{P}(i)$. (33) can potentially reconcile the time invariance of $t(i,z)$ inferred from gravity and the obvious time variation of direct price measures, and also differences in elasticities inferred from the spatial variation of prices and the spatial variation of trade flows. Future research on disaggregated price and trade flow data can test the consistency of (12) and (33).[39]

-

[39] The application of this paper to aggregate manufacturing data is not suitable for use of (33) because composition effects contaminate the spatial variation of manufacturing price indexes.

Appendix B Empirical Analysis

This appendix reports the results from a series of sensitivity experiments that demonstrate the robustness of our main results. Following the development of the exposition in the main text of the manuscript, we present successively robustness experiments for each the two econometric methods that we employ to test the SR gravity theory.

B.1 Sensitivity Experiments: Econometric Approach

We start this section with a brief summary of the main empirical findings from Section 4.2, where we tested the short run gravity theory with an estimation approach that employed international border dummies under the assumption that cross-border trade capacities $\lambda(i, z, \tau)$ are inefficiently small, while domestic capacities $\lambda(i, i, \tau)$ are inefficiently large and, over time, capacity investment evolves toward efficiency so that cross-border investment in trade links rises relative to domestic investment in trade links. The two main findings from the estimates in Section 4.2 were that (i) the estimates of the standard gravity variables remained statistically unchanged once we introduced the bilateral border dummies; and, more importantly for our purposes, (ii) we obtained estimates on the bilateral border dummies that were all positive, statistically significant, and gradually increasing over time, which, consistent with our theory we interpreted as improvement in efficiency. The following experiments confirm the robustness of our main findings from Section 4.2:

- The estimates in Table 4 are obtained with size-adjusted trade, $X_{i,z,\tau}/Y_{i,\tau}$, as dependent variable. The motivation for using size-adjusted trade is twofold. First, such specification is consistent with theory. Second, using size-adjusted trade as dependent variable mitigates heteroskedasticity concerns by construction, cf. Anderson and van Wincoop (2003). The estimates from columns (1)-(4) of Table 4 are obtained with the OLS estimator. Columns (1) and (2) reproduce the estimates with standard gravity variables from columns (1) and (2) of Table 1. Columns (3) and (4) reproduce the estimates with pair fixed effects from columns (4) and (5) of Table 1. Columns (5)-(8) reproduce the estimates from columns (1)-(4) with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Both the OLS and the PPML estimates from Table 4 confirm the two main findings as summarized at the beginning of this section.

- Figure 2 reports estimates of the evolution of the international border estimates from column (2) of Table 1, which are obtained with OLS and standard gravity variables, over time. The figure also obtains the slope of the fitted line, which corresponds to the parameter b_τ from the main text. The estimate of $b_\tau = 0.036$ from Figure 2 is smaller but comparable to the estimate of $b_\tau = 0.046$ from the main text. The results in Figure 3 are constructed based on the estimates from column (5) of Table 2, which are obtained with the PPML estimator and pair fixed effects, over time. Once again, we obtain a similar value for the parameter $b_\tau = 0.028$.

- Table 5 reproduces the main results from Table 1 with data for all years. The estimates from columns (1)-(4) are obtained with the OLS estimator. Columns (1) and (2) reproduce the estimates with standard gravity variables from columns (1) and (2) of Table 1. Columns (3) and (4) reproduce the estimates with pair fixed effects from columns (4) and (5) of Table 1. Columns (5)-(8) of this table reproduce the estimates from columns (1)-(4) of the same table with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. In addition, also for brevity, we have omitted the estimates on the bilateral border dummies for the odd years in the sample. These are available by request. The estimates from Table 5, once again, confirm our main findings as summarized at the beginning of this section.

- Table 6 reproduces the results from column (3) of Table 1 with data for all years. We remind the reader that the estimates from column (3) of Table 1 depicted and resolved the 'distance puzzle' in international trade. The estimates in column (1) of Table 6 are obtained with the OLS estimator and the estimates in column (2) of Table 6 are obtained with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. In addition, also for brevity, we have omitted the estimates on the time-varying distance variables and the estimates on the bilateral border dummies for the odd years in the sample for each group of countries. The omitted estimates are available by request. In each case we allow for time-varying distance effects and, in each case, we find no evidence for time-varying effects of distance. This is consistent with the distance puzzle in international trade. However, similar to the discussion of our main results, our theory suggests that the correct interpretation of the effects of distance should combine the joint impact that is captured by the distance estimates and by the estimates of the efficiency improvements. Thus, combined with the decreasing border effects (captured by the positive, significant and increasing estimates on the border dummies), our theory implies that the impact of distance have steadily decreased over time. In sum, the estimates from Table 6 confirm that our theory can indeed resolve the famous 'distance' and 'missing globalization' puzzles in the international trade literature.

- Table 7 reproduces the results from column (6) of Table 1 with data for all years. We remind the reader that the estimates from column (6) of Table 1 distinguished between capacity improvement for the developed vs. developing countries in our sample. The estimates in column (1) of Table 7 are obtained with the OLS estimator and the estimates in column (2) of Table 7 are obtained with the PPML estimator. In each case we allow for different efficiency improvement rates depending on whether a country is developed or developing. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. In addition, also for brevity, we have omitted the estimates on the bilateral border dummies for the odd years in the sample for each group of countries. The omitted estimates are available by request.

The estimates from columns (1) and (2) of Table 7 deliver results that are very similar to our main corresponding findings from Table 1. Specifically, and most important, we do observe efficiency improvements for both groups of countries. In addition, we see that the low income countries have actually converged faster toward more efficient trade during the period of investigation.

- Table 8 reproduces the main results from Table 1 with with the latest edition of the WIOD data, as described in the data section. The WIOD dataset covers the period 2000-2014 and we use 2-year intervals to obtain the estimates in Table 8. The estimates from columns (1)-(4) are obtained with the OLS estimator. Columns (1) and (2) reproduce the estimates with standard gravity variables from columns (1) and (2) of Table 1. Columns (3) and (4) of Table 8 reproduce the estimates with pair fixed effects from columns (4) and (5) of Table 1. Columns (5)-(8) of Table 8 reproduce the estimates from columns (1)-(4) of the same table with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The estimates from Table 8 confirm our main findings, as summarized at the beginning of this section, with the WIOD data.

- Table 9 reproduces the main results from Table 1 by capitalizing on the sectoral dimension of the latest edition of the WIOD data, as described in the data section. Once again, we use 2-year interval data. For brevity, for each of the six main sectors in the WIOD database (including Crop and Animal Production, Forestry and Logging, Fishing and Aquaculture, Mining and Quarrying, Manufacturing, and Services), we only reproduce the main specifications with pair fixed effects, which correspond to those in columns (4) and (5) of Table 1. All results are obtained with the OLS estimator and with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The main finding from the results in Table 9 is that the estimates of the time-varying bilateral border dummies are positive, statistically significant, and increasing over time for each sector in our sample. These results are consistent with our main findings and support the SR gravity theory. In addition, we note that while there is certain heterogeneity in the bilateral border estimates across sectors, the efficiency improvements have been significant at the sectoral level and similar in magnitude.

In sum, based on the sensitivity experiments that we employed in this section, we conclude that our main findings from Section 4.2 represent robust results.

B.2 Sensitivity Experiments: Structural Approach

We start this section with a brief summary of the main empirical findings from Section 4.3, where we tested the short run gravity theory by developing a formal model of investment, which resulted in the addition of the lagged dependent variable as a covariate in the estimating equation for SR gravity. Three main findings stood out from the estimates in Section 4.3. First, we obtained a large, positive and statistically significant estimate of the coefficient on the structural efficiency term. Second, we recovered estimates of the structural efficiency parameter ranging from 0.202 to 0.371. Third, as predicted by the SR gravity theory, the estimates that controlled for efficiency improvements were significantly smaller in magnitude as compared to their long run counterparts. The following experiments confirm the robustness of our main findings from Section 4.2:

- Table 10 uses 3-year interval data to reproduce the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Column (4) implements an IV LSDV estimation. Finally, column (5) implements the Arellano-Bond estimator. In sum, the estimates from Table 10 are in support of all main findings listed at the beginning of this section. Specifically: (i) We obtain positive and significant estimates of the structural efficiency term; (ii) Taking into account the use of 3-year interval data, which implies $\tilde{\beta}_8 = \frac{(1-\delta)^3(1-\rho)}{1-\delta(1-\rho)}$, we recover $\hat{\rho} \in [0.250, 0.534]$ across the different specifications, which can be found in the bottom panel of Table 10; (iii) The estimates from each specification that controls for efficiency improvements are significantly smaller in magnitude as compared to their long run counterparts from column (1) of Table 10. Finally, we note that the bounds for ρ that we obtain with sample that includes all years are tighter as compared to the bounds with the 3-year interval data. This is evidence in support of using the full sample over using data with intervals.

- Table 11 uses size-adjusted trade as dependent variable and reproduces the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Column (4) implements an IV LSDV estimation. The estimates from Table 11 confirm our main results as summarized at the beginning of this section.

- Table 12 reproduces the results from Table 3 of the main text with 3-year interval data on international trade only. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Column (4) implements an IV LSDV estimation. Once again, the estimates from Table 12 confirm our main results as summarized at the beginning of this section.

- Table 13 uses the system-GMM estimator, which will account for the dynamic features of our model.[40] The estimates from column (1) of Table 13 are obtained with data for all years and pass the Arellano-Bond test for autocorrelation in the disturbances by rejecting (as expected) the null hypothesis of no AR(1) errors with $z = -10.268$, but passing the test for second-order serial correlation AR(2) with $z = 1.302$. However, the instruments do not pass the Sargan test of overidentifcation restrictions with $\chi^2_{15} = 168.99$ (p-value=0.025). Furthermore, we obtain a suspiciously small estimate of the coefficient on $LN\_X_{ij,\tau-1}$ ($\tilde{\beta}_8 = 0.254$, std.err. 0.045), which is clearly outside of the bounds that we established in columns (2) and (4) of Table 3 from th main text. The Arellano-Bond estimates from column (2) of Table 13 are obtained with 3-year interval data and they improve on the estimates from column (1) by delivering an estimate on $LN\_X_{ij,\tau-1}$ ($\tilde{\beta}_8 = 0.603$, std.err. 0.161), from which we recover $\hat{\rho} = 0.303$ (std.err. 0.178), which is not statistically different from our main estimates.

- Table 15 uses the latest edition of the WIOD data and reproduces the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade and we use 2-year interval data. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Columns (3) and

[40] Anderson and Hsiao (1982) are the first to achieve consistency in short time period setting by using appropriate lagged levels and differences of the dependent variable as instruments for the lagged dependent variable. Arellano and Bond (1991) and Arellano and Bover (1995)/Blundell and Bond (1998) extend the Anderson-Hsiao estimator to the difference-GMM and the system-GMM estimators, respectively, which use larger sets of orthogonality conditions in order to obtain consistent estimates in dynamic panels with lagged dependent variable such as ours. We refer the interested reader to Roodman (2009a) for a detailed discussion and step-by-step implementation of alternative dynamic estimators in Stata. Roodman (2009b) An additional advantage of this estimation method for our analysis is that it will enable us to obtain estimates of the coefficient of all gravity variables, which will be used to perform test (27) multiple times. discusses problems associated with the proliferation of instruments in similar settings.

(4) reproduce the results from columns (1) and (2) but with pair fixed effects, thus implementing the LSDV estimator. The estimates with the WIOD data are consistent with the main results and they deliver $\hat{\rho} \in [0.070, 0.598]$. Both the upper and the lower bounds for ρ are wider as compared to the main estimates, but they are within the theoretical limits. We offer two possible explanations for the wider bounds. First, the main sample has longer time coverage, which, according to Roodman (2009a) leads to more precise estimates. Second, similar to the estimates with the 3-year interval sample of the main data, the WIOD estimates from Table 15 are obtained with 2-year interval data.

- Table 16 reproduces the main results from Table 3 by capitalizing on the sectoral dimension of the latest edition of the WIOD data, as described in the data section. For brevity, for each sector we only focus on reproducing the main specifications (i) with lagged trade and standard gravity variables and (ii) with lagged trade and pair fixed effects. All results are obtained with 2-year interval data, with the OLS estimator, and with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The sectoral SR gravity estimates are relatively homogeneous and they confirm our main findings.

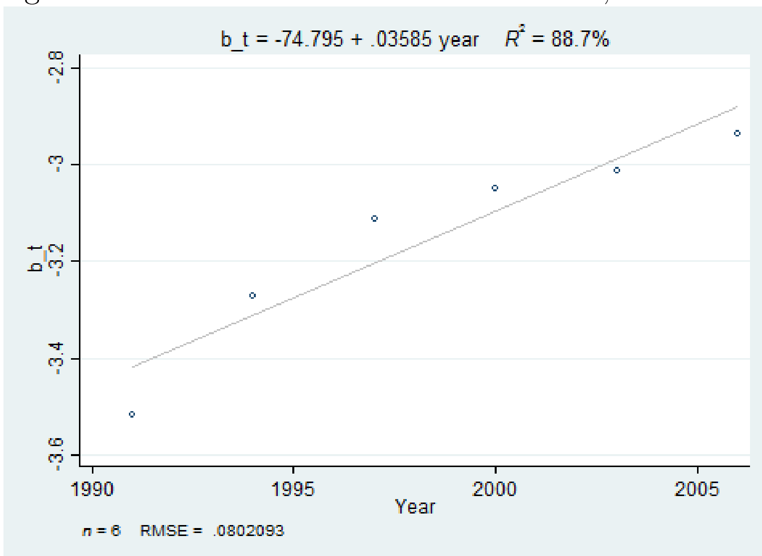
In sum, based on the sensitivity experiments that we employed in this section, we conclude that our main findings from Section 4.3 represent robust results.

Table 4: SR Gravity & Efficiency, Size-adjusted Trade

	OLS Estimator				PPML Estimator			
	Standard	Efficiency	PairFEs	Efficiency2	Standard	Efficiency	PairFEs	Efficiency2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LN_DIST	-1.184	-1.185			-0.711	-0.723		
	(0.049)**	(0.049)**			(0.066)**	(0.067)**		
CNTG	-0.047	-0.048			0.559	0.554		
	(0.189)	(0.189)			(0.150)**	(0.151)**		
CLNY	0.460	0.459			0.060	0.055		
	(0.159)**	(0.159)**			(0.122)	(0.121)		
LANG	0.782	0.782			0.537	0.541		
	(0.107)**	(0.107)**			(0.117)**	(0.117)**		
FTA	0.413	0.406	0.165	0.147	0.122	0.043	0.321	0.108
	(0.062)**	(0.062)**	(0.041)**	(0.041)**	(0.064)$^{+}$	(0.070)	(0.066)**	(0.063)$^{+}$
LN_TARIFF	-1.723	-1.704	-1.311	-1.283	-8.345	-7.981	-2.545	-1.420
	(0.443)**	(0.442)**	(0.308)**	(0.306)**	(0.757)**	(0.794)**	(0.398)**	(0.326)**
INTL_BRDR	-3.235	-3.733			-3.178	-3.501		
	(0.299)**	(0.324)**			(0.179)**	(0.174)**		
INTL_BRDR_1991		0.215		0.231		0.168		0.148
		(0.091)*		(0.092)*		(0.027)**		(0.018)**
INTL_BRDR_1994		0.462		0.511		0.358		0.316
		(0.098)**		(0.098)**		(0.028)**		(0.025)**
INTL_BRDR_1997		0.622		0.713		0.412		0.450
		(0.122)**		(0.121)**		(0.048)**		(0.032)**
INTL_BRDR_2000		0.685		0.812		0.472		0.528
		(0.132)**		(0.131)**		(0.060)**		(0.039)**
INTL_BRDR_2003		0.722		0.864		0.511		0.597
		(0.149)**		(0.144)**		(0.064)**		(0.041)**
INTL_BRDR_2006		0.799		0.965		0.440		0.663
		(0.160)**		(0.158)**		(0.070)**		(0.039)**
N	18345	18345	18344	18344	18928	18928	18928	18928
R^2	0.801	0.801	0.892	0.893	0.974	0.975		

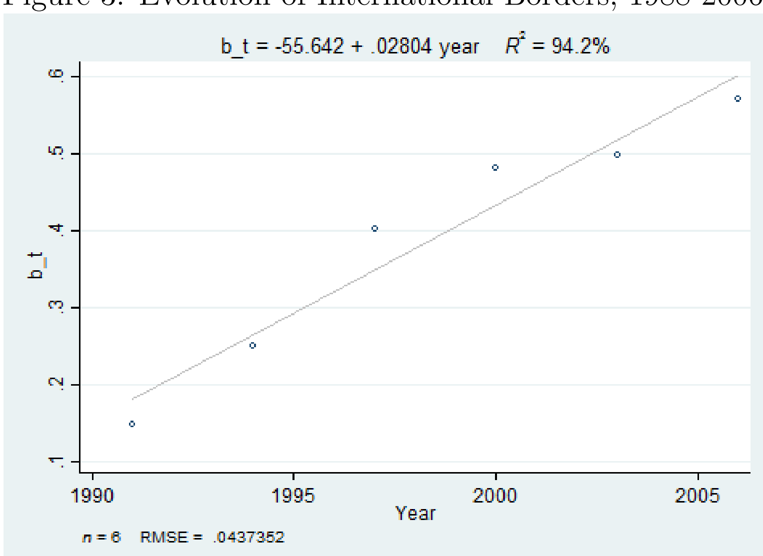
Notes: This table reproduces the main results from Table 1 with size-adjusted trade $X_{i,z,\tau}/Y_{i,\tau}$ as dependent variable. The estimates from columns (1)-(4) are obtained with the OLS estimator. Columns (1) and (2) reproduce the estimates with standard gravity variables from columns (1) and (2) of Table 1. Columns (3) and (4) reproduce the estimates with pair fixed effects from columns (4) and (5) of Table 1. Columns (5)-(8) of this table reproduce the estimates from columns (1)-(4) of the same table with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Figure 2: Evolution of International Borders, 1988-2006.



Note: This figure plots the evolution of the international border estimates from column (2) of Table 1, which are obtained with OLS and standard gravity variables, over time. The figure also reports the slope of the fitted line, which corresponds to the parameter b_T from the main text.

Figure 3: Evolution of International Borders, 1988-2006.



Note: This figure plots the evolution of the international border estimates from column (5) of Table 2, which are obtained with the PPML estimator and pair fixed effects, over time. The figure also reports the slope of the fitted line, which corresponds to the parameter b_T from the main text.

Table 5: SR Gravity & Efficiency, All Years: 1988-2006

	OLS Estimator				PPML Estimator			
	Standard	Efficiency	PairFEs	Efficiency2	Standard	Efficiency	PairFEs	Efficiency2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LN_DIST	-1.186	-1.187			-0.643	-0.656		
	(0.049)**	(0.049)**			(0.065)**	(0.065)**		
CNTG	-0.036	-0.036			0.492	0.500		
	(0.189)	(0.189)			(0.145)**	(0.143)**		
CLNY	0.457	0.457			-0.056	-0.070		
	(0.158)**	(0.158)**			(0.110)	(0.110)		
LANG	0.773	0.773			0.346	0.362		
	(0.105)**	(0.106)**			(0.137)*	(0.136)**		
FTA	0.418	0.412	0.170	0.155	0.127	0.072	0.282	0.102
	(0.062)**	(0.062)**	(0.037)**	(0.036)**	(0.117)	(0.113)	(0.073)**	(0.065)
LN_TARIFF	-1.513	-1.497	-1.114	-1.091	-7.444	-6.894	-4.391	-2.383
	(0.378)**	(0.377)**	(0.259)**	(0.257)**	(1.092)**	(1.157)**	(0.631)**	(0.539)**
INTL_BRDR	-3.229	-3.746			-2.652	-3.043		
	(0.297)**	(0.323)**			(0.154)**	(0.153)**		
INTL_BRDR_1990		0.183		0.195		0.180		0.137
		(0.061)**		(0.060)**		(0.022)**		(0.016)**
INTL_BRDR_1992		0.343		0.381		0.226		0.170
		(0.088)**		(0.091)**		(0.023)**		(0.018)**
INTL_BRDR_1994		0.463		0.515		0.324		0.250
		(0.098)**		(0.098)**		(0.028)**		(0.024)**
INTL_BRDR_1996		0.548		0.633		0.382		0.326
		(0.115)**		(0.116)**		(0.039)**		(0.032)**
INTL_BRDR_1998		0.657		0.769		0.479		0.459
		(0.126)**		(0.122)**		(0.047)**		(0.037)**
INTL_BRDR_2000		0.689		0.816		0.543		0.490
		(0.132)**		(0.130)**		(0.055)**		(0.042)**
INTL_BRDR_2002		0.653		0.789		0.455		0.459
		(0.146)**		(0.139)**		(0.057)**		(0.041)**
INTL_BRDR_2004		0.723		0.883		0.486		0.538
		(0.152)**		(0.148)**		(0.062)**		(0.042)**
INTL_BRDR_2006		0.807		0.975		0.509		0.585
		(0.160)**		(0.157)**		(0.066)**		(0.041)**
N	49916	49916	49916	49916	51376	51376	51376	51376
R^2	0.882	0.882	0.936	0.936				

Notes: This table reproduces the main results from Table 1 with data for all years. The estimates from columns (1)-(4) are obtained with the OLS estimator. Columns (1) and (2) reproduce the estimates with standard gravity variables from columns (1) and (2) of Table 1. Columns (3) and (4) reproduce the estimates with pair fixed effects from columns (4) and (5) of Table 1. Columns (5)-(8) of this table reproduce the estimates from columns (1)-(4) of the same table with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. In addition, also for brevity, we have omitted the estimates on the bilateral border dummies for the odd years in the sample. The omitted estimates are available by request. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 6: SR Gravity & Efficiency, All Years: 1988-2006

	(1) PuzzleOLS	(2) PuzzlePPML
LN_DIST	-1.184	-0.634
	(0.057)**	(0.074)**
LN_DIST_1990	-0.005	0.016
	(0.027)	(0.021)
LN_DIST_1992	-0.061	0.000
	(0.031)*	(0.023)
LN_DIST_1994	-0.027	0.012
	(0.040)	(0.023)
LN_DIST_1996	0.032	0.029
	(0.039)	(0.028)
LN_DIST_1998	0.008	-0.037
	(0.039)	(0.032)
LN_DIST_2000	0.006	-0.032
	(0.038)	(0.037)
LN_DIST_2002	0.013	-0.041
	(0.041)	(0.034)
LN_DIST_2004	0.004	-0.051
	(0.041)	(0.034)
LN_DIST_2006	0.033	-0.063
	(0.042)	(0.036)+
CNTG	-0.035	0.501
	(0.189)	(0.142)**
CLNY	0.458	-0.072
	(0.158)**	(0.110)
LANG	0.772	0.365
	(0.105)**	(0.136)**
FTA	0.425	0.072
	(0.065)**	(0.113)
LN_TARIFF	-1.509	-6.953
	(0.379)**	(1.167)**
INTL_BRDR	-3.754	-3.081
	(0.336)**	(0.164)**
INTL_BRDR_1990	0.200	0.152
	(0.097)*	(0.055)**
INTL_BRDR_1992	0.543	0.225
	(0.129)**	(0.056)**
INTL_BRDR_1994	0.549	0.300
	(0.159)**	(0.055)**
INTL_BRDR_1996	0.443	0.325
	(0.167)**	(0.076)**
INTL_BRDR_1998	0.628	0.546
	(0.170)**	(0.091)**
INTL_BRDR_2000	0.668	0.599
	(0.174)**	(0.108)**
INTL_BRDR_2002	0.608	0.528
	(0.192)**	(0.094)**
INTL_BRDR_2004	0.707	0.579
	(0.201)**	(0.089)**
INTL_BRDR_2006	0.693	0.624
	(0.209)**	(0.091)**
N	49916	51376
R^2	0.882	

Notes: This table reproduces the results from column (3) of Table 1, where we use our theory to resolve the 'distance puzzle' of trade, with data for all years. The estimates in column (1) of this table are obtained with the OLS estimator and the estimates in column (2) are obtained with the PPML estimator. In each case we allow for time-varying distance effects. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. In addition, also for brevity, we have omitted the estimates on the time-varying distance variables and the estimates on the bilateral border dummies for the odd years in the sample for each group of countries. The omitted estimates are available by request. Standard errors are clustered by country pair and are reported in parentheses $^+$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 7: SR Gravity & Efficiency, All Years: 1988-2006

	(1) DvlpmntOLS	(2) DvlpmntPPML
FTA	0.148	0.171
	(0.036)**	(0.049)**
LN_TARIFF	-0.870	-1.826
	(0.227)**	(0.435)**
INTL_BRDR_HIGH_1990	0.329	0.341
	(0.109)**	(0.041)**
INTL_BRDR_HIGH_1992	0.572	0.636
	(0.163)**	(0.085)**
INTL_BRDR_HIGH_1994	0.603	0.774
	(0.185)**	(0.075)**
INTL_BRDR_HIGH_1996	0.759	0.882
	(0.253)**	(0.147)**
INTL_BRDR_HIGH_1998	0.998	1.068
	(0.250)**	(0.184)**
INTL_BRDR_HIGH_2000	1.023	1.019
	(0.258)**	(0.228)**
INTL_BRDR_HIGH_2002	0.904	0.937
	(0.283)**	(0.197)**
INTL_BRDR_HIGH_2004	1.023	1.044
	(0.289)**	(0.187)**
INTL_BRDR_HIGH_2006	1.300	1.168
	(0.304)**	(0.207)**
INTL_BRDR_LOW_1990	0.252	0.339
	(0.160)	(0.067)**
INTL_BRDR_LOW_1992	0.636	0.714
	(0.215)**	(0.098)**
INTL_BRDR_LOW_1994	0.946	1.166
	(0.212)**	(0.123)**
INTL_BRDR_LOW_1996	1.116	1.144
	(0.194)**	(0.132)**
INTL_BRDR_LOW_1998	1.132	1.488
	(0.257)**	(0.160)**
INTL_BRDR_LOW_2000	1.232	1.565
	(0.271)**	(0.172)**
INTL_BRDR_LOW_2002	1.323	1.710
	(0.296)**	(0.193)**
INTL_BRDR_LOW_2004	1.478	2.079
	(0.327)**	(0.189)**
INTL_BRDR_LOW_2006	1.528	2.133
	(0.347)**	(0.193)**
N	49914	51376
R^2	0.956	

Notes: This table reproduces the results from column (6) of Table 1, which distinguished between capacity improvement for the developed vs. developing countries in our sample, with data for all years. The estimates in column (1) of Table 7 are obtained with the OLS estimator and the estimates in column (2) of Table 7 are obtained with the PPML estimator. In each case we allow for different efficiency improvements depending on whether a country is developed or developing. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. In addition, also for brevity, we have omitted the estimates on the bilateral border dummies for the odd years in the sample for each group of countries. The omitted estimates are available by request. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 8: SR Gravity & Efficiency, Aggregate WIOD Data

	OLS Estimator				PPML Estimator			
	Standard	Efficiency	PairFEs	Efficiency2	Standard	Efficiency	PairFEs	Efficiency2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LN_DIST	-1.267	-1.268			-0.845	-0.846		
	(0.055)**	(0.055)**			(0.042)**	(0.042)**		
CNTG	0.343	0.342			0.317	0.317		
	(0.112)**	(0.112)**			(0.106)**	(0.106)**		
CLNY	0.524	0.524			0.244	0.247		
	(0.134)**	(0.134)**			(0.103)*	(0.103)*		
LANG	-0.034	-0.034			0.393	0.395		
	(0.122)	(0.122)			(0.104)**	(0.105)**		
RTA	0.406	0.402	0.172	0.156	0.180	0.180	0.121	0.023
	(0.081)**	(0.081)**	(0.062)**	(0.062)*	(0.065)**	(0.065)**	(0.043)**	(0.045)
INTL_BRDR	-3.553	-3.909			-3.568	-3.621		
	(0.247)**	(0.288)**			(0.107)**	(0.107)**		
INTL_BRDR_2002		0.075		0.081		-0.053		-0.033
		(0.038)*		(0.038)*		(0.011)**		(0.009)**
INTL_BRDR_2004		0.288		0.309		0.008		0.046
		(0.075)**		(0.077)**		(0.018)		(0.016)**
INTL_BRDR_2006		0.410		0.433		0.046		0.104
		(0.095)**		(0.097)**		(0.020)*		(0.018)**
INTL_BRDR_2008		0.503		0.526		0.065		0.147
		(0.103)**		(0.105)**		(0.024)**		(0.020)**
INTL_BRDR_2010		0.487		0.511		0.044		0.130
		(0.107)**		(0.110)**		(0.025)+		(0.022)**
INTL_BRDR_2012		0.577		0.609		0.096		0.185
		(0.120)**		(0.123)**		(0.024)**		(0.021)**
INTL_BRDR_2014		0.548		0.581		0.097		0.189
		(0.126)**		(0.127)**		(0.026)**		(0.021)**
N	14792	14792	14792	14792	14792	14792	14792	14792
R^2	0.890	0.891	0.978	0.978	1.000	1.000		

Notes: This table reproduces the main results from Table 1 with the latest edition of the WIOD data, as described in the data section. The estimates from columns (1)-(4) are obtained with the OLS estimator. Columns (1) and (2) reproduce the estimates with standard gravity variables from columns (1) and (2) of Table 1. Columns (3) and (4) reproduce the estimates with pair fixed effects from columns (4) and (5) of Table 1. Columns (5)-(8) of this table reproduce the estimates from columns (1)-(4) of the same table with the PPML estimator. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 9: SR Gravity & Efficiency, Sectoral WIOD Data

	CropAnimal		ForestryLogging		FishingAquaculture		MiningQuarrying		Manufacturing		Services	
	PairFEs	SR_PairFEs	PairFEs	SR_PairFEs	PairFEs	SR_PairFEs	PairFEs	SR_PairFEs	PairFEs	SR_PairFEs	PairFEs	SR_PairFEs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
RTA	0.130	0.112	0.202	0.186	0.301	0.283	0.136	0.116	0.159	0.141	0.181	0.164
	$(0.072)^{+}$	(0.071)	$(0.078)^{**}$	$(0.077)^{*}$	$(0.085)^{**}$	$(0.084)^{**}$	$(0.079)^{+}$	(0.077)	$(0.059)^{**}$	$(0.058)^{*}$	$(0.067)^{**}$	$(0.066)^{*}$
INTL_BRDR_2002		0.106		0.083		0.091		0.155		0.098		0.087
		$(0.046)^{*}$		$(0.048)^{+}$		$(0.044)^{*}$		$(0.047)^{**}$		$(0.040)^{*}$		$(0.038)^{*}$
INTL_BRDR_2004		0.309		0.331		0.293		0.400		0.322		0.324
		$(0.079)^{**}$		$(0.087)^{**}$		$(0.089)^{**}$		$(0.086)^{**}$		$(0.075)^{**}$		$(0.082)^{**}$
INTL_BRDR_2006		0.445		0.480		0.458		0.568		0.459		0.443
		$(0.098)^{**}$		$(0.115)^{**}$		$(0.113)^{**}$		$(0.107)^{**}$		$(0.098)^{**}$		$(0.101)^{**}$
INTL_BRDR_2008		0.562		0.544		0.586		0.671		0.550		0.552
		$(0.106)^{**}$		$(0.118)^{**}$		$(0.131)^{**}$		$(0.117)^{**}$		$(0.108)^{**}$		$(0.108)^{**}$
INTL_BRDR_2010		0.579		0.525		0.535		0.641		0.563		0.526
		$(0.123)^{**}$		$(0.145)^{**}$		$(0.140)^{**}$		$(0.120)^{**}$		$(0.113)^{**}$		$(0.112)^{**}$
INTL_BRDR_2012		0.680		0.613		0.605		0.746		0.651		0.618
		$(0.132)^{**}$		$(0.165)^{**}$		$(0.153)^{**}$		$(0.136)^{**}$		$(0.127)^{**}$		$(0.124)^{**}$
INTL_BRDR_2014		0.643		0.566		0.635		0.707		0.638		0.594
		$(0.135)^{**}$		$(0.181)^{**}$		$(0.149)^{**}$		$(0.141)^{**}$		$(0.133)^{**}$		$(0.127)^{**}$
N	14792	14792	14104	14104	14104	14104	14792	14792	14792	14792	14792	14792
R^2	0.970	0.970	0.969	0.969	0.966	0.966	0.972	0.972	0.979	0.979	0.976	0.976

Notes: This table reproduces the main results from Table 1 by capitalizing on the sectoral dimension of the latest edition of the WIOD data, as described in the data section. For brevity, for each sector we focus on reproducing the main specifications with pair fixed effects, which are reported in columns (4) and (5) of Table 1. All results are obtained with the OLS estimator and with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Standard errors are clustered by country pair and are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 10: SR Gravity & Efficiency, 3-year Intervals

	(1) Standard	(2) LagTrade	(3) LSDV	(4) LSDV_IV
LN_DIST	-1.184	-0.404		
	(0.049)**	(0.026)**		
CNTG	-0.047	0.019		
	(0.189)	(0.066)		
CLNY	0.460	0.146		
	(0.159)**	(0.058)*		
LANG	0.782	0.238		
	(0.107)**	(0.036)**		
FTA	0.413	0.200	0.111	0.077
	(0.062)**	(0.025)**	(0.033)**	(0.054)
LN_TARIFF	-1.723	-0.802	-0.868	-1.195
	(0.443)**	(0.286)**	(0.260)**	(0.433)**
INTL_BRDR	-3.235	-0.992		
	(0.299)**	(0.105)**		
L.ln_trade		0.651	0.397	0.630
		(0.014)**	(0.017)**	(0.276)*
N	18345	15551	15549	10122
R^2	0.880	0.940	0.953	0.955
$\hat{\rho}$		0.250	0.534	0.273
		(0.015)**	(0.020)**	(0.304)
UnderId χ^2				7.092
p-val				(0.069)
Weak Id χ^2				2.345
p-val				(0.071)
Over Id χ^2				4.056
p-val				(0.131)

Notes: This table uses 3-year interval data to reproduce the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Columns (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Column (4) implements an IV LSDV estimation. Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 11: SR Gravity & Efficiency, Size-adjusted Trade

	(1)	(2)	(3)	(4)
	Standard	LagTrade	LSDV	LSDV_IV
LN_DIST	-1.186	-0.248		
	(0.049)**	(0.016)**		
CNTG	-0.036	0.005		
	(0.189)	(0.039)		
CLNY	0.457	0.092		
	(0.158)**	(0.034)**		
LANG	0.773	0.147		
	(0.105)**	(0.022)**		
FTA	0.418	0.108	0.070	0.029
	(0.062)**	(0.013)**	(0.016)**	(0.017)$^{+}$
LN_TARIFF	-1.513	-0.348	-0.446	-0.579
	(0.378)**	(0.100)**	(0.118)**	(0.141)**
INTL_BRDR	-3.229	-0.635		
	(0.297)**	(0.064)**		
L.ln_trade_y		0.788	0.614	0.747
		(0.009)**	(0.011)**	(0.107)**
N	49916	46868	46868	35411
R^2	0.802	0.932	0.939	0.940
$\hat{\rho}$		0.202	0.371	0.241
		(0.008)**	(0.011)**	(0.104)*
UnderId χ^2				15.863
p-val				(0.003)
Weak Id χ^2				4.135
p-val				(0.003)
Over Id χ^2				0.921
p-val				(0.820)

Notes: This table uses size-adjusted trade as dependent variable to reproduce the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Column (4) implements an IV LSDV estimation. Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 12: SR Gravity & Efficiency, International Trade Data Only

	(1) Standard	(2) LagTrade	(3) LSDV	(4) LSDV_IV
LN_DIST	-1.210	-0.272		
	(0.040)**	(0.014)**		
CNTG	-0.082	-0.006		
	(0.192)	(0.042)		
CLNY	0.452	0.098		
	(0.163)**	(0.037)**		
LANG	0.761	0.155		
	(0.104)**	(0.023)**		
FTA	0.420	0.116	0.067	0.023
	(0.063)**	(0.014)**	(0.017)**	(0.018)
LN_TARIFF	-1.204	-0.304	-0.439	-0.458
	(0.336)**	(0.098)**	(0.118)**	(0.162)**
L.ln_trade		0.772	0.613	0.889
		(0.008)**	(0.011)**	(0.174)**
N	48928	45932	45932	34697
R^2	0.879	0.955	0.960	0.957
$\hat{\rho}$		0.217	0.373	0.105
		(0.000)**	(0.011)**	(0.165)
UnderId χ^2				10.190
p-val				(0.070)
Weak Id χ^2				1.720
p-val				(0.143)
Over Id χ^2				6.875
p-val				(0.143)

Notes: This table uses international trade only as the dependent variable to reproduce the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Column (4) implements an IV LSDV estimation. Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 13: SR Gravity & Efficiency, Arellano-Bond

	(1)	(2)
	All Years	3-Year Intervals
LN_DIST	-0.856	-0.460
	(0.064)**	(0.194)*
CNTG	0.020	0.024
	(0.117)	(0.101)
CLNY	0.324	0.162
	(0.100)**	(0.111)
LANG	0.544	0.277
	(0.077)**	(0.141)*
FTA	0.248	0.184
	(0.038)**	(0.047)**
LN_TARIFF	-0.927	-0.694
	(0.228)**	(0.340)*
INTL_BRDR	-2.352	-1.140
	(0.268)**	(0.578)*
L.LN_TRADE	0.254	0.603
	(0.045)**	(0.161)**
N	46868	15551
Over Id χ^2	168.99	17.530
p-val	(0.025)	(0.041)
AR(1)	-10.268	-4.514
χ^2 p-val	(0.000)	(0.000)
AR(2)	1.302	2.293
χ^2 p-val	(0.193)	(0.022)
AR(3)		-0.476
χ^2 p-val		(0.634)

Notes: This table reports results from the structural investment approach to test SR gravity theory with the Arellano-Bond estimator. All estimates are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) implements the Arellano-Bond estimator with data for all years. Column (2) uses 3-year interval data. Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 14: SR Gravity & Efficiency, PPML Estimations

	3-year Interval Data			Data All Years		
	Standard	SRG	SRG_LSDV	Standard	SRG	SRG_LSDV
	(1)	(2)	(3)	(4)	(5)	(6)
LN_DIST	-0.647	-0.073		-0.643	-0.028	
	(0.065)**	(0.008)**		(0.065)**	(0.002)**	
CNTG	0.489	0.021		0.491	0.008	
	(0.148)**	(0.018)		(0.145)**	(0.007)	
CLNY	-0.049	-0.018		-0.056	-0.007	
	(0.109)	(0.021)		(0.110)	(0.007)	
LANG	0.325	0.006		0.346	0.005	
	(0.139)*	(0.016)		(0.137)*	(0.006)	
FTA	0.142	0.060	0.059	0.126	0.015	0.010
	(0.123)	(0.013)**	(0.024)*	(0.117)	(0.004)**	(0.008)
LN_TARIFF	-7.440	0.524	-0.539	-7.441	0.166	-0.209
	(1.167)**	(0.150)**	(0.138)**	(1.091)**	(0.055)**	(0.057)**
INTL_BRDR	-2.636	-0.124		-2.651	-0.048	
	(0.154)**	(0.025)**		(0.154)**	(0.008)**	
L_trade		0.913	0.691		0.967	0.869
		(0.007)**	(0.014)**		(0.002)**	(0.006)**
N	18928	15665	15657	51376	47236	47231
$\hat{\rho}$		-0.033	0.205		0.031	0.124
		(0.007)**	(0.015)**		(0.002)**	(0.006)**

Notes: This table uses the PPML estimator to reproduce some of the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is nominal bilateral trade. The estimates from columns (1)-(3) are obtained with 3-year interval data and the estimates from columns (4)-(6) use data for all years. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Column (3) reproduces the results from column (2) but with pair fixed effects, thus implementing the LSDV estimator. Columns (4)-(6) reproduce the results from columns (1)-(3). Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 15: SR Gravity & Efficiency, Aggregate WIOD Data

	(1) Standard	(2) SRG_OLS	(3) LSDV	(4) SRG_LSDV
LN_DIST	-1.267	-0.144		
	(0.055)**	(0.014)**		
CNTG	0.343	0.039		
	(0.112)**	(0.017)*		
CLNY	0.524	0.066		
	(0.134)**	(0.025)**		
LANG	-0.034	-0.009		
	(0.122)	(0.023)		
RTA	0.406	0.110	0.172	0.112
	(0.081)**	(0.016)**	(0.062)**	(0.051)*
INTL_BRDR	-3.553	-0.470		
	(0.247)**	(0.045)**		
L.ln_trade		0.870		0.363
		(0.009)**		(0.022)**
N	14792	12943	14792	12943
R^2	0.890	0.973	0.978	0.982
$\hat{\rho}$		0.070		0.598
		(0.009)**		(0.024)**

Notes: This table uses the latest edition of the WIOD data and reproduces the results from Table 3 of the main text. All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The dependent variable in each specification is the log of nominal bilateral trade. Column (1) reports standard (long run) gravity estimates. The estimates from column (2) are obtained with lagged trade added as a covariate. Columns (3) and (4) reproduce the results from columns (1) and (2) but with pair fixed effects, thus implementing the LSDV estimator. Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 16: SR Gravity & Efficiency, Sectoral WIOD Data

	CropAnimal		ForestryLogging		FishingAquaculture		MiningQuarrying		Manufacturing		Services	
	SR_OLS	SR_LSDV	SR_OLS	SR_LSDV	SR_OLS	SR_LSDV	SR_OLS	SR_LSDV	SR_OLS	SR_LSDV	SR_OLS	SR_LSDV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
L.ln_trade	0.827	0.288	0.841	0.327	0.819	0.303	0.844	0.318	0.850	0.327	0.870	0.374
	(0.010)**	(0.021)**	(0.008)**	(0.021)**	(0.010)**	(0.019)**	(0.011)**	(0.024)**	(0.009)**	(0.021)**	(0.008)**	(0.020)**
LN_DIST	-0.227		-0.223		-0.227		-0.194		-0.180		-0.140	
	(0.019)**		(0.019)**		(0.019)**		(0.019)**		(0.015)**		(0.014)**	
CNTG	0.098		0.121		0.120		0.047		0.046		0.031	
	(0.023)**		(0.029)**		(0.029)**		(0.024)*		(0.018)*		(0.017)$^{+}$	
CLNY	0.069		0.112		0.110		0.098		0.070		0.064	
	(0.028)*		(0.038)**		(0.034)**		(0.033)**		(0.026)**		(0.025)**	
LANG	-0.013		-0.068		-0.054		-0.028		-0.002		-0.011	
	(0.027)		(0.033)*		(0.032)$^{+}$		(0.029)		(0.024)		(0.024)	
RTA	0.110	0.122	0.110	0.142	0.123	0.183	0.098	0.118	0.107	0.112	0.106	0.111
	(0.022)**	(0.066)$^{+}$	(0.026)**	(0.076)$^{+}$	(0.026)**	(0.074)*	(0.021)**	(0.067)$^{+}$	(0.017)**	(0.051)*	(0.017)**	(0.053)*
INTL_BRDR	-0.597		-0.580		-0.658		-0.512		-0.438		-0.519	
	(0.060)**		(0.068)**		(0.070)**		(0.059)**		(0.047)**		(0.046)**	
N	12943	12943	12341	12341	12341	12341	12943	12943	12943	12943	12943	12943
R^2	0.961	0.974	0.961	0.974	0.956	0.971	0.965	0.976	0.974	0.983	0.972	0.981
$\hat{\rho}$	0.113	0.679	0.099	0.637	0.121	0.663	0.095	0.647	0.090	0.637	0.069	0.586
	(0.010)**	(0.022)$^{+}$	(0.008)**	(0.022)$^{+}$	(0.010)**	(0.020)*	(0.011)**	(0.026)$^{+}$	(0.009)**	(0.022)*	(0.008)**	(0.022)*

Notes: This table reproduces the main results from Table 3 by capitalizing on the sectoral dimension of the latest edition of the WIOD data, as described in the data section. For brevity, for each sector we focus on reproducing the main specifications with lagged trade and standard gravity variables and with lagged trade and pair fixed effects. All results are obtained with the OLS estimator and with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Robust standard errors are reported in parentheses $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.