The Dynamics of Comparative Advantage*

Gordon H. Hanson ‡ — *UC San Diego and NBER*

Nelson Lind§ — *Emory University*

Marc-Andreas Muendler¶ — *UC San Diego and NBER*

March 18, 2018

Abstract

This paper characterizes the dynamics of comparative advantage and draws implications from these dynamics for quantitative analysis. In cross-section data, we establish that the distribution of export capabilities across industries is approximately log normal. This heavy-tailed shape is similar across 90 countries and stable over 40 years. Over time, there is mean reversion in export capability and this mean reversion, rather than indicating degeneracy, is instead consistent with a stationary stochastic process. We develop a GMM estimator for a Markov process whose stationary distribution nests many commonly studied distributions, and show that the Ornstein-Uhlenbeck (OU) special case closely approximates the dynamics of comparative advantage. The OU process implies a log normal stationary distribution and has a discrete-time representation that can be estimated with simple linear regression. Incorporating stochastic comparative advantage into the counterfactual analysis of changes in trade costs, we document the transitory nature of policy effects: churning causes targeted trade-policy changes to decay markedly, with most impacts fully dissipated within 10 to 20 years. These findings speak to the importance of incorporating dynamic comparative advantage into quantitative trade analysis.

Keywords: International trade; comparative advantage; generalized logistic diffusion; estimation of diffusion process

JEL Classification: F14, F17, C22

---

*Kilian Heilmann, Ina Jäkle, Heea Jung, Chen Liu and Xiao Ma provided excellent research assistance. We thank Sebastian Sotelo for kindly sharing estimation code. We thank Davin Chor, Arnaud Costinot, Dave Donaldson, Jonathan Eaton, Keith Head, Sam Kortum and Ben Moll as well as seminar participants at NBER ITI, Carnegie Mellon U, UC Berkeley, Brown U, the Cowles Foundation Yale and Princeton IES, U Dauphine Paris, the West Coast Trade Workshop UCLA, CESifo Munich, Paris 1 and Sciences Po, Banque de France, FGV São Paulo, U Munich, UC Riverside, Chinese University Hong Kong, Hong Kong University, Hong Kong University of Science and Technology, and the WTO Geneva for helpful discussions and comments. We gratefully acknowledge funding from the National Science Foundation under grant SES-1427021. An Online Supplement with supporting derivations and additional empirical evidence is available at *econ.ucsd.edu/muendler/papers/compadv-suppl.pdf*.

‡GPS, University of California, San Diego, 9500 Gilman Dr MC 0519, La Jolla, CA 92093-0519 (*gohanson@ucsd.edu*)

§Department of Economics, Emory University, 1602 Fishburne Dr, Rich Bldg, 3rd Floor, Atlanta, GA 30322-2240, *sites.google.com/view/nelsonlind* (*nlind@emory.edu*).

¶Department of Economics, University of California, San Diego, 9500 Gilman Dr MC 0508, La Jolla, CA 92093-0508, *econ.ucsd.edu/muendler* (*muendler@ucsd.edu*). Further affiliations: CAGE, CESifo, ifo and IGC.

1 Introduction

Quantitative analysis in international trade has attained new prominence (Costinot and Rodríguez-Clare 2014). Equipped with models that permit computational analysis when estimated or calibrated, much current research in trade focuses on performing counterfactual exercises to evaluate the impacts of trade reform, technological advance, and other shocks on national and global welfare. Work along this line has deepened our understanding of the effects of NAFTA and other trade agreements on real incomes (Caliendo and Parro 2015), the consequences of the global financial crisis for trade flows (Eaton, Kortum, Neiman, and Romalis 2016), and the mechanisms through which trade barriers affect the distribution of income within and between nations (Burstein and Vogel 2017), among a rapidly growing set of topics.[1]

Two advances have helped make the quantitative revolution in trade possible (Arkolakis, Costinot, and Rodríguez-Clare 2012). One is the formulation of trade models that generate gravity in bilateral trade and realistic global specialization patterns, while still characterizing national technological capabilities parsimoniously (e.g., Eaton and Kortum 2002, Melitz 2003). A second advance is the technical insight that "exact hat algebra" permits the measurement of discrete differences between actual and counterfactual equilibria (Dekle, Eaton, and Kortum 2007). This approach collapses the time-invariant features of the environment—including country-specific preference parameters, barriers to trade deriving from geography and related features, and supplies of fixed factors—into initial-period shares of consumer expenditure on goods and of producer expenditure on inputs. One can then conduct counterfactual analysis armed with little more than gravity estimates of trade-cost elasticities (which embody preference or technology parameters) and data on initial expenditure shares.

The standard approach in quantitative trade analysis is to allow specific features of the environment to vary selectively—e.g., home bias in consumption (Costinot, Donaldson, Kyle, and Williams 2016), import tariffs (Caliendo and Parro 2015), or trade imbalances (Eaton, Kortum, Neiman, and Romalis 2016). Comparative advantage, however, is commonly taken as static, even though it is a primary force responsible for trade in the first place. This approach leaves relative national-industry productivities in the background to be absorbed into expenditure shares. Resulting simulations of counterfactual outcomes are therefore conditional on the implicit assumption that shocks to comparative advantage over relevant time horizons are either modest or transitory. Despite the importance of this premise for modern research in trade, the literature lacks an accepted set of facts about the dynamics of comparative advantage. In this paper, we utilize a workhouse framework in trade, the Eaton and Kortum (2002) model of Ricardian comparative advantage (EK hereafter), to guide our analysis of comparative-advantage dynamics and to perform counterfactual exercises that allow us to evaluate how these

[1] Also see Burstein and Cravino (2015), Caliendo, Parro, Rossi-Hansberg, and Sarte (2017), Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2018).

dynamic properties affect predictions of common quantitative trade models.[2]

Our motivations for using EK are that it has strong empirical support in cross-section trade data (Chor 2010; Costinot, Donaldson, and Komunjer 2012) and is a foundational model in quantitative analysis (e.g., Costinot and Rodríguez-Clare 2014, Di Giovanni, Levchenko, and Zhang 2014, Caliendo and Parro 2015). Additionally, the EK model presents a framework to recover measures of comparative advantage from the gravity equation of trade. Using data for 133 industries in 90 countries over the period 1962 to 2007, we estimate gravity equations year by year to extract an exporter-industry fixed effect, which measures the exporting country's export capability in an industry in a given year; an importer-industry fixed effect, which captures the importing country's effective demand for foreign goods in an industry in a given year; and an exporter-importer component, which accounts for bilateral trade frictions (Head and Mayer 2014).[3] In the EK model, the exporter-industry fixed effect embodies national average factor prices and the location parameter of a country's productivity distribution for an industry. By taking the deviation of a country's log export capability from the global industry mean, we obtain a measure of a country's absolute advantage in an industry. Further normalizing this value by its country-wide mean removes the effects of country-level productivity and economy-wide factor prices. We use export capability under this double normalization to measure comparative advantage.

After estimating the gravity model, our analysis proceeds in three parts. First, we document the dynamic empirical properties of comparative advantage. For this purpose, we jointly analyze the time series and the cross-sectional distribution at given moments in time. Strikingly, the cross-industry distribution of absolute advantage for a country in a given year (national industry export capability relative to the global industry mean) is similar across countries (except for the intercept). Its shape is approximately log normal with ratios of the mean to the median of about 11. Importantly, this log-normal shape is stable over time.[4] Temporal stability in the distribution of export advantage makes a second empirical regularity all the more surprising: there is continual and rapid turnover in countries' top export industries. Among the goods that account for the top 5% of a country's current absolute-advantage industries, 60% were not in the top 5% two decades earlier.[5] Such churning is consistent with mean reversion in comparative advantage. In an OLS regression of the ten-year change in log export capability on

[2]Despite our chosen link to EK, the gravity model that we employ is consistent with a large class of trade models (Anderson 1979, Arkolakis, Costinot, and Rodríguez-Clare 2012), of which EK is just one example. The Krugman (1980), Heckscher-Ohlin (Deardorff 1998), Melitz (2003), and Anderson and van Wincoop (2003) models also yield gravity specifications and give alternative interpretations of the exporter-industry fixed effects that we use as measures of absolute advantage in our analysis.

[3]On decomposing sources of changes in bilateral trade, see also Gaubert and Itskhoki (2015). We estimate the gravity equation using both OLS and methods developed by Silva and Tenreyro (2006) and Eaton, Kortum, and Sotelo (2012) to correct for zero bilateral trade flows.

[4]In our data, the median share for the top good (out of 133) in a country's total exports is 23%, for the top 3 goods is 46%, and for the top 7 goods is 64%. See Easterly and Reshef (2010) and Freund and Pierola (2013) on export concentration in low-income countries, and Hidalgo and Hausmann (2009) and Hausmann and Hidalgo (2011) on the link between export concentration and export composition.

[5]On changes in export diversification over time see Imbs and Wacziarg (2003), Cadot, Carrére, and Strauss Kahn (2011), and Sutton and Trefler (2016).

its initial log value and industry-year and country-year fixed effects—to which we refer as a decay regression—we estimate mean reversion at the rate of about one-third per decade.

Levchenko and Zhang (2013) also find evidence of mean reversion in comparative advantage, in their case for 19 aggregate manufacturing sectors. One may be tempted to see mean reversion as evidence of convergence in sectoral productivities across countries, possibly indicating the degeneracy of comparative advantage. Such an interpretation, however, would be subject to the Quah (1993, 1996) critique of cross-country growth regressions: mean reversion in a variable alone is uninformative about the dynamics of its distribution. Depending on the stochastic process, mean reversion may alternatively coexist with a cross-section distribution that is degenerate, non-stationary, or stationary and stable. We find that the latter is the case for comparative advantage. Stability of the heavy-tailed distribution of export advantage over time suggests that, far from being degenerate, a country's distribution of comparative advantage is stationary.

In the second part of our analysis, we estimate a stochastic process that can account for the combination of a stable cross-industry distribution for national export advantage with churning in national industry export ranks. As a mean-reverting AR(1) specification, our OLS decay regression is a discrete-time analogue of a continuous-time Ornstein-Uhlenbeck (OU) process, which is the unique Markov process that has a stationary normal distribution (Karlin and Taylor 1981). The OU process is governed by two parameters, which we recover from our OLS estimates. The *dissipation rate* regulates the speed at which absolute advantage reverts to its long-run mean and determines the shape of its stationary distribution; the *innovation intensity* scales the stochastic shocks to absolute advantage and determines how frequently industries reshuffle along the distribution. Our estimates of the dissipation rate are similar across countries and industries, which confirms that the heavy-tailedness of export advantage is close to universal. To relax the assumption of log normality, which is implied by the OU process, we estimate via GMM a *generalized logistic diffusion* (GLD) for absolute advantage, which has the OU process as a limiting case. The GLD adds an additional parameter to estimate—the *decay elasticity* allows the rate of mean reversion to differ from above and from below the mean. The stationary distribution for the GLD is a generalized gamma distribution, which unifies the extreme-value and gamma families and nests many common distributions (Crooks 2010), including those used to study city size (Gabaix and Ioannides 2004, Luttmer 2007) and firm size (Sutton 1997, Gabaix 1999, Cabral and Mata 2003).

Our estimation approach targets moments of the stochastic process directly. We apply statistical insights from Forman and Sørensen (2008) to our Markov process and develop a GMM time-series estimator of the three global parameters of the GLD: the dissipation rate, the innovation intensity, and the decay elasticity. Beyond earlier applications of the Forman and Sørensen (2008) method, we show that it is possible to perform unit-level estimation in the presence of an aggregate stochastic process and recover country-wide productivity as a by-product. We then predict the cross-section distribution of absolute advantage, which is not targeted in our estimation. Based on just the three parameters (for all industries in all countries and in all years), the predicted values match the individual cross-section distributions with considerable accuracy. The observed churning of industry export ranks within countries over time is matched by the model-predicted transition probabilities between percentiles of the cross-section distribution, except in the very low tail of the distribution. This exercise also reveals that while the data select the GLD over the more restrictive OU form, the two models yield very similar predictions for period-to-period transition probabilities between quantiles of the distribution of export advantage. Thus, in many applications, the OU process may be sufficient to quantify export dynamics.

In the third part of our analysis, we incorporate our findings on the dynamics of comparative advantage into a quantitative EK trade model. We go beyond conventional exact hat algebra, which presumes time-invariant absolute advantage parameters for each country, and account for the dynamic evolution of comparative advantage. Taking as given our parameter estimates, we solve for counterfactual EK equilibria under a stipulated policy shock for a large set of simulated comparative advantage realizations. We show in this context that a country's self trade (its expenditure on own production) can be recovered from gravity estimates in the EK model, without requiring industry-level production data (after standardizing by country-wide output), so that quantitative analysis can be conducted largely with trade data. To illustrate the procedure, we conduct counterfactual exercises in which we permanently reduce the trade costs faced by China's top-5 export industries or top-50 export industries (out of 133) by 10% in 1990 and assess mean simulated outcomes over the ensuing 20 years. The experiments are akin to selective reductions of trade barriers on imports from China by the rest of the world.

For each of the two policy shocks (falling trade barriers for China's top-5 or top-50 industries in 1990), we consider three scenarios. In one, which follows standard practice in the literature (e.g., Alvarez and Lucas 2007; Dekle, Eaton, and Kortum 2007), we hold all fundamentals—including comparative advantage—fixed at their 1990 levels and compute a counterfactual equilibrium resulting from the change in selected trade costs. Not surprisingly, China's real wages and total exports rise (permanently). In alternative scenarios, we allow comparative advantage to be stochastic by initializing the series at 1990 levels and then allowing the series to evolve over time according to our estimated GLD process. With stochastic comparative advantage, the change in equilibrium outcomes due to a reduction in trade costs is a random variable. To characterize the effect of a trade-cost reduction, we solve for equilibrium outcomes across many simulated paths of comparative advantage, where for each simulation we obtain outcomes, first, without the change in trade costs, and next, with the change. We define the treatment effect, or more precisely the treatment-effect path, to be the mean percentage difference in outcomes with and without the trade-cost reduction across simulations at each moment in time. Under stochastic comparative advantage, the mean treatment effect of reduced trade costs on total exports dissipates over time.

Because in expectation China's initial top export industries steadily lose their top ranking—and because heavy tails in export advantage mean that top industries matter for aggregate outcomes disproportionately—trade-cost reductions for these industries steadily lose their importance in the aggregate. When China's top-5 1990 export industries are treated, the impacts on China's real wages and total exports on average fully dissipate within approximately 15 years; when China's top-50 1990 export industries are treated, the impact on China's total exports on average dissipates by almost 10% after 10 years and by more than 50% after 20 years, while some real wage impact is lasting.

The need to reconcile dynamics with well defined cross-sectional distributions arises in many fields. To study economic growth, for example, Quah develops a kernel estimator and analyzes non-parametrically the transitions of countries between percentiles in the cross-country income distribution; his estimates provide evidence of an evolving bimodal cross-sectional distribution of incomes among 105 countries, beyond conventional patterns of convergence or divergence.[6] Taking this kernel estimator to study the evolution of revealed comparative advantage, Proudman and Redding (2000) document for the exports of G-5 countries to the OECD between 1970 and 1993 high degrees of turnover but no marked change in the concentration of comparative advantage. Non-parametric estimation can help identify non-monotonicities in transitions from initial conditions, but a parametric specification is essential for calibrating economic models. We therefore introduce the GLD as a stochastic process that generates a unified family of common cross-sectional distributions and use the estimates to check its predicted transitions against non-parametric kernel measures of the frequencies: except for deviations in the lower-most decile of the comparative advantage distribution, the GLD fits transitions remarkably well.

In macroeconomic labor studies, as another example, Postel-Vinay and Robin (2002) estimate the cross sectional distribution of French wages and back out consistent dynamics from the cross-sectional estimates; Guvenen, Ozkan, and Song (2014) characterize non-parametrically the short-term transitions in the distribution of U.S. household incomes and calibrate a stochastic process that is consistent with these moments. In the literature on firm sizes, Arkolakis (2016) uses data on the cross section of U.S. firm sales and Brazilian exports to the United States to calibrate parameters of a stochastic process for firm productivity. Similarly, Gaubert and Itskhoki (2015) estimate the equilibrium distribution of French firms' domestic and export sales to calibrate a consistent stochastic process for productivity; they find that industry-level comparative advantage accounts for 70 percent and granular sales shocks at dominant firms for another 30 percent of the variation in export shares. Most papers across these fields have in common that they rely on properties of the stationary distribution to calibrate a suitable stochastic process. Instead, we directly estimate the stochastic process itself and then use the predicted stationary distribution to assess the fit of our estimates to the cross section. While entry, exit and

[6] Using Quah's kernel estimator for regional economics, Overman and Ioannides (2001) estimate the dynamics of the U.S. city size distribution and document that U.S cities in a second tier below the top-ten cities exhibit more size turnover than the largest cities.

switching between macroeconomic regimes complicate the stochastic process beyond our GLD in those lines of research, extensions of our proposed GMM approach may also permit the estimation of stochastic processes in other fields.

A broader literature on export dynamics includes reduced-form analysis that characterizes sources of export growth and parametric approaches that are amenable to simulation. Estimates of the exact importance of supply-side exporter components at the industry level differ between varying reduced-form approaches (Daruich, Easterly, and Reshef 2016, Egger and Nigai 2016), but have in common that exporter components account for a substantive part of export growth. Lederman, Pienknagura, and Rojas (2015) point out that export-market exit and re-entry patterns at the product level are suggestive of a latent comparative advantage beyond manifest comparative advantage in current exports, especially in small countries. Through gravity estimation we separate the supply-side components from both bilateral trade cost and demand-side factors in importer markets, and then focus on the realized comparative-advantage contribution to export growth. Exploring sources of comparative advantage empirically, Cameron, Proudman, and Redding (2005) tie trade growth to technology and human capital for the United Kingdom and argue that international trade raises rates of UK productivity growth through technology transfer but not innovation. Cai, Li, and Santacreu (2018) estimate the rate of cross-border patent citations to quantify a multi-sector endogenous growth model, in which knowledge flows across borders are independent of trade but relative sectoral productivities change the allocation of labor to innovative activity; they show in simulations with data on 28 countries and 19 sectors that a reduction in trade costs leads to a reallocation of innovative activity to sectors with stronger comparative advantage and dynamic gains from trade that are multiple times larger than the static gains from specialization.

In addition to knowledge flows that are independent of trade in Cai, Li, and Santacreu (2018), alternative theoretical approaches model the origins of sectoral comparative advantage in production with an instantaneous flow of knowledge across borders but varying adoption depending on relative local factor supplies (Acemoglu and Zilibotti 2001), or with exogenous comparative advantage in innovation and an endogenous home market effect in demand (Somale 2017), or with sectoral productivity growth that results from investments in innovation by incumbent firms with heterogenous innovation capabilities in the presence of varying knowledge spillovers across borders and within and between sectors (Sampson 2017). While multi-sector models of trade are important to address comparative advantage, single-sector models that study growth through learning in the presence of trade—such as Basu and Weil (1998), Lucas and Moll (2014), Perla and Tonetti (2014) and Sampson (2016)—explain the flow of knowledge and technology adoption across countries, or—in Perla, Tonetti, and Waugh (2015)—within countries and across firms, or both—in Buera and Oberfield (2016). While we analyze only the stochastic and stationary properties of comparative advantage, our estimates of the stable cross-sectional concentration and the simultaneous churning of comparative advantages can serve as an empirical benchmark to discipline the study of knowledge flows in the presence of trade.

In Section 2 we present a theoretical motivation for our gravity specification. In Section 3 we describe the data and gravity model estimates, and document stationarity and heavy tails in export advantage as well as churning in top export goods. In Section 4 we introduce a stochastic process that generates a cross-sectional distribution consistent with heavy tails and embeds innovations consistent with churning, and we derive a GMM estimator for this process. In Section 5 we present estimates and evaluate the fit of the model. In Section 6 we turn to simulations that use our estimates and perform counterfactual policy exercises. In Section 7 we conclude.

2 Theoretical Motivation

We use the EK model to motivate our definitions of export capability and absolute advantage, and describe our approach for extracting these measures from the gravity equation of trade.

2.1 Export capability, absolute advantage, and comparative advantage

In EK, an industry consists of many product varieties. The productivity q of a source-country s firm that manufactures a variety in industry i is determined by a random draw from a Fréchet distribution with CDF $F_Q(q) = \exp\{-(q/\underline{q}_{is})^{-\theta}\}$ for $q > 0$. The location parameter $\underline{q}_{is}$ determines the typical productivity level of a firm in the industry while the shape parameter θ controls the dispersion in productivity across firms. Consumers, who have CES preferences over product varieties within an industry, buy from the firm that delivers a variety at the lowest price. With marginal-cost pricing, a higher productivity draw makes a firm more likely to be the lowest-cost supplier of a variety to a given market.

Comparative advantage stems from the location of the industry productivity distribution, given by $\underline{q}_{is}$, which may vary by country and industry. In a country-industry with a higher $\underline{q}_{is}$, firms are more likely to have a high productivity draw, such that in this country-industry a larger fraction of firms succeeds in exporting to multiple destinations. Consider the many-industry version of the EK model in Costinot, Donaldson, and Komunjer (2012). Exports by source country s to destination country d in industry i can be written as,

$$X_{isd} = \frac{\left(w_s \tau_{isd} / \underline{q}_{is}\right)^{-\theta}}{\sum_{\varsigma} \left(w_{\varsigma} \tau_{i\varsigma d} / \underline{q}_{i\varsigma}\right)^{-\theta}} \mu_{id} E_d, \tag{1}$$

where w_s is the unit production cost in source country s, τ_{isd} is the iceberg trade cost between s and d in industry i, μ_{id} is the Cobb-Douglas share of industry i in destination d expenditure, and E_d is national expenditure in country d.[7] Taking logs of (1), we obtain a gravity equation for bilateral trade

$$\ln X_{isd} = k_{is} + m_{id} - \theta \ln \tau_{isd}, \tag{2}$$

where $k_{is} \equiv \theta \ln(\underline{q}_{is}/w_s)$ is source country s's log *export capability* in industry i, which is a function of the country-industry's efficiency ($\underline{q}_{is}$) and the country's unit production cost (w_s),[8] and

$$m_{id} \equiv \ln \left[\mu_{id} E_d / \sum_{\varsigma} \left(w_{\varsigma} \tau_{i\varsigma d} / \underline{q}_{i\varsigma} \right)^{-\theta} \right]$$

is the log of effective import demand by country d in industry i, which depends on national expenditure on goods in the industry divided by an index of the toughness of industry competition in the country.[9]

Looking forward to the estimation, the presence of the importer-industry fixed effect m_{id} in (2) implies that export capability k_{is} is only identified up to an industry normalization. We therefore re-express export capability as the deviation from its global industry mean $(1/S)\sum_{\varsigma=1}^{S} k_{i\varsigma}$, where S is the number of source countries. Exponentiating this value, we measure *absolute advantage* of source country s in industry i as

$$A_{is} \equiv \frac{\exp\{k_{is}\}}{\exp\left\{\frac{1}{S}\sum_{\varsigma=1}^{S} k_{i\varsigma}\right\}} = \frac{(\underline{q}_{is}/w_s)^{\theta}}{\exp\left\{\frac{1}{S}\sum_{\varsigma=1}^{S} \ln(\underline{q}_{i\varsigma}/w_{\varsigma})^{\theta}\right\}}. \tag{3}$$

The normalization in (3) differences out both worldwide industry supply conditions, such as shocks to global total factor productivity, and worldwide industry demand conditions, such as variation in the expenditure share μ_{id}. When A_{is} rises for country-industry is, we say that country s's absolute advantage has increased in industry i even though it is only strictly the case that its export capability has risen relative to the global geometric mean for i. In fact, s's export capability in i may have gone up relative to some countries and fallen relative to others. We use the deviation from the industry geometric mean to define absolute advantage because it simplifies the specification of a stochastic process for export capability. Rather than specifying export capability itself, we

[7] In our simulations we allow for trade imbalances so that $E_d = Y_d - TB_d$, where Y_d is national income and TB_d is the trade balance.

[8] Our assumption that unit production costs w_s are country specific and not also industry specific allows us to difference out this term in the country normalization of export advantage that we apply below. The presence of industry specific production costs would imply that export capability k_{is} depends on endogenously determined factor prices. Although in this case we could no longer interpret export capability as a primitive, it would retain an interpretation as a reduced-form determinant of comparative advantage.

[9] Any trade model that has a gravity structure will generate exporter-industry fixed effects and a reduced-form expression for export capability (k_{is}). In the Armington (1969) model, as applied by Anderson and van Wincoop (2003), export capability is a country's endowment of a good relative to its remoteness from the rest of the world. In Krugman (1980), export capability equals the number of varieties a country produces in an industry times effective industry marginal production costs. In Melitz (2003), export capability is analogous to that in Krugman adjusted by the Pareto lower bound for productivity in the industry. In a Heckscher-Ohlin model (Deardorff 1998), export capability reflects the relative size of a country's industry based on factor endowments and industry-specific factor intensities. The common feature of these models is that export capability is related to a country's productive potential in an industry, be it associated with resource supplies, a home-market effect, or the distribution of firm-level productivity.

model its deviation from a worldwide industry trend, which frees us from having to model the global trend component.

To relate our use of absolute advantage A_{is} to conventional approaches, average (2) over destinations and define (harmonic) log exports from source country s in industry i at the country's industry trade costs as

$$\ln \bar{X}_{is} \equiv k_{is} + \frac{1}{D}\sum_{d=1}^{D} m_{id} - \frac{1}{D}\sum_{d=1}^{D} \theta \ln \tau_{isd}, \tag{4}$$

where D is the number of destination markets. We say that country s has a comparative advantage over country ς in industry i relative to industry j if the following familiar condition holds:

$$\frac{\bar{X}_{is}/\bar{X}_{i\varsigma}}{\bar{X}_{js}/\bar{X}_{j\varsigma}} = \frac{A_{is}/A_{i\varsigma}}{A_{js}/A_{j\varsigma}} > 1. \tag{5}$$

Intuitively, absolute advantage defines country relative exports, once we neutralize the distorting effects of trade costs and proximity to market demand on trade flows, as in (4). In practice, a large number of industries and countries makes it cumbersome to conduct double comparisons of country-industry is to all other industries and all other countries, as suggested by (5). The definition in (3) simplifies this comparison in the *within-industry dimension* by setting the "comparison country" in industry i to be the global mean across countries in i. In the final estimation strategy that we develop in Section 4, we will further normalize the comparison in the *within-country dimension* by estimating the absolute advantage of the "comparison industry" for country s, consistent with an arbitrary stochastic country-wide growth process. Demeaning in the industry dimension and then estimating the most suitable normalization in the country dimension makes our empirical approach consistent with both worldwide stochastic industry growth and stochastic national country growth.

Our concept of export capability k_{is} can be related to the deeper origins of comparative advantage by treating the country-industry specific location parameter $\underline{q}_{is}$ as the outcome of an exploration and innovation process. In Eaton and Kortum (1999, 2010), firms generate new ideas for how to produce existing varieties more efficiently. The efficiency q of a new idea is drawn from a Pareto distribution with CDF $G(q) = (q/\underline{x}_{is})^{-\theta}$, where $\underline{x}_{is} > 0$ is the minimum efficiency. New ideas arrive in continuous time according to a Poisson process, with intensity rate $\rho_{is}(t)$. At date t, the number of ideas with at least efficiency q is then distributed Poisson with parameter $T_{is}(t)\, q^{-\theta}$, where $T_{is}(t)$ is the number of previously discovered ideas that are available to producers and that is in turn a function of $\underline{x}_{is}^{\theta}$ and past realizations of $\rho_{is}(t)$.[10] Setting $T_{is}(t) = \underline{q}_{is}(t)^{\theta}$, this framework yields

[10]Eaton and Kortum (2010) allow costly research effort to affect the Poisson intensity rate and assume that there is "no forgetting" such that all previously discovered ideas are available to firms. In our simple sketch, we abstract away from research effort and treat the stock of knowledge available to firms in a country (relative to the mean across countries) as stochastic. Buera and Oberfield (2016) microfound the innovation process in Eaton and Kortum (2010) by allowing agents to transmit ideas within and across borders through

identical predictions for the volume of bilateral trade as in equation (1). Our empirical approach is to treat the stock of ideas available to a country in an industry $T_{is}(t)$—relative to the global industry mean stock of ideas $(1/S)\sum_{\varsigma=1}^{S} T_{i\varsigma}(t)$—as following a stochastic process.

2.2 Estimating the gravity model

Allowing for measurement error in trade data or unobserved trade costs, we can introduce a disturbance term into the gravity equation (2), converting it into a linear regression model. With data on bilateral industry trade flows for many importers and exporters, we can obtain estimates of the exporter-industry and importer-industry fixed effects from an OLS regression. The gravity model that we estimate is

$$\ln X_{isdt} = k_{ist} + m_{idt} + \mathbf{r}'_{sdt}\mathbf{b}_{it} + v_{isdt}, \tag{6}$$

where we add time subscript t. We include dummy variables to measure exporter-industry-year k_{ist} and importer-industry-year m_{idt} terms. The regressors $\mathbf{r}_{sdt}$ are the determinants of bilateral trade costs, and v_{isdt} is a residual that is mean independent of $\mathbf{r}_{sdt}$. The variables we use to measure trade costs $\mathbf{r}_{sdt}$ in (6) are standard gravity covariates, which do not vary by industry.[11] However, we allow the coefficient vector $\mathbf{b}_{it}$ on these variables to differ by industry and by year.[12] Absent annual measures of industry-specific trade costs for all years, we model these costs via the interaction of country-level gravity variables and time-and-industry-varying coefficients.

The values we use for empirical analysis are deviations of estimated exporter-industry-year dummies from global industry means. The measure of absolute advantage in (3) for source country s in industry i becomes

$$A_{ist} = \frac{\exp\left\{k_{ist}^{\text{OLS}}\right\}}{\exp\left\{\frac{1}{S}\sum_{\varsigma=1}^{S} k_{i\varsigma t}^{\text{OLS}}\right\}} = \frac{\exp\left\{k_{ist}\right\}}{\exp\left\{\frac{1}{S}\sum_{\varsigma=1}^{S} k_{i\varsigma t}\right\}}, \tag{7}$$

where k_{ist}^{OLS} is the OLS estimate of k_{ist} in (6).

As is well known (Silva and Tenreyro 2006, Head and Mayer 2014), the linear regression model (6) is inconsistent with the presence of zero trade flows, which are common in bilateral data. We recast EK to allow for zero trade by following Eaton, Kortum, and Sotelo (2012), who posit that in each industry in each country only a finite number of firms make productivity draws, meaning that in any realization of the data there may be no

trade. A Fréchet distribution for country-industry productivity emerges as an equilibrium outcome in this environment, where the location parameter of this distribution reflects the current stock of ideas in a country.

[11] These include log distance between the importer and exporter, the time difference (and time difference squared) between the importer and exporter, a contiguity dummy, a regional trade agreement dummy, a dummy for both countries being members of GATT, a common official language dummy, a common prevalent language dummy, a colonial relationship dummy, a common empire dummy, a common legal origin dummy, and a common currency dummy.

[12] We estimate (6) separately by industry and by year. Since in each year the regressors are the same across industries for each bilateral exporter-importer pair, there is no gain to pooling data across industries in the estimation.

firms from country s that have sufficiently high productivity to profitably supply destination market d in industry i. Instead of augmenting the expected log trade flow $\mathbb{E}[\ln X_{isd}]$ from gravity equation (2) with a disturbance, Eaton, Kortum, and Sotelo (2012) consider the expected share of country s in the market for industry i in country d, $\mathbb{E}[X_{isd}/X_{id}]$, and write this share in terms of a multinomial logit model. This approach requires that one know total expenditure in the destination market, X_{id}, including a country's spending on its own goods. Since total spending is unobserved in our data, we invoke independence of irrelevant alternatives and specify the dependent variable as the expectation for the share of source country s in import purchases by destination d in industry i:

$$\mathbb{E}\left[\frac{X_{isdt}}{\sum_{\varsigma\neq d} X_{i\varsigma dt}}\right] = \frac{\exp\left\{k_{ist} - \mathbf{r}'_{sdt}\mathbf{b}_{it}\right\}}{\sum_{\varsigma\neq d}\exp\left\{k_{i\varsigma t} - \mathbf{r}'_{\varsigma dt}\mathbf{b}_{it}\right\}}. \tag{8}$$

Since estimation of (8) is well approximated the Poisson pseudo-maximum-likelihood (PPML) gravity model (Silva and Tenreyro 2006), we re-estimate exporter-industry-year fixed effects by applying PPML.[13]

Our baseline measure of absolute advantage uses regression-based estimates of exporter-industry-year fixed effects, which may be imprecise when a country exports a good to few destinations in a given year. As an alternative measure of export performance, we use the Balassa (1965) revealed comparative advantage (RCA) index:

$$RCA_{ist} \equiv \frac{\sum_d X_{isdt}/\sum_\varsigma \sum_d X_{i\varsigma dt}}{\sum_\iota \sum_d X_{\iota sdt}/\sum_\iota \sum_\varsigma \sum_d X_{\iota\varsigma dt}}. \tag{9}$$

The RCA index does not correct for trade costs or proximity to market demand; it uses just raw trade data. Throughout our analysis we will employ OLS and PPML gravity-based measures of absolute advantage (7) alongside the Balassa RCA index (9). Reassuringly, our results for the three measures are quite similar.

3 Data and Main Regularities

The data for our analysis are World Trade Flows from Feenstra, Lipsey, Deng, Ma, and Mo (2005), and their extension to 2007, which are based on SITC revision 1 industries for 1962 to 1983 and SITC revision 2 industries for 1984 to 2007. We create a consistent set of country aggregates in these data by maintaining as single units countries that split up or unite over the sample period.[14] To further maintain consistency in the countries present, we restrict the sample to nations that trade in all years and that exceed a minimal size threshold, which leaves 116 country units.[15] The switch from SITC revision 1 to revision 2 in 1984 led to the creation of many new industry

[13]We thank Sebastian Sotelo for estimation code.

[14]These countries are the Czech Republic, the Russian Federation, and Yugoslavia. We join East and West Germany, Belgium and Luxembourg, as well as North and South Yemen.

[15]This reporting restriction leaves 141 importers (97.7% of world trade) and 139 exporters (98.2% of world trade) and is roughly equivalent to dropping small countries from the sample. For consistency in terms of country size, we drop countries with fewer than 1

categories. To maintain a consistent set of SITC industries over the sample period, we aggregate industries to a combination of two- and three-digit categories.[16] These aggregations and restrictions leave 133 industries in the data. In an extension of our main analysis, we limit the sample to SITC revision 2 data for 1984 forward, so we can check the sensitivity of our results to industry aggregation by using two-digit (60 industries) and three-digit definitions (225 industries), which bracket the industry definitions that we use for the full-sample period.[17]

A further set of country restrictions is required to estimate importer and exporter fixed effects. For coefficients on exporter-industry dummies to be comparable over time, it is important to require that destination countries import a product in all years. Imposing this restriction limits the sample to 46 importers, which account for an average of 92.5% of trade among the 116 country units. In addition, we need that exporters ship to overlapping groups of importing countries. As Abowd, Creecy, and Kramarz (2002) show, such connectedness assures that all exporter fixed effects are separately identified from importer fixed effects. This restriction leaves 90 exporters in the sample that account for an average of 99.4% of trade among the 116 country units. Using our sample of 90 exporters, 46 importers, and 133 industries, we estimate the gravity equation (6) separately by industry i and year t and then extract absolute advantage A_{ist} given by (7). Data on gravity variables are from CEPII.org.

3.1 Stable heavy tails in export advantage

Figure 1 depicts the full distribution of absolute advantage across industries for 12 countries in 2007.[18] The plots show the log number of industries for exporter s that have at least a given level of absolute advantage in year t against the corresponding log level of industry absolute advantage $\ln A_{ist}$. By design, the plots characterize the cumulative distributions of absolute advantage by country and by year (Axtell 2001, Luttmer 2007). Plots for 28 countries in 1967, 1987 and 2007 are shown in Appendix **Figures A1, A2** and **A3**. While the lower cutoff for

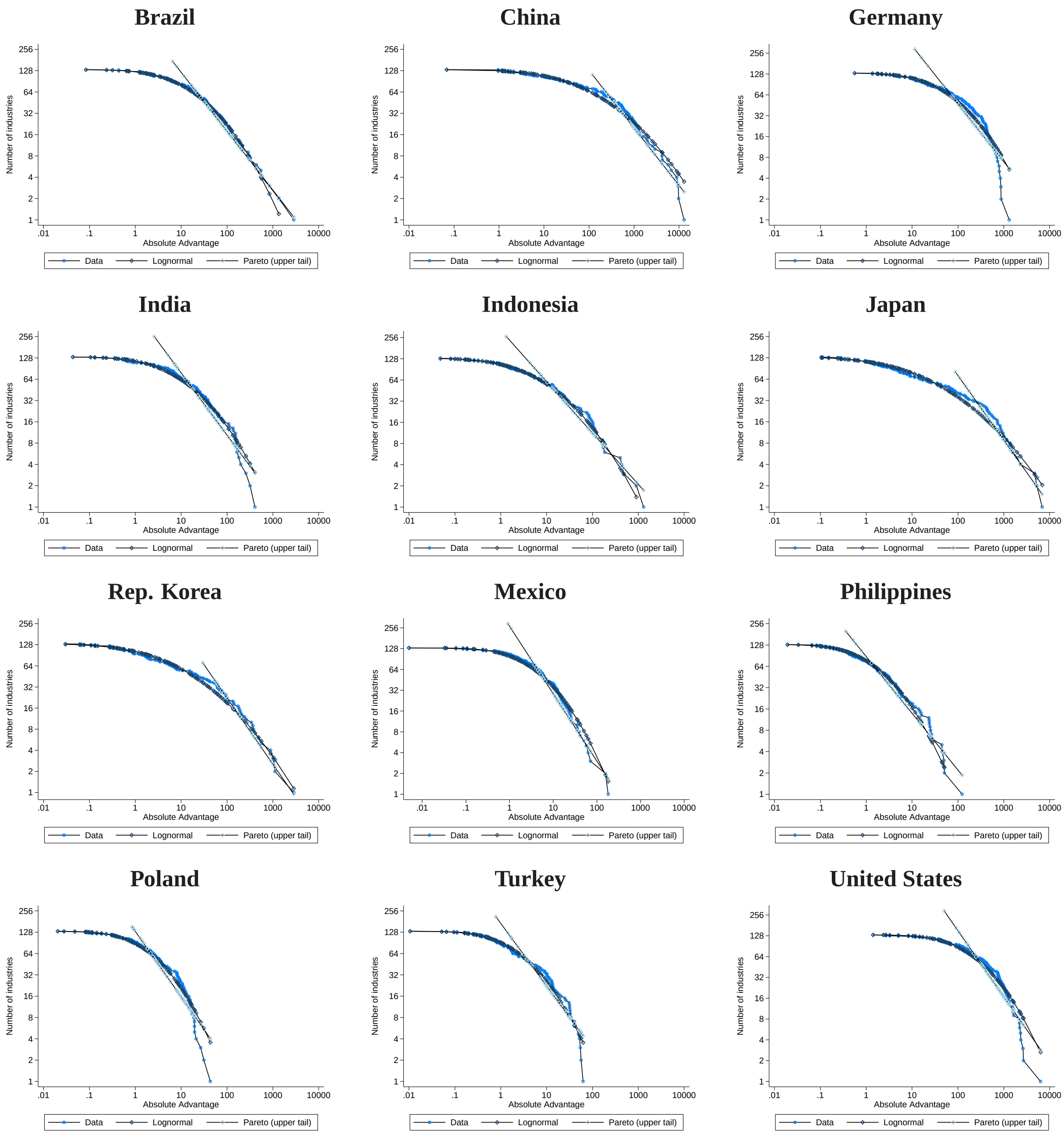
million inhabitants in 1985, reducing the sample to 116 countries (97.4% of world trade).

[16]There are 226 three-digit SITC industries that appear in all years, which account for 97.6% of trade in 1962 and 93.7% in 2007. Some three-digit industries frequently have their trade reported only at the two-digit level (which accounts for the just reported decline in trade shares for three-digit industries). We aggregate over these industries, creating 143 industry categories that are a mix of SITC two and three-digit industries. From this group we drop non-standard industries: postal packages (SITC 911), special transactions (SITC 931), zoo animals and pets (SITC 941), non-monetary coins (SITC 961), and gold bars (SITC 971). We further exclude uranium (SITC 286), coal (SITC 32), petroleum (SITC 33), natural gas (SITC 341), and electrical current (SITC 351), which violate the Abowd, Creecy, and Kramarz (2002) requirement of connectedness for estimating identified exporter fixed effects in many years.

[17]In an earlier version of our paper, we estimated OLS gravity equations for four-digit SITC revision 2 products (682 industries). PPML estimates at the four-digit level turn out to be quite noisy, owing to the many exporters in industries at this level of disaggregation that ship goods to no more than a few importers. Consequently, we exclude data on four-digit industries from the analysis.

[18]In the Online Supplement (Table S1), we show the top two products in terms of $\ln A_{ist}$ for select countries and years. To remove the effect of national market size and make values comparable across countries, we normalize log absolute advantage by its country mean, which produces a double log difference—a country-industry's log deviation from the global industry mean less the country-wide average across all industries—and captures comparative advantage. The magnitudes of export advantage are enormous. In 2007, comparative advantage in the top product is over 300 log points in 88 of the 90 exporting countries. To verify that our measure of export advantage does not peg obscure industries as top industries, in the Online Supplement (Figure S1) we plot $\ln A_{ist}$ against the log of the share of the industry in national exports $\ln(X_{ist}/(\sum_{\iota} X_{\iota st}))$. In all years, there is a strongly positive correlation between log absolute advantage and the log industry share of national exports (0.77 in 1967, 0.78 in 1987, and 0.83 in 2007).

Figure 1: **Cumulative Probability Distribution of Absolute Advantage for Select Countries in 2007**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries in 2005-2007 and CEPII.org; three-year means of OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6). *Note*: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 2007$ (Pareto fit to upper five percentiles only).

absolute advantage shifts right over time, the shape of the cross-section CDF is remarkably stable across countries and years. This shape stability of the cross-sectional absolute advantage distribution suggests that comparative advantage is trend stationary, a robust feature that we will revisit under varying perspectives.

The figures also graph the fit of absolute advantage to a Pareto distribution and to a log normal distribution using maximum likelihood, where each distribution is fit separately for each country in each year. The Pareto and the log normal are common choices in the literatures on the distribution of city and firm sizes (e.g., Sutton 1997). For the Pareto distribution, the cumulative distribution plot is linear in the logs, whereas the log normal distribution generates a relationship that is concave to the origin.

The cumulative distribution plots clarify that the empirical distribution of absolute advantage is not Pareto. The log normal, by contrast, fits the data closely. The concavity of the data plots indicate that gains in absolute advantage fall off progressively more rapidly as one moves up the rank order of absolute advantage, a feature characteristic of the log normal. The upper tails of the distribution are heavy. Across all countries and years, the ratio of the mean to the median is 11.1 for absolute advantage based on our baseline OLS estimates of export capability, 23.5 for absolute advantage based on PPML estimates, and 1.2 for the Balassa RCA index.[19] Though overall the log normal approximates the shape of the distribution for absolute advantage, for some countries the number of industries in the upper tail drops too fast, relative to strict log normality. These discrepancies motivate our specification of a generalized logistic diffusion for absolute advantage in Section 4.

To make sure that our findings are not the byproduct of incompletely modelled zero bilateral trade in the gravity estimation, we also show plots based on PPML estimates of export capability, with similar results. To verify that the graphed cross-section distributions are not a byproduct of specification error in estimating the gravity model, we repeat the plots using the Balassa RCA index in 1987 and 2007, again with similar results. And to verify that the patterns we uncover are not a consequence of industry aggregation, we construct plots at the three-digit level based on SITC revision 2 data in 1987 and 2007, yet again with similar results.[20]

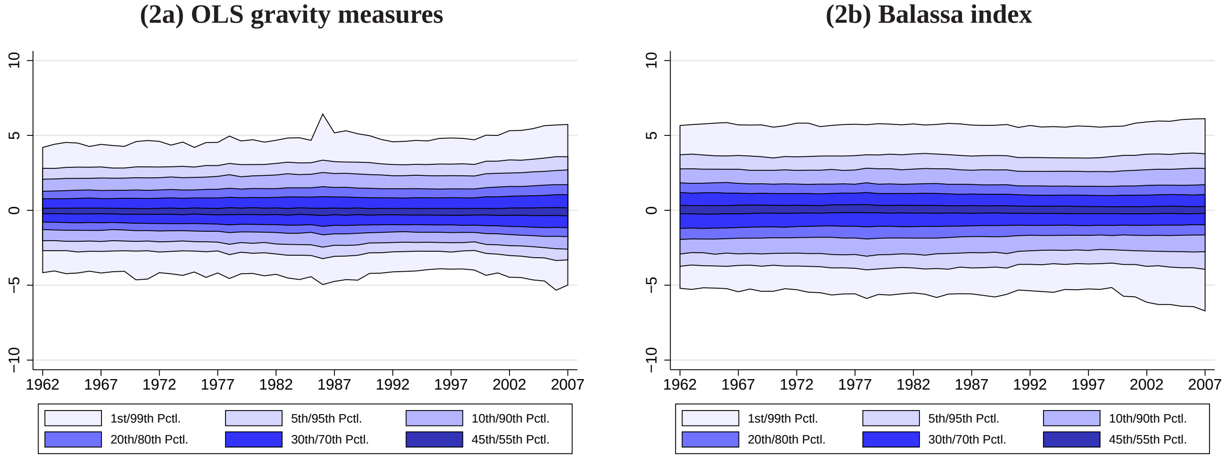
Figures A1, A2 and **A3** in the Appendix provide visual evidence that the heavy-tailed shape of the distribution of absolute advantage for individual countries are stable over time. To substantiate this property of the data, we pool industry-level measures of comparative advantage across countries and plot the percentiles of this global distribution in each year, as shown in **Figure 2** for OLS-based measures of export capability and for Balassa RCA indexes.[21] The plots for the 5th/95th, 20th/80th, 30th/70th, and 45th/55th percentiles are, with minor fluctuation, parallel to the horizontal axis. This is a strong indication that the global distribution of comparative advantage is stationary. If it were the case that comparative advantage degenerated, the percentile lines would

[19] To compute the reported mean-median ratios, we omit outliers and weight by industry counts within country-years.

[20] Each of these additional sets of results is available in the Online Supplement: Figures S2 and S3 for the PPML estimates, Figures S4 and S5 for the Balassa measure, and Figures S6, S7, S8 and S9 for the two- and three-digit industry definitions under SITC revision 2.

[21] The Online Supplement (Figure S10) shows percentile plots for PPML-based measures of export capability.

Figure 2: **Percentiles of Comparative Advantage Distributions by Year**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: We obtain log comparative advantage as the residuals from OLS projections on industry-year and source country-year effects (δ_{it} and δ_{st}) for (a) OLS gravity measures of log absolute advantage $\ln A_{ist}$ and (b) the log Balassa index of revealed comparative advantage $\ln \mathit{RCA}_{ist} = \ln(X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_{\iota} X_{\iota st}/\sum_{\iota}\sum_{\varsigma} X_{\iota\varsigma t})$.

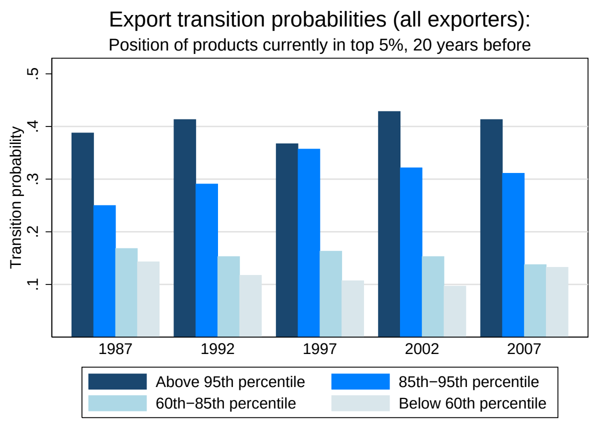
slope downward from above the mean and upward from below the mean, as the distribution became increasingly compressed over time, a pattern clearly not in evidence. If, instead, the distribution of comparative advantage was non-stationary, we would see the upper percentile lines drifting upward and the lower percentile lines drifting downward. There is mild drift only in the extreme tails of the distribution, the 1^{st} and 99^{th} percentiles, and there only during the early 2000s, a pattern which stalls or reverses after 2005.

Before examining the time series of export advantage in more detail, we consider whether a log normal distribution of absolute advantage could be an incidental consequence of the gravity estimation. The exporter-industry fixed effects are estimated sample parameters, which by the Central Limit Theorem converge to being normally distributed around their respective population parameters as the sample size becomes large. However, normality of this log export capability estimator does not imply that the cross-sectional distribution of absolute advantage becomes log normal. If no other element but the residual noise from gravity estimation generated log normality in absolute advantage, then the cross-sectional distribution of absolute advantage between industries in a country would be degenerate around a single mean. The data are clearly in favor of non-degeneracy for the distribution of absolute advantage. **Figure 1** and its counterparts (**Figures A1, A2** and **A3** in the Appendix) document that industries within a country differ markedly in terms of their mean export capability.[22]

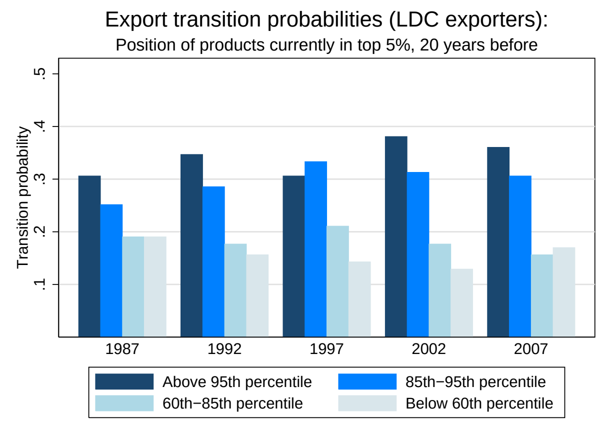
[22]The distribution of Balassa revealed comparative advantage is also approximately log normal, which indicates that non-regression based measures of comparative advantage exhibit similar distributional patterns.

Figure 3: **Absolute Advantage Transition Probabilities**

(3a) All exporters



(3b) LDC exporters



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: The graphs show the percentiles of products is that are currently among the top 5% of products, 20 years earlier. The sample is restricted to products (country-industries) is with current absolute advantage A_{ist} in the top five percentiles $(1 - F_A(A_{ist}) \geq .05)$, and then grouped by frequencies of percentiles twenty years prior, where the past percentile is $1 - F_A(A_{is,t-20})$ of the same product (country-industry) is. For the classification of less developed countries (LDC) see the Supplementary Material (Section S.1).

3.2 Churning in export advantage

The stable distribution plots of absolute advantage give an impression of little variability. The strong concavity in the cross-sectional plots is present in all countries and in all years. Yet, this cross-sectional stability masks considerable turnover in industry rankings of absolute advantage behind the cross-sectional distribution. Of the 90 exporters, 68 have a change in the top comparative-advantage industry between 1987 and 2007.[23] Over this period, Canada's top good switches from sulfur to wheat, China's from fireworks to telecommunications equipment, India's from tea to precious stones, and Poland's from barley to furniture. Moreover, most new top products in 2007 were not the number one or two good in 1987 but from lower down the ranking.

To characterize churning in industry export advantage, in **Figure 3** we calculate the fraction of top products in a given year that were also top products in the past. For each country in each year, we identify where in the distribution the top 5% of absolute-advantage products (in terms of A_{ist}) were 20 years earlier. We then average across outcomes for the 90 export countries. The fraction of top 5% products in a given year that were also top 5% products two decades before ranges from a high of 42.9% in 2002 to a low of 36.7% in 1997. Averaging over all years, the share is 40.2%, indicating a 60% chance that a good in the top 5% in terms of absolute advantage today was not in the top 5% two decades before. On average, 30.6% of new top products come from the 85th to 95th percentiles, 15.5% come from the 60th to 85th percentiles, and 11.9% come from the bottom six deciles.

[23]Evidence of this churning is seen in the Online Supplement (Table S1).

Outcomes are similar when we limit the sample to developing economies.

Turnover in top export goods suggests that over time export advantage dissipates—countries' strong industries weaken and their weak industries strengthen—as would be consistent with mean reversion. We test for mean reversion in export capability by specifying the AR(1) process

$$k^{\mathrm{OLS}}_{is,t+10} - k^{\mathrm{OLS}}_{ist} = \rho\, k^{\mathrm{OLS}}_{ist} + \delta_{it} + \delta_{st} + \varepsilon_{is,t+10}, \tag{10}$$

where k^{OLS}_{ist} is the OLS estimate of log export capability from gravity equation (6). In (10), the dependent variable is the ten-year change in export capability and the predictors are the initial value of export capability and dummies for the industry-year δ_{it} and for the country-year δ_{st}. We choose a long time difference for export capability—a full decade—to help isolate systematic variation in country export advantages. Controlling for industry-year fixed effects converts export capability into a measure of absolute advantage; controlling additionally for country-year fixed effects allows us to evaluate the dynamics of comparative advantage. The coefficient ρ captures the fraction of comparative advantage that decays over ten years. The specification in (10) is similar to the productivity convergence regressions reported in Levchenko and Zhang (2013), except that we use trade data to calculate country advantage in an industry, examine industries at a considerably more disaggregate level, and include both manufacturing and nonmanufacturing industries in the analysis. Because we estimate log export capability k^{OLS}_{ist} from the first-stage gravity estimation in (6), we need to correct the standard errors in (10) for the presence of generated variables. To do so, we apply a generated-variable correction (see Appendix D).

Table 1 presents coefficient estimates for equation (10). The first three columns report results for log export capability based on OLS, the next three for log export capability based on PPML, and the final three for the log Balassa RCA index. Estimates for ρ are uniformly negative, consistent with mean reversion in export advantage. We soundly reject the hypothesis that there is no decay (H_0: $\rho = 0$) and also the hypothesis that there is instantaneous dissipation (H_0: $\rho = -1$) at conventional levels of significance. Estimates for the full sample of countries and industries in columns 1, 4, and 7 are similar in value, equal to -0.35 when using OLS log export capability, -0.32 when using PPML log export capability, and -0.30 when using log RCA. These magnitudes indicate that over the period of a decade the typical country-industry sees approximately one-third of its comparative advantage (or disadvantage) erode. In columns 2, 5, and 8, we present comparable results for the subsample of developing countries. Decay rates for this group are larger than the worldwide averages in columns 1, 4, and 7, indicating that in less-developed economies mean reversion in comparative advantage is more rapid. In columns 3, 6, and 9, we present results for nonmanufacturing industries (agriculture, mining, and other primary commodities). For PPML export capability and Balassa RCA, decay rates for the nonmanufacturing industries are similar to those for the full sample of industries.

Table 1: OLS ESTIMATES OF COMPARATIVE ADVANTAGE DECAY, 10-YEAR TRANSITIONS

	OLS gravity k			PPML gravity k			$\ln$ *RCA*		
	All	LDC	Nonmanf.	All	LDC	Nonmanf.	All	LDC	Nonmanf.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Decay Regression Coefficients									
Decay rate ρ	-0.349	-0.454	-0.450	-0.320	-0.358	-0.322	-0.303	-0.342	-0.293
	(0.002)***	(0.002)***	(0.003)***	(0.0002)***	(0.0003)***	(0.0003)***	(0.01)***	(0.013)***	(0.012)***
Var. of residual s^2	2.089	2.408	2.495	2.709	3.278	3.123	2.318	2.849	2.561
	(0.024)***	(0.026)***	(0.042)***	(0.013)***	(0.018)***	(0.021)***	(0.006)***	(0.009)***	(0.009)***
Implied Ornstein-Uhlenbeck (OU) Parameters									
Dissipation rate η	0.276	0.292	0.280	0.198	0.179	0.173	0.222	0.199	0.195
	(0.003)***	(0.003)***	(0.005)***	(0.0009)***	(0.001)***	(0.001)***	(0.006)***	(0.006)***	(0.006)***
Intensity of innovations σ	0.558	0.644	0.654	0.623	0.703	0.670	0.570	0.648	0.596
	(0.003)***	(0.004)***	(0.006)***	(0.001)***	(0.002)***	(0.002)***	(0.005)***	(0.009)***	(0.006)***
Observations	324,978	202,010	153,768	320,310	199,724	149,503	324,983	202,014	153,773
Adjusted R^2 (within)	0.222	0.267	0.262	0.282	0.290	0.266	0.216	0.224	0.214
Years t	36	36	36	36	36	36	36	36	36
Industries i	133	133	68	133	133	68	133	133	68
Source countries s	90	62	90	90	62	90	90	62	90

Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS and PPML gravity measures of export capability (log absolute advantage) $k = \ln A$ from (8).

Note: Reported figures for ten-year changes. Variables are OLS and PPML gravity measures of log absolute advantage $\ln A_{ist}$ and the log Balassa index of revealed comparative advantage $\ln \mathit{RCA}_{ist} = \ln(X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_{\iota} X_{\iota st}/\sum_{\iota}\sum_{\varsigma} X_{\iota\varsigma t})$. OLS estimation of the ten-year decay rate ρ from

$$k_{is,t+10} - k_{ist} = \rho\, k_{ist} + \delta_{it} + \delta_{st} + \epsilon_{is,t+10},$$

conditional on industry-year and source country-year effects δ_{it} and δ_{st} for the full pooled sample (column 1-2) and subsamples (columns 3-6). The implied dissipation rate η and squared innovation intensity σ^2 are based on the decay rate estimate ρ and the estimated variance of the decay regression residual $\hat{s}^2$ by (13). Less developed countries (LDC) as listed in the Supplementary Material (Section S.1). Nonmanufacturing merchandise spans SITC sector codes 0-4. Robust standard errors, clustered at the industry level and corrected for generated-regressor variation of export capability k, for ρ and s^2, applying the multivariate delta method to standard errors for η and σ. * marks significance at ten, ** at five, and *** at one-percent level.

As an additional robustness check, we re-estimate (10) for the period 1984-2007 using data from the SITC revision 2 sample, reported in Appendix **Table A1**. Estimated decay rates are comparable to those in **Table 1**. At either the two-digit level (60 industries) or three-digit level (224 industries), the decay-rate estimates based on PPML export capability and RCA indexes are similar to those for the baseline combined two- and three-digit level (133 industries), with estimates based on OLS export capability being somewhat more variable. Because these additional samples use data for the 1984-2007 period and the original sample uses the full 1962-2007 period, these results also serve as a robustness check on the stability in coefficient estimates over time.[24]

3.3 Comparative advantage as a stochastic process

On its own, mean reversion in log export capability is uninformative about the dynamics of its distribution.[25] While mean reversion is consistent with a stationary cross-sectional distribution, it is also consistent with a non-ergodic distribution or a degenerate comparative advantage that collapses at a long-term mean of one (log comparative advantage of zero). Degeneracy in comparative advantage is an interpretation that has arisen from the findings in Levchenko and Zhang (2013) of cross-country convergence in industry productivities. Yet, the combination of mean reversion in **Table 1** and temporal stability of the cumulative distribution plots in **Figure 1** suggests a balance between random innovations to export capability and the deterministic dissipation of these capabilities, a balance characteristic of a stochastic process that generates a stationary cross-sectional distribution.

The decay regression in (10) is consistent with the discretized version of a commonly studied stochastic process, the Ornstein-Uhlenbeck (OU) process, which belongs to the family of diffusions (Markov processes for which all realizations of the random variable are continuous functions of time and past realizations). The OU process is the unique non-degenerate diffusion that has a stationary normal distribution (Karlin and Taylor 1981). Consider log comparative advantage $\ln \hat{A}_{is}(t)$—export capability normalized by industry-year and country-year means. Suppose that in continuous time comparative advantage $\hat{A}_{is}(t)$ follows an OU process given by

$$\mathrm{d} \ln \hat{A}_{is}(t) = -\frac{\eta \sigma^2}{2} \ln \hat{A}_{is}(t)\, \mathrm{d}t + \sigma\, \mathrm{d}W_{is}^{\hat{A}}(t), \tag{11}$$

where $W_{is}^{\hat{A}}(t)$ is a Wiener process that induces stochastic innovations in comparative advantage.[26] The parameter

[24] Our finding that decay rates imply incomplete mean reversion is further evidence against absolute advantage being incidental. Suppose that the cumulative distribution plots of log absolute advantage reflected random variation in export capability around a common expected value for each country in each year due to measurement error in trade data. If this measurement error were classical, all within-country variation in the exporter-industry fixed effects would be the result of iid disturbances that were uncorrelated across time. We would then observe no temporal connection between these distributions. When estimating the decay regression in (10), mean reversion would be complete, yielding a value of ρ close to -1. The coefficient estimates are inconsistent with such a pattern.

[25] A case in point is Quah's (1993, 1996) critique of using cross-country regressions to test for convergence in rates of economic growth.

[26] To relate equation (11) to trade theory, our specification for the evolution of export advantage is analogous to the equation of motion for a country's stock of ideas in the dynamic EK model of Buera and Oberfield (2016). In their model, each producer in source country

η regulates the rate at which comparative advantage reverts to its global long-run mean and the parameter σ scales time and therefore the Brownian innovations $\mathrm{d}W_{is}^{\hat{A}}(t)$.[27] Because comparative advantage reflects a double normalization of export capability, it is natural to consider a global mean of zero for $\ln \hat{A}_{is}(t)$. As mentioned, the OU process has a stationary normal distribution, so its specification for log comparative advantage $\ln \hat{A}_{is}(t)$ implies that $\hat{A}_{is}(t)$ has a stationary log normal distribution.

In (11), we refer to the parameter η as the *dissipation rate* of comparative advantage because it contributes to the speed at which $\ln \hat{A}_{is}(t)$ would collapse to a degenerate level of zero if there were no stochastic innovations. The parametrization in (11) implies that η alone determines the shape of the stationary distribution, while σ is irrelevant for the cross section. Our parametrization treats η as a normalized rate of dissipation that measures the "number" of one-standard deviation shocks that dissipate per unit of time. We refer to σ as the *innovation intensity*. It plays a dual role: σ governs volatility by scaling the Wiener innovations, and helps regulate the speed at which time elapses in the deterministic part of the diffusion.

To connect the continuous-time OU process in (11) to our decay regression in (10), we use the fact that the discrete-time process that results from sampling an OU process at a fixed time interval Δ is a Gaussian first-order autoregressive process with autoregressive parameter $\exp\{-\eta\sigma^2\Delta/2\}$ and innovation variance $(1 - \exp\{-\eta\sigma^2\Delta\})/\eta$ (Aït-Sahalia, Hansen, and Scheinkman 2010, Example 13). Applying this insight to the first-difference equation above, we obtain our decay regression:

$$k_{is}(t+\Delta) - k_{is}(t) = \rho\, k_{is}(t) + \delta_i(t) + \delta_s(t) + \varepsilon_{is}(t, t+\Delta), \tag{12}$$

which implies for the reduced-form decay parameter that

$$\rho \equiv -(1 - \exp\{-\eta\sigma^2\Delta/2\}) < 0,$$

for the unobserved country fixed effect $\delta_s(t) \equiv \ln Z_s(t+\Delta) - (1+\rho)\ln Z_s(t)$, where $Z_s(t)$ is an arbitrary time-varying country-specific shock, and for the residual $\varepsilon_{ist}(t, t+\Delta) \sim N\left(0,\ (1 - \exp\{-\eta\sigma^2\Delta\})/\eta\right)$.[28] An

s draws a productivity from a Pareto distribution, where this productivity combines multiplicatively with ideas learned from other firms, either within the same country or in different countries. Learning—or exposure to ideas—occurs at an exogenous rate $\alpha_s(t)$ and the learning of one producer from another depends on the parameter β, which captures the transmissibility of ideas between producers. In equilibrium, the distribution of productivity across suppliers within a country is Fréchet, with location parameter equal to a country's current stock of ideas. The OU process in (11) emerges from the equation of motion for the stock of ideas in Buera and Oberfield (2016, equation (4)) as the limiting case with the transmissibility parameter $\beta \to 1$, provided that the learning rate $\alpha_s(t)$ is subject to random shocks and producers in a country only learn from suppliers within the same country. In Section 7, we discuss how equation (11) could be extended to allow for learning across national borders.

[27] Among possible parameterizations of the OU process, we choose (11) because it is related to our later extension to a generalized logistic diffusion and clarifies that the parameter σ is irrelevant for the shape of the cross-sectional distribution. We deliberately specify η and σ to be invariant over time, industry and country and in Section 5 assess the fit under this restriction.

[28] For theoretical consistency, we state the country fixed effect $\delta_s(t)$ as a function of the shock $Z_s(t)$, which we will formally define as

OU process with $\rho \in (-1, 0)$ generates a log normal stationary distribution in the cross section, with a shape parameter of $1/\eta$ and a zero mean.

The reduced-form decay coefficient ρ in (12) is a function of both the dissipation rate η and the intensity of innovations σ and may differ across samples because either or both of those parameters vary. This distinction is important because ρ may vary even if the shape of the distribution of comparative advantage does not change.[29] From OLS estimation of (12), we can obtain estimates of η and σ^2 using the solutions,

$$\eta = \frac{1 - (1 + \hat{\rho})^2}{\hat{s}^2} \tag{13a}$$

$$\sigma^2 = \frac{\hat{s}^2}{1 - (1 + \hat{\rho})^2} \frac{\ln (1 + \hat{\rho})^{-2}}{\Delta}, \tag{13b}$$

where $\hat{\rho}$ is the estimated decay rate and $\hat{s}^2$ is the estimated variance of the decay regression residual.

Table 1 shows estimates of η and σ^2 implied by the decay regression results, with standard errors obtained using the multivariate delta method.[30] The estimate of η based on OLS export capability, at 0.28 in column 1 of **Table 1**, is larger than those based on PPML export capability, at 0.20 in column 4, or the log RCA index, at 0.22 in column 7, implying that the distribution of OLS export capability will be more concave to the origin. But estimates generally indicate strong concavity, consistent with the visual evidence in **Figure 1**. To gain intuition about η, suppose the intensity of innovations of the Wiener process is unity ($\sigma = 1$). Then a value of η equal to 0.28 means that it will take 5.0 years for half of the initial shock to log comparative advantage to dissipate (and 16.4 years for 90% of the initial shock to dissipate). Alternatively, if η equals 0.20 it will take 6.9 years for half of the initial shock to decay (and 23.0 years for 90% of the initial shock to dissipate).[31]

To see how the dissipation rate and the innovation intensity affect the reduced-form decay parameter ρ, we contrast η and σ^2 estimates across subsamples. First, compare the estimate for ρ in the subsample of developing economies in column 2 of **Table 1** to that in the full sample of countries in column 1. The larger estimate of ρ in the former sample (-0.45 in column 2 versus -0.35 in column 1) implies that *reduced-form* mean reversion is relatively rapid in developing countries. However, this result is silent about how the shape of the distribution of comparative advantage varies across nations. The similarity in the estimated dissipation rate η between the developing-country sample ($\eta = 0.29$) and the full-country sample ($\eta = 0.28$) indicates that

a country-wide stochastic trend in equation (14) below and then identify in subsequent GMM estimation.

[29] The estimated value of ρ is sensitive to the time interval $\triangle$ that we define in (12), whereas the estimated value of η is not. At shorter time differences—for which there may be relatively more noise in export capability—the estimated magnitude of σ is larger and therefore the reduced-form decay parameter ρ is as well. However, the estimated intrinsic speed of mean reversion η is unaffected. In unreported results, we verify these insights by estimating the decay regression in (10) for time differences of 1, 5, 10, and 15 years.

[30] Details on the construction of standard errors for η and σ^2 are available in the Online Supplement (Section S.3).

[31] In the absence of shocks and for $\sigma = 1$, log comparative advantage follows the deterministic differential equation $\mathrm{d} \ln \hat{A}_{is}(t) = -(\eta/2) \ln \hat{A}_{is}(t)\, \mathrm{d}t$ by (16) and Itō's lemma, with the solution $\ln \hat{A}_{is}(t) = \ln \hat{A}_{is}(0) \exp\{-(\eta/2)t\}$. Therefore, the number of years for a dissipation of $\ln \hat{A}_{is}(0)$ to a remaining level $\ln \hat{A}_{is}(T)$ is $T = 2 \log[\ln \hat{A}_{is}(0)/ \ln \hat{A}_{is}(T)]/\eta$.

comparative advantage is similarly heavy-tailed in the two groups. The larger reduced-form decay rate ρ for developing countries results from a larger intensity of innovations ($\sigma = 0.64$ in column 2 versus $\sigma = 0.56$ in column 1, where this difference is statistically significant). While a one-standard-deviation shock to comparative advantage in a developing country dissipates at roughly the same rate as in an industrialized country, because the magnitude of this shock is larger for the developing country, its observed rate of decay will be faster.

Second, compare nonmanufacturing industries in column 3 to the full sample of industries in column 1. Whereas the average nonmanufacturing industry differs from the average overall industry in the reduced-form decay rate ρ (-0.45 in column 3 versus -0.35 in column 1), it shows no difference in the estimated dissipation rate η (0.28 in column 1 versus 0.29 in column 3), implying that comparative advantage has comparably heavy tails inside and outside manufacturing. However, the intensity of innovations σ is larger for nonmanufacturing (0.65 in column 3 versus 0.56 in column 1), due perhaps to higher volatility associated with resource discoveries. These nuances regarding the shape of and the convergence speed toward the cross-sectional distribution of comparative advantage are undetectable when one considers the reduced-form decay rate ρ alone.[32]

The diffusion model in (11) and its discrete-time analogue in (12) reveal a deep connection between heavy tails in export advantage and churning in industry export ranks. Random innovations in absolute advantage cause industries to change positions in the cross-sectional distribution of comparative advantage for a country at a rate of innovation precisely fast enough so that the deterministic dissipation of absolute advantage creates a stable, heavy-tailed distribution of export prowess. We turn next to a generalization of the OU process and a more rigorous characterization of the dynamic behavior.

4 The Diffusion of Comparative Advantage

Over time, the stochastic process must match the cumulative distributions in **Figure 1**. **Figures A1** through **A3** in the Appendix show for more countries, and over time in 1967, 1987 and 2007, that the cross-sectional distributions of absolute advantage shift right for each country, consistent with the series being non-stationary. Yet, the cross-section distributions preserve their shape across periods, suggesting that once we adjust absolute advantage for country-wide productivity growth, the resulting series is stationary. We define this series to be *generalized comparative advantage*, written in continuous time as

$$\hat{A}_{is}(t) \equiv \frac{A_{is}(t)}{Z_s(t)}, \tag{14}$$

[32] Appendix **Table A1** shows results for two- and three-digit industries for the subperiod 1984-2007. Whereas reduced-form decay rates ρ increase in magnitude as one goes from the two- to the three-digit level, dissipation rates η remain stable. The difference in reduced-form decay rates ρ is driven by a higher intensity of innovations σ among the more narrowly defined three-digit industries. Intuitively, the magnitude of shocks to comparative advantage is larger in the more disaggregated product groupings.

where $A_{is}(t)$ is observed absolute advantage and $Z_s(t)$ is an unobserved country-wide stochastic trend (an arbitrary country-specific shock to absolute advantage).[33]

The appealing simplicity of the OU process notwithstanding, a concern for empirically characterizing comparative advantage over time is the strict log-linearity of the deterministic dissipation component: by (11) the change in log comparative advantage depends linearly on the log level of comparative advantage. In order to allow the deterministic component to vary with the level of comparative advantage more generally, we replace the term $\ln \hat{A}_{is}(t)$ with a common transformation:

$$\mathrm{d} \ln \hat{A}_{is}(t) = -\frac{\eta\sigma^2}{2} \frac{\hat{A}_{is}(t)^{\phi} - 1}{\phi} \, \mathrm{d}t + \sigma \, \mathrm{d}W_{is}^{\hat{A}}(t). \tag{15}$$

By L'Hôpital's rule, the generalized term $[\hat{A}_{is}(t)^{\phi} - 1]/\phi$ simplifies to $\ln \hat{A}_{is}(t)$ as ϕ approaches zero.[34] Next, we use (15) to derive a generalized diffusion, which guides the specification and estimation of a stochastic process for comparative advantage that is less restrictive than the OU.

4.1 Generalized logistic diffusion

Using Itō's lemma, we can restate the diffusion of log comparative advantage (15) as the relative change in comparative advantage with

$$\frac{\mathrm{d}\hat{A}_{is}(t)}{\hat{A}_{is}(t)} = \frac{\sigma^2}{2} \left[1 - \eta \, \frac{\hat{A}_{is}(t)^{\phi} - 1}{\phi} \right] \mathrm{d}t + \sigma \, \mathrm{d}W_{is}^{\hat{A}}(t), \tag{16}$$

a diffusion for the real parameters (η, σ, ϕ). The variable $W_{is}^{\hat{A}}(t)$ is the Wiener process. The diffusion (16) nests the OU process as $\phi \to 0$ (with η finite). For the special case of $\phi = 1$, the process is known as the stochastic logistic equation or ordinary logistic diffusion (Leigh 1968). We therefore call (16) a *generalized logistic diffusion* (GLD). While we intentionally stay within the family of diffusions, the GLD allows us to test the OU process against well-defined alternatives, to evaluate the fit of the model to the data, and to characterize the dynamic implications of the model—all of which we undertake in Section 5.[35] The GLD also allows us to make the

[33]This measure satisfies the properties of comparative advantage in (5), which compares country and industry pairs.

[34]The generalized term is a common choice in many fields. In econometrics, it is known as the Box-Cox transformation (Box and Cox 1964), in macroeconomics and decision theory a similar generalization of log utility is called the isoelastic utility function or CRRA (constant relative risk aversion) utility (Pratt 1964), and in statistical mechanics it is referred to as Tsallis entropy (Tsallis 1988).

[35]Returning to the connection between our approach and the dynamic EK model in Buera and Oberfield (2016)—also see footnotes 10 and 26—the specification in (16) is equivalent to their equation of motion for the stock of ideas (Buera and Oberfield 2016, equation (4)) under the assumptions that producers only learn from suppliers within their national borders and the learning rate $\alpha_s(t)$ is constantly growing across industries, countries, and over time but subject to idiosyncratic shocks. The parameter ϕ in (16) is equivalent to the value $\beta - 1$ in their model, where β captures the transmissibility of ideas between producers. Our finding, discussed in Section 5, that ϕ is small and negative implies that the value of β in the Buera and Oberfield (2016) model is large (but just below 1, as they require).

deterministic dissipation of comparative advantage depend on the current level of comparative advantage, so as to provide additional realism for our simulations in Section 6.

The term $(\sigma^2/2)[1 - \eta\{\hat{A}_{is}(t)^{\phi} - 1\}/\phi]$ in (16) is a deterministic drift that regulates the relative change in comparative advantage $\mathrm{d}\hat{A}_{is}(t)/\hat{A}_{is}(t)$. It involves constant parameters (η, σ, ϕ) and a level-dependent component $\hat{A}_{is}(t)^{\phi}$, where ϕ is the elasticity of the mean reversion with respect to the current level of absolute advantage, which we call the *decay elasticity*. For the OU process ($\phi \to 0$), the relative change in absolute advantage is neutral with respect to the current level. If $\phi > 0$, then the drift component $\hat{A}_{is}(t)^{\phi}$ leads to a faster than neutral mean reversion from above than from below the mean, indicating that the loss of absolute advantage in a currently strong industry tends to occur more rapidly than the buildup of absolute disadvantage in a currently weak industry. Conversely, if $\phi < 0$ then mean reversion tends to occur more slowly from above than below the mean. The parameters η and σ in (16) inherit their interpretations from the OU process in (11) as the dissipation rate and the innovation intensity. As before, the innovation intensity σ regulates the speed of convergence to the stationary distribution but has no effect on its shape. Under the GLD, the dissipation rate η and decay elasticity ϕ jointly determine the heavy tail of the cross-sectional distribution, to which we turn now.

4.2 Cross-sectional distribution of comparative advantage

For real parameters (η, σ, ϕ), the GLD (16) has a stationary distribution that is generalized gamma. We provide a derivation in Appendix A and restrict our discussion here to a description of the main properties. The generalized gamma distribution unifies the gamma and extreme-value distributions, as well as many others (Crooks 2010), and has the log normal, the Pareto, and other commonly used distributions as special or limiting cases. To motivate our choice of the GLD, and hence of the generalized gamma as the cross-sectional distribution for comparative advantage, consider the graphs in **Figure 1** (as well as **Figures A1** through **A3** in the Appendix). These figures are broadly consistent with comparative advantage being log normal in the cross section. But they also indicate that for many countries the number of industries drops off more quickly or more slowly in the upper tail than the log normal distribution can capture. The generalized gamma distribution accommodates such kurtosis.[36]

Formally, after arbitrarily much time has passed under the GLD, a cross section of the data has the generalized gamma pdf for a realization $\hat{a}_{is}$ of the random variable comparative advantage $\hat{A}_{is}$, given by:

$$f_{\hat{A}}(\hat{a}_{is}|\hat{\theta}, \kappa, \phi) = \frac{1}{\Gamma(\kappa)} \left|\frac{\phi}{\hat{\theta}}\right| \left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi\kappa-1} \exp\left\{-\left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi}\right\} \qquad \text{for} \quad \hat{a}_{is} > 0, \tag{17}$$

[36]Our implementation of the generalized gamma uses three parameters, as in Stacy (1962). In their analysis of the firm size distribution, Cabral and Mata (2003) also use a version of the generalized gamma distribution.

where $\Gamma(\cdot)$ denotes the gamma function and $(\hat{\theta}, \kappa, \phi)$ are real parameters with[37]

$$\hat{\theta} = \left(\phi^2/\eta\right)^{1/\phi} > 0 \quad \text{and} \quad \kappa = 1/\hat{\theta}^{\phi} > 0.$$

The generalized gamma nests the ordinary gamma distribution for $\phi = 1$ and the log normal or Pareto distributions when ϕ tends to zero.[38] A non-degenerate stationary distribution exists only if $\eta > 0$.[39]

4.3 Cross-sectional distributions of absolute advantage

Absolute advantage, defined as in (3), is measurable by exporter-industry-year fixed effects estimated from the gravity model in (6). By contrast, generalized comparative advantage, as defined in (14), has an unobserved country-specific stochastic trend $Z_s(t)$, and lacks a direct empirical counterpart. We therefore need to identify $Z_s(t)$ in estimation. Intuitively, identification of $Z_s(t)$ is possible because we can observe the evolving position of the cumulative absolute advantage distribution over time and, as we now show, the evolving position is the only difference between the cumulative distributions of absolute and comparative advantage.

The stationary distribution of absolute advantage is closely related to that of comparative advantage under the maintained assumption that comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion given by (16). As stated before, the GLD of comparative advantage implies that the stationary distribution of comparative advantage $\hat{A}_{is}(t)$ is generalized gamma with the CDF

$$F_{\hat{A}}(\hat{a}_{is} \big| \hat{\theta}, \phi, \kappa) = G\left[\left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi}; \kappa\right],$$

where $G[x;\kappa] \equiv \gamma_x(\kappa;x)/\Gamma(\kappa)$ is the ratio of the lower incomplete gamma function and the gamma function. We show in Appendix A.3 that then the cross-sectional distribution of absolute advantage $A_{is}(t)$ is also generalized gamma, but with the CDF

$$F_A(a_{is} \big| \theta_s(t), \phi, \kappa) = G\left[\left(\frac{a_{is}}{\theta_s(t)}\right)^{\phi}; \kappa\right]$$

[37] We allow ϕ to take any real value (see Crooks 2010), including a strictly negative ϕ for a generalized inverse gamma distribution. Crooks (2010) shows that this generalized gamma distribution (Amoroso distribution) nests the Fréchet, Weibull, gamma, inverse gamma and numerous other distributions as special cases and yields the normal, log normal and Pareto distributions as limiting cases.

[38] As ϕ goes to zero, it depends on the limiting behavior of κ whether a log normal distribution or a Pareto distribution results (Crooks 2010, Table 1). The parameter restriction $\phi = 1$ clarifies that the generalized gamma distribution results when one takes an ordinary gamma distributed variable and raises it to a finite power $1/\phi$.

[39] In the estimation, we will impose the constraint that $\eta > 0$. If η were negative, comparative advantage would collapse over time for $\phi < 0$ or explode for $\phi \geq 0$. We do not constrain η to be finite.

for the strictly positive parameters

$$\hat{\theta} = \left(\phi^2/\eta\right)^{1/\phi}, \quad \theta_s(t) = \hat{\theta} Z_s(t) \quad \text{and} \quad \kappa = 1/\hat{\theta}^{\phi}.$$

These cumulative distribution functions follow from Kotz, Johnson, and Balakrishnan (1994, Ch. 17, Section 8.7).

The cross-section distributions of comparative and absolute advantage differ only in the scale parameter. For comparative advantage, the scale parameter $\hat{\theta}$ is time invariant; for absolute advantage, the scale parameter $\hat{\theta} Z_s(t)$ is time varying but country specific. Empirically, $\hat{\theta} Z_s(t)$ increases over time so that, visually, the plotted cumulative distributions of absolute advantage shift rightward over time (as can bee seen from a comparison of the cumulative distribution plots for 1967, 1987 and 2007 in Appendix **Figures A1, A2** and **A3**).

This connection between the cumulative distributions of absolute and comparative advantage allows us to estimate a GLD for generalized comparative advantage based on data for absolute advantage alone. The mean of the log of the distribution of absolute advantage is as a function of the model parameters, enabling us to identify the trend from the relation $\mathbb{E}_{st}[\ln \hat{A}_{is}(t)] = \mathbb{E}_{st}[\ln A_{is}(t)] - \ln Z_s(t)$, which follows by definition (14).[40] As we show in Appendix B, if comparative advantage $\hat{A}_{is}(t)$ follows the GLD (16), then the country specific stochastic trend $Z_s(t)$ can be identified from the first moment of the logarithm of absolute advantage using

$$Z_s(t) = \exp\left\{\mathbb{E}_{st}[\ln A_{is}(t)] - \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\}, \tag{18}$$

where $\Gamma'(\kappa)/\Gamma(\kappa)$ is the digamma function. Crucially, we can obtain detrended comparative advantage measures based on the sample analog of equation (18):

$$\hat{A}_{is}(t) = \exp\left\{\ln A_{is}(t) - \frac{1}{I}\sum_{j=1}^{I} \ln A_{js}(t) + \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\}, \tag{19}$$

which permits us to use (the observed series of) absolute advantage $A_{is}(t)$ to estimate the GLD of (the unobserved series) comparative advantage $\hat{A}_{is}(t)$.

4.4 A GMM estimator

The generalized logistic diffusion model (16) has no known closed-form transition density when $\phi \neq 0$. We therefore cannot evaluate the likelihood of the observed data and cannot perform maximum likelihood estimation. However, an attractive feature of the GLD is that it can be transformed into a stochastic process that belongs to the

[40]The expectations operator $\mathbb{E}_{st}[\cdot]$ denotes the conditional expectation for source country s at time t.

Pearson-Wong family, for which closed-form solutions of the conditional moments do exist.[41] As documented in detail in Appendix C, we construct a consistent GMM estimator based on the conditional moments of a transformation of comparative advantage, using results from Forman and Sørensen (2008).

Formally, if comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion (16) with real parameters η, σ, ϕ ($\eta > 0$), then the transformed variable

$$\hat{B}_{is}(t) = [\hat{A}_{is}(t)^{-\phi} - 1]/\phi \tag{20}$$

follows the diffusion

$$\mathrm{d}\hat{B}_{is}(t) = -\frac{\sigma^2}{2}\left[\left(\eta - \phi^2\right)\hat{B}_{is}(t) - \phi\right]\mathrm{d}t + \sigma\sqrt{\phi^2\hat{B}_{is}(t)^2 + 2\phi\hat{B}_{is}(t) + 1}\,\mathrm{d}W_{is}^{\hat{B}}(t)$$

and belongs to the Pearson-Wong family (see Appendix C.1 for the derivation). As elaborated in Appendix C.2, it is then possible to recursively derive the n-th conditional moment of the transformed process $\hat{B}_{is}(t)$ and to calculate a closed form for the conditional moments of the transformed process at time t_τ given the information set at time $t_{\tau-1}$. If we use these conditional moments to forecast the m-th power of $\hat{B}_{is}(t_\tau)$ with time $t_{\tau-1}$ information, the resulting forecast errors are uncorrelated with any function of past $\hat{B}_{is}(t_{\tau-1})$. We can therefore construct a GMM criterion for estimation. Denote the forecast error with

$$U_{is}(m, t_{\tau-1}, t_\tau) = \hat{B}_{is}(t_\tau)^m - \mathbb{E}\left[\hat{B}_{is}(t_\tau)^m \,\middle|\, \hat{B}_{is}(t_{\tau-1})\right].$$

This random variable represents an unpredictable innovation in the m-th power of $\hat{B}_{is}(t_\tau)$. As a result, the forecast error $U_{is}(m, t_{\tau-1}, t_\tau)$ is uncorrelated with any measurable transformation of $\hat{B}_{is}(t_{\tau-1})$.

A GMM criterion function based on these forecast errors is

$$g_{is\tau}(\eta, \sigma, \phi) \equiv [h_1(\hat{B}_{is}(t_{\tau-1}))U_{is}(1, t_{\tau-1}, t_\tau), \ldots, h_M(\hat{B}_{is}(t_{\tau-1}))U_{is}(M, t_{\tau-1}, t_\tau)]',$$

where each h_m is a row vector of measurable functions specifying instruments for the m-th moment condition. This criterion function has mean zero due to the orthogonality between the forecast errors and the time $t_{\tau-1}$ instruments. Implementing GMM requires a choice of instruments. Computational considerations lead us to choose polynomial vector instruments of the form $h_m(\hat{B}_{is}(t)) = (1, \hat{B}_{is}(t), \ldots, \hat{B}_{is}(t)^{K-1})'$ to construct K instruments for each of the M moments that we include in our GMM criterion.[42] In the estimation, we use $K = 2$

[41] Pearson (1895) first studied the family of distributions now called Pearson distributions. Wong (1964) showed that the Pearson distributions are stationary distributions of a specific class of stochastic processes, for which conditional moments exist in closed form.

[42] We work with a suboptimal estimator because the optimal-instrument GMM estimator considered by Forman and Sørensen (2008)

instruments and $M = 2$ conditional moments, providing us with $K \cdot M = 4$ equations and overidentifying the three parameters (η, σ, ϕ). Appendix C.3 gives further details on our GMM routine.

Standard errors of our estimates need to account for the preceding estimation of our absolute advantage $\ln A_{is}(t)$ measures. Newey and McFadden (1994) present a two-step estimation method for GMM, which accounts for generated (second-stage) variables that are predicted (from a first stage). However, our absolute advantage $\ln A_{is}(t)$ measures are not predicted variables but parameter estimates from a gravity equation: $\ln A_{is}(t)$ is a normalized version of the estimated exporter-industry-year fixed effect in equations (6) and (8). Whereas the Newey-McFadden results require a constant number of first-stage parameters, the number of parameters we estimate in our first stage increases with our first-stage sample size. Moreover, the moments in GMM time series estimation (just as the variables in OLS decay estimation in Section 3.2) involve pairs of parameter estimates from different points in time—$\ln A_{is}(t)$ and $\ln A_{is}(t + \Delta)$—and thus require additional treatments of induced covariation in the estimation. In Appendix D, we extend Newey and McFadden (1994), which leads to an alternative two-step estimation method to compute standard errors. We then use the multivariate delta method to calculate standard errors for transformed functions of the estimated parameters.

5 Estimates

Following the GMM procedure described in Section 4.4, we estimate the dissipation rate η, innovation intensity σ, and decay elasticity ϕ in the diffusion of comparative advantage, subject to an estimated country-specific stochastic trend $Z_s(t)$. The trend allows absolute advantage to be non-stationary but, because it is common to all industries in a country, the trend has no bearing on comparative advantage. Estimating the GLD permits us to test the strong distributional assumptions implicit in the OLS estimation of the discretized OU process and to evaluate the fit of the model, with or without the OU restrictions applied.

5.1 GMM results for the Generalized Logistic Diffusion

Table 2 presents our baseline GMM estimation results using moments on five-year intervals. We move to a five-year horizon, from the ten-year horizon in the OLS decay regressions in **Table 1**, to allow for a more complete description of the time-series dynamics of comparative advantage. For robustness, we also report GMM results using moments on ten-year intervals (see the Online Supplement, Table S3). Similar to the OLS decay regressions, we use measures of export advantage based on OLS gravity estimates of export capability, PPML gravity

requires the inversion of a matrix for each observation. Given our large sample, this task is numerically expensive. Also, we found local minima in our GMM criterion. At the cost of additional computation, we use a global optimization algorithm to find our estimates of ϕ, η, and σ^2. Specifically, we use Matlab's Genetic Algorithm included in the Global Optimization Toolbox.

Table 2: GMM ESTIMATES OF COMPARATIVE ADVANTAGE DIFFUSION, 5-YEAR TRANSITIONS

	OLS gravity k			PPML gravity k			$\ln$ *RCA*		
	All	LDC	Nonmanf.	All	LDC	Nonmanf.	All	LDC	Nonmanf.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimated Generalized Logistic Diffusion Parameters									
Dissipation rate η	0.256	0.270	0.251	0.180	0.166	0.147	0.212	0.194	0.174
	(0.004)***	(0.006)***	(0.005)***	(0.006)***	(0.004)***	(0.005)***	(0.006)***	(0.012)***	(0.008)***
Intensity of innovations σ	0.739	0.836	0.864	0.767	0.863	0.852	0.713	0.789	0.722
	(0.010)***	(0.017)***	(0.017)***	(0.037)***	(0.03)***	(0.045)***	(0.051)***	(0.082)***	(0.042)***
Elasticity of decay ϕ	-0.041	-0.071	-0.033	-0.009	-0.002	-0.006	0.006	-0.011	-0.045
	(0.017)**	(0.027)***	(0.018)*	(0.035)	(0.028)	(0.038)	(0.053)	(0.083)	(0.039)
Implied Parameters									
Log gen. gamma scale $\ln\hat{\theta}$	121.94	56.50	164.79	900.95	6,122.90	1,425.40	-1,410.50	708.56	99.83
	(71.526)*	(32.175)*	(120.946)	(4581.812)	(113520.900)	(11449.450)	(14980.320)	(7069.866)	(126.167)
Log gen. gamma shape $\ln\kappa$	5.017	3.991	5.439	7.788	10.873	8.360	8.641	7.467	4.464
	(0.842)***	(0.76)***	(1.077)***	(8.062)	(31.199)	(12.926)	(17.289)	(15.685)	(1.714)***
Mean/median ratio	8.203	8.203	8.293	16.897	20.469	31.716	10.256	13.872	25.286
Observations	392,850	250,300	190,630	389,290	248,360	187,390	392,860	250,300	190,630
Industry-source obs. $I \times S$	11,542	7,853	5,845	11,531	7,843	5,835	11,542	7,853	5,845
Root mean sq. forecast error	1.851	2.028	1.958	1.898	2.026	2.013	1.760	1.930	1.965
Min. GMM obj. ($\times$ 1,000)	3.27e-13	7.75e-13	7.37e-13	2.56e-12	7.82e-12	5.53e-12	6.79e-12	2.16e-11	1.65e-11

Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS and PPML gravity measures of export capability (log absolute advantage) $k = \ln A$ from (8).
Note: GMM estimation at the five-year horizon for the generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$,

$$\mathrm{d}\ln\hat{A}_{is}(t) = -\frac{\eta\sigma^2}{2}\frac{\hat{A}_{is}(t)^{\phi}-1}{\phi}\,\mathrm{d}t + \sigma\,\mathrm{d}W^{\hat{A}}_{is}(t)$$

using absolute advantage $A_{is}(t) = \hat{A}_{is}(t)Z_s(t)$ based on OLS and PPML gravity measures of export capability k from (6) and (8), and the Balassa index of revealed comparative advantage $RCA_{ist} = (X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_{\iota} X_{\iota st}/\sum_{\iota}\sum_{\varsigma} X_{\iota\varsigma t})$. Parameters η, σ, ϕ are estimated under the constraints $\ln\eta, \ln\sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (20), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (A.10). Less developed countries (LDC) as listed in the Supplementary Material (Section S.1). The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

estimates of export capability, and the Balassa RCA index.

The key distinction between the OU process in (11) and the GLD in (15) is the presence of the decay elasticity ϕ, which allows for asymmetry in mean reversion from above versus below the mean. Using OLS gravity estimates of comparative advantage (columns 1 to 3 in **Table 2**), the GMM estimate of ϕ is negative and statistically significantly different from zero at conventional levels. Negativity in ϕ implies that comparative advantage reverts to the long-run mean more slowly from above than from below. Industries that randomly churn into the upper tail of the cross section will tend to retain their comparative advantage for longer than those below the mean, affording high-advantage industries with opportunities to reach higher levels of comparative advantage as additional innovations arrive. Thus, we reject log normality in favor of the generalized gamma distribution.

The rejection of log normality, however, is not robust across measures of comparative advantage. In **Table 2**, using PPML gravity estimates of comparative advantage (columns 4 to 6) or the Balassa RCA index (columns 7 to 9) produces GMM estimates of ϕ that are not statistically significantly different from zero at conventional levels and small in magnitude.[43] These results are an initial indication that imposing log normality on comparative advantage may not strongly misrepresent reality. A second indication is that GMM estimates of the dissipation rate η for the GLD in **Table 2** are similar to those derived from the OLS decay regression in **Table 1**. In both sets of results, η takes a value of about one-quarter for OLS gravity comparative advantage, about one-sixth for PPML gravity comparative advantage, and about one-fifth for the Balassa RCA index.

To make precise comparisons of parameter estimates under alternative distributional assumptions for comparative advantage, in **Table 3** we report GMM results (for OLS gravity estimates of comparative advantage) with and without imposing the restriction that $\phi = 0$. Without this restriction (columns 1, 3, 5 and 7), we allow comparative advantage to have a generalized gamma distribution; with this restriction (columns 2, 4, 6, and 8), we impose log normality. Estimates for the dissipation rate η and the innovation intensity σ are nearly identical in each pair of columns. Parameter stability implies that the special case of the OU process captures the broad persistence and overall variability of comparative advantage. Because the decay elasticity ϕ also determines the shape of the stationary distribution of the GLD, two processes that have identical values of η but distinct values of ϕ will differ in the shape of their generalized gamma distributions. We see in **Table 3** that the implied mean/median ratios are modestly higher for columns where ϕ is unrestricted (and found to be small and negative) versus columns in which ϕ is set to zero. The estimated mean-median ratio increases from $6.2 - 7.0$ under the constrained estimation of the OU process to $8.2 - 8.3$ under the unconstrained case. The extension to a GLD thus appears to help explain the export concentration in the upper tail documented in subsection 3.1 above.

Table 3 also allows us to see the impact on the GMM parameter estimates of altering the time interval on

[43] As shown in Appendix **Table A2**, we obtain similar results for the 1984 to 2007 period when we use two- or three-digit SITC revision 2 industries, thus establishing the robustness of the GMM results under alternative industry aggregation.

Table 3: GMM ESTIMATES OF COMPARATIVE ADVANTAGE DIFFUSION, UNRESTRICTED AND RESTRICTED

	OLS gravity k, 5-year transitions						OLS k, 10-yr. trans.	
	Full sample		LDC exp.		Non-manuf.		Full sample	
		$\phi = 0$		$\phi = 0$		$\phi = 0$		$\phi = 0$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Generalized Logistic Diffusion Parameters								
Dissipation rate η	0.256	0.263	0.270	0.274	0.251	0.256	0.264	0.265
	(0.004)***	(0.003)***	(0.006)***	(0.003)***	(0.005)***	(0.004)***	(0.004)***	(0.003)***
Intensity of innovations σ	0.739	0.736	0.836	0.831	0.864	0.860	0.569	0.568
	(0.01)***	(0.008)***	(0.017)***	(0.012)***	(0.017)***	(0.014)***	(0.007)***	(0.006)***
Elasticity of decay ϕ	-0.041		-0.071		-0.033		-0.029	
	(0.017)**		(0.027)***		(0.018)*		(0.014)**	
Implied Parameters								
Log gen. gamma scale $\ln \hat{\theta}$	121.940		56.502		164.790		202.430	
	(71.526)*		(32.175)*		(120.946)		(131.469)	
Log gen. gamma shape $\ln \kappa$	5.017		3.991		5.439		5.781	
	(0.842)***		(0.76)***		(1.077)***		(0.968)***	
Mean/median ratio	8.203	6.691	8.203	6.222	8.293	7.036	7.281	6.588
Observations	392,850	392,850	250,300	250,300	190,630	190,630	335,820	335,820
Industry-source obs. $I \times S$	11,542	11,542	7,853	7,853	5,845	5,845	11,213	11,213
Root mean sq. forecast error	1.851	1.726	2.028	1.821	1.958	1.859	1.876	1.799
Min. GMM obj. ($\times$ 1,000)	3.27e-13	2.87e-12	7.75e-13	2.08e-11	7.37e-13	8.99e-12	3.03e-12	5.92e-12

Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).

Note: GMM estimation at the five-year (ten-year) horizon for the generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$,

$$\mathrm{d} \ln \hat{A}_{is}(t) = -\frac{\eta\sigma^2}{2} \frac{\hat{A}_{is}(t)^{\phi} - 1}{\phi} \, \mathrm{d}t + \sigma \, \mathrm{d}W_{is}^{\hat{A}}(t)$$

using absolute advantage $A_{is}(t) = \hat{A}_{is}(t) Z_s(t)$, unrestricted and restricted to $\phi = 0$. Parameters η, σ, ϕ are estimated under the constraints $\ln \eta, \ln \sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (20), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (A.10). Less developed countries (LDC) as listed in the Supplementary Material (Section S.1). The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

which moment conditions are based. Columns 7 and 8 show results for ten-year intervals, which compare to the preceding columns whose results are for five-year intervals.[44] Whereas estimates for the dissipation rate η are nearly identical for the two time horizons, estimates for the innovation intensity σ become smaller when we move from five- to ten-year intervals. Similar to attenuation bias driving estimates of persistence to zero in auto-regression models, measurement error may deliver larger values of σ at shorter horizons.[45]

5.2 Model fit I: Matching dynamic transition probabilities

We next evaluate the performance of the model by assessing how well the GLD replicates the churning of export industries in the data. Using estimates based on the five-year horizon from column 1 in **Table** 2, we simulate trajectories of the GLD. In the simulations, we predict the model's transition probabilities over the one-year horizon across percentiles of the cross-section distribution. We deliberately use a shorter time horizon for the simulation than the five-year horizon used for estimation to assess moments that we did not target in GMM. We then compare the model-based predictions to the empirical transition probabilities at the one-year horizon.

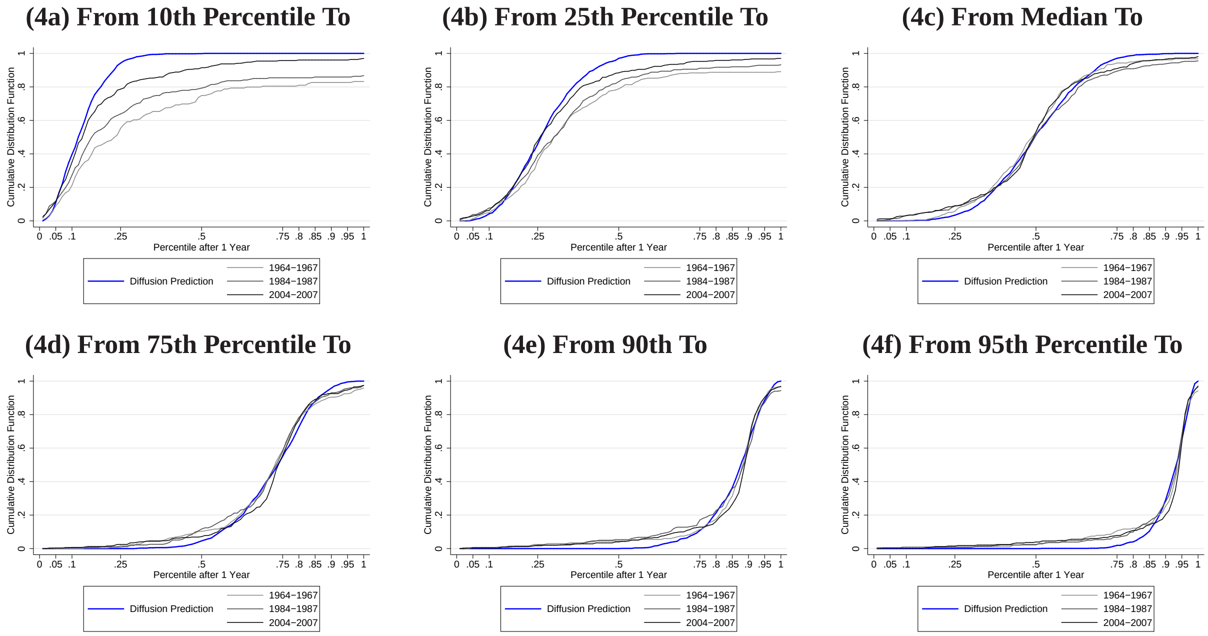
Figure 4 shows empirical and model-predicted conditional cumulative distribution functions for annual transitions of comparative advantage. We select percentiles in the start year: the 10th and 25th percentile, the median, the 75th, 90th and 95th percentile. The left-most upper panel in **Figure 4**, for example, considers industries that were at the 10th percentile of the cross-section distribution of comparative advantage in the start year; panel **Figure 4c** shows industries that were at the median of the distribution in the start year. Each curve in a panel then plots the conditional CDF for the transitions from the given percentile in the start year to any percentile of the cross section one year later. By design, data that are re-sampled under an iid distribution would show up at a 45-degree line, while complete persistence of comparative advantage would make the CDF a step function. To characterize the data, we use three windows of annual transitions: the mean annual transitions during the years 1964-67 at the beginning of our sample period, the mean annual transitions during the years 1984-87 at the middle of our sample, and the mean annual transitions during the years 2004-07 at the end of the sample. These transitions are shown in gray. Our GLD estimation constrains parameters to be constant over time, so the model predicted transition probabilities give rise to a time-invariant CDF shown in blue.

The five-year GLD performs well in capturing the annual dynamics of comparative advantage for most industries. As **Figure 4** shows, the model-predicted conditional CDF's tightly fit their empirical counterparts for industries at the median and higher percentiles in the start year. It is only in the lower tail, in particular around the 10th percentile, that the fit of the GLD model becomes less close, though the model predictions are more

[44]The Online Supplement (Table S3) presents GMM results for moments on ten-year intervals using PPML gravity estimates of comparative advantage and the Balassa RCA index.

[45]In the limit when σ becomes arbitrarily large, the GLD would exhibit no persistence, converging to an iid process.

Figure 4: **Diffusion Predicted Annual Transitions**



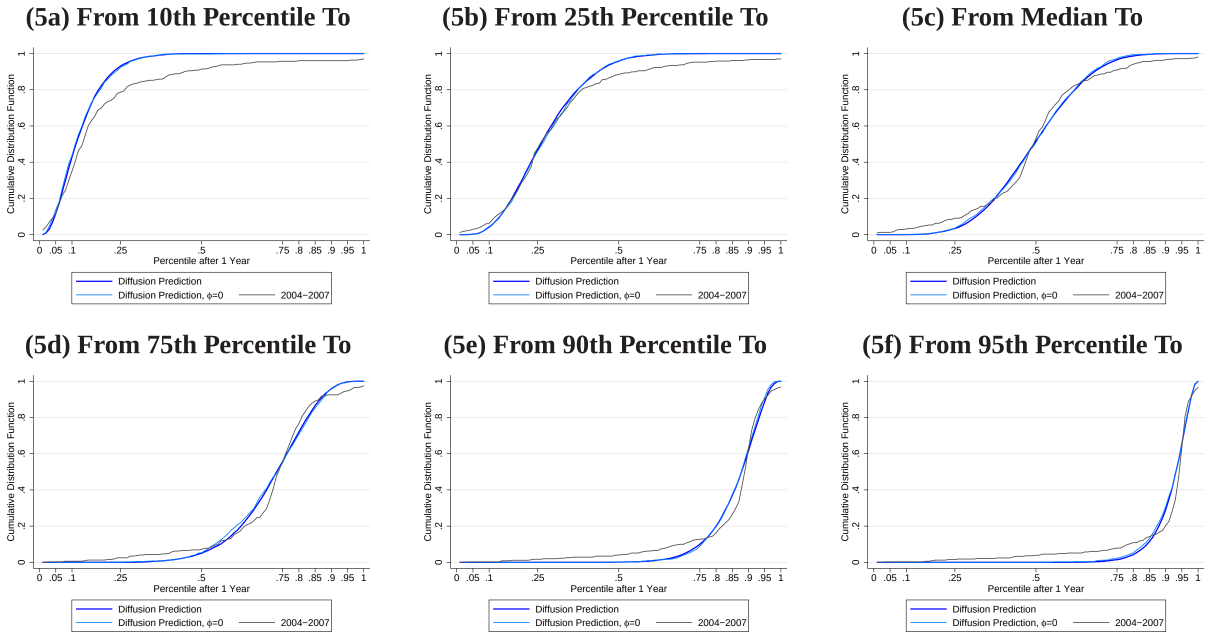
Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: Predicted cumulative distribution function of comparative advantage $\hat{A}_{is,t+1}$ after one year, given the percentile (10th, 25th, median, 75th, 90th, 95th) of current comparative advantage $\hat{A}_{is,t}$. Predictions based on simulations using estimates from Table 2 (column 1). Observed cumulative distribution function from mean annual transitions during the periods 1964-1967, 1984-1987, and 2004-2007.

comparable to the data in later than in earlier periods. Country-industries in the bottom tail have low trade volumes, especially in the early sample period, meaning that estimates of the empirical transition probabilities in the lower tail are not necessarily precisely estimated and may fluctuate more over time. **Figure 4** indicates that the dynamic fit becomes relatively close for percentiles at around the 25th percentile. The discrepancies in the lowest tail notwithstanding, for industries with moderate to high trade values, which account for the bulk of global trade, the model succeeds in matching empirical transition probabilities.

The transition probabilities implied by the GLD also allow us to assess how well a simple OU process approximates trade dynamics. In a statistical horse race between the unconstrained GLD and the OU process, the former wins—at least for OLS gravity estimates of comparative advantage—because we reject that $\phi = 0$ in **Table 3**, columns 1 to 3. Yet, estimating the GLD is substantially more burdensome than estimating the simple discretized linear form of the OU process. For both empirical and theoretical modeling, it is helpful to understand how much is lost by imposing log normality on comparative advantage.

Following **Figure 4**, we simulate trajectories of the GLD, once from estimates with ϕ unconstrained and once from estimates with $\phi = 0$, using coefficients from columns 1 and 2 in **Table 3**. The simulations predict the tran-

Figure 5: **Diffusion Predicted Annual Transitions, Constrained and Unconstrained ϕ**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: Predicted cumulative distribution function of comparative advantage $\hat{A}_{is,t+1}$ after one year, given the percentile (10th, 25th, median, 75th, 90th, 95th) of current comparative advantage $\hat{A}_{is,t}$. Predictions based on simulations using estimates from Table 2 (column 1) and Table 3 (column 2, $\phi = 0$). Observed cumulative distribution function from mean annual transitions during the period 2006-2009.

sition probabilities over the one-year horizon across percentiles of the cross-section distribution. **Figure 5** shows the empirical cumulative distribution functions for annual transitions of comparative advantage over the full sample period 1962-2007 (in gray) and compares the empirical distribution to the two model-predicted cumulative distribution functions (light and dark blue), where the fit of the unconstrained GLD model (dark blue) is the same as depicted in **Figure 4** above. As in **Figure 4**, each panel in **Figure 5** considers industries that were at a given percentile of the cross-section distribution of comparative advantage in the start year. Each curve in a panel shows the conditional CDF for the transitions from the given percentile in the start year to any percentile of the cross section one year later. For all start-year percentiles, the model-predicted transitions hardly differ between the constrained specification (light blue) and the unconstrained specification (dark blue). When alternating between the two models, the shapes of the model-predicted conditional CDF's are very similar, even in the upper tail. In the lower tail, where the GLD produces the least tight dynamic fit, the constrained OU specification performs no worse than the unconstrained GLD. The simple OU process thus appears to approximate the empirical dynamics of trade in a manner that is very close to the GLD extension.

5.3 Model fit II: Matching the empirical cross-section distribution

Next, we evaluate the fit of our GLD by examining how well the GMM parameter estimates describe the cross-section distribution of comparative advantage. We have given the GMM estimator a heavy burden: to fit the export dynamics across 90 countries for 46 years using only three time-invariant parameters (η, σ, ϕ), conditional on stochastic country-wide growth trends. Because the moments we use in GMM estimation reflect the time-series behavior of country-industry exports, our estimator fits the diffusion of comparative advantage but not its stationary cross-section distribution. We can therefore use the stationary generalized gamma distribution implied by the GLD to assess how well our model captures the stability of the heavy tails of export advantage observed in the repeated cross-section data. For this comparison, we use the benchmark estimates from **Table 2** in column 1. (We obtain similar results for ϕ constrained to zero in column 2 of **Table 3**.)

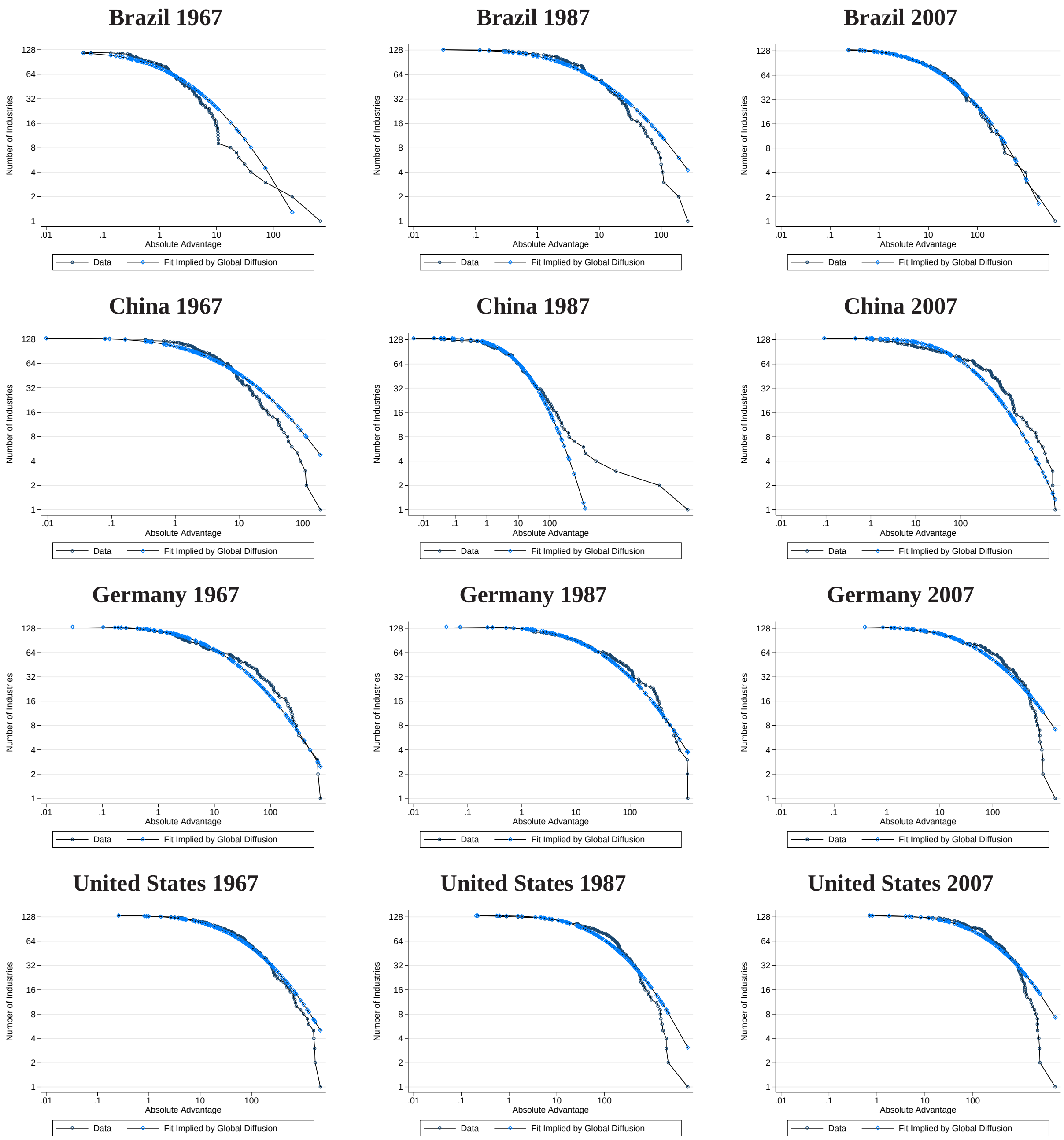
For each country in each year, we project the cross-section distribution of comparative advantage implied by the parameters estimated from the diffusion and compare it to the empirical distribution. To implement this validation exercise, we need a measure of $\hat{A}_{ist}$ in (14), the value of which depends on the unobserved country-specific stochastic trend Z_{st}. This trend accounts for the observed horizontal shifts in distribution of log absolute advantage over time, which may result from country-wide technological progress, factor accumulation, or other sources of aggregate growth. In the estimation, we concentrate out Z_{st} by (18), which allows us to estimate its realization for each country in each year. Combining observed absolute advantage A_{ist} with the stochastic-trend estimate allows us to compute realized values of comparative advantage $\hat{A}_{ist}$.

To gauge the goodness of fit of our specification, we first plot our empirical measure of absolute advantage A_{ist}. To do so, following the earlier exercise in **Figure 1**, we rank order the data and plot for each country-industry observation the level of absolute advantage (in log units) against the log number of industries with absolute advantage greater than this value, which is equal to the log of one minus the empirical CDF. To obtain the simulated distribution resulting from the parameter estimates, we plot the global diffusion's implied stationary distribution for the same series. The diffusion implied values are constructed, for each level of A_{ist}, by evaluating the log of one minus the predicted generalized gamma CDF at $\hat{A}_{ist} = A_{ist}/Z_{st}$. The implied distribution thus uses the global diffusion parameter estimates (to project the scale and shape of the CDF) as well as the identified country-specific trend Z_{st} (to project the position of the CDF).

Figure 6 compares plots of the actual data against the GLD-implied distributions for four countries in three years, 1967, 1987, 2007. **Figures A4, A5** and **A6** in the Appendix present plots in these years for the 28 countries that are also shown in **Figures A1, A2** and **A3**.[46] While **Figures A1** to **A3** depicted Pareto and log normal

[46]Because the country-specific trend Z_{st} shifts the implied stationary distribution horizontally, we clarify fit by cutting the depicted part of that single distribution at the lower and upper bounds of the specific country's observed support in a given year.

Figure 6: **Diffusion Predicted and Observed Cumulative Probability Distributions of Absolute Advantage for Select Countries in 1967, 1987 and 2007**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: The graphs show the observed and the predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (15) in Table 2 (parameters η and ϕ in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (18), which horizontally shifts the distributions but does not affect their shape.

maximum likelihood estimates of each individual country's cross-sectional distribution by year (number of parameters estimated = number of countries × number of years), our exercise now is vastly more parsimonious and based on a fit of the time-series evolution, not the observed cross sections. **Figure 6** and Appendix **Figures A4** to **A6** show that the empirical distributions and the GLD-implied distributions have the same concave shape and horizontally shifting position. Considering that the shape of the distribution depends on only two parameters for all country-industries and years, the GLD-predicted distributions are remarkably accurate. There are important differences between the actual and predicted plots in only a few countries and a few years, including China in 1987, Russian Federation in 1987 and 2007, Taiwan in 1987, and Vietnam in 1987 and 2007. Three of these cases involve countries transitioning away from central planning, suggesting periods of economic disruption.

There are other, minor discrepancies between the empirical distributions and the GLD-implied distributions that merit further attention. In 2007 in a handful of countries in East and Southeast Asia—China, Japan, Rep. Korea, Malaysia, Taiwan, and Vietnam—the empirical distributions exhibit less concavity than the generalized gamma distributions (or the log normal for that matter). These countries show more mass in the upper tail of comparative advantage than they ought, implying that they excel in too many industries, relative to the norm. It remains to be investigated whether these differences in fit are associated with conditions in the countries themselves or with the particular industries in which these countries tend to specialize.

The noticeable deviations for some countries in certain years notwithstanding, across countries and for the full sample period the percentiles of the country-level distributions of comparative advantage are remarkably stable for each of our three measures of comparative advantage. This stability suggests that there is a unifying global and stationary distribution of comparative advantage. Our estimates of the GLD time series imply shape-parameter values of a generalized gamma CDF, and those predicted shape parameters tightly fit the relevant percentiles of the global comparative-advantage distribution.[47]

6 Simulations

Having established the dynamic properties of comparative advantage, we next consider their relevance for quantitative trade analysis. We examine how accounting for these dynamics affects the results of a common type of counterfactual exercise. A distinguishing feature of quantitative trade models is that they allow for shocks that are asymmetric across industries—the existence of such shocks is in part what motivates multi-sector trade models in the first place. If there is churning in comparative advantage, the impacts of such industry-specific treatments may be fleeting. This impermanence arises because prominent industries that are treated today are likely to be-

[47] The Online Supplement (Figure S11) shows percentile plots for OLS- and PPML-based measures of export capability over time and the fit of our according GLD estimates to those percentiles.

come less prominent tomorrow, and the heavy tails of the distribution can dictate that it is mostly shocks to the prominent industries that matter in the aggregate.[48] The long-run effect of a permanent reduction in trade costs may therefore differ substantially from the short-run effect.

6.1 Counterfactual exercises

As an application, we consider a counterfactual in which China's top export industries in 1990—either its top-5 or top-50 industries—have their export costs reduced permanently by 10%, which is equivalent to countries in the rest of the world lowering their barriers on selected imports from China. In the presence of stochastic comparative advantage, the change in equilibrium outcomes due to this reduction in trade costs becomes a random variable. Unlike standard counterfactuals considered in the trade literature (e.g., Alvarez and Lucas 2007 and Dekle, Eaton, and Kortum 2007), we must now solve for equilibrium repeatedly across many simulated potential paths for comparative advantage in order to characterize the effect of a trade-cost reduction.

To measure the "typical" impact of a change in trade costs, we compute an average treatment effect, or more precisely an average path for the treatment effect. Specifically, we solve for the counterfactual equilibrium at each moment in time across 10,000 simulated paths of comparative advantage, with and without the change in trade costs. For each simulated path of comparative advantage, we compute the percent difference in equilibrium outcomes between the counterfactual with the trade-cost reduction and the counterfactual without the trade-cost reduction. This percent change measures the effect of the trade-cost drop conditional on the simulated path of comparative advantage. Our measure of the average treatment effect is then the average of this percent change over many simulations.

To prepare our simulations of comparative-advantage paths for a balanced group of countries and industries, we need to construct a set of unobserved variables that are now required because we vary the productivity fundamentals of the EK model, consistent with a GLD of comparative advantage. Appendix E presents the procedure in detail. We need to infer self trade at the level of industries and countries to compute industry-level expenditure for our balanced group of importers and exporters. Existing data sets do not offer production information at the level of disaggregated industries. However, we show that a country-industry's share in the country's total self trade can be inferred from gravity fixed-effect estimates under the EK model. We can then combine the estimates of the shares of self trade with the observed country-level self trade level using UNIDO and WIOD data and obtain measures of self trade at the level of industries and countries in the base year 1990. To account for a

[48]The importance of shocks to prominent industries, or firms, for aggregate outcomes has been called "granularity." While most results on granularity are stated for power-law distributions, they arguably carry over to our case of a log normal cross sectional distribution of industry capabilities. Gabaix (2011, p. 744) states: "Though the benchmark case of Zipf's law is empirically relevant, and theoretically clean and appealing, many arguments [about granularity] do not depend on it. ... For instance, if the distribution of firm sizes were lognormal with a sufficiently high variance, then quantitatively very little would change."

few missing gravity fixed-effect estimates, we interpolate and extrapolate to balance the group of importers and exporters. To balance global trade, we need a rest-of-world entity in addition, for which we also have to construct self trade by industry. We follow the same idea as for the gravity-sample countries and construct self trade from available gravity estimates. We consider the rest of the world a synthetic entity, and predict a complete hypothetical set of bilateral gravity covariates (distance, common language, common border, and so forth) as if the rest of the world were a single entity. For this purpose, we linearly project exporter capabilities, import propensities and trade costs for the single rest-of-world entity on convex combinations of the covariates for the countries behind the rest-of-world entity; from those projections we obtain synthetic gravity fixed-effect estimates, form which we then build self trade shares and ultimately industry-level self trade in the rest of the world as described above. Given simulated comparative advantages (and the resulting Fréchet location parameters) and given trade cost changes, we then solve for equilibrium year by year and path by path in wages, expenditure shares by industry and country, and prices. For given realizations of comparative advantage, we can now characterize the difference between the initial (in 1990) and the counterfactual equilibrium using the exact hat algebra of Dekle, Eaton, and Kortum (2007) path by path.

To isolate how churning in comparative advantage influences the effect of a reduction in trade costs, we consider three scenarios for relative industry productivity. The first, which we call the *static equilibrium*, represents the usual exercise in the trade literature. We hold all fundamentals—including comparative advantage—fixed at their 1990 levels and compute a counterfactual equilibrium in which the only change is the reduction in trade costs. The second scenario, which we describe as the *transition path*, initializes comparative advantage at 1990 levels, and then allows it to evolve stochastically over time according to our estimated GLD process. This exercise permits us to see how churning in comparative advantage affects equilibrium outcomes over time relative to the standard counterfactual captured by our static-equilibrium scenario. Finally, we consider a *steady state* scenario, in which we sample initial conditions from the stationary distribution of comparative advantage and then again allow comparative advantage to evolve stochastically over time according to our GLD process.

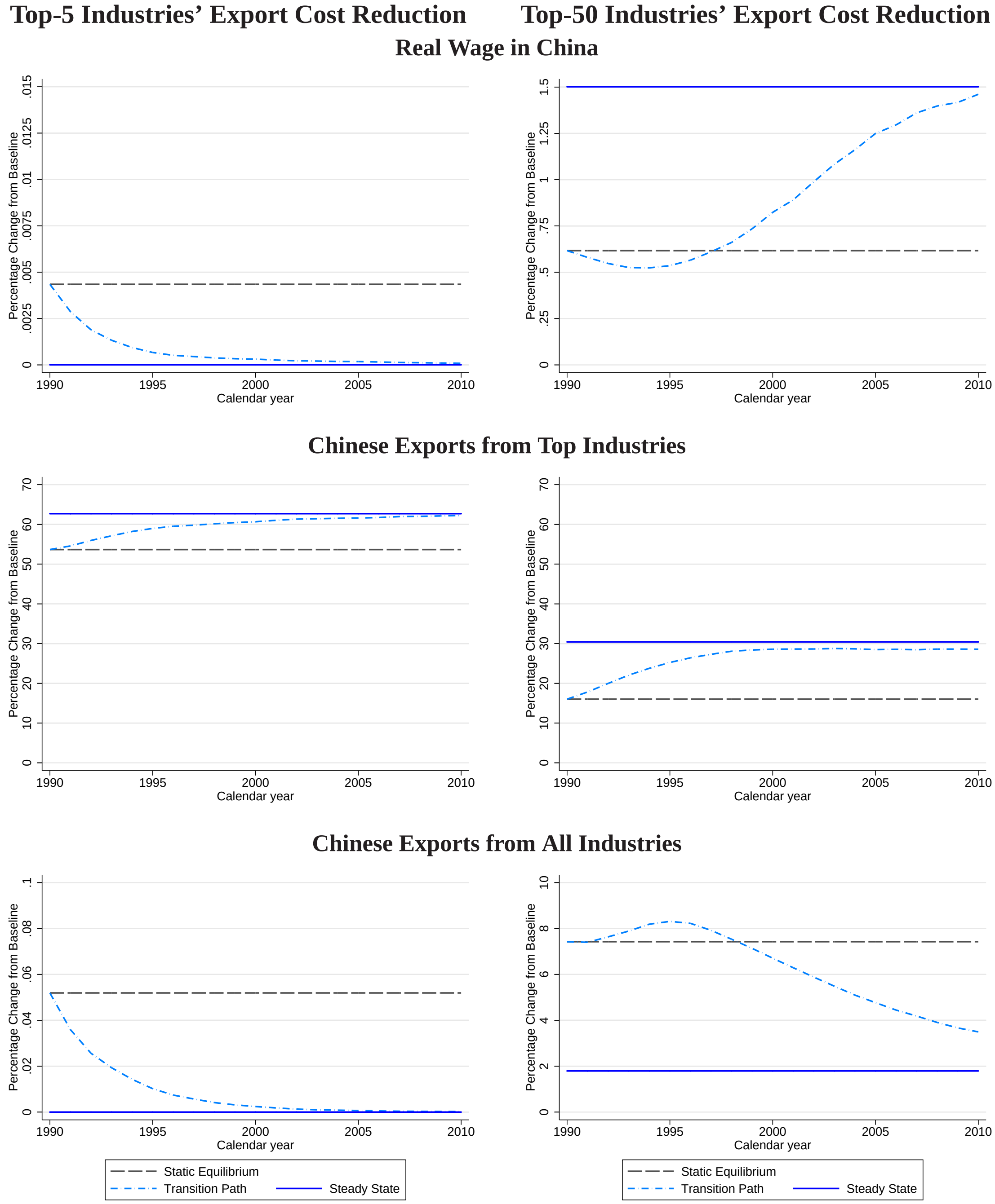
Drawing initial values from the distribution for each simulation eliminates the influence of initial conditions—that is China's top export industries in 1990 will not be its top industries when averaging over many draws from the distribution—and allows us to characterize long-run outcomes in the presence of stochastic comparative advantage. We emphasize that this long-run equilibrium is far from static. Comparative advantage continues to evolve dynamically. Averaging over many initial draws and the period-by-period change in comparative advantage following each draw causes the average treatment effect to be stable. Even though for each simulation the equilibrium differs period by period, on average there is no variation in the treatment effect across time. For all simulations, we hold trade balances fixed at their 1990 levels in order to isolate the importance of stochastic comparative advantage.

Figure 7 shows the average percent change in equilibrium outcomes due to a reduction in Chinese export costs in each of these relative productivity scenarios. The first row shows the effect on real wages in China, the second row shows the effect on exports in treated industries, and the final row shows the effect on aggregate Chinese exports. The left column shows the impact of a narrow trade-cost reduction that affects only the top-5 export industries in 1990, while the right column shows the impact of a broad trade-cost reduction that affects the top-50 industries (out of 133). Within each panel, the black dashed line corresponds to the static equilibrium, the light blue dash-dot line corresponds to the transition path, and the solid blue line corresponds to the steady state. Whereas static-equilibrium values are constant over time, steady-state values appear to be constant because of the averaging over simulations. The transition path shows the average path from the initial static equilibrium to the steady-state equilibrium, or how long it takes for the dynamic evolution of comparative advantage to wash out the impact of initial conditions on the average treatment effect.

We see immediately that the short-run impact of the trade cost reduction, as captured by the static equilibrium, can differ substantially from the long-run impact, as captured by the steady-state scenario. Consider first the narrow trade-cost reduction in the left column. On impact, China's real wage rises, exports of treated industries increase, and aggregate exports expand. This initial impact is shown both by the values for the static equilibrium in all periods and by the values of the transition path in the initial period. Under stochastic comparative advantage, the treatment effect on macro outcomes decays over time: in the transition-path scenario, the real wage and aggregate exports decline. The effect on both outcomes becomes negligible in the long run, convergence to which is largely complete after 10 years and fully complete after 20 years. When the shock initially arrives, it is targeted towards high comparative-advantage industries which make up a large portion of Chinese exports. However, churning in comparative advantage implies that the reduction in trade costs becomes less targeted over time. An industry that was initially in the top of the comparative advantage distribution will tend to shuffle to a new position in the distribution, which makes long-run rankings independent of initial rankings. It is this reshuffling that makes the steady-state scenario immune to the treatment, since the initial draw of comparative advantage pays no heed to the industries that topped China's export rankings in 1990. Note that, although the effect is fleeting at the macro level, there is a permanent effect on exports within those industries for which trade costs were reduced. In fact, the impact on exports within these industries is increasing over time, since the decline of the real wage implies falling export prices and hence rising exports.

Consider next the broad trade-cost reduction in the right column. The impact of the falling trade costs on treated-industry exports in the second row and on aggregate exports in the third row are qualitatively similar to those for the more targeted reduction in industry trade costs in the left column. The impact on exports by treated

Figure 7: **Simulated Outcomes in China after 10-percent Export Trade Cost Reduction**



Note: Simulations of GLD after a variable trade cost reduction by 10% for Chinese exports (but not Chinese imports) in top-5 (left column) or top-50 (right column) industries by comparative advantage in 1990 China. Graphs show the average over 10,000 simulations for percentage deviations in equilibrium outcomes between a counterfactual with reduced trade costs and a counterfactual with constant trade costs (*Baseline*). The *Static Equilibrium* curves show the counterfactual impact of the trade cost reduction if comparative advantage remained at the 1990 equilibrium levels in all countries. The *Transition Path* curves show the average counterfactual effect of the trade cost reduction starting from observed 1990 comparative advantages and simulating forward under the estimated GLD. The *Steady State* curves show the counterfactual effect of the trade cost reduction when initial conditions are drawn from the stationary comparative advantage distribution implied by the estimated GLD. To isolate the effects of the GLD, all computations take the empirically observed Chinese trade balance as exogenous and compute the implied Chinese wages and price levels.

industries rises over time, while the macro impact on aggregate exports dissipates over time. By contrast, the path for the real wage differs sharply between the two experiments. Whereas real-wage impacts decay over time when the narrow set of industries is treated, real-wage impacts actually increase over time when the broad set of industries is treated. Next, we unpack the forces behind these rich transition dynamics, which are of course absent in conventional applications of quantitative trade analysis.

6.2 The dissipation of treatment effects in a quantitative trade model

To obtain intuition as to why these treatment effects decay over time, assume that s is a small country and therefore has a negligible influence on competitiveness indices and on aggregate expenditure in country d. Let X_{isd} be observed expenditure by destination country d on goods from source country s in industry i in 1990 and $\pi_{isd} \equiv X_{isd}/\sum_{\varsigma} X_{i\varsigma d}$ be the expenditure share within industry i. Let a hat on a variable denote the ratio of the counterfactual outcome and the variable's observed 1990 level. We denote the permanent change in trade costs with $\hat{\tau}_{isd} \equiv \hat{\tau}_{isdt}$ since it is constant over time. Counterfactual trade flows relative to their initial levels satisfy

$$\hat{X}_{isdt} = \frac{(\hat{\tau}_{isd}\hat{w}_{st})^{-\theta}\hat{A}_{ist}}{\hat{\Phi}_{idt}}\hat{E}_{dt},$$

where $\hat{A}_{ist} = \hat{\underline{q}}_{ist}^{\theta}$ is the change in comparative advantage and $\hat{\Phi}_{idt} \equiv \sum_s \pi_{isd}(\hat{\tau}_{isd}\hat{w}_{st})^{-\theta}\hat{A}_{ist}$ is the change in competitiveness in industry i within destination d.

Given that country s is small, $\hat{\Phi}_{idt} = 1$ and $\hat{E}_{dt} = 1$, so the counterfactual level of exports at time t is

$$X'_{st} = \sum_{d \neq s}\sum_{i} \pi_{isd}(\hat{\tau}_{isd}\hat{w}_{st})^{-\theta}\hat{A}_{ist}\mu_{id}E_d$$

where $\mu_{id} = \sum_{\varsigma} X_{i\varsigma d}/\sum_{\iota}\sum_{\varsigma} X_{\iota\varsigma d}$ is the (constant) Cobb-Douglas expenditure share. Using the sample analog of the expectations operator $\mathbb{E}[\cdot] \equiv (1/I)\sum_i(\cdot)$ over industries, we can invoke the properties of the covariance to decompose exports as

$$X'_{st} = \sum_{d \neq s}\left[\underbrace{\mathbb{E}\left[\pi_{isd}\hat{\tau}_{isd}^{-\theta}\right]\ \mathbb{E}\left[\hat{A}_{ist}\right]}_{\text{Direct effect on impact}} + \underbrace{\mathbb{C}ov\left(\pi_{isd}\hat{\tau}_{isd}^{-\theta}, \hat{A}_{ist}\right)}_{\text{Change in direct effect}}\right]\underbrace{\hat{w}_{st}^{-\theta}}_{\text{Indirect wage effect}}\mu_{id}E_d,$$

where $\mathbb{C}ov(\cdot)$ is the associated covariance operator.

There are two effects that determine the counterfactual level of exports. The first is a direct effect, which is summarized by the expression in brackets. The product of expectations inside the brackets captures the direct effect on impact and the covariance term captures the change in the direct effect over time. The product of expectations in the bracket is a constant because the change in trade costs $\hat{\tau}_{isd}$ is constant over time, and so are initial expenditure shares π_{isd}; mean comparative advantage $\mathbb{E}[\hat{A}_{ist}]$ is also constant since the distribution of comparative advantage is stationary. The covariance term in the brackets is equal to zero on impact because initially there is no change in productivities and therefore no correlation with the term $\pi_{isd}\hat{\tau}_{isd}^{-\theta}$. However, industries with high comparative advantage will tend to lose comparative advantage over time. These industries will also have reduced trade costs and high initial expenditure shares. As a result, this covariance term—the change in the direct effect—is negative and increasingly negative over time, which implies that the overall direct effect is strongest on impact and decays over time.

The second effect—captured by the term outside of the brackets—is an indirect general-equilibrium effect that operates through the wage. Through the direct effect, a trade-cost reduction drives up exports and therefore increases the demand for labor in country s. If this increase in labor demand is large, the wage will rise. But a rising wage increases the cost of country s exports and leads to an offsetting reduction in exports. That is, if the trade-cost change affects a large portion of the economy, it will lead to a rising wage in general equilibrium, which will dampen the increase in exports. Over time, as the direct effect decays due to churning, the wage will tend to fall and this indirect wage effect will tend to raise exports. These dynamics are such that exports can rise or fall over time, depending on the relative importance of the direct and indirect effects. Since the change in trade costs underlying the results in the left column of **Figure 7** is narrowly limited to the top-5 industries in 1990, the indirect wage effect is small. The dynamics are driven primarily by decay in the direct effect due to churning in comparative advantage. As a result, the impact of the trade-cost reduction mainly reflects the direct effect and decays over time as comparative advantage churns.

By contrast, the right column of **Figure 7** (for the broad reduction in trade costs that affects the top-50 industries in 1990) shows how outcomes can change when the secondary indirect effect on the wage is large. This shock impacts a large portion of industries, so it has a non-negligible indirect wage effect. The implied increase in demand for labor within China drives up the wage and, because so many industries are treated, continues to do so even as churning in comparative advantage alters the composition of top industries. Although rising wages imply rising prices, because China is an open economy with an aggregate self-trade share less than one, the net result is an increase in the real wage. The real wage rises on impact, and rises further over time. Aggregate exports therefore increase on impact and then decay by about 75% in the long run. Over time, the combination of decay in the direct effect (due to churning in comparative advantage) and the lack of decay in the offsetting indirect wage effect creates a hump shape in the time path of aggregate exports.

This conceptual exercise demonstrates that our main conclusion—namely that the effects from permanent changes in trade costs can be fleeting—is robust to the industry-wide breadth of the change in trade costs, despite the presence of general equilibrium effects that generate rich transition dynamics. These transition dynamics are missing from standard quantitative trade analysis, as is a mechanism that would generate differences between impacts in the static equilibrium.

In summary, our simulation results show that conclusions from standard counterfactual exercises in trade can change significantly once we account for randomness in comparative advantage. The effect of an uneven reduction in industry trade costs—which describes many trade liberalization episodes—can be transitory given the perpetual churning in comparative advantage. When performing counterfactual analysis in quantitative trade models, it is therefore crucial to account for the dynamics of comparative advantage.

7 Conclusion

Quantitative analysis of global general equilibrium models is a vibrant area of research, due in part to the success of the Eaton and Kortum (2002) model of Ricardian trade. The primitives in the EK model are the parameters of the distribution for industry productivity, which pin down country export capabilities and hence comparative advantage. Despite the importance of these primitives in driving international trade, much current analysis of changes in trade policy leaves comparative advantage in the background by treating it as static. Our goals in this paper are, first, to characterize the dynamic empirical properties of comparative advantage; second, to show that these properties are consistent with a unifying family of estimable stochastic processes; and third, to demonstrate how the stochastic nature of comparative advantage materially affects the counterfactual policy exercises that have become central to quantitative trade modeling.

Our analysis starts from two strong empirical regularities in trade that economists have studied mostly in isolation. Many papers have noted the tendency for countries to concentrate their exports in a relatively small number of industries. Our first contribution is to show that this concentration arises from a heavy-tailed distribution of industry export capability that is approximately log normal and whose shape is stable across countries, industries, and time. Likewise, the trade literature has detected in various forms a tendency for mean reversion in national industry productivities. Our second contribution is to establish that mean reversion in export capability, rather than indicative of convergence in productivities and degeneracy in comparative advantage, is instead consistent with a stationary stochastic process, whose properties are common across borders and industries. In literatures on the growth of cities and the growth of firms, economists have used stochastic processes to study the determinants of the long-run distribution of sizes. Our third contribution is to develop an analogous empirical framework for identifying the parameters that govern the stationary distribution of export capability. One result of this analysis is that log normality offers a reasonable approximation and a discrete-time version for the analogous stochastic process can be estimated with straightforward linear regression. A fourth contribution is to quantify the time horizon at which policy or cost shocks to a country's exports dissipate: even substantive interventions in targeted industries become largely irrelevant for export flows in a matter of a decade. Allowing comparative advantage to be stochastic differs strongly from most current approaches in the literature. Our fifth contribution is to show that when one incorporates stochastic comparative advantage into standard counterfactual exercises the impact of industry-specific treatments (such as changes in trade policy that favor some industries over others) can be fleeting, with initial impacts decaying substantially within 10 to 20 years.

In the stochastic process that we estimate, country export capabilities evolve independently across industries, subject to controls for aggregate country growth, and independently across countries, subject to controls for global industry growth. Recent work in trade theory examines how innovations to productivity are transmitted across space and time. Our analysis can be extended straightforwardly to allow for such interactions. The Ornstein-Uhlenbeck process generalizes to a multivariate diffusion, in which stochastic innovations to an industry in one country also affect related industries in the same economy or the same industry in a nation's trading partners. Because of the linearity of the discretized OU process, it is feasible to estimate such interactions while still identifying the parameters that characterize the stationary distribution of comparative advantage. An obvious next step is to model diffusions that allow for such intersectoral and international productivity linkages.

References

ABOWD, J. M., R. H. CREECY, AND F. KRAMARZ (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," *LEHD Technical Paper*, 2002-06.

ACEMOGLU, D., AND F. ZILIBOTTI (2001): "Productivity Differences," *Quarterly Journal of Economics*, 116(2), 563–606.

AÏT-SAHALIA, Y., L. P. HANSEN, AND J. A. SCHEINKMAN (2010): "Operator Methods for Continuous-Time Markov Processes," in *Handbook of Financial Econometrics*, ed. by Y. Aït-Sahalia, and L. P. Hansen, vol. 1 of *Handbooks in Finance*, chap. 1, pp. 1–66. North-Holland/Elsevier, Amsterdam.

ALVAREZ, F., AND R. E. J. LUCAS (2007): "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, 54(6), 1726–68.

ANDERSON, J. E. (1979): "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69(1), 106–16.

ANDERSON, J. E., AND E. VAN WINCOOP (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93(1), 170–92.

ARKOLAKIS, C. (2016): "A Unified Theory of Firm Selection and Growth," *Quarterly Journal of Economics*, pp. 89–155.

ARKOLAKIS, C., A. COSTINOT, D. DONALDSON, AND A. RODRÍGUEZ-CLARE (2018): "The Elusive Pro-Competitive Effects of Trade," *Review of Economic Studies*, forthcoming.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

ARMINGTON, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *International Monetary Fund Staff Papers*, 16(1), 159–178.

AXTELL, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293(5536), 1818–20.

BALASSA, B. (1965): "Trade Liberalization and Revealed Comparative Advantage," *Manchester School of Economic and Social Studies*, 33, 99–123.

BASU, S., AND D. N. WEIL (1998): "Appropriate Technology and Growth," *Quarterly Journal of Economics*, 113(4), 1025–54.

BOX, G. E. P., AND D. R. COX (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B*, 26(2), 211–252.

BUERA, F. J., AND E. OBERFIELD (2016): "The Global Diffusion of Ideas," *NBER Working Paper*, 21844.

BURSTEIN, A., AND J. CRAVINO (2015): "Measured Aggregate Gains from International Trade," *American Economic Journal: Macroeconomics*, 7(2), 181–218.

BURSTEIN, A. T., AND J. VOGEL (2017): "International Trade, Technology, and the Skill Premium," *Journal of Political Economy*, 125(5), 1356–1412.

CABRAL, L. M. B., AND J. MATA (2003): "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, 93(4), 1075–1090.

CADOT, O., C. CARRÉRE, AND V. STRAUSS KAHN (2011): "Export Diversification: What's behind the Hump?," *Review of Economics and Statistics*, 93(2), 590–605.

CAI, J., N. LI, AND A. M. SANTACREU (2018): "Knowledge Diffusion, Trade and Innovation Across Countries and Sectors," Shanghai University of Finance and Economics, unpublished manuscript.

CALIENDO, L., AND F. PARRO (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 82(1), 1–44.

CALIENDO, L., F. PARRO, E. ROSSI-HANSBERG, AND P. D. SARTE (2017): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *Review of Economic Studies*, forthcoming.

CAMERON, G., J. PROUDMAN, AND S. REDDING (2005): "Technological Convergence, R&D, Trade and Productivity Growth," *European Economic Review*, 49(3), 775–807.

CHOR, D. (2010): "Unpacking Sources of Comparative Advantage: A Quantitative Approach," *Journal of International Economics*, 82(2), 152–67.

COSTINOT, A., D. DONALDSON, AND I. KOMUNJER (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79(2), 581–608.

COSTINOT, A., D. DONALDSON, M. KYLE, AND H. WILLIAMS (2016): "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect," *NBER Working Paper*, 22538.

COSTINOT, A., AND A. RODRÍGUEZ-CLARE (2014): "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in *Handbook of International Economics*, ed. by E. Helpman, K. Rogoff, and G. Gopinath, vol. 4, chap. 4, pp. 197–261. Elsevier, Amsterdam.

CROOKS, G. E. (2010): "The Amoroso Distribution," *ArXiv*, 1005(3274), 1–27.

DARUICH, D., W. EASTERLY, AND A. RESHEF (2016): "The Surprising Instability of Export Specializations," *NBER Working Paper*, 22869.

DEARDORFF, A. V. (1998): "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?," in *The Regionalization of the World Economy*, ed. by J. A. Frankel, chap. 1, pp. 7–32. University of Chicago Press, Chicago.

DEKLE, R., J. EATON, AND S. KORTUM (2007): "Unbalanced Trade," *American Economic Review: Papers and Proceedings*, 97(2), 351–55.

DENNIS, B. (1989): "Stochastic Differential Equations As Insect Population Models," in *Estimation and Analysis of Insect Populations*, ed. by L. L. McDonald, B. F. J. Manly, J. A. Lockwood, and J. A. Logan, vol. 55 of *Lecture Notes in Statistics*, chap. 4, pp. 219–238. Springer, New York, Proceedings of a Conference held in Laramie, Wyoming, January 25-29, 1988.

DI GIOVANNI, J., A. A. LEVCHENKO, AND J. ZHANG (2014): "The Global Welfare Impact of China: Trade Integration and Technological Change," *American Economic Journal: Macroeconomics*, 6(3), 153–83.

EASTERLY, W., AND A. RESHEF (2010): "African Export Successes: Surprises, Stylized Facts, and Explanations," *NBER Working Paper*, 16597.

EATON, J., AND S. KORTUM (1999): "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40(3), 537–70.

——— (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–79.

——— (2010): "Technology in the Global Economy: A Framework for Quantitative Analysis," University of Chicago, unpublished manuscript.

EATON, J., S. KORTUM, B. NEIMAN, AND J. ROMALIS (2016): "Trade and the Global Recession," *American Economic Review*, 106(11), 3401–38.

EATON, J., S. S. KORTUM, AND S. SOTELO (2012): "International Trade: Linking Micro and Macro," *NBER Working Paper*, 17864.

EGGER, P., AND S. K. NIGAI (2016): "World-Trade Growth Accounting," *CESifo Group Munich CESifo Working Paper*, 5831.

FEENSTRA, R. C., R. E. LIPSEY, H. DENG, A. C. MA, AND H. MO (2005): "World Trade Flows: 1962-2000," *NBER Working Paper*, 11040.

FORMAN, J. L., AND M. SØRENSEN (2008): "The Pearson Diffusions: A Class of Statistically Tractable Diffusion Processes," *Scandinavian Journal of Statistics*, 35(3), 438–465.

FREUND, C. L., AND M. D. PIEROLA (2013): "Export Superstars," The World Bank, unpublished manuscript.

GABAIX, X. (1999): "Zipf's Law For Cities: An Explanation," *Quarterly Journal of Economics*, 114(3), 739–767.

——— (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79(3), 733–72.

GABAIX, X., AND Y. M. IOANNIDES (2004): "The Evolution of City Size Distributions," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. F. Thisse, vol. 4 of *Handbooks in Economics*, chap. 53, pp. 2341–78. Elsevier, Amsterdam, San Diego and Oxford.

GAUBERT, C., AND O. ITSKHOKI (2015): "Granular Comparative Advantage," University of California, Berkeley, unpublished manuscript.

GUVENEN, F., S. OZKAN, AND J. SONG (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122(3), 621–60.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–54.

HAUSMANN, R., AND C. A. HIDALGO (2011): "The Network Structure of Economic Output," *Journal of Economic Growth*, 16(4), 309–342.

HEAD, K., AND T. MAYER (2014): "Gravity Equations: Workhorse, Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by E. Helpman, K. Rogoff, and G. Gopinath, vol. 4, chap. 3, pp. 131–195. Elsevier, Amsterdam.

HIDALGO, C. A., AND R. HAUSMANN (2009): "The Building Blocks of Economic Complexity," *Proceedings of the National Academy of Sciences*, 106(2), 10570–10575.

IMBS, J., AND R. WACZIARG (2003): "Stages of Diversification," *American Economic Review*, 93(1), 63–86.

KARLIN, S., AND H. M. TAYLOR (1981): *A second course in stochastic processes*. Academic Press, New York, 1st edn.

KESSLER, M., AND M. SØRENSEN (1999): "Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process," *Bernoulli*, 5(2), 299–314.

KOTZ, S., N. L. JOHNSON, AND N. BALAKRISHNAN (1994): *Continuous Univariate Distributions*, vol. 1 of *Wiley Series in Probability and Statistics*. Wiley, New York, 2nd edn.

KRUGMAN, P. R. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950–59.

LEDERMAN, D., S. PIENKNAGURA, AND D. ROJAS (2015): "Latent Trade Diversification and Its Relevance for Macroeconomic Stability," *World Bank Policy Research Working Paper*, 7332.

LEIGH, E. G. (1968): "The Ecological Role of Volterra's Equations," in *Some mathematical problems in biology*, ed. by M. Gerstenhaber, vol. 1 of *Lectures on mathematics in the life sciences*, pp. 19–61. American Mathematical Society, Providence, Rhode Island, Proceedings of the Symposium on Mathematical Biology held at Washington, D.C., December 1966.

LEVCHENKO, A. A., AND J. ZHANG (2013): "The Evolution of Comparative Advantage: Measurement and Welfare Implications," University of Michigan, unpublished manuscript.

LUCAS, R. E. J., AND B. MOLL (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122(1), 1–51.

LUTTMER, E. G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122(3), 1103–44.

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

NEWEY, W. K., AND D. L. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden, vol. 4, chap. 23, pp. 2111–2245. Elsevier Science, Amsterdam.

OVERMAN, H. G., AND Y. M. IOANNIDES (2001): "Cross-Sectional Evolution of the U.S. City Size Distribution," *Journal of Urban Economics*, 49(3), 543–66.

PEARSON, K. (1895): "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material," *Philosophical Transactions of the Royal Society of London. A*, 186, 343–414.

PERLA, J., AND C. TONETTI (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122(1), 52–76.

PERLA, J., C. TONETTI, AND M. E. WAUGH (2015): "Equilibrium Technology Diffusion, Trade, and Growth," *NBER Working Paper*, 20881.

POSTEL-VINAY, F., AND J. M. ROBIN (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

PRAJNESHU (1980): "Time-Dependent Solution of the Logistic Model for Population Growth in Random Environment," *Journal of Applied Probability*, 17(4), 1083–1086.

PRATT, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32(1/2), 122–136.

PROUDMAN, J., AND S. J. REDDING (2000): "Evolving Patterns of International Trade," *Review of International Economics*, 8(3), 373–96.

QUAH, D. T. (1993): "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics*, 95(4), 427–43.

——— (1996): "Empirics for Economic Growth and Convergence," *European Economic Review*, 40(6), 1353–75.

——— (1997): "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, 2(1), 27–59.

SAMPSON, T. (2016): "Dynamic Selection: An Idea Flows Theory of Entry, Trade, and Growth," *Quarterly Journal of Economics*, 131(1), 315–380.

——— (2017): "The Global Productivity Distribution and Ricardian Comparative Advantage," London School of Economics, unpublished manuscript.

SILVA, J. M. C. S., AND S. TENREYRO (2006): "The Log of Gravity," *Review of Economics and Statistics*, 88(4), 641–58.

SOMALE, M. A. (2017): "Comparative Advantage in Innovation and Production," *Board of Governors of the Federal Reserve System, International Finance Discussion Paper*, 1206.

STACY, E. W. (1962): "A Generalization of the Gamma Distribution," *Annals of Mathematical Statistics*, 33(3), 1187–1192.

SUTTON, J. (1997): "Gibrat's Legacy," *Journal of Economic Literature*, 35(1), 40–59.

SUTTON, J., AND D. TREFLER (2016): "Capabilities, Wealth, and Trade," *Journal of Political Economy*, 124(3), 826–878.

TSALLIS, C. (1988): "Possible Generalization of Boltzmann-Gibbs Statistics," *Journal of Statistical Physics*, 52(1), 479–487.

WONG, E. (1964): "The Construction of a Class of Stationary Markoff Processes," in *Stochastic processes in mathematical physics and engineering*, ed. by R. Bellman, vol. 16 of *Proceedings of symposia in applied mathematics*, chap. 12, pp. 264–276. American Mathematical Society, Providence, R.I.

Appendix

A Generalized Logistic Diffusion

The principal insights of Subsections 4.1 and 4.3 are based on the following relationship.

Lemma 1. *The generalized logistic diffusion*

$$\frac{\mathrm{d}\hat{A}_{is}(t)}{\hat{A}_{is}(t)} = \frac{\sigma^2}{2}\left[1 - \eta\,\frac{\hat{A}_{is}(t)^{\phi} - 1}{\phi}\right]\mathrm{d}t + \sigma\,\mathrm{d}W_{is}^{\hat{A}}(t) \tag{A.1}$$

for real parameters (η, σ, ϕ) *has a stationary distribution that is generalized gamma with a probability density* $f_{\hat{A}}(\hat{a}_{is}|\hat{\theta}, \kappa, \phi)$ *given by* (17) *and the real parameters*

$$\hat{\theta} = \left(\phi^2/\eta\right)^{1/\phi} > 0 \quad \textit{and} \quad \kappa = 1/\hat{\theta}^{\phi} > 0.$$

A non-degenerate stationary distribution exists only if $\eta > 0$.

Equation (A.1) restates equation (16) from the text.

A.1 Derivation of the generalized logistic diffusion

We now establish Lemma 1. As a starting point, note that the ordinary gamma distribution is known to be the stationary distribution of the stochastic logistic equation (Leigh 1968). We generalize this ordinary logistic diffusion to yield a generalized gamma distribution as the stationary distribution in the cross section. Our (three-parameter) generalization of the gamma distribution relates back to the ordinary (two-parameter) gamma distribution through a power transformation. Take an ordinary gamma distributed random variable X with two parameters $\bar{\theta}, \kappa > 0$ and the density function

$$f_X(x|\bar{\theta}, \kappa) = \frac{1}{\Gamma(\kappa)}\frac{1}{\bar{\theta}}\left(\frac{x}{\bar{\theta}}\right)^{\kappa-1}\exp\left\{-\frac{x}{\bar{\theta}}\right\} \qquad \text{for} \quad x > 0. \tag{A.2}$$

Then the transformed variable $\hat{A} = X^{1/\phi}$ has a generalized gamma distribution under the accompanying parameter transformation $\hat{\theta} = \bar{\theta}^{1/\phi}$ because

$$\begin{aligned}
f_{\hat{A}}(\hat{a}|\hat{\theta}, \kappa, \phi) &= \tfrac{\partial}{\partial \hat{a}}\Pr(\hat{A} \leq \hat{a}) = \tfrac{\partial}{\partial \hat{a}}\Pr(X^{1/\phi} \leq \hat{a}) \\
&= \tfrac{\partial}{\partial \hat{a}}\Pr(X \leq \hat{a}^{\phi}) = f_X(\hat{a}^{\phi}|\hat{\theta}^{\phi}, \kappa)\cdot|\phi\hat{a}^{\phi-1}| \\
&= \frac{\hat{a}^{\phi-1}}{\Gamma(\kappa)}\left|\frac{\phi}{\hat{\theta}^{\phi}}\right|\left(\frac{\hat{a}^{\phi}}{\hat{\theta}^{\phi}}\right)^{\kappa-1}\exp\left\{-\frac{\hat{a}^{\phi}}{\hat{\theta}^{\phi}}\right\} = \frac{1}{\Gamma(\kappa)}\left|\frac{\phi}{\hat{\theta}}\right|\left(\frac{\hat{a}}{\hat{\theta}}\right)^{\phi\kappa-1}\exp\left\{-\left(\frac{\hat{a}}{\hat{\theta}}\right)^{\phi}\right\},
\end{aligned}$$

which is equivalent to the generalized gamma probability density function (17), where $\Gamma(\cdot)$ denotes the gamma function and $\hat{\theta}, \kappa, \phi$ are the three parameters of the generalized gamma distribution in our context ($\hat{a} > 0$ can be arbitrarily close to zero).

The ordinary logistic diffusion of a variable X follows the stochastic process

$$\mathrm{d}X(t) = \left[\bar{\alpha} - \bar{\beta}\,X(t)\right]X(t)\,\mathrm{d}t + \bar{\sigma}\,X(t)\,\mathrm{d}W(t) \qquad \text{for} \quad X(t) > 0, \tag{A.3}$$

where $\bar{\alpha}, \bar{\beta}, \bar{\sigma} > 0$ are parameters, t denotes time, $W(t)$ is the Wiener process (standard Brownian motion) and a reflection ensures that $X(t) > 0$. The stationary distribution of this process (the limiting distribution of $X = X(\infty) = \lim_{t\to\infty} X(t)$) is known to be an ordinary gamma distribution (Leigh 1968):

$$f_X(x|\bar{\theta}, \kappa) = \frac{1}{\Gamma(\kappa)} \left| \frac{1}{\bar{\theta}} \right| \left(\frac{x}{\bar{\theta}} \right)^{\kappa-1} \exp\left\{ -\frac{x}{\bar{\theta}} \right\} \qquad \text{for} \quad x > 0, \tag{A.4}$$

as in (A.2) with

$$\begin{aligned} \bar{\theta} &= \bar{\sigma}^2/(2\bar{\beta}) > 0, \\ \kappa &= 2\bar{\alpha}/\bar{\sigma}^2 - 1 > 0 \end{aligned} \tag{A.5}$$

under the restriction $\bar{\alpha} > \bar{\sigma}^2/2$. The ordinary logistic diffusion can also be expressed in terms of infinitesimal parameters as

$$\mathrm{d}X(t) = \mu_X(X(t))\,\mathrm{d}t + \sigma_X(X(t))\,\mathrm{d}W(t) \qquad \text{for} \quad X(t) > 0,$$
$$\mu_X(X) = (\bar{\alpha} - \bar{\beta}\,X)X \quad \text{and} \quad \sigma_X^2(X) = \bar{\sigma}^2\,X^2.$$

Now consider the diffusion of the transformed variable $\hat{A}(t) = X(t)^{1/\phi}$. In general, a strictly monotone transformation $\hat{A} = g(X)$ of a diffusion X is a diffusion with infinitesimal parameters

$$\mu_{\hat{A}}(\hat{A}) = \frac{1}{2}\sigma_X^2(X)g''(X) + \mu_X(X)g'(X) \quad \text{and} \quad \sigma_{\hat{A}}^2(\hat{A}) = \sigma_X^2(X)g'(X)^2$$

(see Karlin and Taylor 1981, Section 15.2, Theorem 2.1). Applying this general result to the specific monotone transformation $\hat{A} = X^{1/\phi}$ yields our specification of a *generalized logistic diffusion*:

$$\mathrm{d}\hat{A}(t) = \left[\alpha - \beta\hat{A}(t)^{\phi}\right]\hat{A}(t)\,\mathrm{d}t + \sigma\hat{A}(t)\,\mathrm{d}W(t) \qquad \text{for} \quad \hat{A}(t) > 0. \tag{A.6}$$

with the parameters

$$\alpha \equiv \left[\frac{1-\phi}{2}\frac{\bar{\sigma}^2}{\phi^2} + \frac{\bar{\alpha}}{\phi}\right], \qquad \beta \equiv \frac{\bar{\beta}}{\phi}, \qquad \sigma \equiv \frac{\bar{\sigma}}{\phi}. \tag{A.7}$$

The term $-\beta\hat{A}(t)^{\phi}$ now involves a power function and the parameters of the generalized logistic diffusion collapse to the parameters of the ordinary logistic diffusion for $\phi = 1$.

We infer that the stationary distribution of $\hat{A}(\infty) = \lim_{t\to\infty} \hat{A}(t)$ is a generalized gamma distribution by (17) and by the derivations above:

$$f_{\hat{A}}(\hat{a}|\hat{\theta}, \kappa, \phi) = \frac{1}{\Gamma(\kappa)} \left| \frac{\phi}{\hat{\theta}} \right| \left(\frac{\hat{a}}{\hat{\theta}} \right)^{\phi\kappa-1} \exp\left\{ -\left(\frac{\hat{a}}{\hat{\theta}} \right)^{\phi} \right\} \qquad \text{for} \quad x > 0,$$

with

$$\begin{aligned} \hat{\theta} &= \bar{\theta}^{1/\phi} = [\bar{\sigma}^2/(2\bar{\beta})]^{1/\phi} = [\phi\sigma^2/(2\beta)]^{1/\phi} > 0, \\ \kappa &= 2\bar{\alpha}/\bar{\sigma}^2 - 1 = [2\alpha/\sigma^2 - 1]/\phi > 0 \end{aligned} \tag{A.8}$$

by (A.5) and (A.7).

A.2 Existence and parametrization

Existence of a non-degenerate stationary distribution with $\hat{\theta}, \kappa > 0$ circumscribes how the parameters of the diffusion α, β, σ and ϕ must relate to each other. A strictly positive $\hat{\theta}$ implies that $\text{sign}(\beta) = \text{sign}(\phi)$. Second, a strictly positive κ implies that $\text{sign}(\alpha - \sigma^2/2) = \text{sign}(\phi)$. The latter condition is closely related to the requirement that comparative advantage neither collapse nor explode. If the level elasticity of dissipation ϕ is strictly positive ($\phi > 0$) then, for the stationary probability density $f_{\hat{A}}(\cdot)$ to be non-degenerate, the offsetting constant drift parameter α needs to strictly exceed the variance of the stochastic innovations: $\alpha \in (\sigma^2/2, \infty)$. Otherwise absolute advantage would "collapse" as arbitrarily much time passes, implying industries die out. If $\phi < 0$ then the offsetting positive drift parameter α needs to be strictly less than the variance of the stochastic innovations: $\alpha \in (-\infty, \sigma^2/2)$; otherwise absolute advantage would explode.

Our preferred parametrization of the generalized logistic diffusion is (A.1) in Lemma 1 for real parameters η, σ, ϕ. That parametrization can be related back to the parameters in (A.6) by setting $\alpha = (\sigma^2/2) + \beta$ and $\beta = \eta\sigma^2/(2\phi)$. In this simplified formulation, the no-collapse and no-explosion conditions are satisfied for the single restriction that $\eta > 0$. The reformulation in (A.1) also clarifies that one can view our generalization of the drift term $[\hat{A}_{is}(t)^{\phi} - 1]/\phi$ as a conventional Box-Cox transformation of $\hat{A}_{is}(t)$ to model the level dependence.

The non-degenerate stationary distribution accommodates both the log normal and the Pareto distribution as limiting cases. When $\phi \to 0$, both α and β tend to infinity; if β did not tend to infinity, a drifting random walk would result in the limit. A stationary log normal distribution requires that $\alpha/\beta \to 1$, so $\alpha \to \infty$ at the same rate with $\beta \to \infty$ as $\phi \to 0$. For existence of a non-degenerate stationary distribution, in the benchmark case with $\phi \to 0$ we need $1/\alpha \to 0$ for the limiting distribution to be log normal. In contrast, a stationary Pareto distribution with shape parameter p would require that $\alpha = (2-p)\sigma^2/2$ as $\phi \to 0$ (see e.g. Crooks 2010, Table 1; proofs are also available from the authors upon request).

A.3 From comparative to absolute advantage

If comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion by (A.1), then the stationary distribution of comparative advantage is a generalized gamma distribution with density (17) and parameters $\hat{\theta} = \left(\phi^2/\eta\right)^{1/\phi} > 0$ and $\kappa = 1/\hat{\theta}^{\phi} > 0$ by Lemma 1. From this stationary distribution of comparative advantage $\hat{A}_{is}$, we can infer the cross-sectional distribution of absolute advantage $A_{is}(t)$. Note that, by definition (14), absolute advantage is not necessarily stationary because the stochastic trend $Z_s(t)$ may not be stationary.

Absolute advantage is related to comparative advantage through a country-wide stochastic trend by definition (14). Plugging this definition into (17), we can infer that the probability density of absolute advantage must be proportional to

$$f_A(a_{is}|\hat{\theta}, Z_s(t), \kappa, \phi) \propto \left(\frac{a_{is}}{\hat{\theta} Z_s(t)}\right)^{\phi\kappa - 1} \exp\left\{-\left(\frac{a_{is}}{\hat{\theta} Z_s(t)}\right)^{\phi}\right\}.$$

It follows from this proportionality that the probability density of absolute advantage must be a generalized gamma distribution with $\theta_s(t) = \hat{\theta} Z_s(t) > 0$, which is time varying because of the stochastic trend $Z_s(t)$. We summarize these results in a lemma.

Lemma 2. *If comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion* (A.1) *with real parameters η, σ, ϕ ($\eta > 0$), then the cross-sectional distribution of absolute advantage $A_{is}(t)$ is generalized gamma with the CDF*

$$F_A(a_{is}|\theta_s(t), \phi, \kappa) = G\left[\left(\frac{a_{is}}{\theta_s(t)}\right)^{\phi}; \kappa\right] \tag{A.9}$$

for the strictly positive parameters

$$\hat{\theta} = \left(\phi^2/\eta\right)^{1/\phi}, \quad \theta_s(t) = \hat{\theta} Z_s(t) \quad \textit{and} \quad \kappa = 1/\hat{\theta}^{\phi}.$$

Proof. Derivations above establish that the cross-sectional distribution of absolute advantage is generalized gamma. The cumulative distribution function follows from Kotz, Johnson, and Balakrishnan (1994, Ch. 17, Section 8.7). □

Lemma 2 establishes that the diffusion and cross-sectional distribution of absolute advantage inherit all relevant properties of comparative advantage after adjustment for an (arbitrary) country-level growth trend. Equation (A.9) predicts cumulative probability distributions of absolute advantage such as those in **Figure 1** (and in Appendix **Figures A1, A2** and **A3**). The lower cutoff for absolute advantage shifts right over time, but the shape of the cross sectional CDF is stable across countries and years. We will document in Appendix B how the trend can be recovered from estimation of the comparative-advantage diffusion using absolute advantage data.

A.4 Moments and the mean-median ratio

As a prelude to the GMM estimation, the r-th raw moments of the ratios $a_{is}/\theta_s(t)$ and $\hat{a}_{is}/\hat{\theta}$ are

$$\mathbb{E}\left[\left(\frac{a_{is}}{\theta_s(t)}\right)^r\right] = \mathbb{E}\left[\left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^r\right] = \frac{\Gamma(\kappa + r/\phi)}{\Gamma(\kappa)}$$

and identical because both $[a_{is}/\theta_s(t)]^{1/\phi}$ and $[\hat{a}_{is}/\hat{\theta}]^{1/\phi}$ have the same standard gamma distribution (Kotz, Johnson, and Balakrishnan 1994, Ch. 17, Section 8.7). As a consequence, the raw moments of absolute advantage A_{is} are scaled by a country-specific time-varying factor $Z_s(t)^r$ whereas the raw moments of comparative advantage are constant over time if comparative advantage follows a diffusion with three constant parameters $(\hat{\theta}, \kappa, \phi)$:

$$\mathbb{E}\left[(a_{is})^r \middle| Z_s(t)^r\right] = Z_s(t)^r \cdot \mathbb{E}\left[(\hat{a}_{is})^r\right] = Z_s(t)^r \cdot \hat{\theta}^r \frac{\Gamma(\kappa + r/\phi)}{\Gamma(\kappa)}.$$

By Lemma 2, the median of comparative advantage is $\hat{a}_{.5} = \hat{\theta}(G^{-1}[.5;\kappa])^{1/\phi}$. A measure of concentration in the right tail is the ratio of the mean and the median, which is independent of $\hat{\theta}$ and equals

$$\textit{Mean/median ratio} = \frac{\Gamma(\kappa + 1/\phi)/\Gamma(\kappa)}{(G^{-1}[.5;\kappa])^{1/\phi}}. \tag{A.10}$$

We report this measure of concentration to characterize the curvature of the stationary distribution.

B Identification of the Generalized Logistic Diffusion

Our implementation of the Generalized Logistic Diffusion requires not only identification of the three time-invariant real parameters (η, σ, ϕ)—or equivalently $(\hat{\theta}, \kappa, \phi)$—, but also identification of a stochastic trend: the country-specific time-varying factor $Z_s(t)$.

Proposition 1. *If comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion* (A.1) *with real parameters η, σ, ϕ ($\eta > 0$), then the country specific stochastic trend $Z_s(t)$ is recovered from the first moment of the logarithm of absolute advantage as:*

$$Z_s(t) = \exp\left\{\mathbb{E}_{st}[\ln A_{is}(t)] - \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\} \tag{B.11}$$

where $\Gamma'(\kappa)/\Gamma(\kappa)$ is the digamma function.

Equation (B.11) restates equation (18) from the text. For a proof of Proposition 1, first consider a random variable X that has a gamma distribution with scale parameter θ and shape parameter κ. For any power $n \in \mathbb{N}$ we have

$$\begin{aligned}
\mathbb{E}\left[\ln(X^n)\right] &= \int_0^\infty \ln(x^n)\frac{1}{\Gamma(\kappa)}\frac{1}{\theta}\left(\frac{x}{\theta}\right)^{\kappa-1}\exp\left\{-\frac{x}{\theta}\right\}\mathrm{d}x \\
&= \frac{n}{\Gamma(\kappa)}\int_0^\infty \ln(\theta z)z^{\kappa-1}e^{-z}\mathrm{d}z \\
&= n\ln\theta + \frac{n}{\Gamma(\kappa)}\int_0^\infty \ln(z)z^{\kappa-1}e^{-z}\mathrm{d}z \\
&= n\ln\theta + \frac{n}{\Gamma(\kappa)}\frac{\partial}{\partial\kappa}\int_0^\infty z^{\kappa-1}e^{-z}\mathrm{d}z \\
&= n\ln\theta + n\frac{\Gamma'(\kappa)}{\Gamma(\kappa)},
\end{aligned}$$

where $\Gamma'(\kappa)/\Gamma(\kappa)$ is the digamma function.

From Appendix A (Lemma 1) we know that raising a gamma random variable to the power $1/\phi$ creates a generalized gamma random variable $X^{1/\phi}$ with shape parameters κ and ϕ and scale parameter $\theta^{1/\phi}$. Therefore

$$\mathbb{E}\left[\ln(X^{1/\phi})\right] = \frac{1}{\phi}\mathbb{E}\left[\ln X\right] = \frac{\ln(\theta) + \Gamma'(\kappa)/\Gamma(\kappa)}{\phi}$$

This result allows us to identify the country specific stochastic trend $X_s(t)$.

For $\hat{A}_{is}(t)$ has a generalized gamma distribution across i for any given s and t with shape parameters ϕ and η/ϕ^2 and scale parameter $(\phi^2/\eta)^{1/\phi}$ we have

$$\mathbb{E}_{st}\left[\ln\hat{A}_{is}(t)\right] = \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}.$$

From definition (14) and $\hat{A}_{is}(t) = A_{is}(t)/Z_s(t)$ we can infer that $\mathbb{E}_{st}[\ln\hat{A}_{is}(t)] = \mathbb{E}_{st}[\ln A_{is}(t)] - \ln Z_s(t)$. Re-arranging and using the previous result for $\mathbb{E}[\ln\hat{A}_{is}(t) \mid s, t]$ yields

$$Z_s(t) = \exp\left\{\mathbb{E}_{st}[\ln A_{is}(t)] - \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\}$$

as stated in the text.

C GMM Estimation of the Associated Pearson-Wong Process

GMM estimation of the Generalized Logistic Diffusion requires conditional moments, which we obtain from a Pearson-Wong transformation.

Proposition 2. *If comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion* (A.1) *with real parameters η, σ, ϕ ($\eta > 0$), then the following two statements are true.*

- *The transformed variable*

$$\hat{B}_{is}(t) = [\hat{A}_{is}(t)^{-\phi} - 1]/\phi \tag{C.12}$$

follows the diffusion

$$\mathrm{d}\hat{B}_{is}(t) = -\frac{\sigma^2}{2}\left[\left(\eta-\phi^2\right)\hat{B}_{is}(t)-\phi\right]\mathrm{d}t + \sigma\sqrt{\phi^2\hat{B}_{is}(t)^2+2\phi\hat{B}_{is}(t)+1}\,\mathrm{d}W_{is}^{\hat{B}}(t)$$

and belongs to the Pearson-Wong family.

- *For any time t, time interval $\Delta > 0$, and integer $n \leq M < \eta/\phi^2$, the n-th conditional moment of the transformed process $\hat{B}_{is}(t)$ satisfies the recursive condition:*

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^n \,\middle|\, \hat{B}_{is}(t)=b\right] = \exp\left\{-a_n\Delta\right\}\sum_{m=0}^{n}\pi_{n,m}b^m - \sum_{m=0}^{n-1}\pi_{n,m}\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^m \,\middle|\, \hat{B}_{is}(t)=b\right], \tag{C.13}$$

for coefficients a_n and $\pi_{n,m}$ ($n, m = 1, \ldots, M$) as defined below.

Equation (C.12) restates equation (20) in the text.

C.1 Derivation of the Pearson-Wong transform

To establish Proposition 2, first consider a random variable X with a standard logistic diffusion (the $\phi = 1$ case). The Bernoulli transformation $1/X$ maps the standard logistic diffusion into the Pearson-Wong family (see e.g. Prajneshu 1980, Dennis 1989). Similar to our derivation of the generalized logistic diffusion in Appendix A, we follow up on that transformation with an additional Box-Cox transformation and apply $\hat{B}_{is}(t) = [\hat{A}_{is}(t)^{-\phi}-1]/\phi$ to comparative advantage, as stated in (C.12). Define $W_{is}^{\hat{B}}(t) \equiv -W_{is}^{\hat{A}}(t)$. Then $\hat{A}_{is}^{-\phi} = \phi\hat{B}_{is}(t)+1$ and, by Itō's lemma,

$$\begin{aligned}
\mathrm{d}\hat{B}_{is}(t) &= \mathrm{d}\left(\frac{\hat{A}_{is}(t)^{-\phi}-1}{\phi}\right) \\
&= -\hat{A}_{is}(t)^{-\phi-1}\,\mathrm{d}\hat{A}_{is}(t) + \frac{1}{2}(\phi+1)\hat{A}_{is}(t)^{-\phi-2}(\mathrm{d}\hat{A}_{is}(t))^2 \\
&= -\hat{A}_{is}(t)^{-\phi-1}\left[\frac{\sigma^2}{2}\left(1-\eta\,\frac{\hat{A}_{is}(t)^{\phi}-1}{\phi}\right)\hat{A}_{is}(t)\,\mathrm{d}t + \sigma\hat{A}_{is}(t)\,\mathrm{d}W_{is}^{\hat{A}}(t)\right] \\
&\qquad + \frac{1}{2}(\phi+1)\hat{A}_{is}(t)^{-\phi-2}\sigma^2\hat{A}_{is}(t)^2\,\mathrm{d}t \\
&= -\frac{\sigma^2}{2}\left[\left(1+\frac{\eta}{\phi}\right)\hat{A}_{is}(t)^{-\phi}-\frac{\eta}{\phi}\right]\mathrm{d}t - \sigma\hat{A}_{is}(t)^{-\phi}\,\mathrm{d}W_{is}^{\hat{A}}(t) + \frac{\sigma^2}{2}(\phi+1)\hat{A}_{is}(t)^{-\phi}\,\mathrm{d}t \\
&= -\frac{\sigma^2}{2}\left[\left(\frac{\eta}{\phi}-\phi\right)\hat{A}_{is}(t)^{-\phi}-\frac{\eta}{\phi}\right]\mathrm{d}t - \sigma\hat{A}_{is}(t)^{-\phi}\,\mathrm{d}W_{is}^{\hat{A}}(t) \\
&= -\frac{\sigma^2}{2}\left[\left(\frac{\eta}{\phi}-\phi\right)(\phi\hat{B}_{is}(t)+1)-\frac{\eta}{\phi}\right]\mathrm{d}t + \sigma(\phi\hat{B}_{is}(t)+1)\,\mathrm{d}W_{is}^{\hat{B}}(t) \\
&= -\frac{\sigma^2}{2}\left[\left(\eta-\phi^2\right)\hat{B}_{is}(t)-\phi\right]\mathrm{d}t + \sigma\sqrt{\phi^2\hat{B}_{is}(t)^2+2\phi\hat{B}_{is}(t)+1}\,\mathrm{d}W_{is}^{\hat{B}}(t).
\end{aligned}$$

The mirror diffusion $\hat{B}_{is}(t)$ is therefore a Pearson-Wong diffusion of the form:

$$\mathrm{d}\hat{B}_{is}(t) = -q(\hat{B}_{is}(t)-\bar{B})\,\mathrm{d}t + \sqrt{2q(a\hat{B}_{is}(t)^2+b\hat{B}_{is}(t)+c)}\,\mathrm{d}W_{is}^{\hat{B}}(t),$$

where $q = (\eta - \phi^2)\sigma^2/2$, $\bar{B} = \sigma^2\phi/(2q)$, $a = \phi^2\sigma^2/(2q)$, $b = \phi\sigma^2/q$, and $c = \sigma^2/(2q)$.

To construct a GMM estimator based on this Pearson-Wong representation, we apply results in Forman and Sørensen (2008) to construct closed form expressions for the conditional moments of the transformed data and then use these moment conditions for estimation. This technique relies on the convenient structure of the Pearson-Wong class and a general result in Kessler and Sørensen (1999) on calculating conditional moments of diffusion processes using the eigenfunctions and eigenvalues of the diffusion's infinitesimal generator.[49]

A Pearson-Wong diffusion's drift term is affine and its dispersion term is quadratic. Its infinitesimal generator must therefore map polynomials to equal or lower order polynomials. As a result, solving for eigenfunctions and eigenvalues amounts to matching coefficients on polynomial terms. This key observation allows us to estimate the mirror diffusion of the generalized logistic diffusion model and to recover the generalized logistic diffusion's parameters.

Given an eigenfunction and eigenvalue pair (h_s, λ_s) of the infinitesimal generator of $\hat{B}_{is}(t)$, we can follow Kessler and Sørensen (1999) and calculate the conditional moment of the eigenfunction:

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)\,\middle|\,\hat{B}_{is}(t)\right] = \exp\{\lambda_s t\}\, h(\hat{B}_{is}(t)). \tag{C.14}$$

Since we can solve for polynomial eigenfunctions of the infinitesimal generator of $B_{is}(t)$ by matching coefficients, this results delivers closed form expressions for the conditional moments of the mirror diffusion for $\hat{B}_{is}(t)$.

To construct the coefficients of these eigen-polynomials, it is useful to consider the case of a general Pearson-Wong diffusion $X(t)$. The stochastic differential equation governing the evolution of $X(t)$ must take the form:

$$\mathrm{d}X(t) = -q(X(t) - \bar{X}) + \sqrt{2(aX(t)^2 + bX(t) + c)\Gamma'(\kappa)/\Gamma(\kappa)}\,\mathrm{d}W^X(t).$$

A polynomial $p_n(x) = \sum_{m=0}^{n} \pi_{n,m} x^m$ is an eigenfunction of the infinitesimal generator of this diffusion if there is some associated eigenvalue $\lambda_n \neq 0$ such that

$$-q(x-\bar{X})\sum_{m=1}^{n}\pi_{n,m} m x^{m-1} + \theta(ax^2+bx+c)\sum_{m=2}^{n}\pi_{n,m} m(m-1)x^{m-2} = \lambda_n \sum_{m=0}^{n}\pi_{n,m}x^m$$

We now need to match coefficients on terms.

From the x^n term, we must have $\lambda_n = -n[1-(n-1)a]q$. Next, normalize the polynomials by setting $\pi_{m,m} = 1$ and define $\pi_{m,m+1} = 0$. Then matching coefficients to find the lower order terms amounts to backward recursion from this terminal condition using the equation

$$\pi_{n,m} = \frac{b_{m+1}}{a_m - a_n}\pi_{n,m+1} + \frac{k_{m+2}}{a_m - a_n}\pi_{n,m+2} \tag{C.15}$$

with $a_m \equiv m[1-(m-1)a]q$, $b_m \equiv m[\bar{X} + (m-1)b]q$, and $c_m \equiv m(m-1)cq$. Focusing on polynomials with order of $n < (1 + 1/a)/2$ is sufficient to ensure that $a_m \neq a_n$ and avoid division by zero.

Using the normalization that $\pi_{n,n} = 1$, equation (C.14) implies a recursive condition for these conditional

[49] For a diffusion

$$\mathrm{d}X(t) = \mu_X(X(t))\,\mathrm{d}t + \sigma_X(X(t))\,\mathrm{d}W^X(t)$$

the infinitesimal generator is the operator on twice continuously differentiable functions f defined by $A(f)(x) = \mu_X(x)\,\mathrm{d}/\mathrm{d}x + \frac{1}{2}\sigma_X(x)^2\,\mathrm{d}^2/\mathrm{d}x^2$. An eigenfunction with associated eigenvalue $\lambda \neq 0$ is any function h in the domain of A satisfying $Ah = \lambda h$.

moments:

$$\mathbb{E}\left[X(t+\Delta)^n\right)|X(t)=x] = \exp\{-a_n\Delta\}\sum_{m=0}^{n}\pi_{n,m}x^m - \sum_{m=0}^{n-1}\pi_{n,m}\mathbb{E}\left[X(t+\Delta)^m\,|X(t)=x\right].$$

These moments exist if we restrict ourselves to the first $N < (1+1/a)/2$ moments.

C.2 Conditional moment recursion

To arrive at the result in the second part of Proposition 2, set the parameters as $q_s = \sigma^2(\eta-\phi^2)/2$, $\bar{X}_s = \phi/(\eta-\phi^2)$, $a_s = \phi^2/(\eta-\phi^2)$, $b_s = 2\phi/(\eta-\phi^2)$, and $c_s = 1/(\eta-\phi^2)$. From these parameters, we can construct eigenvalues and their associated eigenfunctions using the recursive condition (C.15). For any time t, time interval $\Delta > 0$, and integer $n \leq M < \eta/\phi^2$, these coefficients correspond to the n-th conditional moment of the transformed process $\hat{B}_{is}(t)$ and satisfy the recursive moment condition

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^n\,\middle|\hat{B}_{is}(t)=b\right] = \exp\left\{-a_n\Delta\right\}\sum_{m=0}^{n}\pi_{n,m}b^m - \sum_{m=0}^{n-1}\pi_{n,m}\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^m\,\middle|\hat{B}_{is}(t)=b\right],$$

where the coefficients a_n and $\pi_{n,m}$ $(n, m = 1, \ldots, M)$ are defined above. This equation restates (C.13) in Proposition 2 and is n-th conditional moment recursion referenced in Subsection 4.4.

In practice, it is useful to work with a matrix characterization of these moment conditions by stacking the first N moments in a vector $\mathbf{Y}_{is}(t)$:

$$\mathbf{\Pi}\cdot\mathbb{E}\left[\mathbf{Y}_{is}(t+\Delta)\,\middle|\hat{B}_{is}(t)\right] = \mathbf{\Lambda}(\Delta)\cdot\mathbf{\Pi}\cdot\mathbf{Y}_{is}(t) \tag{C.16}$$

with $\mathbf{Y}_{is}(t) \equiv (1, \hat{B}_{is}(t), \ldots, \hat{B}_{is}(t)^M)'$ and the matrices $\mathbf{\Lambda}(\Delta) = \textbf{diag}(e^{-a_1\Delta}, e^{-a_2\Delta}, \ldots, e^{-a_M\Delta})$ and $\mathbf{\Pi} = (\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_M)'$, where $\boldsymbol{\pi}_m \equiv (\pi_{m,0}, \ldots, \pi_{m,m}, 0, \ldots, 0)'$ for each $m = 1, \ldots, M$. In our implementation of the GMM criterion function based on forecast errors, we work with the forecast errors of the linear combination $\mathbf{\Pi}\cdot\mathbf{Y}_{is}(t)$ instead of the forecast errors for $\mathbf{Y}_{is}(t)$. Either estimator is numerically equivalent since the matrix $\mathbf{\Pi}$ is triangular by construction and therefore invertible.

C.3 GMM minimization problem

To derive the GMM estimator (stated in Subsection 4.4), let T_{is} denote the number of time series observations available in industry i and country s. Given sample size of $N = \sum_i\sum_s T_{is}$, our GMM estimator solves the minimization problem

$$(\eta^*, \sigma^*, \phi^*) = \arg\min_{(\eta,\sigma,\phi)}\left(\frac{1}{N}\sum_i\sum_s\sum_\tau g_{is\tau}(\eta,\sigma,\phi)\right)'\mathbf{W}\left(\frac{1}{N}\sum_i\sum_s\sum_\tau g_{is\tau}(\eta,\sigma,\phi)\right) \tag{C.17}$$

for a given weighting matrix $\mathbf{W}$. Being overidentified, we adopt a two-step estimator. On the first step we compute an identity weighting matrix, which provides us with a consistent initial estimate. On the second step we update the weighting matrix to an estimate of the optimal weighting matrix by setting the inverse weighting matrix to $\mathbf{W}^{-1} = (1/N)\sum_i\sum_s\sum_\tau g_{is\tau}(\eta,\sigma,\phi)g_{is\tau}(\eta,\sigma,\phi)'$, which is calculated at the parameter value from the first step. Forman and Sørensen (2008) establish asymptotics for a single time series as $T \to \infty$.[50] For

[50] Our estimator would also fit into the standard GMM framework of Hansen (1982), which establishes consistency and asymptotic normality of our second stage estimator as $IS \to \infty$. To account for the two-step nature of our estimator, we use an asymptotic

estimation, we impose the constraints that $\eta > 0$ and $\sigma^2 > 0$ by reparametrizing the model in terms of $\ln \eta > -\infty$ and $2 \ln \sigma > -\infty$. We evaluate the objective function (C.17) at values of (η, σ, ϕ) by detrending the data at each iteration to obtain $\hat{A}_{is}^{\text{GMM}}(t)$ from equation (19), transforming these variables into their mirror variables $\hat{B}_{is}^{\text{GMM}}(t) = [\hat{A}_{is}^{\text{GMM}}(t)^{-\phi} - 1]/\phi$, and using equation (C.13) to compute forecast errors. Then we calculate the GMM criterion function for each industry and country pair by multiplying these forecast errors by instruments constructed from $\hat{B}_{is}^{\text{GMM}}(t)$, and finally sum over industries and countries to arrive at the value of the GMM objective.

D Correction for Generated Variables in GMM Estimation

D.1 Sampling variation in estimated absolute and comparative advantage

Let $\mathbf{k}_{i \cdot t}$ denote the vector of export capabilities of industry i at time t across countries and $\mathbf{m}_{i \cdot t}$ the vector of importer fixed effects. Denote the set of exporters in the industry in that year with $\mathcal{S}_{it}$ and the set of destinations, to which a country-industry is ships in that year, with $\mathcal{D}_{ist}$. The set of industries active as exporters from source country s in a given year is denoted with $\mathcal{I}_{st}$. Consider the gravity regression (6)

$$\ln X_{isdt} = k_{ist} + m_{idt} + \mathbf{r}'_{sdt}\mathbf{b}_{it} + v_{isdt}.$$

Stacking observations, the regression can be expressed more compactly in matrix notation as

$$\mathbf{x}_{i \cdot \cdot t} = \mathbf{J}_{it}^{S}\mathbf{k}_{i \cdot t} + \mathbf{J}_{it}^{D}\mathbf{m}_{i \cdot t} + \mathbf{R}_{\cdot \cdot t}\mathbf{b}_{it} + \mathbf{v}_{i \cdot \cdot t},$$

where $\mathbf{x}_{i \cdot \cdot t}$ is the stacked vector of *log* bilateral exports, $\mathbf{J}_{it}^{S}$ and $\mathbf{J}_{it}^{D}$ are matrices of indicators reporting the exporter and importer country by observation, $\mathbf{R}_{\cdot \cdot t}$ is the matrix of bilateral trade cost regressors and $\mathbf{v}_{i \cdot \cdot t}$ is the stacked vector of residuals.

We assume that the two-way least squares dummy variable estimator for each industry time pair it is consistent and asymptotically normal for an individual industry i shipping from source country s to destination d at time t,[51] and state this assumption formally.

Assumption 1. *If $\mathbf{k}_{i \cdot t}^{\text{OLS}}$ is the OLS estimate of $\mathbf{k}_{i \cdot t}$, then*

$$\sqrt{\bar{D}_{it}}(\mathbf{k}_{i \cdot t}^{\text{OLS}} - \mathbf{k}_{i \cdot t}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{it}) \text{ as } \bar{D}_{it} \to \infty,$$

where $\bar{D}_{it} \equiv (1/|\mathcal{S}_{it}|) \sum_{s \in \mathcal{S}_{it}} |\mathcal{D}_{ist}|$ is the source-country-average number of countries importing industry i goods in year t and

$$\mathbf{\Sigma}_{it} = \sigma_{it}^2 \left[\lim_{\bar{D}_{it} \to \infty} \frac{1}{\bar{D}_{it}} \left(\mathbf{J}_{it}^{S}\right)' \mathbf{M}_{it} \left(\mathbf{J}_{it}^{S}\right) \right]^{-1}$$

with $\sigma_{it}^2 \equiv \mathbb{E}_{it} v_{isdt}^2$,

$$\mathbf{M}_{it} \equiv \mathbf{I}_{|\mathcal{S}_{it}|\bar{D}_{it}} - [\mathbf{J}_{it}^{D}, \mathbf{R}_{\cdot \cdot t}]\{[\mathbf{J}_{it}^{D}, \mathbf{R}_{\cdot \cdot t}]'[\mathbf{J}_{it}^{D}, \mathbf{R}_{\cdot \cdot t}]\}^{-1}[\mathbf{J}_{it}^{D}, \mathbf{R}_{\cdot \cdot t}]',$$

and $\mathbf{I}_{|\mathcal{S}_{it}|\bar{D}_{it}}$ the identity matrix.

In finite samples, uncertainty as captured by $\mathbf{\Sigma}_{it}$ can introduce sampling variation in second-stage estimation

approximation where each dimension of our panel data gets large simultaneously (see Appendix D).

[51] This high-level assumption can be justified by standard missing-at-random assumptions on the gravity model.

because $\mathbf{k}^{\text{OLS}}_{i\cdot t}$ is a generated variable. To perform an according finite sample correction, we use

$$\boldsymbol{\Sigma}^{\text{OLS}}_{it} = (\sigma^{\text{OLS}}_{it})^2 \left[\frac{1}{\bar{D}_{it}} \left(\mathbf{J}^S_{it}\right)' \mathbf{M}_{it} \left(\mathbf{J}^S_{it}\right) \right]^{-1}$$

with $(\sigma^{\text{OLS}}_{it})^2 = (1/|\mathcal{S}_{it}|\bar{D}_{it})(\mathbf{v}^{\text{OLS}}_{i\cdot\cdot t})'\mathbf{v}^{\text{OLS}}_{i\cdot\cdot t}$ to consistently estimate the matrix $\boldsymbol{\Sigma}_{it}$.

Our second stage estimation uses demeaned first-stage estimates of export capability. For the remainder of this Appendix, we define *log* absolute advantage and *log* comparative advantage in the population as

$$a_{ist} \equiv \ln A_{ist} = k_{ist} - \frac{1}{|\mathcal{S}_{it}|} \sum_{\varsigma \in \mathcal{S}_{it}} k_{i\varsigma t} \quad \text{and} \quad \hat{a}_{ist} \equiv \ln \hat{A}_{ist} = a_{ist} - \frac{1}{|\mathcal{I}_{st}|} \sum_{j \in \mathcal{I}_{st}} a_{jst}. \tag{D.18}$$

Correspondingly, we denote their estimates with a^{OLS}_{ist} and $\hat{a}^{\text{OLS}}_{ist}$.

For each year, let $\mathbf{K}^{\text{OLS}}_t$ denote an $I \times S$ matrix with entries equal to estimated export capability whenever available and equal to zero otherwise, let $\mathbf{H}_t$ record the pattern of non-missing observations and $\mathbf{K}_t$ collect the population values of export capability:

$$[\mathbf{K}^{\text{OLS}}_t]_{is} = \begin{cases} k^{\text{OLS}}_{ist} & s \in \mathcal{S}_{it} \\ 0 & s \notin \mathcal{S}_{it} \end{cases}, \quad [\mathbf{H}_t]_{is} = \begin{cases} 1 & s \in \mathcal{S}_{it} \\ 0 & s \notin \mathcal{S}_{it} \end{cases}, \quad [\mathbf{K}_t]_{is} = \begin{cases} k_{ist} & s \in \mathcal{S}_{it} \\ 0 & s \notin \mathcal{S}_{it}, \end{cases}.$$

where $[\cdot]_{is}$ denotes the specific entry is. Similarly, collect estimates of log absolute advantage into the matrix $\mathbf{A}^{\text{OLS}}_t$ and estimates of log comparative advantage into the matrix $\hat{\mathbf{A}}^{\text{OLS}}_t$:

$$[\mathbf{A}^{\text{OLS}}_t]_{is} = \begin{cases} \ln A^{\text{OLS}}_{ist} & s \in \mathcal{S}_{it} \\ 0 & s \notin \mathcal{S}_{it} \end{cases}, \quad [\hat{\mathbf{A}}^{\text{OLS}}_t]_{is} = \begin{cases} \ln \hat{A}^{\text{OLS}}_{ist} & s \in \mathcal{S}_{it} \\ 0 & s \notin \mathcal{S}_{it} \end{cases}.$$

We maintain the OLS superscripts to clarify that absolute advantage A^{OLS}_{ist} and comparative advantage $\hat{A}^{\text{OLS}}_{ist}$ are generated variables.

The two matrices $\mathbf{A}^{\text{OLS}}_t$ and $\hat{\mathbf{A}}^{\text{OLS}}_t$ are linearly related to the matrix containing our estimates of export capability $\mathbf{K}^{\text{OLS}}_t$. From equation (D.18), the matrix $\mathbf{A}^{\text{OLS}}_t$ is related to $\mathbf{K}^{\text{OLS}}_t$ and $\mathbf{H}_t$ by

$$\mathbf{vec}(\mathbf{A}^{\text{OLS}}_t) = \mathbf{Trans}(I,S) \underbrace{\begin{pmatrix} \mathbf{I}_S - \frac{[\mathbf{H}_t]'_{1\cdot}[\mathbf{H}_t]_{1\cdot}}{[\mathbf{H}_t]_{1\cdot}[\mathbf{H}_t]'_{1\cdot}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I}_S - \frac{[\mathbf{H}_t]'_{I\cdot}[\mathbf{H}_t]_{I\cdot}}{[\mathbf{H}_t]_{I\cdot}[\mathbf{H}_t]'_{I\cdot}} \end{pmatrix}}_{\equiv \mathbf{Z}_{IS}(\mathbf{H}_t)} \mathbf{vec}[(\mathbf{K}^{\text{OLS}}_t)']. \tag{D.19}$$

Here $\mathbf{vec}(\cdot)$ stacks the columns of a matrix into a vector and $\mathbf{Trans}(I,S)$ is a vectorized-transpose permutation matrix.[52] The function $\mathbf{Z}_{IS}(\mathbf{H}_t)$ maps the matrix $\mathbf{H}_t$ into a block diagonal $IS \times IS$ matrix, which removes the

[52] The vectorized-transpose permutation matrix of type (m,n) is uniquely defined by the relation

$$\mathbf{vec}(\mathbf{B}) = \mathbf{Trans}(m,n)\mathbf{vec}(\mathbf{B}') \quad \forall \mathbf{B} \in \mathbb{R}^{m \times n}.$$

The (ij)-th entry of this matrix is equal to 1 if $j = 1 + m(i-1) - (mn-1)\text{floor}((i-1)/n)$ and 0 otherwise.

global industry average across countries. The matrix of comparative advantage estimates is then:

$$\mathbf{vec}(\hat{\mathbf{A}}_t^{\text{OLS}}) = \underbrace{\begin{pmatrix} \mathbf{I}_I - \frac{[\mathbf{H}_t']_{1\cdot}'[\mathbf{H}_t']_{1\cdot}}{[\mathbf{H}_t']_{1\cdot}[\mathbf{H}_t']_{1\cdot}'} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I}_I - \frac{[\mathbf{H}_t']_{S\cdot}'[\mathbf{H}_t']_{S\cdot}}{[\mathbf{H}_t']_{S\cdot}[\mathbf{H}_t']_{S\cdot}'} \end{pmatrix}}_{\equiv \mathbf{Z}_{SI}(\mathbf{H}_t')} \mathbf{vec}(\mathbf{A}_t^{\text{OLS}}) \equiv \mathbf{Z}_{SI}(\mathbf{H}_t')\mathbf{vec}(\mathbf{A}_t^{\text{OLS}}). \tag{D.20}$$

The function $\mathbf{Z}_{SI}(\mathbf{H}_t')$ maps the matrix $\mathbf{H}_t$ into a block diagonal $SI \times SI$ matrix, which removes the national average across industries.

For simplicity, we assume that the sampling variation in export capability estimates is uncorrelated across industries and years.

Assumption 2. *For any $(it) \neq (jT)$, $\mathbb{E}(\mathbf{k}_{i\cdot t}^{\text{OLS}} - \mathbf{k}_{i\cdot t})(\mathbf{k}_{j\cdot T}^{\text{OLS}} - \mathbf{k}_{j\cdot T})' = \mathbf{0}$.*

We then have the following result.

Lemma 3. *Suppose Assumptions 1 and 2 hold and that there is an $\omega_{it} > 0$ for each (it) so that $\lim_{D\to\infty} \bar{D}_{it}/D = \omega_{it}$. Then*

$$\sqrt{D}[\mathbf{vec}(\mathbf{A}_t^{\text{OLS}}) - \mathbf{Trans}(I,S)\mathbf{Z}_{IS}(\mathbf{H}_t)\mathbf{vec}[(\mathbf{K}_t^{\text{OLS}})']] \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{Trans}(I,S)\mathbf{Z}_{IS}(\mathbf{H}_t)\,\mathbf{\Sigma}_t^*\,\mathbf{Z}_{IS}(\mathbf{H}_t)'\mathbf{Trans}(I,S)')$$

and

$$\begin{aligned}\sqrt{D}\{&\mathbf{vec}(\hat{\mathbf{A}}_t^{\text{OLS}}) - \mathbf{Z}_{SI}(\mathbf{H}_t')\mathbf{Trans}(I,S)\mathbf{Z}_{IS}(\mathbf{H}_t)\mathbf{vec}[(\mathbf{K}_t^{\text{OLS}})']\\ &\xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{Z}_{SI}(\mathbf{H}_t')\mathbf{Trans}(I,S)\mathbf{Z}_{IS}(\mathbf{H}_t)\,\mathbf{\Sigma}_t^*\,\mathbf{Z}_{IS}(\mathbf{H}_t)'\mathbf{Trans}(I,S)'\mathbf{Z}_{SI}(\mathbf{H}_t')')\end{aligned}$$

with

$$\mathbf{\Sigma}_t^* \equiv \begin{pmatrix} \omega_{1t}^{-1}\mathbf{\Sigma}_{1t}^* & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \omega_{It}^{-1}\mathbf{\Sigma}_{It}^* \end{pmatrix}$$

where the s-th column of $\mathbf{\Sigma}_{it}^*$ is equal to country s's corresponding column in $\mathbf{\Sigma}_{it}$ whenever export capability is estimated for (ist) and is a vector of zeros otherwise.

Proof. Assumptions 1 and 2 along with $\bar{D}_{it} \to D \to \infty$ for all (it) implies that $\sqrt{D}(\mathbf{vec}[(\mathbf{K}_t^{\text{OLS}})'] - \mathbf{vec}[\mathbf{K}_t']) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_t^*)$. The results then follow from equation (D.19) and equation (D.20). $\square$

D.2 Second-stage generated variable correction

We estimate two time series models which both can be implemented as GMM estimators. For brevity, we focus on GLD estimation here. (We present the case of OLS estimation of the decay regression in the Online Supplement (Section S.2), which simply uses a different GMM criterion and absolute advantage as data instead of comparative advantage.) GLD estimation is based on a conditional moment of the form:

$$\mathbf{0} = \mathbb{E}_{is,t-\Delta}\mathbf{g}\left(\boldsymbol{\theta}, \hat{a}_{ist}, \hat{a}_{is,t-\Delta}\right), \tag{D.21}$$

where $\boldsymbol{\theta} = (\eta, \sigma, \phi)'$ is the vector of parameters. In our overidentified GMM estimator, $\mathbf{g}$ is a column vector of known continuously differentiable functions (moment conditions) for any time lag $\Delta > 0$.

The moment conditions apply to any instant in continuous time, but our data come in discrete annual observations for a finite period of years. To account for missing data, let $\mathcal{S}_{it}^P \subset \mathcal{S}_{it}$ denote the set of countries that were *previously* observed to export good i and that are still exporting good i at current time t: $\mathcal{S}_{it}^P \equiv \{s \in \mathcal{S}_{it} \mid \exists\, \tau^P < t \,\text{s.t.}\, s \in \mathcal{S}_{i\tau^P}\}$. Similarly, let $\mathcal{S}_{it}^F \equiv \{s \in \mathcal{S}_{it} \mid \exists\, \tau^F > t \,\text{s.t.}\, s \in \mathcal{S}_{i\tau^F}\}$ be current exporter countries that ship good i to at least one destination also some *future* year. Denote the most recent prior period in which s exported in industry i by $\tau_{ist}^P \equiv \sup\{\tau^P < t \mid s \in \mathcal{S}_{i\tau^P}\}$ and the most recent future period in which s will export by $\tau_{ist}^F \equiv \inf\{\tau^F > t \mid s \in \mathcal{S}_{i\tau^F}\}$. We will use these objects to keep track of timing.

For instance, for each $i = 1, \ldots, I$, $t = 2, \ldots, T$, and $s \in \mathcal{S}_{it}^P$ we can design a GMM criterion based on the following conditional moment:

$$\mathbb{E}_{i,s,\tau_{ist}^P} \mathbf{g}\left(\theta, \hat{a}_{ist}, \hat{a}_{is\tau_{ist}^P}\right) = \mathbf{0}.$$

Our finite sample analog for second-stage estimation is:

$$\frac{1}{I(T-1)} \sum_{i=1}^{I} \sum_{t=2}^{T} \frac{1}{|\mathcal{S}_{it}^P|} \sum_{s \in \mathcal{S}_{it}^P} \mathbf{g}_{ist}(\theta) \quad \text{with } \mathbf{g}_{ist}(\theta) \equiv \mathbf{g}\left(\theta, \hat{a}_{ist}^{\text{OLS}}, \hat{a}_{is\tau_{ist}^P}^{\text{OLS}}\right),$$

where $|\mathcal{S}_{it}^P|$ is the number of exporters in industry i at time t that were also observed exporting good i at a previous time.

The effective sample size for the second stage is $N \equiv \sum_{t=1}^{I} \sum_{t=1}^{T-1} |\mathcal{S}_{it}|$ and the GMM criterion can be expressed as

$$Q_N(\theta; \mathbf{W}) = \left(\frac{1}{N} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \mathbf{g}_{ist}(\theta)\right)' \mathbf{W} \left(\frac{1}{N} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \mathbf{g}_{ist}(\theta)\right)$$

where $\mathbf{W}$ is a weighting matrix.

In order to get consistency, we assume that all dimensions of our data are large as N gets large.

Assumption 3. *As $N \to \infty$ we have*

1. $D \to \infty$*;*
2. $\forall (it)\ \exists \omega_{it} > 0$ *so that* $\bar{D}_{it}/D \to \omega_{it}$*,* $N/[I|\mathcal{S}_{it}^P|(T-1)] \to 1$*, and* $|\mathcal{S}_{it}| \to \infty$*;*
3. $\forall (st)\ |\mathcal{I}_{st}| \to \infty$*;*
4. $T \to \infty$.

Letting $D \to \infty$ and $\bar{D}_{it}/D \to \omega_{it} > 0$ ensures that we consistently estimate $\mathbf{k}_{i \cdot t}$ on the first stage and we can use Lemma 3 for the first stage sampling distribution of comparative advantage. Then, letting $|\mathcal{S}_{it}| \to \infty$ ensures that we consistently estimate absolute advantage and $|\mathcal{I}_{st}| \to \infty$ lets us consistently estimate comparative advantage. The asymptotic results of Forman and Sørensen (2008) apply under the assumption that $T \to \infty$.

Under the maintained assumptions, we get the following consistency result.

Proposition 3. *Suppose that*

1. $\theta \in \Theta$ *for some compact set* Θ*;*
2. *for any* $\Delta > 0$*, there is a unique* $\theta_0 \in \Theta$ *such that*

$$\mathbf{0} = \mathbb{E}\mathbf{g}\left(\theta_0, \hat{a}_{ist}, \hat{a}_{is,t-\Delta}\right);$$

3. *for any given positive definite matrix W and for each N, there is a unique minimizer of $Q_N(\theta;\mathbf{W})$ given by $\hat{\theta}_N$;*

4. *both $\mathbb{E}_{it}k_{ist}$ and $\mathbb{E}_{st}k_{ist}$ exist and are finite.*

Then, under Assumptions 1 and 3, we have $\hat{\theta}_N \xrightarrow{p} \theta_0$.

Proof. The proof follows from a standard consistency argument for extremum estimators (see e.g. Newey and McFadden 1994). Given (a) compactness of the parameter space, (b) the continuity of the GMM objective, and (c) the existence of moments as in Forman and Sørensen (2008), we get a uniform law of large numbers for the objective function on the parameter space as $N \to \infty$. The GLD estimator is then consistent under the assumption that the model is identified, provided that we consistently estimate comparative advantage. The consistency of our comparative advantage estimates follows from the strong law of large numbers given Assumption 3 and the existence and finiteness of $\mathbb{E}_{it}k_{ist}$ and $\mathbb{E}_{st}k_{ist}$. $\square$

Proposition 4. *Under the conditions of Proposition 3 and Assumptions 1, 2, and 3 we have*

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, (\mathbf{\Lambda}'\mathbf{W}\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{W}(\mathbf{\Xi}+\mathbf{\Omega})\,\mathbf{W}\,\mathbf{\Lambda}(\mathbf{\Lambda}'\mathbf{W}\mathbf{\Lambda})^{-1}),$$

where

$$\begin{aligned}
\mathbf{\Lambda} &= \mathbb{E}\frac{\partial}{\partial\theta}\mathbf{g}\left(\theta_0, \hat{a}_{ist}, \hat{a}_{is\tau^P_{ist}}\right),\\
\mathbf{\Xi} &= \mathbb{E}\mathbf{g}\left(\theta_0, \hat{a}_{ist}, \hat{a}_{is\tau^P_{ist}}\right)\mathbf{g}\left(\theta_0, \hat{a}_{ist}, \hat{a}_{is\tau^P_{ist}}\right)',\\
\mathbf{\Omega} &= \lim_{N\to\infty}\frac{1}{ND}\sum_{t=1}^{T}\mathbf{G}_t\mathbf{Z}_{SI}(\mathbf{H}_t')\mathbf{Trans}(I,S)\mathbf{Z}_{IS}(\mathbf{H}_t)\,\mathbf{\Sigma}_t^*\,\mathbf{Z}_{IS}(\mathbf{H}_t)'\mathbf{Trans}(I,S)'\mathbf{Z}_{SI}(\mathbf{H}_t')'\mathbf{G}_t'
\end{aligned}$$

for a $\mathbf{G}_t$ matrix of weighted Jacobians of $\mathbf{g}_{ist}(\theta)$, as defined below.

Proof. To get a correction for first stage sampling variation, we use a mean-value expansion of the GMM criterion. Given continuous differentiability of the moment function $\mathbf{g}_{ist}(\theta)$ and the fact that $\hat{\theta}_N$ maximizes $Q_N(\theta;\mathbf{W})$ we must have

$$\begin{aligned}
\mathbf{0} &= \frac{\partial}{\partial\theta}Q_N(\hat{\theta}_N;\mathbf{W})\\
&= \left(\frac{1}{N}\sum_{i=1}^{I}\sum_{t=2}^{T}\sum_{s\in\mathcal{S}^P_{it}}\frac{N}{I|\mathcal{S}^P_{it}|(T-1)}\frac{\partial}{\partial\theta}\mathbf{g}_{ist}(\hat{\theta}_N)\right)'\mathbf{W}\left(\frac{1}{N}\sum_{i=1}^{I}\sum_{t=2}^{T}\sum_{s\in\mathcal{S}^P_{it}}\frac{N}{I|\mathcal{S}^P_{it}|(T-1)}\mathbf{g}_{ist}(\hat{\theta}_N)\right).
\end{aligned}$$

The criterion function $\mathbf{g}$ is continuously differentiable. Therefore, by the mean value theorem, there exist random variables $\tilde{\theta}_N$ and $\tilde{a}_{ist}$ such that $|\tilde{\theta}_N - \theta_0| \le |\hat{\theta}_N - \theta_0|$, $|\tilde{a}_{ist} - \hat{a}_{ist}| \le |\hat{a}^{\text{OLS}}_{ist} - \hat{a}_{ist}|$, and

$$\begin{aligned}
\mathbf{g}(\hat{\theta}_N; ist) = {} & \underbrace{\mathbf{g}\left(\theta_0, \hat{a}_{ist}, \hat{a}_{is\tau^P_{ist}}\right)}_{\equiv\mathbf{G}^0_{ist}} + \underbrace{\frac{\partial}{\partial\theta}\mathbf{g}\left(\theta, \tilde{a}_{ist}, \tilde{a}_{is\tau^P_{ist}}\right)\bigg|_{\theta=\tilde{\theta}_N}}_{\equiv\tilde{\mathbf{G}}^1_{ist}}(\hat{\theta}_N - \theta_0)\\
& + \underbrace{\frac{\partial}{\partial a}\mathbf{g}\left(\tilde{\theta}_N, a, \tilde{a}_{is\tau^P_{ist}}\right)\bigg|_{a=\tilde{a}_{ist}}}_{\equiv\tilde{\mathbf{G}}^2_{ist}}(\hat{a}^{\text{OLS}}_{ist} - \hat{a}_{ist}) + \underbrace{\frac{\partial}{\partial a^P}\mathbf{g}\left(\tilde{\theta}_N, \tilde{a}_{ist}, a^P\right)\bigg|_{a^P=\tilde{a}_{is\tau^P_{ist}}}}_{\equiv\tilde{\mathbf{G}}^3_{ist}}\left(\hat{a}^{\text{OLS}}_{is\tau^P_{ist}} - \hat{a}_{is\tau^P_{ist}}\right).
\end{aligned}$$

Then,

$$\mathbf{0} = \tilde{\mathbf{\Lambda}}'_N \mathbf{W} \frac{1}{N} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \left[\mathbf{G}_{ist}^0 + \tilde{\mathbf{G}}_{ist}^1 (\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) + \tilde{\mathbf{G}}_{ist}^2 (\hat{a}_{ist}^{\text{OLS}} - \hat{a}_{ist}) + \tilde{\mathbf{G}}_{ist}^3 (\hat{a}_{is\tau_{ist}^P}^{\text{OLS}} - \hat{a}_{is\tau_{ist}^P}) \right]$$

where $\tilde{\mathbf{\Lambda}}_N = \frac{1}{I(T-1)} \sum_{i=1}^{I} \sum_{t=2}^{T} \frac{1}{|\mathcal{S}_{it}^P|} \sum_{s \in \mathcal{S}_{it}^P} \tilde{\mathbf{G}}_{ist}^1$.

Solving for $\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0$ and multiplying by $\sqrt{N}$, we obtain

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) =$$

$$-\left[\tilde{\mathbf{\Lambda}}'_N \mathbf{W} \tilde{\mathbf{\Lambda}}_N \right]^{-1} \tilde{\mathbf{\Lambda}}'_N \mathbf{W} \frac{1}{\sqrt{N}} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \left[\mathbf{G}_{ist}^0 + \tilde{\mathbf{G}}_{ist}^2 (\hat{a}_{ist}^{\text{OLS}} - \hat{a}_{ist}) + \tilde{\mathbf{G}}_{ist}^3 (\hat{a}_{is\tau_{ist}^P}^{\text{OLS}} - \hat{a}_{is\tau_{ist}^P}) \right].$$

Note that the set $\mathcal{S}_{i1}^P$ is empty since no country is observed exporting in years before the first sample year and $\mathcal{S}_{iT}^F$ is empty since no country is observed exporting after the final sample year. Moreover,

$$\begin{aligned}
\tilde{\mathbf{\Lambda}}_N &\xrightarrow{p} \Lambda \equiv \mathbb{E} \frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{g} \left(\boldsymbol{\theta}_0, \hat{a}_{ist}, \hat{a}_{is\tau_{ist}^P} \right) \\
\tilde{\mathbf{G}}_{ist}^2 &\xrightarrow{p} \mathbf{G}_{ist}^2 \equiv \left. \frac{\partial}{\partial a} \mathbf{g} \left(\boldsymbol{\theta}_0, a, \hat{a}_{is\tau_{ist}^P} \right) \right|_{a = \hat{a}_{ist}} \\
\tilde{\mathbf{G}}_{ist}^3 &\xrightarrow{p} \mathbf{G}_{ist}^3 \equiv \left. \frac{\partial}{\partial a^P} \mathbf{g} \left(\boldsymbol{\theta}_0, \hat{a}_{ist}, a^P \right) \right|_{a^P = \hat{a}_{is\tau_{ist}^P}}
\end{aligned}$$

because $\hat{\boldsymbol{\theta}}_N$ and $\hat{a}_{ist}^{\text{OLS}}$ are consistent and $\mathbf{g}$ is the continuously differentiable.

As a result, we can re-write the sum as

$$\begin{aligned}
&\frac{1}{\sqrt{N}} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \left[\mathbf{G}_{ist}^0 + \tilde{\mathbf{G}}_{ist}^2 (\hat{a}_{ist}^{\text{OLS}} - \hat{a}_{ist}) + \tilde{\mathbf{G}}_{ist}^3 (\hat{a}_{is\tau_{ist}^P}^{\text{OLS}} - \hat{a}_{is\tau_{ist}^P}) \right] \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \left[\mathbf{G}_{ist}^0 + \mathbf{G}_{ist}^2 (\hat{a}_{ist}^{\text{OLS}} - \hat{a}_{ist}) + \mathbf{G}_{ist}^3 (\hat{a}_{is\tau_{ist}^P}^{\text{OLS}} - \hat{a}_{is\tau_{ist}^P}) \right] + o_p(1) \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}_{it}^P} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \mathbf{G}_{ist}^0 + o_p(1) \\
&\quad + \frac{1}{\sqrt{N}} \sum_{t=1}^{T} \underbrace{\sum_{i=1}^{I} \sum_{s=1}^{S} \left[\mathbf{1}\{s \in \mathcal{S}_{it}^P\} \frac{N}{I|\mathcal{S}_{it}^P|(T-1)} \mathbf{G}_{ist}^2 + \mathbf{1}\{s \in \mathcal{S}_{i\tau_{ist}^F}^0\} \frac{N}{I|\mathcal{S}_{i\tau_{ist}^F}^0|(T-1)} \mathbf{G}_{is\tau_{ist}^F}^3 \right] (\hat{a}_{ist}^{\text{OLS}} - \hat{a}_{ist})}_{\equiv \mathbf{L}_t},
\end{aligned}$$

using the fact that $\tau^F = \tau_{ist}^F \Leftrightarrow \tau_{is\tau^F}^P = t$.

The term $\mathbf{L}_t$ is a vector and a linear function of the entries of the matrix $\hat{\mathbf{A}}_t^{\text{OLS}} - \hat{\mathbf{A}}_t$. This vector can also be expressed as

$$\mathbf{L}_t = \mathbf{G}_t \mathbf{vec}(\hat{\mathbf{A}}_t^{\text{OLS}} - \hat{\mathbf{A}}_t),$$

and the matrix $\mathbf{G}_t$ has entries

$$[\mathbf{G}_t]_{\cdot j} = \mathbf{1}\left\{s(j) \in \mathcal{S}^P_{i(j),t}\right\} \frac{N}{I\left|\mathcal{S}^P_{i(j),t}\right|(T-1)} \mathbf{G}^2_{i(j),s(j),t} + \mathbf{1}\left\{s(j) \in \mathcal{S}^F_{i(j),\tau^F_{i(j),s(j),t}}\right\} \frac{N}{I\left|\mathcal{S}^F_{i(j),\tau^F_{i(j),s(j),t}}\right|(T-1)} \mathbf{G}^3_{i(j),s(j),\tau^F_{i(j),s(j),t}}$$

for

$$i(j) = 1 + (j \mod S), \quad s(j) = 1 + \text{floor}((j-1)/S).$$

We can now re-write the sum as

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}^P_{it}} \frac{N}{I|\mathcal{S}^P_{it}|(T-1)} \left[\mathbf{G}^0_{ist} + \tilde{\mathbf{G}}^2_{ist}(\hat{a}^{\text{OLS}}_{ist} - \hat{a}_{ist}) + \tilde{\mathbf{G}}^3_{ist}(\hat{a}^{\text{OLS}}_{is\tau^P_{ist}} - \hat{a}_{is\tau^P_{ist}})\right]$$
$$= \frac{1}{\sqrt{N}} \sum_{i=1}^{I} \sum_{t=2}^{T} \sum_{s \in \mathcal{S}^P_{it}} \frac{N}{I|\mathcal{S}^P_{it}|(T-1)} \mathbf{G}^0_{ist} + \frac{1}{\sqrt{ND}} \sum_{t=1}^{T} \mathbf{G}_t \sqrt{D}\,\mathbf{vec}(\hat{\mathbf{A}}^{\text{OLS}}_t - \hat{\mathbf{A}}_t) + o_p(1).$$

The first term is asymptotically normal under the results of Forman and Sørensen (2008). The second term is asymptotically normal because $\hat{\mathbf{A}}^{\text{OLS}}_t$ is asymptotically normal by Lemma 3. □

For an adaption of the GMM generated-variable correction to second-stage OLS estimation, see the Online Supplement (Section S.2).

E Simulations

We perform simulations to explore how churning in comparative advantage alters standard counterfactual exercises in international trade. Answering this question requires a more involved simulation procedure than is common in the literature (see Alvarez and Lucas (2007) and Dekle, Eaton, and Kortum (2007)). Typical counterfactual exercises solve for changes in equilibrium outcomes as functions of changes in trade costs, so the treatment effect is deterministic. To answer our question of interest, we need to account for stochastic comparative advantage. Equilibrium outcomes are random variables driven by churning in comparative advantage and their stochastic properties depend on exogenous changes in trade costs.

Our approach measures the average treatment effect of a given change in trade costs. For this purpose, we repeatedly simulate the economy—drawing samples of comparative advantage based on our estimated GLD process—and calculate the cross-simulation average of the change in outcomes attributable to a change in trade costs. There are three steps: (1) simulate many sample paths for comparative advantage; (2) for each sample and year, solve for equilibrium outcomes with and without the change in trade costs and compute the percent difference; and (3) for each year, average this percent difference over all simulations. This appendix describes the procedure and the data requirements.

E.1 Inference of self trade

In order to perform the counterfactual analysis for the global economy, starting from initial conditions in 1990, we need a balanced dataset of comparative advantage estimates and industry-level expenditure across a consistent set of importers and exporters in 1990. In particular, we need to construct estimates of self trade by industry.

We do not have production data at the level of industries that we use in the empirical analysis, so we cannot compute industry-level self trade from production data. Instead, we infer the distribution of self trade at the industry level using our estimated fixed effects and discipline aggregate self trade using country-level production data. The set of importers that we analyze is smaller than the set of exporters, so we build our dataset using the sample of importers from our empirical analysis and aggregate all remaining countries (including countries that are exporters but not importers in our estimation as well as countries excluded from the estimation) into a single rest-of-world entity. In the process, we need to construct self-trade estimates also for this rest-of-world entity.

Our simulations are based on the following CES demand system for the expenditure by destination market d on goods from source country s within industry i during year t:

$$X_{isdt} = \frac{\left(w_{st}\tau_{isdt}/\underline{q}_{ist}\right)^{-\theta}}{\sum_{\varsigma}\left(w_{\varsigma t}\tau_{i\varsigma dt}/\underline{q}_{i\varsigma t}\right)^{-\theta}}\,\mu_{idt}E_{dt}, \tag{E.22}$$

where μ_{idt} is the Cobb-Douglas expenditure share of destination d on industry i.

In our empirical analysis, we assume that trade costs take the following log-linear form

$$-\theta\ln\tau_{isdt} = -c_{dt} + \mathbf{r}'_{sdt}\mathbf{b}_{it} + v_{isdt} \quad \text{for} \quad s \neq d,$$

where c_{dt} is an unobserved destination-year component of trade costs that captures the *closedness* of the destination market. This parameter determines d's aggregate self-trade share, which we will infer from production data. The term ϵ_{isdt} captures unobserved idiosyncratic trade costs that we assume are mean zero over source countries within any given industry, destination, and year, conditional on the gravity covariate vector $\mathbf{r}_{sdt}$.

For self trade, there are no trade costs: $\tau_{idd} = 1$. Both this restriction and the lack of self-trade data imply that we must exclude self trade from gravity regressions.[53] For $s \neq d$, our specification of trade costs implies the following regression at the industry-year level across sources and destinations:

$$\ln X_{isdt} = k_{ist} + m_{idt} + \mathbf{r}'_{sdt}\mathbf{b}_{it} + v_{isdt}, \tag{E.23}$$

where

$$k_{ist} \equiv \theta\ln(\underline{q}_{ist}/w_{st})$$

is our measure of export capability and

$$m_{idt} \equiv -c_{dt} + \ln\left[\mu_{idt}Y_{dt}/\sum_{\varsigma}\left(w_{\varsigma t}\tau_{i\varsigma dt}/\underline{q}_{i\varsigma t}\right)^{-\theta}\right]$$

is our measure of import propensity. We will use our estimates of these fixed effects to infer self trade by industry.

The normalization $\tau_{idd} = 1$ implies that industry-level self trade is directly related to the gravity fixed effect

[53] The normalization of $\tau_{idd} = 1$ means that the regression equation does not hold for self-trade observations. The logarithm for self trade satisfies

$$\ln X_{iddt} = k_{idt} + m_{idt} + c_{dt}$$

instead of the gravity specification. Given the normalization of $\tau_{idd} = 1$, the model implies a structural relationship between self trade, export capability, import propensity, and closedness. Intuitively, the normalization $\tau_{idd} = 1$ means that trade costs are defined relative to internal trade costs. That is, normalizing τ_{idd} forces estimation to be in units relative (within an industry across sources) to the destination's local covariates (such as common language or internal distance).

estimates (export capabilities k_{idt} and import propensities m_{idt}) by

$$X_{iddt} = \frac{\left(w_{dt}/\underline{q}_{idt}\right)^{-\theta}}{\sum_{\varsigma}\left(w_{\varsigma t}\tau_{i\varsigma dt}/\underline{q}_{i\varsigma t}\right)^{-\theta}}\,\mu_{idt}Y_{dt} = e^{k_{idt}+m_{idt}+c_{dt}}.$$

The self trade of a country in industry i is increasing in the country-industry's export capability and import propensity, and in the country's closedness c_{dt}. As a country closes to trade, it reallocates expenditure from the products of other countries towards its own products. Note that, although the two fixed effects are only identified up to a global normalization (and we normalize import propensities in the United States to zero), the sum of fixed effects is always identified. As a result, we can infer industry-level self trade from fixed effect estimates up to a destination market's closedness c_{dt}.

Closedness is common across industries within a destination market, so an industry i's share of self trade can be recovered using

$$\frac{X_{iddt}}{\sum_i X_{iddt}} = \frac{e^{k_{idt}+m_{idt}+c_{dt}}}{\sum_i e^{k_{idt}+m_{idt}+c_{dt}}} = \frac{e^{k_{idt}+m_{idt}}}{\sum_i e^{k_{idt}+m_{idt}}}.$$

That is, we can compute the distribution of self trade across industries directly from our estimates of export capabilities and import propensities.

We then use aggregate production data to discipline the overall level of self trade, which is

$$\sum_i X_{iddt} = E_{dt} - \sum_i \sum_{s\neq d} X_{isdt}.$$

Expenditure is related to aggregate production through $E_{dt} = Y_{dt} - TB_{dt}$ where $TB_{dt} \equiv \sum_i \sum_{d'\neq d} X_{idd't} - \sum_i \sum_{s\neq d} X_{isdt}$ is the trade balance. We therefore have

$$\sum_i X_{iddt} = Y_{dt} - \sum_i \sum_{d'\neq d} X_{idd't}.$$

Self trade equals total aggregate production net of exports. We use UNIDO data on production from 1977 to 2004 as well as WIOD data for China and Taiwan in 1995. We calculate aggregate exports from the trade flow data we used in our empirical analysis.

From the individual self-trade shares by industry and country-level self trade, we compute industry-level self trade as

$$X_{iddt} = \frac{X_{iddt}}{\sum_i X_{iddt}}\sum_i X_{iddt} = \frac{e^{k_{idt}+m_{idt}}}{\sum_i e^{k_{idt}+m_{idt}}}\left(Y_{dt} - \sum_i \sum_{d'\neq d} X_{idd't}\right).$$

Note that the inclusion of the closedness parameter in our specification for trade costs allows us to rationalize any level of aggregate self trade because

$$\sum_i X_{iddt} = \sum_i e^{k_{idt}+m_{idt}+c_{dt}} = e^{c_{dt}}\sum_i e^{k_{idt}+m_{idt}}.$$

E.1.1 Accounting for missing gravity fixed effect estimates

Some estimates of gravity fixed effects (export capabilities and import propensities) are missing. We use the following interpolation and extrapolation procedure to fill in the missing estimates.

First, we calculate country-year means of export capabilities $(1/I)\sum_{i=1} k_{ist}$ as well as import propensities

$(1/I)\sum_{i=1} m_{ist}$. Note that $m_{i\text{US}t} = 0$ due to our omission of the U.S. importer fixed effect in gravity regressions. These country-year means capture overall economic growth within a country over time. Time interpolation and extrapolation without accounting for these aggregate trends can generate patterns of inferred comparative advantage and inferred self trade driven by patterns of missing observations—particularly in the early and late parts of the sample.[54]

Second, to account for aggregate trends, we remove country-year means and calculate the residuals $k_{ist} - (1/I)\sum_{i=1} k_{ist}$ and $m_{idt} - (1/I)\sum_{i=1} m_{ist}$. We interpolate and extrapolate these residuals over years within each industry-country.[55] Finally, we add back in the country-year means to obtain interpolated fixed effect estimates.

E.1.2 Constructing a rest-of-world aggregate

We can perform those calculations only for the set of importers in our estimation sample. To balance global trade, we need an aggregate entity for the rest of the world (ROW). To incorporate a rest-of-world entity into our counterfactual analysis, we require estimates of rest-of-world self trade. We use the following aggregation procedure based on constructing synthetic gravity covariates for the rest-of-world entity that best explain the observed aggregate trade flows.

Order the sample so that the first M countries are the importer subsample. For the remaining $S-M$ countries, we cannot infer self trade because we cannot identify importer fixed effects. Define the aggregate exports from ROW (indexed with $s = 0$) as

$$X_{i0dt} = \sum_{s=M+1}^{S} X_{isdt} \quad \text{for each} \quad d = 1, \ldots, M$$

and aggregate imports to ROW (indexed with $d = 0$) by

$$X_{is0t} = \sum_{d=M+1}^{S} X_{isdt} \quad \text{for each} \quad s = 1, \ldots, M.$$

Sectoral self trade of ROW is defined as

$$X_{i00t} = \sum_{s=M+1}^{S} \sum_{d=M+1}^{S} X_{isdt}.$$

This quantity is unobserved because we do not know industry-level self trade for the countries within the ROW aggregate. However, instead of having to infer self trade for all $S - M$ ROW countries, we only need to infer self trade for the aggregate ROW entity.

To do so, we can choose the "location" of this synthetic ROW country relative to each $s = 1, \ldots, M$ (captured by its bilateral gravity covariates) in order to rationalize the observed ROW aggregate trade flows. As before, we assume an idiosyncratic component for trade costs and choose CES import demand for the synthetic ROW country. The gravity equation (2) therefore holds for each $s, d = 1, \ldots, M$. As a result, flows from ROW to each

[54] If we were to use time extrapolation without adjusting for country trends, then any estimates that are missing for multiple observations at the start or end of the sample would be constant. As a result, those estimates would be systematically above trend or systematically below trend.

[55] The exercise is within industry-country because we use time interpolation. Possible alternatives include averaging over years in either the industry or country dimension, or using a forecasting approach such as Kalman filtering to obtain maximum likelihood estimates of missing observations.

$d = 1, \ldots, M$ satisfy

$$\ln X_{i0dt} - m_{idt} = k_{i0t} + \mathbf{r}'_{0dt}\mathbf{b}_{it} + v_{i0dt},$$

where the gravity covariate vector r_{0dt} represents how "far" the synthetic ROW country is from each importer d. We are free to choose this vector to best explain the ROW aggregate imports.

We choose to construct the synthetic country's gravity covariates as a weighted average of the covariates of the underlying countries

$$\mathbf{r}_{0dt} = \sum_{s=1+M}^{S} \zeta_{sdt}\,\mathbf{r}_{sdt},$$

where $\sum_{s=1+M}^{S} \zeta_{st} = 1$. Note that we can incorporate the constraint by writing

$$\mathbf{r}_{0dt} = \mathbf{r}_{M+1,dt} + \sum_{s=M+2}^{S} \zeta_{sdt}(\mathbf{r}_{sdt} - \mathbf{r}_{M+1,dt}).$$

We then use a regression to find the weights that best explain the observed trade flows. Define

$$y_{idt} \equiv \ln X_{i0dt} - m_{idt} - \mathbf{r}'_{M+1,dt}\mathbf{b}_{it}$$

and

$$x_{isdt} \equiv (\mathbf{r}_{\varsigma dt} - \mathbf{r}_{M+1,dt})'\mathbf{b}_{it} \quad \text{for each } \varsigma = M+2, \ldots, S.$$

Then regress y_{idt} on $\mathbf{x}_{idt} = (x_{i,M+2,dt}, \ldots, x_{iSdt})'$ for each country $d = 1, \ldots, M$ and year t using variation across i. Note that

$$\begin{aligned} k_{i0t} + v_{i0dt} &= \ln X_{i0dt} - m_{idt} - \mathbf{r}'_{0dt}\mathbf{b}_{it} \\ &= y_{idt} - \mathbf{x}'_{idt}\boldsymbol{\zeta}_{dt} \\ &= \text{constant}_{dt} + \text{residual}_{idt}, \end{aligned}$$

where $\boldsymbol{\zeta}_{dt} = (\zeta_{M+2,dt}, \ldots, \zeta_{Sdt})'$ is the vector of coefficients from this regression. We can therefore infer the export capability of ROW in industry i to be

$$k_{i0t} = \frac{1}{M}\sum_{d=1}^{M}(\text{constant}_{dt} + \text{residual}_{idt}).$$

Similarly, we can infer import propensities and trade costs from a regression (for each $s = 1, \ldots, M$ and t) of $\ln X_{is0t} - k_{ist} - \mathbf{r}'_{s,M+1,t}\mathbf{b}_{it}$ on $(\mathbf{r}'_{sd't} - \mathbf{r}'_{s,M+1,t})\mathbf{b}_{it}$ for $d' = M+2, \ldots, S$ using variation across i.[56] From this regression we obtain

$$m_{i0t} = \frac{1}{M}\sum_{s=1}^{M}(\text{constant}_{st} + \text{residual}_{ist}).$$

Having values for k_{i0t} and m_{i0t}, we use the previous procedure to compute industry-level self-trade shares for the ROW aggregate using production data for the ROW. Combined with the inferred industry-level self trade for the sample of importers, we have a complete set of trade flows (including self trade) for all years, industries, and country pairs (sd) with $s, d = 0, \ldots, M$. This information is sufficient to conduct counterfactual analysis

[56] When there are industries with missing export capability, those industries drop from the regression. If we wanted to impose the same weights between the two regressions (so that we could interpret the procedure more directly as defining a synthetic country whose characteristics are a weighted average of the underlying countries), we could stack the data and run a single regression.

since we only need initial expenditure shares and initial comparative advantage (which we calculate from interpolated export capabilities) to be able to simulate complete GLD paths for comparative advantage and calculate counterfactuals using exact hat algebra methods.

E.2 Solving for equilibrium

In any year t and for given trade costs $\{\{\tau_{isdt}\}_{i=1}^{I}\}\}_{s\neq d}$, productivities $\{\{\underline{q}_{ist}\}_{i=1}^{I}\}\}_{s=1}^{S}$, preference weights $\{\{\mu_{idt}\}_{i=1}^{I}\}\}_{d=1}^{S}$, endowments $\{L_{st}\}_{s=1}^{S}$, and trade balances $\{TB_{dt}\}_{d=1}^{N}$, a *competitive equilibrium* is a collection of wages $\{w_{st}\}_{s=1}^{N}$, incomes $\{Y_{st}\}_{s=1}^{N}$, and expenditures $\{E_{dt}\}_{d=1}^{S}$ such that:

i. the labor market in source country s clears

$$w_{st}L_{st} = \sum_{d} \frac{\left(w_{st}\tau_{isdt}/\underline{q}_{ist}\right)^{-\theta}}{\sum_{\varsigma}\left(w_{\varsigma t}\tau_{i\varsigma dt}/\underline{q}_{i\varsigma t}\right)^{-\theta}}\,\mu_{idt}E_{dt} \quad \text{and}$$

ii. the goods market in each destination country d clears

$$w_{dt}L_{dt} = Y_{dt} = E_{dt} + TB_{dt}.$$

Denote the industry-level expenditure share with

$$\pi_{isdt} \equiv \frac{\left(w_{st}\tau_{isdt}/\underline{q}_{ist}\right)^{-\theta}}{\sum_{\varsigma}\left(w_{\varsigma t}\tau_{i\varsigma dt}/\underline{q}_{i\varsigma t}\right)^{-\theta}}.$$

For any quantity x, let $\hat{x} \equiv x'/x$ denote the proportional change to some counterfactual x'. From observed equilibrium expenditure shares of $\{\pi_{isdt}\}_{i,s,d}$, we can solve for the percent change in equilibrium wages due to the combination of a change in productivities $\{\hat{\underline{q}}_{ist}\}_{i,s}$ and a change in trade costs $\{\hat{\tau}_{isdt}\}_{i,s,d}$ by finding the change in wages $\{\hat{w}_{st}\}_{s}$ such that

$$\hat{w}_{st}Y_{st} = \sum_{d} \hat{\pi}_{isdt}\pi_{isdt}\mu_{idt}(\hat{w}_{dt}Y_{dt} - TB_{dt}),$$

where the change in industry-level trade shares is

$$\hat{\pi}_{isdt} \equiv \frac{\left(\hat{w}_{st}\hat{\tau}_{isdt}/\hat{\underline{q}}_{ist}\right)^{-\theta}}{\sum_{\varsigma}\pi_{i\varsigma dt}\left(\hat{w}_{\varsigma t}\hat{\tau}_{i\varsigma dt}/\hat{\underline{q}}_{i\varsigma t}\right)^{-\theta}}.$$

Note that we can compute all necessary initial equilibrium quantities (incomes, trade balances, and preferences) from (square) expenditure matrices across industries, $\{X_{isdt}\}_{isd}$, as

$$E_{dt} = \sum_{i}\sum_{s} X_{isdt}, \quad Y_{st} = \sum_{i}\sum_{d} X_{isdt}, \quad TB_{dt} = Y_{dt} - E_{dt}, \quad \text{and} \quad \mu_{idt} = \frac{\sum_{s} X_{isdt}}{E_{dt}}.$$

Similar to Dekle, Eaton, and Kortum (2007), we use the tâtonnement algorithm of Alvarez and Lucas (2007) to solve for the equilibrium change in wages each period while accounting for non-zero trade balances.

E.3 Counterfactuals

We want to assess how churning in comparative advantage influences the conclusions from common counterfactual exercises under trade-cost changes. In particular, we consider how a 10% decrease in trade costs for top comparative advantage industries in 1990 China impacts the equilibrium real wage and exports from China. In our exercise, we hold trade balances fixed at their 1990 levels from twenty years of simulations and solve for equilibrium outcomes for the whole world under various comparative advantage scenarios—that is, different stochastic processes for comparative advantage.

We consider three comparative advantage scenarios:

1. **Static Equilibrium**: hold the distribution of comparative advantage fixed (in all industries and countries) at 1990 levels.

2. **Transition Path**: initialize comparative advantage at 1990 levels and allow the distribution to evolve over time based on our estimated GLD process.

3. **Steady State**: sample initial comparative advantages from the stationary distribution implied by GLD estimates and then allow it to evolve over time.

The first scenario, static equilibrium, captures a typical exercise in the trade literature—productivities are held fixed and not allowed to evolve stochastically. The second scenario, transition path, allows us to visualize how our estimated GLD process implies transition dynamics—how the influence of initial comparative advantages changes as churning leads to convergence to the stationary distribution of comparative advantage. The third scenario, steady state, allows us to remove the effect of initial conditions. When we sample from the stationary distribution in this third scenario and simulate comparative advantage over time, we converge to the stationary distribution (for a sufficiently large number of simulations). This scenario captures the long-run impact of a permanent trade cost change since the distribution of comparative advantage converges to the stationary distribution in the long run, while the influence of initial conditions fades.

To assess the average impact of a given trade cost change, we simulate a large number of paths for comparative advantage. For each t and simulation sample j, we compute the change in equilibrium wages $\{\hat{w}_{st}^{(j)}\}_s$ where the superscript indexes the simulation sample. We can then compute the implied change in real wages, and the level of trade flows as

$$\frac{\hat{w}_{dt}^{(j)}}{\hat{p}_{dt}^{(j)}} = \frac{\hat{w}_{dt}^{(j)}}{\prod_i \left[\sum_s \hat{\pi}_{isdt}^{(j)} (\hat{\tau}_{isdt} \hat{w}_{st}^{(j)} / \hat{\underline{q}}_{ist}^{(j)})^{-\theta}\right]^{-\frac{\mu_{idt}}{\theta}}}$$

and

$$X_{isdt}^{(j)\prime} = \frac{\pi_{isdt} \left(\hat{w}_{st}^{(j)} \hat{\tau}_{isdt} / \hat{\underline{q}}_{ist}^{(j)}\right)^{-\theta}}{\sum_\varsigma \pi_{i\varsigma dt} \left(\hat{w}_{\varsigma t}^{(j)} \hat{\tau}_{i\varsigma dt} / \hat{\underline{q}}_{i\varsigma t}^{(j)}\right)^{-\theta}} \mu_{idt} \left(\hat{w}_{dt}^{(j)} Y_{dt} - TB_{dt}\right).$$

For each source country, we aggregate these trade flows across destination markets and industries to get exports in those "treated"" industries (those where trade costs are reduced), and also to get the level of total exports.

We compute these quantities for each simulation sample $j = 1, \ldots, J$ for a baseline counterfactual ($\hat{\tau}_{isdt} = 1$), where trade costs do not change (but equilibrium does change because comparative advantages change) and for a treatment counterfactual where trade costs also change ($\hat{\tau}_{isdt} = 0.9$). We then compute the percent difference between the two counterfactuals to get the within-sample-j treatment effect of the trade cost change. Finally, we average over samples to get a measure of the average treatment effect of the trade cost change across simulations within a given productivity scenario (static equilibrium, transition, or steady state).

F Additional Evidence

In this Appendix, we report additional evidence to complement the reported findings in the text.

F.1 Cumulative probability distribution of absolute advantage

Figures A1, A2 and **A3** extend **Figure 1** in the text and plot, for 28 countries in 1967, 1987 and 2007, the log number of a source country s's industries that have at least a given level of absolute advantage in year t against that log absolute advantage level $\ln A_{ist}$ for industries i. The figures also graph the fit of absolute advantage in the cross section to a Pareto distribution and to a log normal distribution using maximum likelihood, where each cross sectional distribution is fit separately for each country in each year (such that the number of parameters estimated equals the number of parameters for a distribution $\times$ number of countries $\times$ number of years). In the Online Supplement (Section S.6) we show comparable cumulative probability distributions of log absolute advantage for PPML-based exporter capability, the Balassa RCA index, and varying industry aggregates of OLS-based exporter capability.

F.2 GLD predicted cumulative probability distributions of absolute advantage

Figures A4, A5 and **A6** present plots for the same 28 countries in 1967, 1987 and 2007 as shown before (in **Figures A1, A2** and **A3**), using log absolute advantage from OLS-based exporter capability. **Figures A4** through **A6** contrast graphs of the actual data with the GLD implied predictions and show a lose fit.

F.3 Comparative advantage at varying industry aggregates

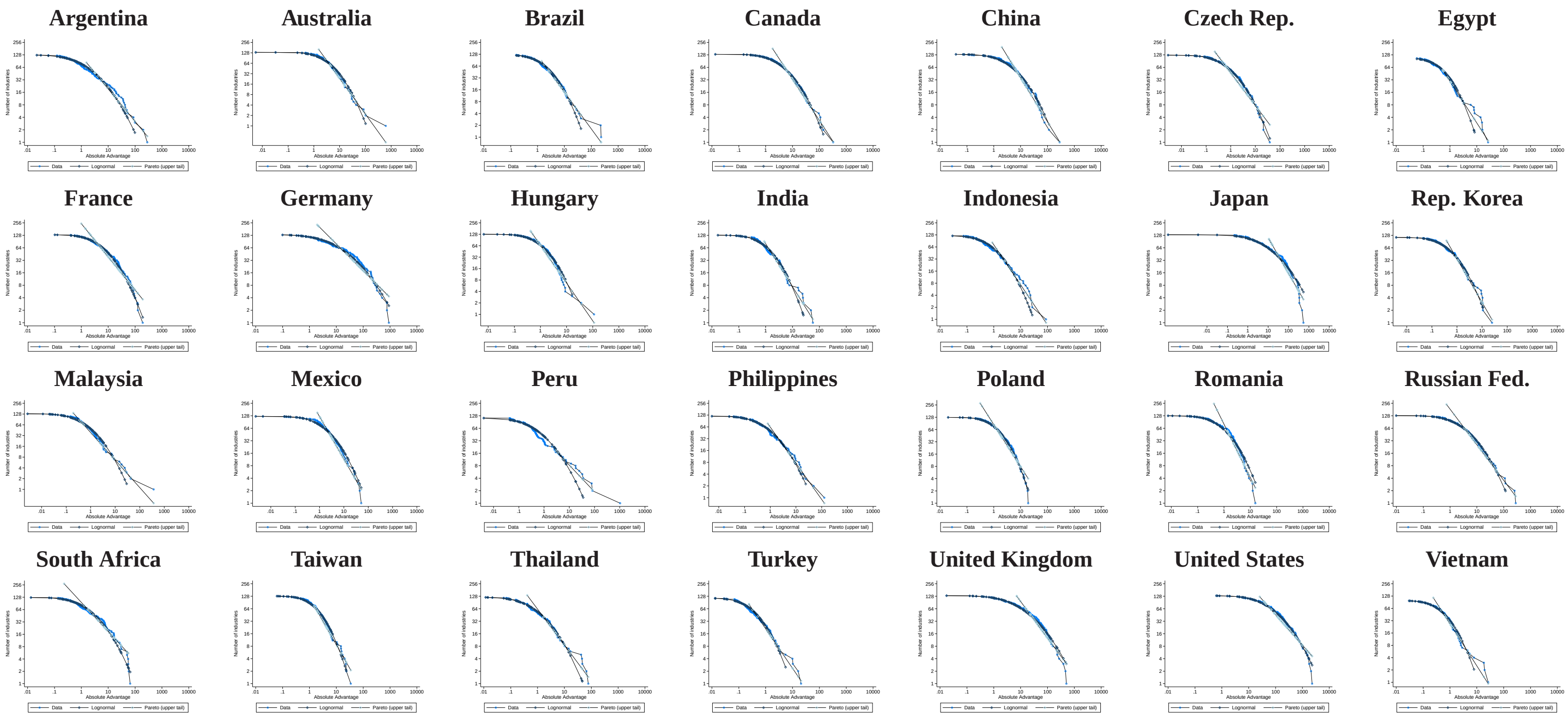
As a robustness check, we restrict the sample to the period 1984-2007 with industry aggregates from the SITC revision 2 classification. Data in this late period allow us to construct varying industry aggregates. We first obtain gravity-based estimates of log absolute advantage from OLS (6) at the refined industry aggregates. Following our benchmark specifications in the text, we then estimate the decay regression (10) at ten-year intervals and the GLD model (C.17) using GMM at five-year intervals.

For the decay regression, **Table A1** repeats in columns 1, 4 and 7 the estimates from **Table 1** for our benchmark industry-level aggregates at the SITC 2-3 digit level (133 industries) during the full sample period 1962-2007. **Table A1** presents in the remaining columns estimates for the SITC revision 2 two-digit level (60 industries) and the three-digit level (224 industries) during the late period 1984-2007. At the two-digit level (60 industries), the ten-year decay rate for absolute advantage using all countries and industries is -0.26, at the three-digit level (224 industries) it is -0.37. When using PPML-based log absolute advantage or the log RCA index, decay rates vary less across aggregation levels, ranging from -0.31 at the two-digit level for PPML-based log absolute advantage to -0.34 at the three-digit level for log RCA. The qualitative similarity in decay rates across definitions of export advantage and levels of industry aggregation suggest that our results are neither the byproduct of sampling error nor the consequence of industry definitions.

For the GLD model under the GMM procedure, Table A2 confirms that results remain largely in line with those in **Table 2** before, for the benchmark aggregates at the SITC 2-3 digit level (133 industries) during 1962-2007. The benchmark estimates are repeated in columns 1, 4 and 7. In the other columns, **Table A2** presents estimates for the SITC revision 2 two-digit level (60 industries) and the three-digit level (224 industries) during the late period 1984-2007.

Estimates of the dissipation rate η are slightly larger during the post-1984 period than over the full sample period and, similar to the implied η estimate in the decay regressions above, become smaller as we move from broader to finer classifications of industries. Estimates of the elasticity of decay ϕ are statistically significantly

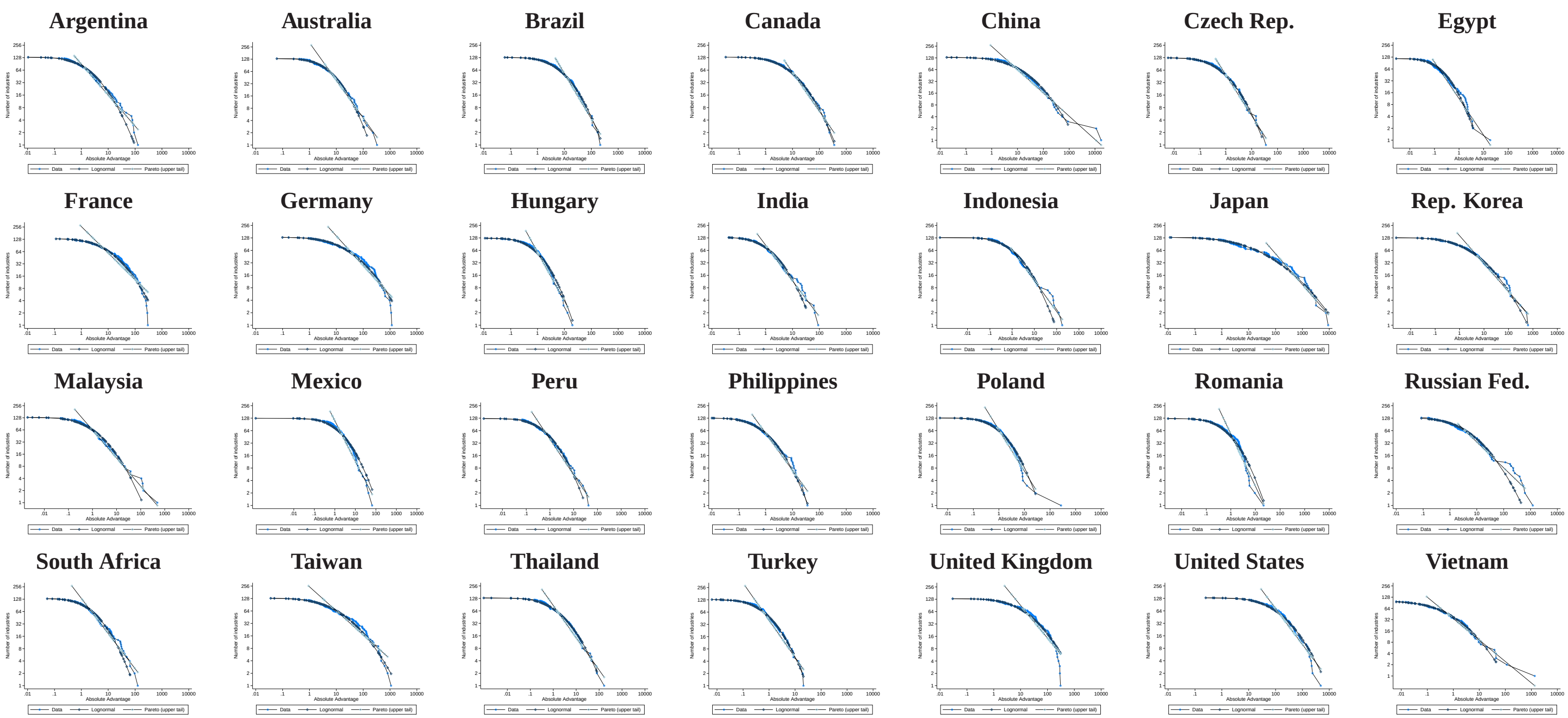
Figure A1: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 1967**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1965-1967 and CEPII.org; three-year means of OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).

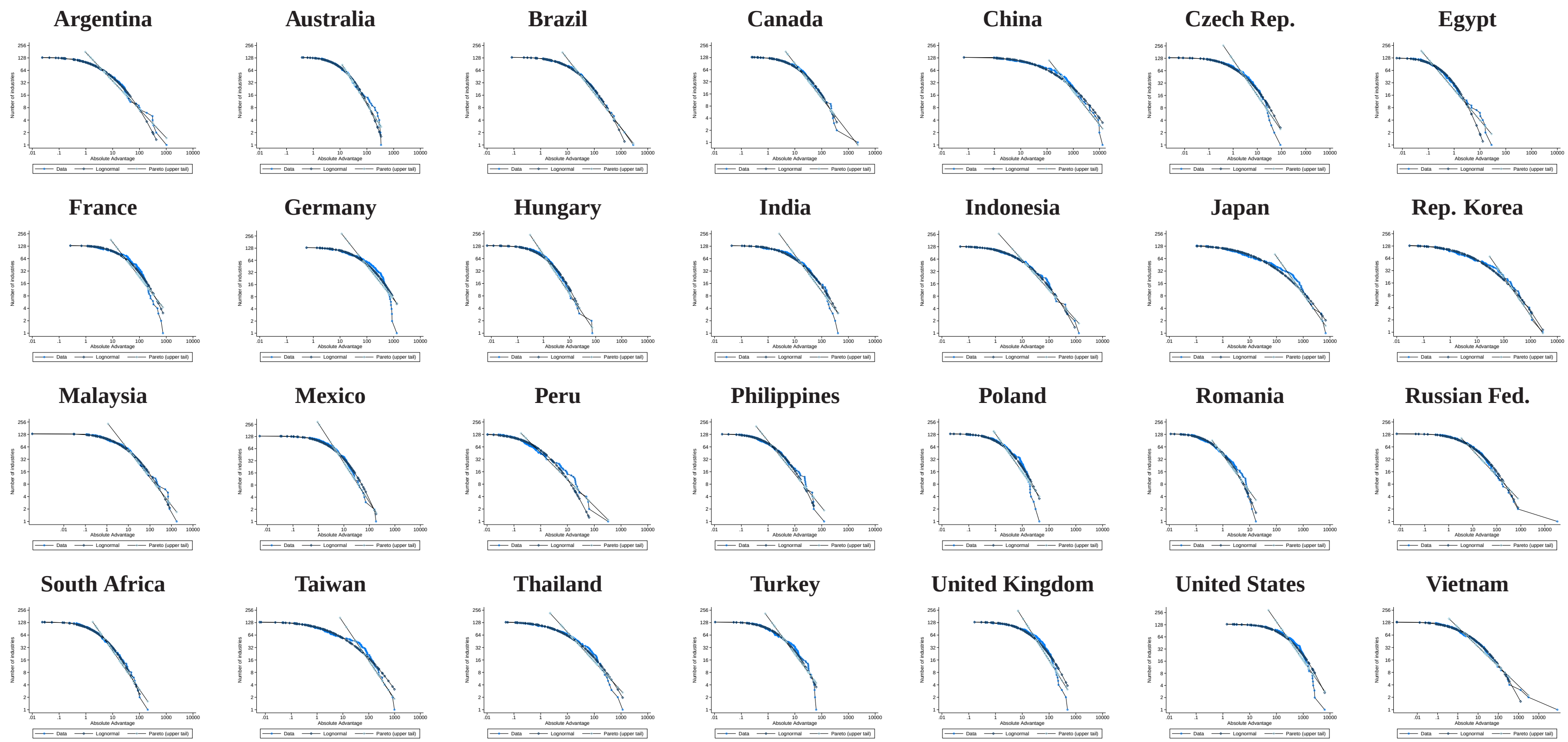
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 1967$ (Pareto fit to upper five percentiles only).

Figure A2: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 1987**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1985-1987 and CEPII.org; three-year means of OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 1987$ (Pareto fit to upper five percentiles only).

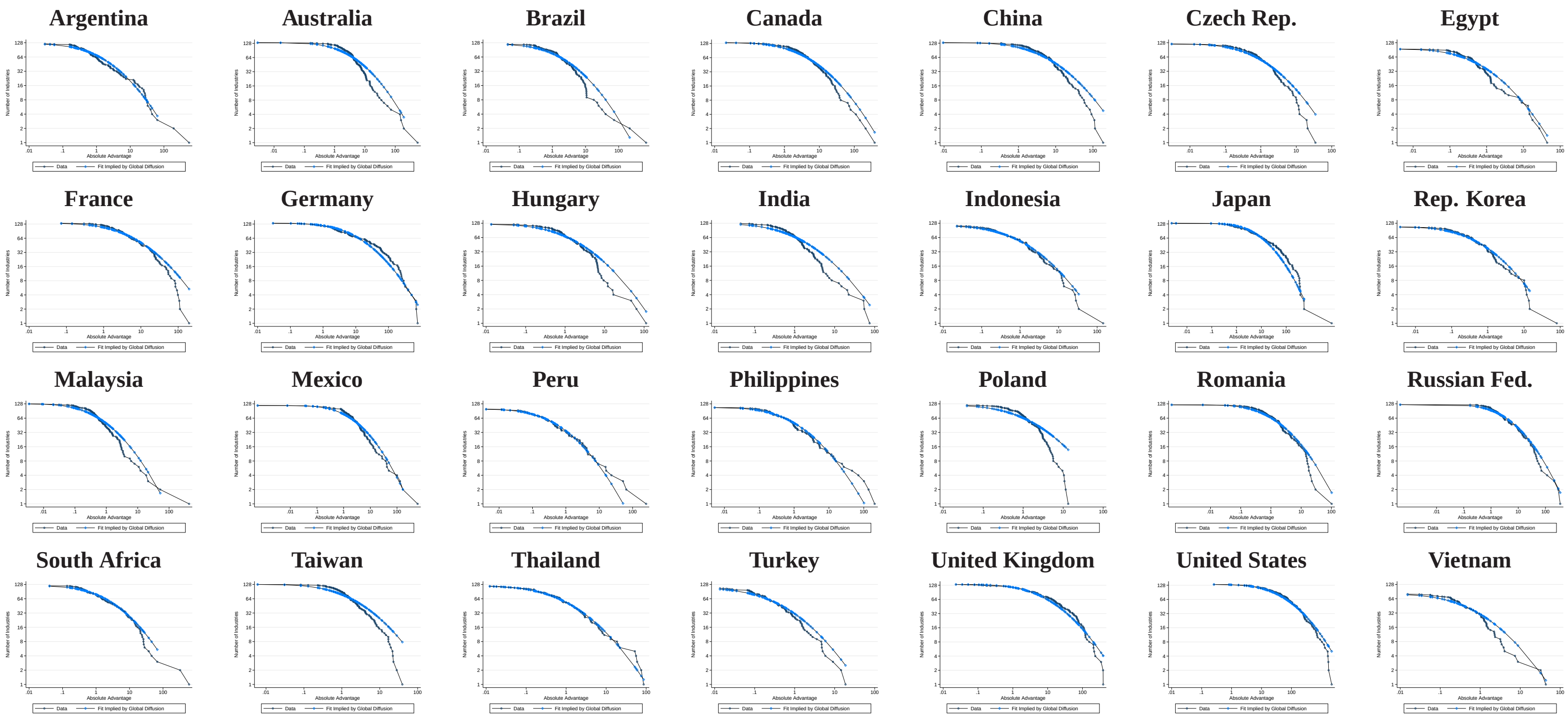
Figure A3: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 2007**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 2005-2007 and CEPII.org; three-year means of OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).

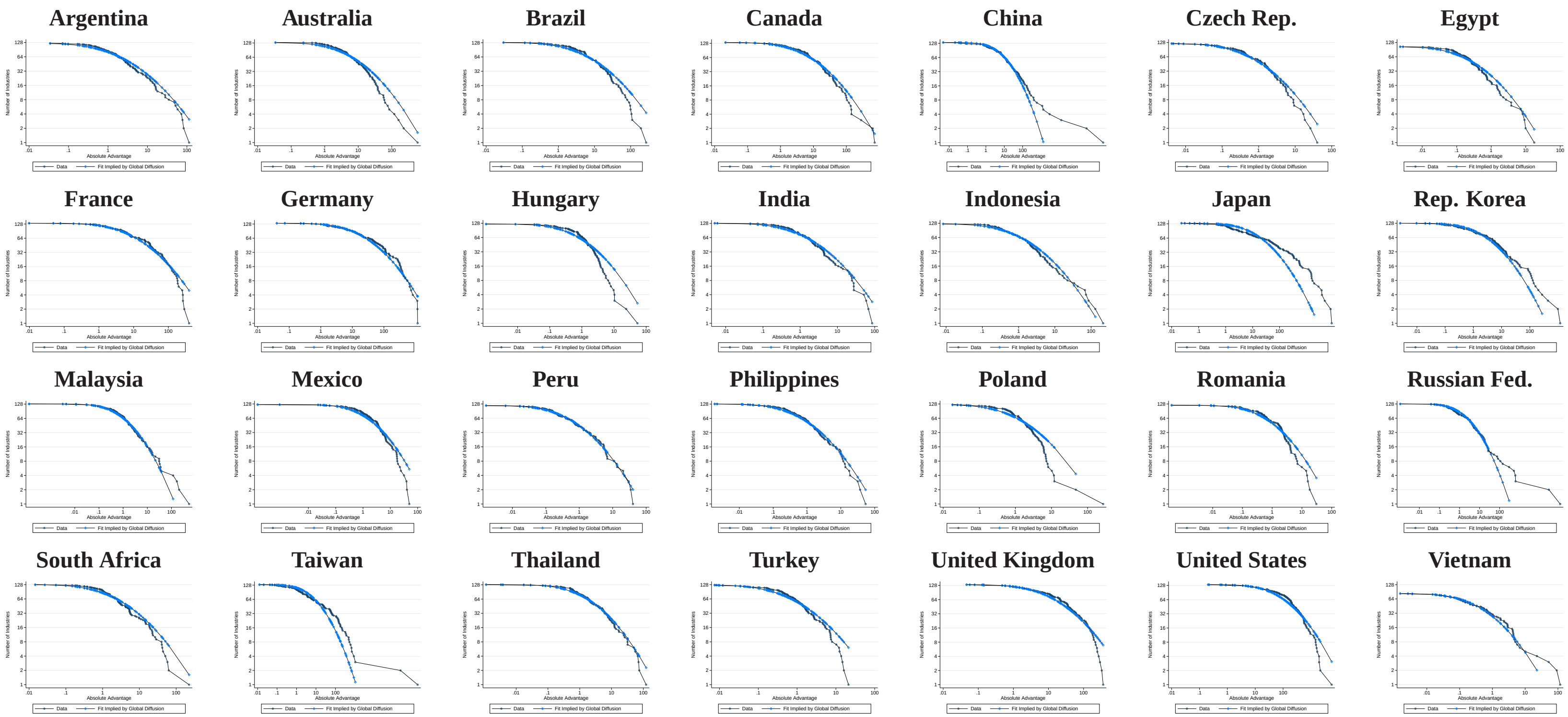
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 2007$ (Pareto fit to upper five percentiles only).

Figure A4: **Diffusion Predicted and Observed Cumulative Probability Distributions of Absolute Advantage in 1967**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).
Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (15) in Table 2 (parameters η and ϕ in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (18), which horizontally shifts the distributions but does not affect their shape.

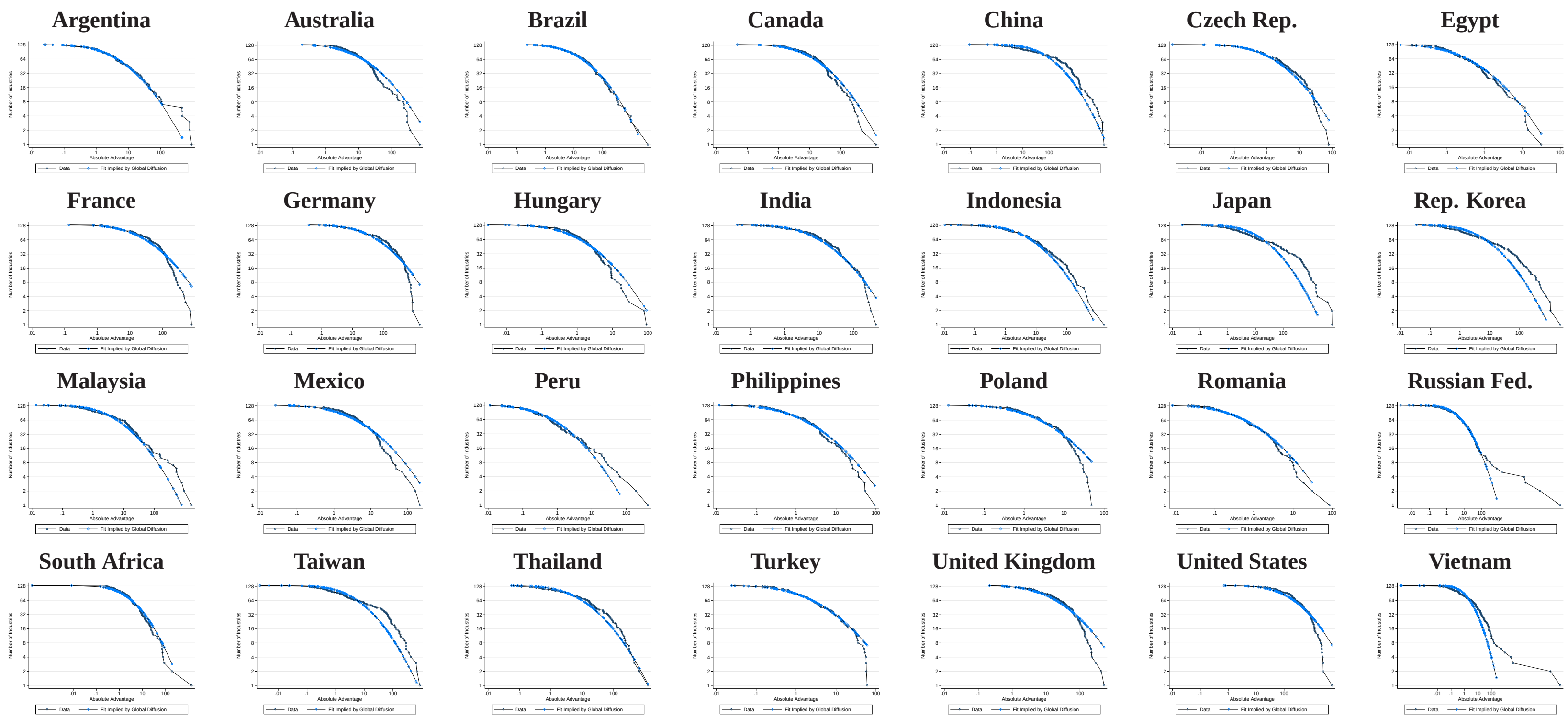
Figure A5: **Diffusion Predicted and Observed Cumulative Probability Distributions of Absolute Advantage in 1987**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).

Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis, for the year $t = 1987$. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (15) in Table 2 (parameters η and ϕ in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (18), which horizontally shifts the distributions but does not affect their shape.

Figure A6: **Diffusion Predicted and Observed Cumulative Probability Distributions of Absolute Advantage in 2007**



Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; OLS gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6).

Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (15) in Table 2 (parameters η and ϕ in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (18), which horizontally shifts the distributions but does not affect their shape.

Table A1: DECAY REGRESSIONS FOR COMPARATIVE ADVANTAGE, VARYING INDUSTRY AGGREGATES

	OLS gravity k			PPML gravity k			$\ln$ *RCA*		
		2-dgt.	3-dgt.		2-dgt.	3-dgt.		2-dgt.	3-dgt.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Decay Regression Coefficients									
Decay rate ρ	-0.349	-0.257	-0.370	-0.320	-0.320	-0.343	-0.303	-0.307	-0.326
	(0.002)***	(0.003)***	(0.002)***	(0.0002)***	(0.0003)***	(0.0003)***	(0.01)***	(0.017)***	(0.01)***
Var. of residual s^2	2.089	1.463	2.005	2.709	1.889	2.583	2.318	1.678	2.267
	(0.024)***	(0.027)***	(0.023)***	(0.013)***	(0.024)***	(0.017)***	(0.006)***	(0.009)***	(0.007)***
Implied Ornstein-Uhlenbeck (OU) Parameters									
Dissipation rate η	0.276	0.306	0.301	0.198	0.284	0.220	0.222	0.310	0.241
	(0.003)***	(0.006)***	(0.004)***	(0.0009)***	(0.004)***	(0.001)***	(0.006)***	(0.014)***	(0.006)***
Intensity of innovations σ	0.558	0.441	0.554	0.623	0.52	0.618	0.570	0.486	0.572
	(0.003)***	(0.004)***	(0.003)***	(0.001)***	(0.003)***	(0.002)***	(0.005)***	(0.008)***	(0.006)***
Observations	324,978	70,609	230,395	320,310	70,457	227,061	324,983	70,609	230,396
Adjusted R^2 (within)	0.222	0.241	0.265	0.282	0.315	0.295	0.216	0.233	0.224
Years t	36	14	14	36	14	14	36	14	14
Industries i	133	60	224	133	60	224	133	60	224
Source countries s	90	90	90	90	90	90	90	90	90

Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) in 90 countries for 133 time-consistent industries from 1962-2007, for 60 time-consistent industries at the 2-digit SITC level from 1984-2007, and 224 industries at at the 3-digit SITC level from 1984-2007, and CEPII.org; OLS and PPML gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6) and (8).
Note: Reported figures for ten-year changes. Variables are OLS and PPML gravity measures of log absolute advantage $\ln A_{ist}$ and the log Balassa index of revealed comparative advantage $\ln \mathit{RCA}_{ist} = \ln(X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_{\iota} X_{\iota st}/\sum_{\iota}\sum_{\varsigma} X_{\iota\varsigma t})$. OLS estimation of the ten-year decay rate ρ from

$$k_{is,t+10} - k_{ist} = \rho\, k_{ist} + \delta_{it} + \delta_{st} + \epsilon_{is,t+10},$$

conditional on industry-year and source country-year effects δ_{it} and δ_{st} for 1962-2007 (column 1-2) and 1984-2007 (columns 3-6). The implied dissipation rate η and squared innovation intensity σ^2 are based on the decay rate estimate ρ and the estimated variance of the decay regression residual $\hat{s}^2$ by (13). Robust standard errors, clustered at the industry level and corrected for generated-regressor variation of export capability k, for ρ and s^2, applying the multivariate delta method to standard errors for η and σ. * marks significance at ten, ** at five, and *** at one-percent level.

Table A2: GMM Estimates of Comparative Advantage Diffusion, Varying Industry Aggregates

	OLS gravity k			PPML gravity k			$\ln$ *RCA*		
		2-dgt.	3-dgt.		2-dgt.	3-dgt.		2-dgt.	3-dgt.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimated Generalized Logistic Diffusion Parameters									
Dissipation rate η	0.256	0.297	0.287	0.180	0.289	0.205	0.212	0.332	0.204
	(0.004)***	(0.014)***	(0.004)***	(0.006)***	(0.107)***	(0.053)***	(0.006)***	(0.02)***	(0.035)***
Intensity of innovations σ	0.739	0.558	0.715	0.767	0.613	0.798	0.713	0.574	0.678
	(0.01)***	(0.011)***	(0.006)***	(0.037)***	(0.454)	(0.336)**	(0.051)***	(0.09)***	(0.42)
Elasticity of decay ϕ	-0.041	-0.070	-0.023	-0.009	-0.035	0.024	0.006	-0.014	-0.008
	(0.017)**	(0.024)***	(0.01)**	(0.035)	(0.593)	(0.272)	(0.053)	(0.119)	(0.404)
Implied Parameters									
Log gen. gamma scale $\ln\hat{\theta}$	121.94	59.09	281.50	900.95	155.22	-239.56	-1,410.50	548.19	1,083.20
	(71.526)*	(31.021)*	(161.258)*	(4581.812)	(3570.434)	(3595.553)	(14980.320)	(6069.679)	(72277.940)
Log gen. gamma shape $\ln\kappa$	5.017	4.115	6.338	7.788	5.456	5.842	8.641	7.484	8.181
	(0.842)***	(0.724)***	(0.875)***	(8.062)	(33.387)	(22.563)	(17.289)	(17.421)	(107.014)
Mean/median ratio	8.203	6.597	6.087	16.897	6.222	10.293	10.256	4.643	12.085
Observations	392,850	96,989	322,860	389,290	96,828	319,140	392,860	96,989	322,860
Industry-source obs. $I \times S$	11,542	5,332	19,160	11,531	5,331	19,118	11,542	5,332	19,160
Root mean sq. forecast error	1.851	1.690	1.737	1.898	1.664	1.817	1.760	1.560	1.768
Min. GMM obj. ($\times$ 1,000)	3.27e-13	1.82e-12	9.14e-13	2.56e-12	3.53e-11	9.91e-12	6.79e-12	1.61e-10	6.01e-11

Source: WTF (Feenstra, Lipsey, Deng, Ma, and Mo 2005, updated through 2008) in 90 countries for 133 time-consistent industries from 1962-2007, for 60 time-consistent industries at the 2-digit SITC level from 1984-2007, and 224 industries at at the 3-digit SITC level from 1984-2007, and CEPII.org; OLS and PPML gravity measures of export capability (log absolute advantage) $k = \ln A$ from (6) and (8).
Note: GMM estimation at the five-year horizon for the generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$,

$$\mathrm{d}\ln\hat{A}_{is}(t) = -\frac{\eta\sigma^2}{2}\,\frac{\hat{A}_{is}(t)^{\phi}-1}{\phi}\,\mathrm{d}t + \sigma\,\mathrm{d}W^{\hat{A}}_{is}(t)$$

using absolute advantage $A_{is}(t) = \hat{A}_{is}(t)Z_s(t)$ based on OLS and PPML gravity measures of export capability k from (6), and the Balassa index of revealed comparative advantage $RCA_{ist} = (X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_{\iota} X_{\iota st}/\sum_{\iota}\sum_{\varsigma} X_{\iota\varsigma t})$. Parameters η, σ, ϕ for 1962-2007 (column 1-2) and 1984-2007 (columns 3-6) are estimated under the constraints $\ln\eta, \ln\sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (20), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (A.10). Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

negative across all industry aggregates for the OLS-based absolute advantage measures but statistically indistinguishable from zero for PPML-based log absolute advantage and the log RCA index, again regardless of industry aggregation.