The Dynamics of Comparative Advantage*

Gordon H. Hanson† UC San Diego and NBER
Nelson Lind§ UC San Diego
Marc-Andreas Muendler¶ UC San Diego and NBER

June 20, 2014

Abstract

We characterize the evolution of country export performance over the last five decades. Using the gravity model of trade, we extract a measure of country export capability by industry which we use to evaluate how absolute advantage changes over time for 135 industries in 90 countries. We alternatively use the Balassa RCA index as a measure of comparative advantage. Part I of the analysis documents two empirical regularities in country export behavior. One is hyperspecialization: in the typical country, export success is concentrated in a handful of industries. Hyperspecialization is consistent with a heavy upper tail in the distribution of absolute advantage across industries within a country, which is well approximated by a generalized gamma distribution whose shape is stable both across countries and over time. The second empirical regularity is a high rate of turnover in a country's top export industries. Churning in top exports reflects mean reversion in a typical country's absolute advantage, which we estimate to be on the order of 30% per decade. Part II of the analysis reconciles hyperspecialization in exports with high decay rates in export capability by modeling absolute advantage as a stochastic process. We specify a generalized logistic diffusion for absolute advantage that allows for Brownian innovations (accounting for surges in a country's export prowess), a country-wide stochastic trend (flexibly transforming absolute into comparative advantage), and deterministic mean reversion (permitting export surges to be impermanent). To gauge the fit of the model, we take the parameters estimated from the pooled time series and project the cross-sectional distribution of absolute advantage for each country in each year. Based on just three global parameters, the simulated values match the cross-sectional distributions—which are not targeted in the estimation—with considerable accuracy. Our results provide an empirical road map for dynamic theoretical models of the determinants of comparative advantage.

*Ina Jäkle and Heea Jung provided excellent research assistance. We thank Dave Donaldson, Davin Chor, Sam Kortum and Ben Moll as well as seminar participants at NBER ITI, Carnegie Mellon U, UC Berkeley, Brown U, U Dauphine Paris, the West Coast Trade Workshop UCLA, CESifo Munich, and the Cowles Foundation Yale for helpful discussions and comments.

†IR/PS 0519, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093-0519 (*gohanson@ucsd.edu*)

§*econweb.ucsd.edu/ nrlind*, Dept. of Economics, University of California, San Diego, 9500 Gilman Dr MC 0508, La Jolla, CA 92093-0508 (*nrlind@ucsd.com*).

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

1 Introduction

Comparative advantage has made a comeback in international trade. After a long hiatus during which the Ricardian model was universally taught to undergraduates but rarely used in quantitative research, the role of comparative advantage in explaining trade flows is again at the center of inquiry. Its resurgence is due in part to the success of the Eaton and Kortum (2002) model (EK hereafter), which gives a probabilistic structure to firm productivity and allows for settings with many countries and many goods.[1] On the empirical side, Costinot et al. (2012) uncover strong support for a multi-sector version of EK in cross-section data for OECD countries. Another source of renewed interest in comparative advantage comes from the dramatic recent growth in North-South and South-South trade (Hanson 2012). The emerging-economy examples of China and Mexico specializing in labor-intensive manufactures, Brazil and Indonesia concentrating in agricultural commodities, and Peru and South Africa shipping out large quantities of minerals give the strong impression that resource and technology differences between countries have a prominent role in determining current global trade flows.

In this paper, we characterize the evolution of country export advantages over the last five decades. Using the gravity model of trade, we extract a measure of country export capability which we use to evaluate how export performance changes over time for 135 industries in 90 countries between 1962 and 2007. Distinct from Costinot et al. (2012) and Levchenko and Zhang (2013), our gravity-based approach does not use industry production or price data to evaluate countries' export prowess. Instead, we rely on trade data only, which allows us to impose less theoretical structure on the determinants of trade, examine industries at a fine degree of disaggregation and over a long time span, and include both manufacturing and non-manufacturing sectors in our analysis. These features help in identifying the stable and heretofore underappreciated patterns of export dynamics that we uncover.

The gravity model is consistent with a large class of trade models (Anderson 1979, Anderson and van Wincoop 2003, Arkolakis et al. 2012). These have in common an equilibrium relationship in which bilateral trade in a particular industry and year can be decomposed into three components (Anderson 2011): an *exporter-industry fixed effect*, which captures the exporting country's average export capability in an industry; an *importer-industry fixed effect*, which captures the importing country's effective demand for foreign goods in an industry; and an *exporter-importer component*, which captures bilateral trade costs between pairs of exporting and importing countries. We estimate these components for each year in our data, with and without correcting for zero trade flows.[2] In the EK model, the exporter-industry fixed effect is the prod-

[1]Shikher (2011, 2012) expand EK to a multi-industry setting.

[2]See Silva and Tenreyro (2006), Helpman et al. (2008), Eaton et al. (2012), and Fally (2012) for alternative econometric approaches to account for zero trade between countries.

uct of a country's overall efficiency in producing goods and its unit production costs. In the Krugman (1980), Heckscher-Ohlin (Deardorff 1998), Melitz (2003), and Anderson and van Wincoop (2003) models, which also yield gravity specifications, the form of the exporter-industry component differs but its interpretation as a country-industry's export capability still applies. By taking the deviation of a country's export capability from the global mean for the industry, we obtain a measure of a country's absolute advantage in an industry. This definition is equivalent to a country's share of world exports in an industry that we would obtain were trade barriers in importing countries non-discriminating across exporters. By further normalizing absolute advantage by a country-wide term, we remove the effects of aggregate country growth, focusing attention on how the ranking of a country's export performance across industries changes over time. We refer to export capability after its double normalization by global-industry and country-wide terms as a measure of comparative advantage.

The aim of our analysis is to identify the dynamic empirical properties of absolute and comparative advantage that any theory of their determinants must explain. Though we motivate our approach using EK, we remain agnostic about the origins of a country's export strength. Export capability may depend on the accumulation of ideas (Eaton and Kortum 1999), home-market effects (Krugman 1980), relative factor supplies (Trefler 1995, Davis and Weinstein 2001, Romalis 2004, Bombardini et al. 2012), the interaction of industry characteristics and country institutions (Levchenko 2007, Costinot 2009, Cuñat and Melitz 2012), or some combination of these elements. Rather than search for cross-section covariates of export capability, as in Chor (2010), we seek the features of its distribution across countries, industries, and time. For robustness, we repeat the analysis by replacing our gravity-based measure of export capability with Balassa's (1965) index of revealed comparative advantage (RCA) and obtain similar results. We further restrict the period to 1984 and later, when more detailed industry data are available. This more recent period allows us to vary industry aggregation from two-digit to four-digit sectors, and we demonstrate that our results are not a byproduct of sector definitions.

After estimating country-industry export capabilities, our analysis proceeds in two stages. First, we document two strong empirical regularities in country export behavior that are seemingly in opposition to one another but whose synthesis reveals stable underlying patterns in the evolution of export advantage. One regularity is hyperspecialization in exporting.[3] In any given year, exports in the typical country tend to be highly concentrated in a small number of industries. Across the 90 countries in our data, the median share for the single top good (out of 135) in a country's total exports is 21%, for the top 3 goods is 45%, and for the top 7 goods is 64%. Consistent with strong concentration, the cross-industry distribution of absolute

[3] See Easterly and Reshef (2010), Hanson (2012), and Freund and Pierola (2013) for related findings.

advantage for a country in a given year is heavy tailed and approximately log normal, with ratios of the mean to the median of about 7. Strikingly, this approximation applies to countries specializing in distinct types of goods and at diverse stages of economic development. The Balassa RCA index is similarly heavy tailed.

Stability in the shape of the distribution of absolute advantage makes the second empirical regularity regarding exports all the more surprising: there is steady turnover in a country's top export products. Among the goods that account for the top 5% of a country's absolute-advantage industries in a given year, nearly 60% were not in the top 5% two decades earlier. Such churning is consistent with mean reversion in export superiority, which we confirm by regressing the change in a country-industry's absolute advantage on its initial value, obtaining decadal decay rates on the order of 25% to 30%. These regressions control for country-time fixed effects, and so may be interpreted as summarizing the dissipation of comparative advantage. The mutability of a country's relative export capabilities is consistent with Bhagwati's (1994) description of comparative advantage as "kaleidoscopic," with the dominance of a country's top export products often being short lived.

A concern about log normality in absolute advantage is whether it may be a byproduct of the estimation of the exporter-industry fixed effects. If these fixed effects varied randomly around a common mean for a country, they would be approximately normally distributed around a constant expected value, making absolute advantage tend toward log normality. Such logic, however, rests on the exporter-industry fixed effects having a common country mean. Our central focus is precisely on how mean export capability varies across industries for a country and how this variation progresses over time. Incidental log normality—resulting, say, from classical measurement error in trade data—would imply that in our decay regressions mean reversion in log absolute advantage from one period to the next would be more or less complete. Yet, this is not what we find. Mean reversion is partial, with estimated annual decay rates being similar whether based on 5, 10, or 20-year changes. Moreover, subsequent shocks to absolute advantage preserve the shape of its cross sectional distribution within a country. This subtle balance between mean reversion and random innovation, which also holds for the RCA index, is highly suggestive of a stochastic growth process at work for individual industries.

In the second stage of our analysis, we seek to characterize the stochastic process that guides export capability and thereby reconcile hyperspecialization in exports with mean reversion in export advantage. We specify a generalized logistic diffusion for absolute advantage that allows for Brownian innovations (accounting for surges in a country's relative export prowess), a country-wide stochastic trend (flexibly transforming absolute into comparative advantage), and deterministic mean reversion (permitting export surges to be impermanent). The generalized logistic diffusion that we specify has the generalized gamma as a stationary distribution.[4] The generalized gamma unifies the gamma and extreme-value families (Crooks 2010) and therefore flexibly nests many common distributions. To gauge the fit of the model, we take the three global parameters estimated from the pooled country *time series* and project the *cross-section* distribution of absolute advantage, which is not targeted in the estimation, for each country in each year. Based on just these three parameters (and controlling for a country-wide stochastic trend), the simulated values match the cross-sectional distributions, country-by-country and period-by-period, with considerable accuracy. The stochastic nature of absolute advantage implies that, at any moment in time, a country is especially strong at exporting in only a few industries and that, over time, this strength is temporary, with the identity of top industries churning perpetually.

We then allow model parameters to vary by groups of countries and by broad industry and estimate them for varying levels of industry aggregation. The three parameters of the generalized gamma govern the rate at which the process reverts to the global long-run mean (the dissipation of comparative advantage), the degree of asymmetry in mean reversion from above versus below the mean (the stickiness of comparative advantage), and the rate at which industries are reshuffled within the distribution (the intensity of innovations in comparative advantage). The first two parameters alone determine the shape of the stationary cross sectional distribution, with the third determining how quickly convergence to the long-run distribution is achieved. The intensity of innovations is stronger for developing than for developed economies. Whereas comparative advantage dissipates more quickly for manufacturing than for non-manufacturing industries, it is also relatively sticky for manufacturing, implying that industries revert towards the long-term mean more slowly from a position of comparative advantage than from a position of disadvantage.

A growing literature, to which our work contributes, employs the gravity model of trade to estimate the determinants of comparative advantage.[5] In exercises based on cross-section data, Chor (2010) explores whether the interaction of industry factor intensity with national characteristics can explain cross-industry variation in export volume and Waugh (2010) identifies asymmetries in trade costs between rich and poor countries that contribute to cross-country differences in income. In exercises using data for multiple years, Fadinger and Fleiss (2011) find that the implied gap in countries' export capabilities vis-a-vis the United States closes as countries' per capita GDP converges to U.S. levels,[6] and Levchenko and Zhang (2013), who calibrate the EK model to estimate overall sectoral efficiency levels by country, find that these efficiency

[4]Kotz et al. (1994) present properties of the generalized gamma distribution. Cabral and Mata (2003) use the generalized gamma distribution to study firm-size distributions. The finance literature considers a wide family of stochastic asset price processes with linear drift and power diffusion terms (see, e.g., Chan et al. 1992, on interest rate movements). Those specifications nest neither an ordinary nor a generalized logistic diffusion.

[5]On changes in export diversification over time see see Imbs and Wacziarg (2003) and Cadot et al. (2011).

[6]Related work on gravity and industry-level productivity includes Finicelli et al. (2009, 2013) and Kerr (2013).

levels converge across countries over time, weakening comparative advantage in the process.[7]

Our approach differs from the literature in two notable respects. By not using functional forms specific to EK or other trade models, we free ourselves from having to use industry production data (which is necessary to pin down model parameters) and are thus able to examine all merchandise sectors, including non-manufacturing, at the finest level of industry disaggregation possible. We gain from this approach a perspective on hyperspecialization in exporting and churning in top export goods that is less apparent in data limited to manufacturing or based on more aggregate industry categories. We lose, however, the ability to evaluate the welfare consequence of changes in comparative advantage (as in Levchenko and Zhang 2013). A second distinctive feature of our approach is that we treat export capability as being inherently dynamic. Previous work tends to study comparative advantage by comparing repeated static outcomes over time. We turn the empirical approach around, and estimate the underlying stochastic process itself. The virtue is that we can then predict the distribution of export advantage in the cross section, which our estimator does not target, and use the the cross-section projections as a check on the goodness of fit.

Section 2 of the paper presents a theoretical motivation for our gravity specification. Section 3 describes the data and our estimates of country export capabilities, and documents empirical regularities regarding comparative advantage, hyperspecialization in exporting and churning in countries' top export goods. Section 4 describes a stochastic process that has a cross sectional distribution consistent with hyperspecialization and a drift consistent with turnover, and introduces a GMM estimator to identify the fundamental parameters. Section 5 presents the estimates and evaluates the fit of the diffusion. Section 6 concludes.

2 Theoretical Motivation

In this section, we use the EK model to motivate our definitions of export capability and absolute advantage and then describe our approach for extracting these values from the gravity model of trade.

2.1 Export capability and comparative advantage

In the EK model, 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 $F_Q(q) = \exp\{-(q/\underline{q}_{is})^{-\theta_i}\}$ for $q > 0$. Consumers, who have CES preferences over product varieties within an industry, buy from the firm that is able to deliver a variety at the lowest price. With firms

[7] Other related literature includes dynamic empirical analyses of the Heckscher-Ohlin model that examine how trade flows change in response to changes in country factor supplies (Schott 2003, Romalis 2004) and work by Hausmann et al. (2007) on how the composition of exports relates to the pace of economic growth.

pricing according to marginal cost, a higher productivity draw makes a firm more likely to be the low-priced supplier of a variety to a given market.

Comparative advantage stems from the position of the industry productivity distribution, given by $\underline{q}_{is}$. The position can differ across source countries s and industries i. In countries with a higher $\underline{q}_{is}$, firms are more likely to have a higher productivity draw, creating cross-country variation in the fraction of firms that succeed within an industry in being low-cost suppliers to different destination markets.[8] Consider the many-industry version of the EK model in Costinot et al. (2012). Exports by source country s to destination country d in industry i can be written as,

$$X_{isd} = \frac{\left(w_s \tau_{isd}/\underline{q}_{is}\right)^{-\theta_i}}{\sum_{s'} \left(w_{s'} \tau_{is'd}/\underline{q}_{is'}\right)^{-\theta_i}} \, \mu_i Y_d, \tag{1}$$

where w_s is the unit production cost for country s, τ_{isd} is the iceberg trade cost between s and d in industry i, μ_i is the Cobb-Douglas share of expenditure on industry i, and Y_d is total expenditure in country d. Taking logs of (1), we obtain a gravity equation for bilateral trade

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

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's overall efficiency in the industry ($\underline{q}_{is}$) and its unit production costs (w_s), and

$$m_{id} \equiv \ln \left[\mu_i Y_d \Big/ \sum_{s'} \left(w_{s'} d_{is'd}/\underline{q}_{is'} \right)^{-\theta_i} \right]$$

is the log of *effective import demand* by country d in industry i, which depends on the country's expenditure on goods in the industry divided by an index of the toughness of competition for the country in the industry.

Export capability is a function of a primitive country characteristic—the position of a country's productivity distribution—and of endogenously determined unit production costs. EK does not yield a closed-form solution for wages, we can therefore not solve for export capabilities as explicit functions of the $\underline{q}_{is}$'s. Yet, in a model with a single factor of production the $\underline{q}_{is}$'s are the only country-specific variable for the industry (other than population and trade costs) that may determine factor prices, meaning that the w_s's are implicit functions of these parameters. Our concept of export capability k_{is} can further be related to the

[8] The importance of the position of the productivity distribution for trade depends in turn on the shape of the distribution, given by θ_i. Lower dispersion in productivity draws (a higher value of θ_i) elevates the role of the distribution's position in determining a country's strength in an industry. These two features—the country-industry position parameter $\underline{q}_{is}$ and the industry dispersion parameter θ_i—pin down a country's export capability.

deeper origins of comparative advantage by modeling the country-industry-specific Fréchet position parameter $T_{is} \equiv (\underline{q}_{is})^{\theta_i}$ as the outcome of an exploration and innovation process, similar to Eaton and Kortum (1999), a connection we sketch in Appendix D.

Any trade model that has a gravity structure will generate exporter-industry fixed effects and a reduced-form expression for exporter capability. 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, with the added difference that bilateral trade is a function of both variable and fixed trade costs. And in a Heckscher-Ohlin model (Deardorff 1998), export capability reflects the relative size of a country's industry based on factor endowments and sectoral 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.

The principle of comparative advantage requires that a country-industry's export capability $K_{is} \equiv \exp\{k_{is}\}$ be compared to both the same industry across countries and to other industries within the same country. This double comparison of a country-industry's export capability to other countries and other industries is also at the core of measures of revealed comparative advantage (Balassa 1965) and recent implementations of comparative advantage, as in Costinot et al. (2012). Consider two exporters s and s' and two industries i and i', and define geography-adjusted trade flows as

$$\tilde{X}_{isd} \equiv X_{isd} \left(\tau_{isd}\right)^{\theta_i} = \left(w_s/\underline{q}_{is}\right)^{-\theta_i} \exp\{m_{id}\}.$$

The correction of observed trade X_{isd} by trade costs $(\tau_{isd})^{\theta}$ removes the distortion that geography exerts on export capability when trade flows are realized.[9] When compared to any country s', country s has a comparative advantage in industry i relative to industry i' if the following condition holds:

$$\frac{\tilde{X}_{isd}/\tilde{X}_{is'd}}{\tilde{X}_{i'sd}/\tilde{X}_{i's'd}} = \frac{K_{is}/K_{is'}}{K_{i's}/K_{i's'}} > 1. \tag{3}$$

The comparison of a country-industry to the same industry in other source countries makes the measure independent of destination-market characteristics m_{id} because the standardization $\tilde{X}_{isd}/\tilde{X}_{is'd}$ removes the destination-market term. In practice, a large number of industries and countries makes it cumbersome to

[9]This adjustment ignores any impact of trade costs on equilibrium factor prices w_s.

conduct double comparisons of a country-industry is to all other industries and all other countries. Our gravity-based correction of trade flows for geographic frictions gives rise to a natural alternative summary measure.

2.2 Estimating the gravity model

By allowing for measurement error in trade data or unobserved trade costs, we introduce a disturbance term into (2), converting it into a 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 via OLS. The gravity model that we estimate is

$$\ln X_{isdt} = k_{ist} + m_{idt} - b_{it} D_{sdt} + \epsilon_{isdt}, \tag{4}$$

where we have added a time subscript t, we include dummy variables to measure exporter-industry-year k_{ist} and importer-industry-year m_{idt} terms, D_{sdt} represents the determinants of bilateral trade costs, and ϵ_{isdt} is a residual that is mean independent of D_{sdt}. The variables we use to measure trade costs D_{sdt} in (4) are standard gravity covariates, which do not vary by industry.[10] However, we do allow the coefficients b_{it} on these variables to differ by industry and by year.[11] Absent annual measures of industry-specific trade costs for the full sample period, we model these costs via the interaction of country-level gravity variables and time-and-industry-varying coefficients.

In the estimation, we exclude a constant term, include an exporter-industry-year dummy for every exporting country in each industry, and include an importer-industry-year dummy for every importing country except for one, which we select to be the United States. The exporter-industry-year dummies we estimate thus equal

$$k_{ist}^{\mathrm{OLS}} = k_{ist} + m_{i\mathrm{US}\,t}, \tag{5}$$

where k_{ist}^{OLS} is the estimated exporter-industry dummy for country s in industry i and year t, $m_{i\mathrm{US}\,t}$ is the U.S. importer-industry-year fixed effect, and k_{ist} is the underlying log export capability. The estimator of the exporter-industry variables is therefore meaningful only up to an industry normalization.

The values that we will use for empirical analysis are the deviations of the estimated exporter-industry-

[10] 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.

[11] We estimate (4) separately by industry and by year. Since the regressors are the same across industries for each bilateral pair, there is no gain to pooling data across industries in the estimation, which helps reduce the number of parameters to be estimated in each regression.

year dummies from the global industry means:

$$\hat{k}_{ist} = k_{ist}^{\text{OLS}} - \frac{1}{S} \sum_{s'=1}^{N} k_{is't}^{\text{OLS}}, \tag{6}$$

where the deviation removes the excluded importer-industry-year term as well as any global industry-specific term. This normalization obviates the need to account for worldwide industry TFP growth, demand changes, or producer price index movements, allowing us to conduct analysis of comparative advantage with trade data exclusively.

From this exercise, we take as a measure of *absolute advantage* of country s's industry i,

$$A_{ist} \equiv \exp\{\hat{k}_{ist}\} = \frac{\exp\left\{k_{ist}^{\text{OLS}}\right\}}{\exp\left\{\frac{1}{S}\sum_{s'=1}^{S} k_{is't}^{\text{OLS}}\right\}} = \frac{\exp\left\{k_{ist}\right\}}{\exp\left\{\frac{1}{S}\sum_{s'=1}^{S} k_{is't}\right\}}. \tag{7}$$

By construction, this measure is unaffected by the choice of the omitted importer-industry-year fixed effect. As the final equality in (7) shows, the measure is equivalent to the comparison of underlying exporter capability K_{ist} to the geometric mean of exporter capability across countries in industry i.

There is some looseness in our measure of absolute advantage. When A_{ist} rises for country-industry is, we say that its absolute advantage has risen even though it is only strictly true that its export capability has increased relative to the global industry geometric mean. In truth, the country's export capability may have risen relative to some countries and fallen relative to others. Our motivation for using the deviation from the geometric mean to define absolute advantage is twofold. One is that our statistic removes the global industry component of estimated export capability, making our measure immune to the choice of normalization in the gravity estimation. Two is that removing the industry-year component relates naturally to specifying a stochastic process for export capability. Rather than modeling export capability itself, we model its deviation from an industry trend, which simplifies the estimation by freeing us from having to model the trend component that will reflect global industry demand and supply. We establish the main regularities regarding the cross section and the dynamics of exporter performance using absolute advantage A_{ist} in Section 3. In Section 4, we let the stochastic process that is consistent with the empirical regularities of absolute advantage determine the remaining country-level standardization that transforms absolute advantage A_{ist} into a measure of comparative advantage.

As is well known, the gravity model in (2) and (4) 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 the approach in Eaton et al. (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 firms from country s that have sufficiently high productivity to profitably supply destination market d in industry i. In their framework, the analogue to equation (1) is an expression for the expected share of country s in the market for industry i in country d, $\mathbb{E}\left[X_{isd}/X_{id}\right]$, which can be written as a multinomial logit. This approach, however, requires that one know total expenditure in the destination market, X_{id}, including a country's spending on its own goods. Since total expenditure is unobserved in our data, we apply the independence of irrelevant alternatives and specify the dependent variable as the expectation for an exporting country's share of total import purchases in the destination market:

$$\mathbb{E}\left[\frac{X_{isd}}{\sum_{s'\neq d} X_{is'd}}\right] = \frac{\exp\left(k_{ist} - b_{it} D_{isdt}\right)}{\sum_{s'\neq d} \exp\left(k_{is't} - b_{it} D_{is'dt}\right)}. \tag{8}$$

We re-estimate exporter-industry-year fixed effects by applying multinomial pseudo-maximum likelihood to (8).[12]

Our baseline measure of absolute advantage relies on regression-based estimates of exporter-industry-year fixed effects. Even when following the approach in Eaton et al. (2012), estimates of these fixed effects may become imprecise when a country exports a good to only a few destinations. As an alternative measure of export performance, we use the Balassa (1965) measure of revealed comparative advantage, defined as,

$$\mathit{RCA}_{ist} = \frac{\sum_d X_{isdt} / \sum_{i'} \sum_{d'} X_{i's'd't}}{\sum_{i'} \sum_d X_{i'sdt} / \sum_{s'} \sum_{i'} \sum_{d'} X_{i's'd't}} \tag{9}$$

While the RCA index is ad hoc and does not correct for distortions in trade flows introduced by trade costs or proximity to market demand, it has the appealing attribute of being based solely on raw trade data. Throughout our analysis we will employ the gravity-based measure of absolute advantage alongside the Balassa RCA measure. Reassuringly, our results for the two measures are quite similar.

3 Data and Main Regularities

The data for our analysis are World Trade Flows from Feenstra et al. (2005),[13] which are based on SITC revision 1 industries for 1962 to 1983 and SITC revision 2 industries for 1984 and later.[14] We create a consistent set of country aggregates in these data by maintaining as single units countries that divide over the sample period.[15] To further maintain consistency in the countries present, we restrict the sample to

[12]We thank Sebastian Sotelo for estimation code.

[13]We use a version of these data that have been extended to 2007 by Robert Feenstra and Gregory Wright.

[14]A further source of observed zero trade is that for 1984 and later bilateral industry trade flows are truncated below $100,000.

[15]These are the Czech Republic, the Russian Federation, and Yugoslavia. We also join East and West Germany, Belgium and Luxembourg, and North and South Yemen.

nations that trade in all years and that exceed a minimal size threshold, which leaves 116 country units.[16] The switch from SITC revision 1 to revision 2 in 1984 led to the creation of many new industry categories. To maintain a consistent set of SITC industries over the sample period, we aggregate industries from the four-digit to three-digit level.[17] These aggregations and restrictions leave 135 industries in the data. In an extension of our main results, we limit the sample to SITC revision 2 data for 1984 forward, alternatively using two-digit (61 industries), three-digit (227 industries), or four-digit (684 industries) sector definitions.

A further set of country restrictions are required to estimate importer and exporter fixed effects. For coefficients on exporter-industry dummies to be comparable over time, the countries that import a good must do so 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. We also need that exporters ship to overlapping groups of importing countries. As Abowd et al. (2002) show, such connectedness assures that all exporter fixed effects are separately identified from importer fixed effects.[18] 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 135 industries, we estimate the gravity equation (4) 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 Hyperspecialization in exporting

We first characterize export behavior in the cross section of industries for each country at a given moment of time. For an initial take on the concentration of exports in leading products, we tabulate the share of a country-industry's exports $X_{ist}/(\sum_{i'} X_{i'st})$ in the country's total exports across the 135 industries. We then average these shares across the current and preceding two years to account for measurement error and cyclical fluctuations. In **Figure 1a**, we display median export shares across the 90 countries in our sample for the top export industry as well as the top three, top seven, and top 14 industries, which roughly translate into the top 1%, 3%, 5% and 10% of products.

For the typical country, a handful of industries dominate exports.[19] The median export share of just

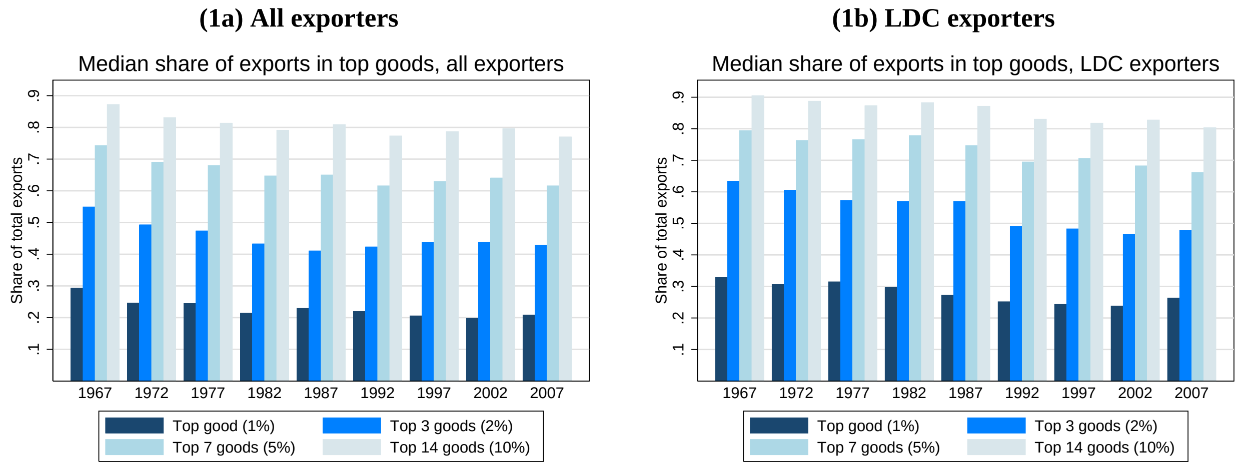
[16] This reporting restriction leaves 141 importers (97.7% of world trade) and 139 exporters (98.2% of world trade) and is roughly equivalent to dropping small countries from the sample. For consistency in terms of country size, we drop countries with fewer than 1 million inhabitants in 1985 (42 countries had 1985 population less than 250,000, 14 had 250,000 to 500,000, and 9 had 500,000 to 1 million), which reduces the sample to 116 countries (97.4% of world trade).

[17] 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 products. From this group we drop nonstandard industries (postal packages, coins, gold bars, DC current) and three industries that are always reported as one-digit aggregates in the US data. We further exclude oil and natural gas, which in some years have estimated exporter-industry fixed effects that are erratic.

[18] Countries that export to mutually exclusive sets of destinations would not allow us to separately identify the exporter fixed effect from the importer fixed effects.

[19] In analyses of developing-country trade, Easterly and Reshef (2010) document the tendency of a small number of bilateral-

Figure 1: **Concentration of Exports**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007.
Note: Shares of industry i's export value in country s's total export value: $X_{ist}/(\sum_{i'} X_{i'st})$. For the classification of less developed countries (LDC) see Appendix E.

the top export good is 24% in 1972, which declines modestly over time to 20% by 2007. Over the full period, the median export share of the top good averages 21%. For the top three products, the median export share declines slightly from the 1960s to the 1970s and then is stable from the early 1980s onward at approximately 42%. The median export shares of the top seven and top 14 products display a similar pattern, stabilizing by the early 1980s at around 62% and 77%, respectively. Thus, the bulk of a country's exports tend to be accounted for by the top 10% of its goods. In **Figure 1b**, we repeat the exercise, limiting the sample to less developed countries (see Appendix E). The patterns are quite similar to those for all countries, though median export shares for LDCs are modestly higher in the reported quantiles.

One concern about using export shares to measure export concentration is that these values may be distorted by demand conditions. Exports in some industries may be large simply because these industries capture a relatively large share of global expenditure, leading the same industries to be top export industries in *all* countries. In 2007, for instance, the top export industry in Great Britain, France, Germany, Japan, and Mexico is road vehicles. In the same year in Korea, Malaysia, the Philippines, Taiwan, and the United States the top industry is electric machinery. One would not want to conclude from this fact that each of these countries has an advantage in exporting one of these two products.

To control for variation in industry size that is associated with preferences, we turn to our measure of

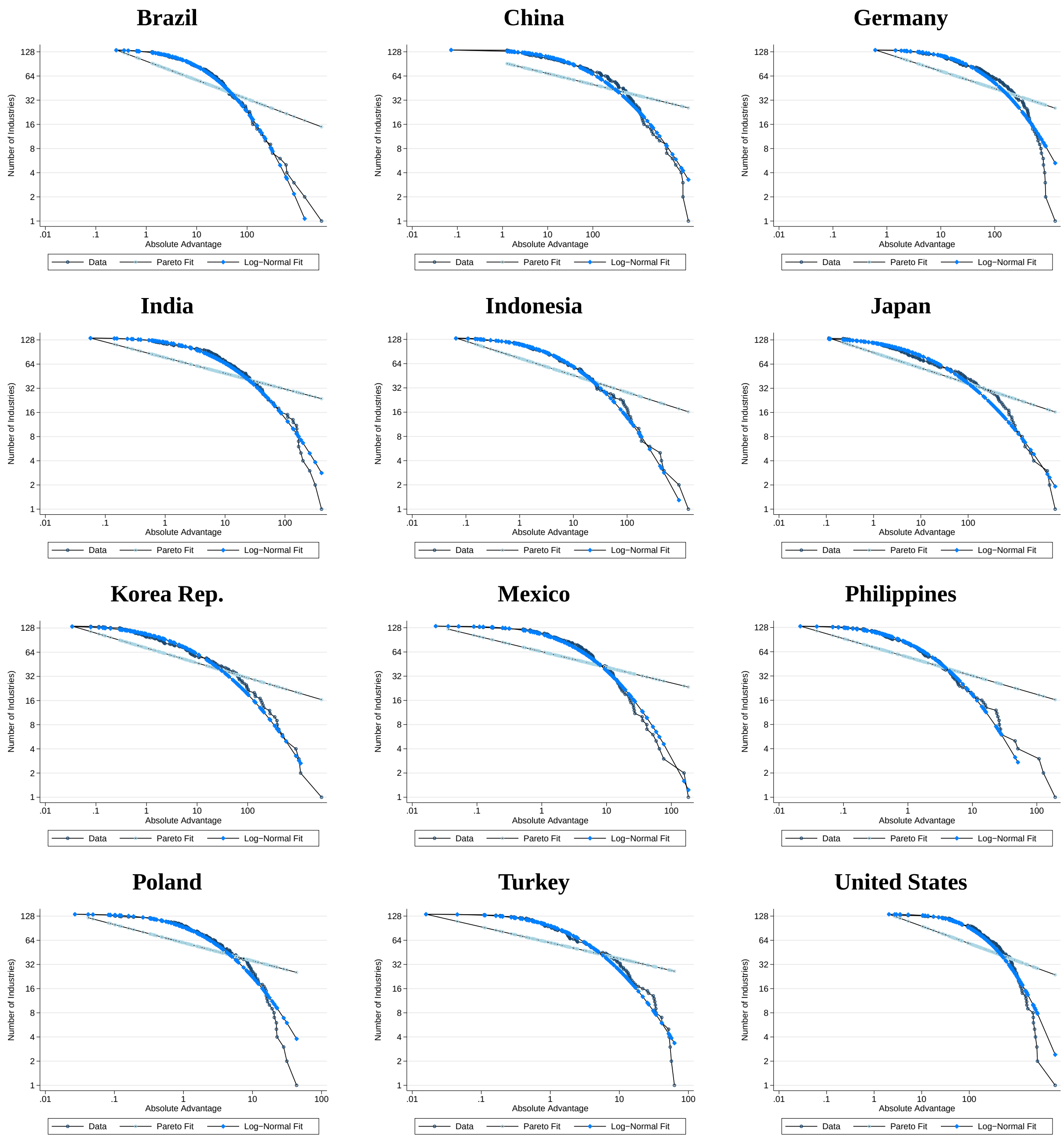
industry relationships to dominate national exports and Freund and Pierola (2013) describe the prominent role of the largest few firms in countries' total foreign shipments.

absolute advantage in (7) expressed in logs as $\ln A_{ist} = \hat{k}_{ist}$. As this value is the log industry export capability in a country minus global mean log industry export capability, industry characteristics that are common across countries—including the state of global demand—are differenced out. To provide a sense of the identities of absolute-advantage goods and the magnitudes of their advantages, we show in Appendix **Table A1** the top two products in terms of A_{ist} for 28 of the 90 exporting countries, using 1987 and 2007 as representative years. To remove the effect of overall market size and thus make values comparable across countries, we normalize log absolute advantage by its country mean, such that the value we report for country-industry is is $\ln A_{ist} - (1/I)\sum_{i'}^{I} \ln A_{i'st}$. The country normalization yields a double log difference—a country's log deviation from the global industry mean minus its average log deviation across all industries—which is a measure of comparative advantage.

There is considerable variation across countries in the top advantage industries. In 2007, comparative advantage in Argentina is strongest in maize, in Brazil it is iron ore, in Canada it is wheat, in Germany it is road vehicles, in Indonesia it is rubber, in Japan it is telecommunications equipment, in Poland it is furniture, in Thailand it is rice, Turkey it is glassware, and in the United States it is other transport equipment (mainly commercial aircraft). The implied magnitudes of these advantages are enormous. Among the 90 countries in 2007, comparative advantage in the top product—i.e., the double log difference—is over 400 log points in 76 of the cases. Further, the top industries in each country by and large correspond to those one associates with national export advantages, suggesting that the observed rankings of export capability are not simply a byproduct of measurement error in trade values.

To characterize the full distribution of absolute advantage across industries for a country, we next plot the log number of a source country s's industries that have at least a given level of absolute advantage in a year t against that log absolute advantage level $\ln A_{ist}$ for industries i. By design, the plot characterizes the cumulative distribution of absolute advantage by country and by year (Axtell 2001, Luttmer 2007). **Figure 2** shows the distribution plots of log absolute advantage for 12 countries in 2007. Plots for 28 countries in 1967, 1987 and 2007 are shown in Appendix **Figures A1, A2** and **A3**. 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 (such that the number of parameters estimated equals the number of parameters for a distribution $\times$ number of countries $\times$ number of years). We choose the Pareto and the log normal as comparison cases because these are the standard options in the literature on firm size (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. Relevant to our later analysis, each is a special case of the generalized gamma distribution. To verify that the graphed

Figure 2: **Cumulative Probability Distribution of Absolute Advantage for Select Countries in 2007**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).

Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 for absolute advantage A_{ist} are based on maximum likelihood estimation by country s in year $t = 2007$.

cross-sectional distributions are not a byproduct of specification error in estimating export capabilities from the gravity model, we repeat the plots using the Balassa (1965) RCA index, with similar results. And to verify that the patterns we uncover are not a consequence of arbitrary industry aggregations we construct plots at the two, three, and four-digit level based on SITC revision 2 data in 1987 and 2007, again with similar results.[20]

The cumulative distribution plots clarify that the empirical distribution of absolute advantage is decidedly not Pareto. The log normal, in contrast, fits the data closely. The concavity of the cumulative distribution plots drawn for the data indicate that gains in absolute advantage fall off progressively more rapidly as one moves up the rank order of absolute advantage, a feature absent from the scale-invariant Pareto but characteristic of the log normal. This concavity could indicate limits on industry export size associated with resource depletion, congestion effects, or general diminishing returns. Though the log normal is a rough approximation, there are noticeable discrepancies between the fitted log normal plots and the raw data plots. For some countries, we see that compared to the log normal the number of industries in the upper tail drops too fast (i.e., is more concave), relative to what the log normal distribution implies. These discrepancies motivate our specification of a generalized logistic diffusion for absolute advantage in Section 4, which is consistent with a generalized gamma distribution in the cross section.

Overall, we see that in any year countries have a strong export advantage in just a few industries, where this pattern is stable both across countries and over time. Before examining the time series of comparative advantage in more detail, we consider whether log normality in absolute advantage could be merely incidental. The exporter-industry fixed effects are estimated mean values, which by the Central Limit Theorem will converge to being normally distributed as the sample size becomes large. Incidental log normality in absolute advantage could result if the estimated exporter-industry fixed effects varied randomly around a common expected value for a given country. Our preferred view is that log normality in absolute advantage results instead from differences in the *industry means* of export capability by country, where these industry means determine comparative advantage. Indeed, if absolute advantage did have a common expected value across industries for each country there would be no basis for comparative advantage at the industry level. From the cross sectional distribution of absolute advantage alone, however, one cannot differentiate between random variation in industry fixed effects around a common mean for each exporter and variation in each exporter's industry means. Examining how absolute advantage changes over time will help resolve this issue.[21]

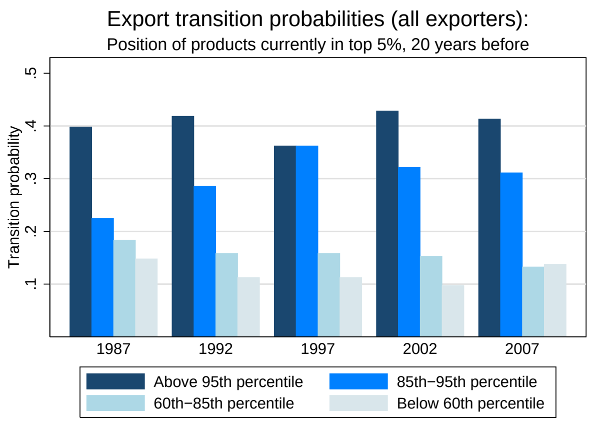
[20]Each of these additional sets of results is available in an online appendix.

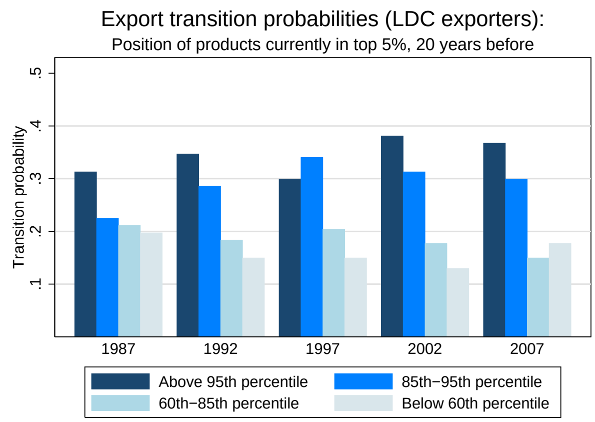
[21]It is worth noting that the hypothesis of incidental normality in the estimated exporter-industry fixed effects applies just as readily to the estimated importer-industry fixed effects. As an instructive exercise, we also constructed cumulative distribution plots, analogous to those in Appendix **Figures A1, A2** and **A3,** for the estimated importer-industry fixed effects, which involves plotting

Figure 3: **Absolute Advantage Transition Probabilities**

(3a) All exporters

(3b) LDC exporters





Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007. *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 Appendix E.

3.2 The dissipation of comparative advantage

The distribution plots of absolute advantage give an impression of stability. The strong concavity in the plots is present in all countries and in all years. Yet, this stability masks considerable industry churning in the distribution of absolute advantage, which we investigate next. Initial evidence of churning is evident in Appendix **Table A1**. Between 1987 and 2007, Canada's top good switches from sulfur to wheat, China's from explosives (fireworks) to telecommunications equipment, Egypt's from cotton to crude fertilizers, India's from tea to precious stones, Malaysia's from rubber to radios, the Philippine's from vegetable oils to office machines, and Romania's from furniture to footwear. Of the 90 total exporters, 70 exhibit a change in the top comparative-advantage industry between 1987 and 2007. Moreover, most new top products in 2007 were not the number two product in 1987, but from lower down in the distribution. Churning thus appears to be both pervasive and disruptive.

To characterize turnover in industry export advantage more completely, in **Figure 3** we calculate the fraction of top products in a given year that were also top products in previous years. We identify for each country in each year where in the distribution the top 5% of absolute-advantage products (in terms of A_{ist})

$\exp\{m_{idt}\}$, the exponentiated importer-industry fixed effect in equation (4), across industries for each country in representative years. The plots show little evidence of log normality for these values. In particular, the distribution of the exponentiated importer-industry fixed effects are much less concave to the origin than log normality would imply. These results are in the online appendix.

were 20 years before, with the options being top 5% of products, next 10%, next 25% or bottom 60%. We then average across outcomes for the 90 exporters. The fraction of top 5% products in a given year that were also top 5% products two decades before ranges from a high of 43% in 2002 to a low of 37% in 1997. Averaging over all years, the share is 41%. There is thus nearly a 60% chance that a good in the top 5% in terms of absolute advantage today was not in the top 5% two decades earlier. On average, 30% of new top products come from the 85th to 95th percentiles, 16% come from the 60th to 85th percentiles, and 13% come from the bottom six deciles. Figures are similar when we limit the sample to just developing economies.

Turnover in top export goods suggests that over time absolute advantage dissipates—countries' strong sectors weaken and some weak sectors strengthen. To evaluate this impermanence, we test for mean reversion in log absolute advantage by estimating regressions of the form

$$\ln A_{is,t+10} - \ln A_{ist} = \rho \ln A_{ist} + \delta_{st} + \varepsilon_{ist}. \tag{10}$$

In (10), the dependent variable is the ten-year change in log absolute advantage and the predictors are the initial value of log absolute advantage and dummy variables for the country-year δ_{st}. Absolute advantage represents the deviation in industry export capability for a country relative to the global mean. The inclusion of country-year dummies introduces a further level of differencing from the country-year mean, so that the regression in (10) evaluates the dynamics of comparative advantage. The coefficient ρ captures the fraction of comparative advantage that dissipates over the time interval of one decade, either decaying towards a log level of zero when currently above or strengthening towards a log level of zero when currently below.

Table 1 presents coefficient estimates for equation (10). The first two columns report results for all countries and industries, first for log absolute advantage in column 1 and next for the log RCA index in column 2. Subsequent pairs of columns show results separately for less development countries and non-manufacturing industries. Estimates for ρ are uniformly negative and precisely estimated, consistent with mean reversion in comparative advantage. For the sample of all industries and countries, estimates for ρ in columns 1 and 2 are similar in value, equal to -0.24 when using log absolute advantage and -0.30 when using log RCA. These magnitudes indicate that over the period of a decade the typical country-industry sees one-quarter to three-tenths of its comparative advantage (or disadvantage) erode. In columns 3 and 4 we present comparable results for the subsample of developing countries. Decay rates appear to be larger for this group of countries than worldwide average, indicating that in less developed economies mean reversion in comparative advantage is more rapid. In columns 5 and 6 we present results for only for non-manufacturing industries, but all countries. For both measures of comparative advantage decay rates are larger in absolute value for non-manufacturing industries (agriculture and mining), but the difference in decay rates between

Table 1: DECAY REGRESSIONS FOR COMPARATIVE ADVANTAGE

	Full sample		LDC exporters		Non-manufacturing	
	Exp. cap. k	RCA $\ln \hat{X}$	Exp. cap. k	RCA $\ln \hat{X}$	Exp. cap. k	RCA $\ln \hat{X}$
	(1)	(2)	(3)	(4)	(5)	(6)
Decay rate ρ	-0.237	-0.300	-0.338	-0.352	-0.359	-0.315
	(0.018)**	(0.013)**	(0.025)**	(0.015)**	(0.025)**	(0.013)**
Dissipation rate η	0.114	0.115	0.121	0.104	0.120	0.103
	(0.007)**	(0.004)**	(0.007)**	(0.003)**	(0.007)**	(0.004)**
Innov. intens. σ^2	0.476	0.618	0.683	0.836	0.741	0.737
	(0.011)**	(0.011)**	(0.023)**	(0.017)**	(0.026)**	(0.014)**
Obs.	66,276	67,901	39,937	41,103	30,942	32,390
Adj. R^2	0.114	0.125	0.129	0.133	0.124	0.126

Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org.
Note: Reported figures for five-year decadalized changes. Variables are OLS-estimated gravity measures of export capability k by (5) and the log Balassa index of revealed comparative advantage $\ln \hat{X}_{ist} = \ln(X_{ist}/\sum_{s'} X_{is't})/(\sum_{i'} X_{i'st}/\sum_{i'}\sum_{s'} X_{i's't})$. OLS estimation of the decadal decay rate ρ from

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

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 innovation intensity σ^2 are based on the decadal 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 Appendix E. Nonmanufacturing merchandise spans SITC sector codes 0-4. Standard errors (reported below coefficients) for ρ are clustered by country and for η and σ are calculated using the delta method; ** indicates significance at the 1% level.

non-manufacturing industries and the average industry is particularly pronounced for the log absolute advantage measure.[22]

As an additional robustness check on the decay regressions, we re-estimate (10) for the period 1984-2007 using data from the SITC revision 2 sample. This allows us to perform regressions for log absolute advantage and the log RCA index at the two, three and four-digit level. Results are reported in Appendix **Table A2**. Estimated decay rates are comparable to those in **Table 1**. At the two-digit level (61 industries), the decadal decay rate for absolute advantage using all countries and industries is 19%, at the three-digit level (226 industries) it is 24%, and at the four-digit level (684 industries) it is 37%. When using the log RCA index, decay rates vary less across aggregation levels, ranging from 26% at the two-digit level to 32% at the four-digit level. The similarity in decay rates across definitions of comparative advantage and levels of industry aggregation suggest that our results are neither merely generated by econometric estimation nor the consequence of arbitrary industry definitions.

Our finding that decay rates imply less than complete mean reversion is evidence against the log normality of absolute advantage being incidental. Suppose the cumulative distribution plots in **Figure 2** reflected

[22] In the next section, we offer further interpretation of these results.

random variation in log absolute advantage around a common expected value for each country in each year, due say to measurement error in trade data. Under the assumption that this measurement error was classical, all within-country variation in the exporter-industry fixed effects would be the result of iid disturbances that were uncorrelated across time. In the cross section, we would observe a log normal distribution for absolute advantage—and possibly also for the RCA index—for each country in each year, with no temporal connection between these distributions. When estimating the decay regression in (10), mean reversion in absolute advantage would be complete, yielding a value of ρ equal or close to -1. The coefficient estimates in **Table 1** are strongly inconsistent with such a pattern. Instead, as we document next, the results reveal that the stable cross sectional distribution of absolute absolute and the churn of industry export rankings are intimately related phenomena.

3.3 Comparative advantage as a stochastic process

On its own, the finding that comparative advantage reverts to a long-term mean is uninformative about the cross sectional distribution.[23] While mean reversion is consistent with a stationary cross sectional distribution, mean reversion is also consistent with a non-ergodic distribution and consistent with degenerate comparative advantage that collapses at a long-term mean. Yet, the combination of mean reversion in **Table 1** and temporal stability in the cumulative distribution plots in **Figure 2** are strongly suggestive of a balance between random innovations to export capability and the dissipation of these capabilities, a balance characteristic of the class of stochastic processes that generate a stationary cross sectional distribution.

In exploring the dynamics of comparative advantage here and in Section (4) we limit ourselves to diffusions: Markov processes for which all realizations of the random variable are continuous functions of time and past realizations. As a preliminary exercise, we exploit the fact that the decay regression in (10) is consistent with the discretized version of a commonly studied diffusion, the Ornstein-Uhlenbeck (OU) process. Suppose that comparative advantage, which we express in continuous time as $\hat{A}_{is}(t)$, follows an OU process given by

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

where $W^{\hat{A}}_{is}(t)$ is a Wiener process that induces stochastic changes in comparative advantage. The parameter η regulates the rate of convergence at which comparative advantage reverts to its global long-run mean and the parameter σ scales time and therefore the Brownian innovations $\mathrm{d}W^{\hat{A}}_{is}(t)$ in addition to regulating the rate of convergence.[24] Comparative advantage reflects a double normalization of export capability—by

[23]This point is analogous to critiques of using cross-country regressions to test for convergence in rates of economic growth (see e.g. Quah 1996).

[24]Among possible parametrizations of the OU process, we choose (11) because it is closely related to our later extension to

the global industry-year and by the country-year. It is therefore natural to consider a global mean of one, implying a global mean of zero for $\ln \hat{A}_{is}(t)$.

The OU process is the unique non-degenerate Markov process that has a stationary normal distribution (Karlin and Taylor 1981, ch. 15, proposition 5.1).[25] The OU process of log comparative advantage $\ln \hat{A}_{is}(t)$ has therefore as its stationary distribution a log normal distribution of comparative advantage $\hat{A}_{is}(t)$. In other words, if we observed comparative advantage $\hat{A}_{is}(t)$ and plotted it with graphs like those in **Figure 2**, we would find a log normal shape if and only if the underlying Markov process of log comparative advantage $\ln \hat{A}_{is}(t)$ is an OU process. In **Figure 2**, we only observe absolute advantage, however, so it remains for us to relate the two cross sectional distributions of comparative and absolute advantage.

In (11), we refer to the parameter η as the *rate of dissipation* of comparative advantage because it contributes to the speed with which log comparative advantage would collapse to a degenerate level of zero in all industries and all countries if there were no stochastic innovations. The parametrization in (11) implies that η alone determines the shape and heavy tail of the resulting stationary distribution, while σ is irrelevant for the cross sectional distribution. Our parametrization is akin to a standardization by which η is a normalized rate of dissipation that measures the "number" of typical (one-standard deviation) shocks that dissipate per unit of time. We refer to the parameter σ as the *intensity of innovations*. Under our parametrization of η, σ plays a dual role: on the one hand magnifying volatility by scaling up the Wiener innovations and on the other hand contributing to 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 from 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 et al. 2010, Example 13).[26] Applying this insight to the first-difference equation above, we obtain our decay regression:

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

a generalized logistic diffusion and because it clarifies that the parameter σ is irrelevant for the cross sectional distribution. We deliberately specify parameters η and σ that are invariant over time, industry and country and will explore the goodness of fit under that restriction.

[25] The Ornstein-Uhlenbeck process is a continuous-time analogue to a mean reverting AR(1) process in discrete time. It is a baseline stochastic process in the natural sciences and finance (see e.g. Vasicek 1977, Chan et al. 1992).

[26] Concretely, $\ln \hat{A}_{is}(t+\Delta) = \exp\{-\eta\sigma^2\Delta/2\} \ln \hat{A}_{is}(t) + \varepsilon_{ist}(t, t+\Delta)$ for a disturbance $\varepsilon_{ist}(t, t+\Delta) \sim \mathcal{N}(0, [1 - \exp\{-\eta\sigma^2\Delta\}]/\eta)$.

implying reduced-form expressions for the decay parameter

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

and the unobserved fixed effect $\delta_s(t) \equiv \ln Z_s(t+\Delta) - (1+\rho)\ln Z_s(t)$, where the residual $\varepsilon_{ist}(t, t+\Delta)$ is normally distributed with mean zero and variance $(1-\exp\{-\eta\sigma^2\Delta\})/\eta$. An OU process with $\rho \in (0,1)$ generates a log normal stationary distribution of absolute advantage in the cross section, with a shape parameter of $1/\eta$ and a mean of zero.

The estimated dissipation coefficient ρ is a function both of the dissipation rate η and the intensity of innovations σ and therefore may vary across samples because either or both of these parameters vary. This distinction is important because ρ may change even though the heavy tail of the distribution of comparative advantage does not. From OLS estimation of the decay regression in (12), we can obtain estimates of η and σ^2 using the solutions,

$$\begin{aligned} \eta &= \frac{1-(1+\hat{\rho})^2}{\hat{s}^2} \\ \sigma^2 &= \frac{\hat{s}^2}{1-(1+\hat{\rho})^2}\frac{\ln(1+\hat{\rho})^{-2}}{\Delta}, \end{aligned} \qquad (13)$$

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 delta method. Across samples, the estimate of η based on log absolute advantage is very similar to that based on the log RCA index, implying that the two measures of comparative advantage have a cross sectional distribution of similar shape. Patterns of interest emerge when we compare η and σ^2 across subsamples.

First, consider the subsample of developing economies in columns 3 and 4 of **Table 1** and compare the estimates to those for the average country in the full sample (columns 1 and 2). The larger estimates of ρ in absolute value imply that mean reversion is more rapid in the developing-country group. However, this result is silent about any underlying country differences in the cross sectional distribution of comparative advantage. We see that the estimated dissipation rate η among developing countries is not markedly different from that in the average country; in fact the η estimates are not statistically significantly different from each other for the exporter capability measure k. This similarity in the estimated dissipation rate η indicates that comparative advantage is similarly heavy-tailed in the group of developing countries as in the sample of all countries. The faster reduced-form decay rate ρ for developing countries results mainly from their having a larger intensity of innovations σ. In other words, a typical comparative-advantage innovation (a one-standard-deviation shock) in a developing country dissipates at roughly the same rate as in an industrialized country but the typical innovation is larger in a developing country.

Second, we can compare non-manufacturing industries in columns 5 and 6 to the average industry in columns 1 and 2. Whereas non-manufacturing industries differ considerably from the average industry in measured decay rates ρ, there is no such marked difference in the estimated dissipation rates η. For either measure of comparative advantage, the dissipation rate η is more similar between a non-manufacturing industry and the average industry than the measured decay rates ρ would appear. This implies that comparative advantage is similarly heavy-tailed among non-manufacturing industries as in the sample of all industries. However, the intensity of innovations is much larger in non-manufacturing industries than in the average industry, perhaps due to higher volatility in commodity output or commodity prices. These nuances regarding the implied shape of, and the convergence speed towards, the cross sectional distribution of comparative advantage are not apparent when one focuses only on the reduced-form decay rates themselves.

Finally, we compare results across two, three, and four-digit industries in Appendix **Table A2** for the subperiod 1984-2007 when a more detailed industry classification becomes available. Whereas reduced-form decay rates ρ increase in magnitude as one goes from the two to four-digit level, dissipation rates η tend to move in the opposite direction and fall as one goes from the more aggregate to the more detailed industry classification. For the exporter capability measure of comparative advantage, the drop in η between the two and the four-digit level is not statistically significant in the full sample of industries and countries—indicating intuitively that the shape of the cross sectional distribution of comparative advantage remains similar at varying levels of industry aggregation.[27] The difference in reduced-form decay rates ρ is largely driven by a larger intensity of innovations σ among the more narrowly defined industries at the four-digit level.

The diffusion model in (11) and its discrete analogue in (12) reveal a deep connection between hyperspecialization in exporting and churning in industry export ranks. Random innovations in absolute advantage cause industries to alternate places in the cross sectional distribution of comparative advantage for a country, while the dissipation of absolute advantage creates a stable, heavy-tailed distribution of export prowess. Having established a connection between hyperspecialization and industry churning, we turn next to a more rigorous analysis of its origins.

[27] In the subsamples of less developed countries and non-manufacturing industries, however, the dissipation rates η fall more pronounced as one goes from the more aggregate to the more detailed industry classification. Those findings imply that, in those subsamples, the cross sectional distribution of comparative advantage is more widely dispersed and thus more heavy tailed for the more detailed industry classification. Intuitively, at finer levels of aggregation, the few top industries carry more weight.

4 The Diffusion of Comparative Advantage

We now search in a more general setting for a parsimonious stochastic process that characterizes the dynamics of comparative advantage. In Figure 2, the cross-sectional distributions of absolute advantage drift rightward perpetually, implying that absolute advantage is not stationary. However, the cross-sectional distributions preserve their shape over time. We therefore consider absolute advantage as a proportionally scaled outcome of an underlying stationary and ergodic variable: comparative advantage. One candidate stationary and ergodic variable is the Balassa RCA index because it removes a specific type of country-wide trend. Instead of limiting ourselves to a narrowly imposed form, we specify *generalized comparative advantage* in continuous time as

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

where $A_{is}(t)$ is observed absolute advantage and $Z_s(t)$ is an unobserved country-wide stochastic trend. It follows directly that this measure satisfies the properties of the comparative advantage statistic in (3) that compares individual country and industry pairs.

To find a well-defined stochastic process that is consistent with the churning of absolute advantage over time and with heavy tails in the cross section, we implement a generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$, which has a generalized gamma as its stationary distribution. Comparative advantage in the cross section is then denoted with $\hat{A}_{is}$ and understood to have a time-invariant distribution. Absolute advantage $A_{is}(t)$, in contrast, has a trend-scaled generalized gamma as its cross-sectional distribution, with stable shape but moving position as in Figure 2.[28]

The attractive feature of the generalized gamma is that it nests many distributions as special or limiting cases, making the diffusion we employ flexible in nature. We construct a GMM estimator by working with a mirror diffusion, which is related to the generalized logistic diffusion through an invertible transformation. Our estimator uses the conditional moments of the mirror diffusion and accommodates the fact that we observe absolute advantage only at discrete points in time. After estimating the stochastic process from the time series of absolute advantage in Section 5, we explore how well the implied cross-sectional distribution fits the actual cross-section data, which we do not target in estimation.

[28] In log terms, the non-stationary trend becomes an additive component that continually shifts the stationary distribution of comparative advantage: $\ln A_{is}(t) = \ln Z_s(t) + \ln \hat{A}_{is}$.

4.1 Generalized logistic diffusion

The regularities in Section 3.1 indicate that the log normal distribution is a plausible benchmark distribution for the cross section of absolute advantage.[29] But the graphs in **Figure 2** (and their companion graphs in **Figures A1** through **A3**) also suggest that for many countries and years, the number of industries drops off faster or more slowly in the upper tail than the log normal distribution can capture. We require a distribution that generates kurtosis that is not simply a function of the lower-order moments, as would be the case in the two-parameter log normal. The generalized gamma distribution, which unifies the gamma and extreme-value distributions as well as many other distributions (Crooks 2010), offers a candidate family.[30] Our implementation of the generalized gamma uses three parameters, as in Stacy (1962).[31]

In a cross section of the data, after arbitrarily much time has passed, the proposed relevant generalized gamma probability density function for a realization $\hat{a}_{is}$ of the random variable comparative advantage $\hat{A}_{is}$ is given by:

$$f_{\hat{A}}(\hat{a}_{is};\hat{\theta},\kappa,\phi) = \frac{1}{\Gamma(\kappa)}\left|\frac{\phi}{\hat{\theta}}\right|\left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi\kappa-1}\exp\left\{-\left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi}\right\} \qquad \text{for} \quad \hat{a}_{is} > 0, \tag{15}$$

where $\Gamma(\cdot)$ denotes the gamma function and $(\hat{\theta},\kappa,\phi)$ are real parameters with $\hat{\theta},\kappa > 0$.[32] The generalized gamma nests as special cases, among several others, the ordinary gamma distribution for $\phi = 1$ and the log normal or Pareto distributions when ϕ tends to zero.[33] 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 $1/\phi$. The exponentiated random variable is then generalized gamma distributed, a result that points to a candidate stochastic process that has a stationary generalized gamma distribution. The ordinary *logistic diffusion*, a widely used stochastic process, generates an ordinary gamma as its stationary

[29] A log normal distribution also approximates the firm size distribution reasonably well (Sutton 1997). For the United States, Axtell (2001) argues that a Pareto distribution offers a tight fit to firm sizes but also documents that, in the upper and lower tails of the cumulative distribution, the data exhibit curvature consistent with a log normal distribution and at variance with a Pareto distribution.

[30] In their analysis of the firm size distribution by age, Cabral and Mata (2003) also use a version of the generalized gamma distribution with a support bounded below by zero and document a good fit.

[31] In the original Amoroso (1925) formulationthe generalized gamma distribution has four parameters. One of the four parameters is the lower bound of the support. However, our measure of absolute advantage A_{is} can be arbitrarily close to zero by construction (because the exporter-industry fixed effect in gravity estimation is not bounded below so that by (7) $\log A_{is}$ can be negative and arbitrarily small). As a consequence, the lower bound of the support is zero in our application. This reduces the relevant generalized gamma distribution to a three-parameter function.

[32] We do not restrict ϕ to be strictly positive (as do e.g. Kotz et al. 1994, ch. 17). 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 gamma, inverse gamma, Fréchet, Weibull and numerous other distributions as special cases and yields the normal, log normal and Pareto distributions as limiting cases.

[33] 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).

distribution (Leigh 1968). By extension, the *generalized* logistic diffusion has a *generalized* gamma as its stationary distribution.

Lemma 1. *The generalized logistic diffusion*

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

for real parameters η, ϕ, σ has a stationary distribution that is generalized gamma with a probability density $f_{\hat{A}}(\hat{a}_{is}; \hat{\theta}, \kappa, \phi)$ given by (15)*, for $\hat{A}_{is}$ (understood to have a time-invariant cross sectional distribution) 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 $\eta > 0$.

Proof. See Appendix A. □

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 $\mathrm{d}\hat{A}_{is}(t)/\hat{A}_{is}(t)$. The variable $W_{is}^{\hat{A}}(t)$ is the Wiener process. The generalized logistic diffusion nests the Ornstein-Uhlenbeck process ($\phi \to 0$), leading to a log normal distribution in the cross section. In the estimation, we will impose the condition that $\eta > 0$.[34]

The deterministic drift involves two types of components: constant parameters (η, ϕ, σ) on the one hand, and a level-dependent component $\hat{A}_{is}(t)^{\phi}$ on the other hand, where ϕ is the elasticity of the mean reversion with respect to the current level of absolute advantage. We call ϕ the *level elasticity of dissipation*. The ordinary logistic diffusion has a unitary level elasticity of dissipation ($\phi = 1$). In our benchmark case of 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 level-dependent 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 tends to occur more rapidly than elimination of absolute disadvantage. Conversely, if $\phi < 0$ then mean reversion tends to occur more slowly from above than below the long-run mean, indicating that absolute advantage is sticky. Only in the level neutral case of $\phi \to 0$ is the rate of mean reversion from above and below the mean the same.

The parameters η and σ in the generalized logistic diffusion in (16) inherit their interpretations from the OU process in (11) as the rate of dissipation and the intensity of innovations, respectively. The intensity of innovations σ again plays a dual role: on the one hand magnifying volatility by scaling up the Wiener innovations and on the other hand regulating how fast time elapses in the deterministic part of the diffusion.

[34] If η where negative, comparative advantage would collapse over time for $\phi < 0$ or explode for $\phi \geq 0$.

This dual role now guarantees that the diffusion will have a non-degenerate stationary distribution. Scaling the deterministic part of the diffusion by $\sigma^2/2$ ensures that stochastic deviations of comparative advantage from the long-run mean do not persist and that dissipation occurs at precisely the right speed to offset the unbounded random walk that the Wiener process would otherwise induce for each country-industry.

Under the generalized logistic diffusion, the dissipation rate η and dissipation elasticity ϕ jointly determine the heavy tail of the cross sectional distribution of comparative advantage, with the intensity of innovations σ determining the speed of convergence to this distribution but having no effect on its shape.

For subsequent derivations, it is convenient to restate the generalized logistic diffusion (16) more compactly in terms of log changes as,

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

which follows from (16) by Itō's lemma.

4.2 The cross sectional distributions of comparative and absolute advantage

If comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion by (16), then the stationary distribution of comparative advantage is a generalized gamma distribution with density (15) 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 $\hat{A}_{is}(t)$, we can infer the cross distribution of absolute advantage $A_{is}(t)$. Note that, by definition (14), absolute advantage is not necessarily stationary because the stochastic trend may not be stationary.

Absolute advantage is related to comparative advantage through a country-wide stochastic trend by definition (14). Plugging this definition into (15), 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*

$$\mathrm{d}\ln \hat{A}_{is}(t) = -\frac{\eta\sigma^2}{2}\frac{\hat{A}_{is}(t)^{\phi}-1}{\phi}\,\mathrm{d}t + \sigma\,\mathrm{d}W_{is}^{\hat{A}}(t) \qquad (17)$$

with real parameters η, σ, ϕ *(*$\eta > 0$*), then the stationary distribution of comparative advantage* $\hat{A}_{is}(t)$ *is generalized gamma with the CDF*

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

where $G[x; \kappa] \equiv \gamma_x(\kappa; x)/\Gamma(\kappa)$ *is the ratio of the lower incomplete gamma function and the gamma function, and the cross sectional distribution of absolute advantage* $A_{is}(t)$ *is generalized gamma with the CDF*

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

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 distributions are generalized gamma. The cumulative distribution functions follow from Kotz et al. (1994, Ch. 17, Section 8.7). □

The graphs in **Figure 2** plot the frequency of industries, that is the probability $1 - F_A(a; \theta_s(t), \phi, \kappa)$ times the total number of industries ($I = 135$), on the vertical axis against the level of absolute advantage a (such that $A \geq a$) on the horizontal axis. Both axes have a log scale. Lemma 2 clarifies that a country-wide stochastic trend $Z_s(t)$ shifts log absolute advantage in the graph horizontally but the shape related parameters ϕ and κ are not country specific if comparative advantage follows a diffusion with a common set of three deep parameters $\hat{\theta}, \kappa, \phi$ worldwide.

Finally, as a prelude to the GMM estimation we note that 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 et al. 1994, Ch. 17, Section 8.7), where $\Gamma(\cdot)$ denotes the gamma function. 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 deep parameters $\hat{\theta}, \kappa, \phi$:

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

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 (*mean/median ratio*), which is independent of $\hat{\theta}$ and equals

$$\textit{Mean/median ratio} = \frac{\Gamma(\kappa + 1/\phi)/\Gamma(\kappa)}{(G^{-1}[.5;\kappa])^{1/\phi}}. \tag{18}$$

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

4.3 Implementation

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 generalized logistic diffusion is that it can be transformed into a diffusion that belongs to the Pearson-Wong family, for which closed-form solutions of the conditional moments exist.[35] We construct a consistent GMM estimator based on the conditional moments of a transformation of comparative advantage, using results from Forman and Sørensen (2008).

Our model depends implicitly on the unobserved stochastic trend $Z_s(t)$. We use a closed form expression for the mean of a log-gamma distribution to identify and concentrate out this trend. For a given country and year, the cross-section of the data across industries has a generalized gamma distribution. The mean of the log of this distribution can be calculated explicitly as a function of the model parameters, enabling us to identify the trend from the relation that $\mathbb{E}_{st}[\ln \hat{A}_{is}(t)] = \mathbb{E}_{st}[\ln A_{is}(t)] - \ln Z_s(t)$ by definition (14). We adopt the convention that the expectations operator $\mathbb{E}_{st}[\cdot]$ denotes the conditional expectation for source country s at time t. This result is summarized in the following proposition:

Proposition 1. *If comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion* (16) *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:*

$$Z_s(t) = \exp\left\{\mathbb{E}_{st}[\ln A_{is}(t)] - \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\} \tag{19}$$

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

Proof. See Appendix B. □

[35] 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.

This proposition implies that for any GMM estimator, we can concentrate out the stochastic trend in absolute advantage and work with an estimate of comparative advantage directly. Concretely, we obtain detrended data based on the sample analog of equation (19):

$$\hat{A}_{is}^{GMM}(t) = \exp\left\{ \ln A_{is}(t) - \frac{1}{I}\sum_{j=1}^{I} \ln A_{js}(t) + \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi} \right\} \tag{20}$$

Detrending absolute advantage to arrive at an estimate of comparative advantage completes the first step in implementing model (16).

Next, we perform a change of variable to recast our model as a Pearson-Wong diffusion. Rewriting our model as a member of the Pearson-Wong family allows us to apply results in Kessler and Sørensen (1999) and construct closed-form expressions for the conditional moments of comparative advantage. This approach, introduced by Forman and Sørensen (2008), enables us to estimate the model using GMM.[36] The following proposition presents an invertible transformation of comparative advantage that makes estimation possible.

Proposition 2. *If comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion* (16) *with real parameters η, σ, ϕ ($\eta > 0$), then:*

1. *The transformed variable*

$$\hat{B}_{is}(t) = [\hat{A}_{is}(t)^{-\phi} - 1]/\phi \tag{21}$$

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.

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

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^n \,\middle|\, \hat{B}_{is}(t) = b\right] = \exp\{-a_n\Delta\}\sum_{m=0}^{n}\pi_{n,m}b^m - \sum_{m=0}^{n-1}\pi_{n,m}\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^m \,\middle|\, \hat{B}_{is}(t) = b\right] \tag{22}$$

where the coefficients a_n and $\pi_{n,m}$ ($n, m = 1, \ldots, M$) are defined in Appendix C.

[36] More generally, our approach fits into the general framework of prediction-based estimating functions reviewed in Sørensen (2011) and discussed in Bibby et al. (2010). These techniques have been previously applied in biostatistics (e.g., Forman and Sørensen 2013) and finance (e.g., Lunde and Brix 2013).

Proof. See Appendix C. □

Transformation (21) converts the diffusion of comparative advantage $\hat{A}_{is}(t)$ into a mirror specification that has closed form conditional moments. This central result enables us to construct a GMM estimator.

Consider time series observations for $\hat{B}_{is}(t)$ at times $t_1, \ldots, t_T$. By equation (22) in Proposition 2, we can calculate a closed form for the conditional moments of the transformed diffusion at time t_τ conditional on the information set at time $t_{\tau-1}$. We then compute the forecast error based on using these conditional moments to forecast the m-th power of $\hat{B}_{is}(t_\tau)$ with time $t_{\tau-1}$ information. These forecast errors must be 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, $U_{is}(m, t_{\tau-1}, t_\tau)$ is uncorrelated with any measurable transformation of $\hat{B}_{is}(t_{\tau-1})$. A GMM criterion function based on these forecast errors is

$$g_{is}(\phi, \eta, \sigma^2) \equiv \frac{1}{T-1} \sum_{\tau=2}^{T} [h_1(\hat{B}_{is}(t_{\tau-1}))U_{is}(1, t_{\tau-1}, t_\tau), \ldots, h_M(\hat{B}_{is}(t_{\tau-1}))U_{is}(M, t_{\tau-1}, t_\tau)]'$$

where each h_m is a row vector of measurable functions specifying instruments for the m-th moment condition. This criterion function is 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.[37]

For observations from I industries across S source countries, our GMM estimator solves the minimization problem

$$(\phi^*, \eta^*, \sigma^{2*}) = \arg \min_{(\phi,\eta,\sigma^2)} \left(\frac{1}{IS} \sum_i \sum_s g_{is}(\phi, \eta, \sigma^2)\right)' W \left(\frac{1}{IS} \sum_i \sum_s g_{is}(\phi, \eta, \sigma^2)\right)$$

for a given weighting matrix W.

[37] We work with a sub-optimal estimator because the optimal-instrument GMM estimator considered by Forman and Sørensen (2008) requires the inversion of a matrix for each observation. Given our large sample, this task is numerically expensive. Moreover, our ultimate GMM objective is ill-conditioned and optimization to find our estimates of ϕ, η, and σ^2 requires the use of an expensive global numerical optimization algorithm. For these computational concerns we sacrifice efficiency and use sub-optimal instruments.

We evaluate this objective function at values of ϕ, η, and σ^2 by detrending the data to obtain $\hat{A}_{is}^{GMM}(t)$ from equation (20), transforming these variables into their mirror variables $\hat{B}_{is}^{GMM}(t) = [\hat{A}_{is}^{GMM}(t)^{-\phi} - 1]/\phi$, and using equation (22) 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 $\hat{B}_{is}^{GMM}(t)$, and finally sum over industries and countries to arrive at the value of the GMM objective.

For estimation we use two conditional moments and three instruments, leaving us with six equations for three parameters. 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 $W^{-1} = (1/IS)\sum_i \sum_s g_{is}(\phi, \eta, \sigma^2) g_{is}(\phi, \eta, \sigma^2)'$, which is calculated at the parameter value from the first step. Forman and Sørensen (2008) establish asymptotics as $T \to \infty$.[38] We impose the constraints that $\eta > 0$ and $\sigma^2 > 0$ by reparameterizing the model in terms of $\ln \eta > -\infty$ and $\ln(\sigma^2) > -\infty$, and use the delta method to calculate standard errors for functions of the transformed parameters.

5 Estimates

Following the GMM procedure described in Section 4.3, we proceed to estimate the parameters for the global diffusion of comparative advantage (η, σ, ϕ). It is worthy of note that, subject to a country-specific stochastic trend, we are attempting to describe the global evolution of comparative advantage using just three time-invariant parameters, which by implication must apply to all industries in all countries and in all time periods. This approach contrasts sharply with our initial descriptive exercise in **Figure 2**, which fits cumulative distribution plots to the log normal based on distribution parameters estimated separately for each country and each year. **Table 2** presents the estimation results. To verify that the results are not a byproduct of specification error in estimating export capabilities from the gravity model, we also perform GMM estimation using the Balassa (1965) RCA index.

The magnitude of the estimate of η, which captures the dissipation of comparative advantage, is somewhat difficult to evaluate on its own. In its combination with the level elasticity of dissipation ϕ, η controls both the magnitude of the long-run mean and the curvature of the cross-sectional distribution. The sign of ϕ captures the stickiness of comparative advantage. The parameter estimate of ϕ is robustly negative (and

[38] Our estimator would also fit into the standard GMM framework of Hansen (1982), which establishes consistency and asymptotic normality of our estimator for the product $IS \to \infty$. Given the dynamic nature of our times series exercise, we base the GMM weighting matrix and computations of standard errors on the asymptotics under $T \to \infty$. Results under the alternative asymptotics of $IS \to \infty$ are available from the authors upon request; those asymptotics tend to lead to less stable estimates across specifications.

Table 2: GMM ESTIMATES OF COMPARATIVE ADVANTAGE DIFFUSION

	Full sample		Subsamples: Absolute advantage A			
	Abs. adv.	Rev. adv.	Exporter countries		Sectors	
	A	$\hat{X}$	LDC	Non-LDC	Manuf.	Nonmanf.
	(1)	(2)	(3)	(4)	(5)	(6)
Estimated Generalized Logistic Diffusion Parameters						
Dissipation rate η	0.265	0.190	0.265	0.264	0.416	0.246
	(0.004)**	(.002)**	(0.004)**	(0.008)**	(0.005)**	(0.004)**
Intensity of innovations σ	1.396	1.189	1.587	0.971	1.157	1.625
	(0.039)**	(.026)**	(0.045)**	(0.073)**	(0.011)**	(0.042)**
Level elast. of dissipation ϕ	-.034	-.030	-.027	-.036	-.064	-.028
	(0.004)**	(.004)**	(0.005)**	(0.01)**	(0.012)**	(0.004)**
Implied Parameters						
Log gen. gamma scale $\ln\hat{\theta}$	160.7	177.8	223.8	146.4	72.6	201.6
	(25.3)**	(28.6)**	(52.5)**	(57.8)*	(20.3)**	(42.3)**
Log gen. gamma shape $\ln\kappa$	5.443	5.349	5.933	5.302	4.628	5.722
	(0.236)**	(0.226)**	(0.355)**	(0.585)**	(0.395)**	(0.316)**
Mean/median ratio	7.375	16.615	7.181	7.533	3.678	8.457
Obs.	459,680	459,680	296,060	163,620	230,890	228,790
Root mean sq. forecast error	1.250	1.090	1.381	.913	1.022	1.409
Min. GMM obj. ($\times$ 1,000)	0.411	0.012	1.040	1.072	0.401	0.440

Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: GMM estimation of 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 annual absolute advantage measures $A_{is}(t) = \hat{A}_{is}(t)Z_s(t)$ on the full pooled sample (column 1) and subsamples (columns 3-6), and using the Balassa index of revealed comparative advantage $\hat{X}_{ist} = (X_{ist}/\sum_{s'} X_{is't})/(\sum_{i'} X_{i'st}/\sum_{i'}\sum_{s'} X_{i's't})$ instead of absolute advantage (column 2). Parameters η, σ, ϕ are estimated under the restrictions $\ln\eta, \ln\sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (21), 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 (18). Less developed countries (LDC) as listed in Appendix E. The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Standard errors in parentheses: * marks significance at five and ** at one percent level. Standard errors of transformed and implied parameters are computed using the delta method.

precisely estimated), so we reject log normality in favor of the generalized gamma distribution. Negativity in ϕ implies that comparative advantage reverts to the long-run mean more slowly from above than from below. However, the value of ϕ is not far from zero, suggesting that in practice deviations in comparative advantage from log normality may be modest, as the plots in **Figure 2** suggest.

The parameter σ regulates the intensity of innovations and captures both the volatility of the Wiener innovations to comparative advantage and the a speed of convergence on the deterministic decay. This dual role binds the parameter estimate of σ to a level precisely such that a non-degenerate stationary distribution exists. The intensity of innovations therefore does not play a role in determining the cross-sectional distribution's shape. That job is performed by κ and $\hat{\theta}$, which exclusively depend on η and ϕ, so we are effectively describing the shape of the cross-sectional distribution with just two parameters.

The parameters η and ϕ together imply a shape of the distribution with a strong concentration of absolute and comparative advantage in the upper tail. The mean exceeds the median by a factor of more than seven, both among developing and industrialized countries. This considerable concentration is mainly driven by industries in the non-manufacturing merchandise sector, which exhibit a mean/median ratio of more than eight (column 6), whereas the ratio is less than four for industries in the manufacturing sector (column 5). When we use the Balassa (1965) RCA index, the mean/median ratio more than doubles to 16 (column 2). One interpretation of the greater concentration in revealed comparative advantage relative to our geography-adjusted absolute advantage measure is that geography reinforces comparative advantage by making countries appear overspecialized in the goods in which their underlying advantage is strong.

In Appendix Table A3 (to be included) we repeat the GMM procedure using data for the post-1984 period on SITC revision 2 industries at the two, three, or four-digit level. The results are largely in line with those in **Table 2**. Estimates of the dissipation rate η are slightly larger for the post-1984 period than for the full sample period, and, similar to what we found in the decay regressions in **Table 1**, become larger as one moves from higher to lower levels of industry disaggregation. Estimates of the elasticity of dissipation ϕ are negative in all cases except one—when we measure export prowess using log absolute advantage (based on the gravity fixed effects) at the four-digit SITC level. As mentioned in Section 3.1, with nearly 700 four-digit SITC rev. 2 industries we frequently have few destination markets per exporter-industry with which to estimate the gravity fixed effects, contributing to noise in the estimated exporter-industry coefficients.

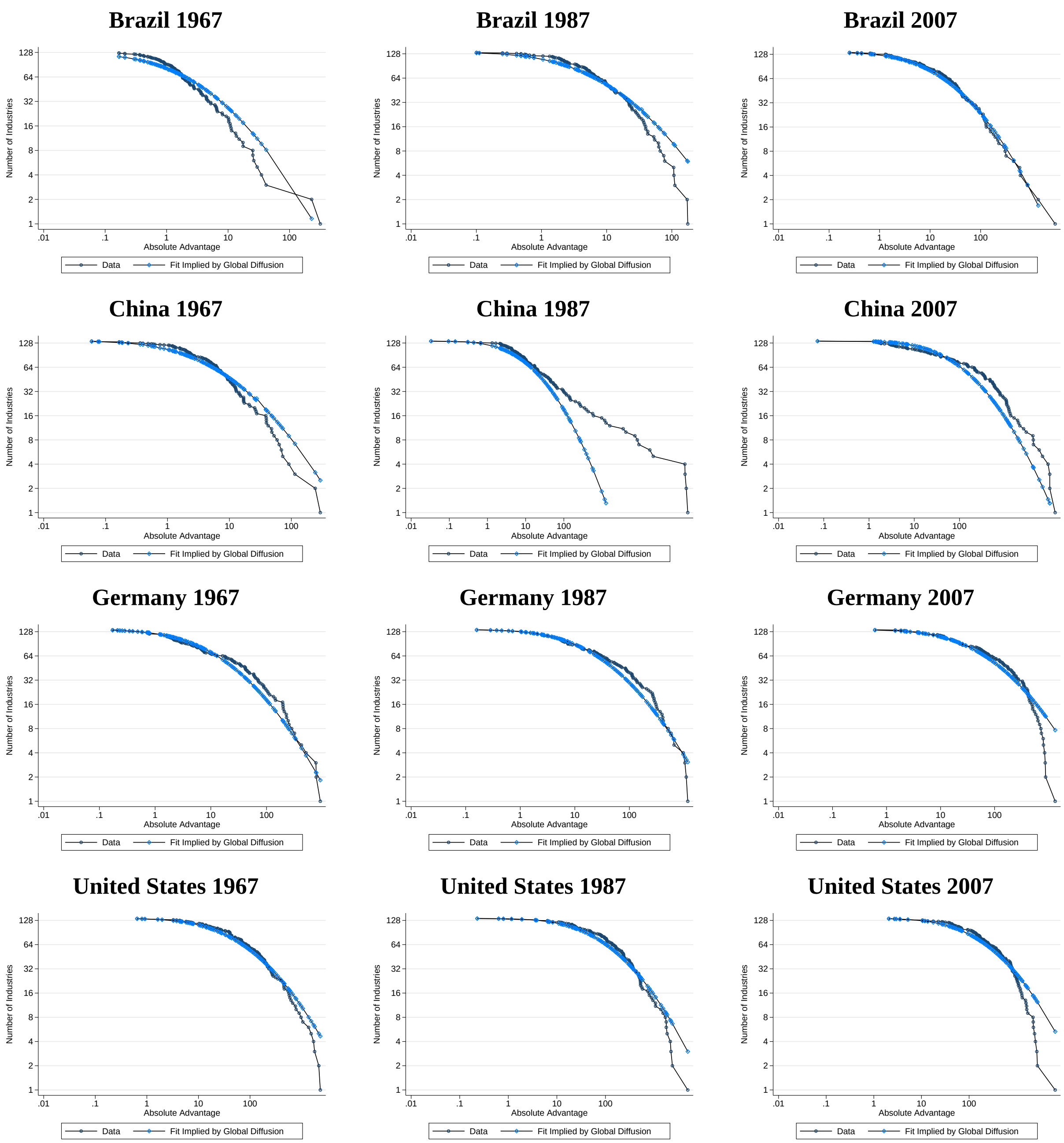
The parameters themselves give no indication of the fit of the model. To evaluate fit, we exploit the fact that our GMM estimation targets exclusively the diffusion of comparative advantage—that is, the time series behavior for country-industries—and not its cross-sectional dimension. Thus, the cross sectional distribution of comparative advantage for a given country at a given moment in time provides a means of validating our estimation procedure. For each country in each year, we project the cross sectional distribution of comparative advantage implied by the parameters estimated from the diffusion and compare it to the distribution based on the raw data.

To implement our validation exercise, we need a measure of $\hat{A}_{ist}$ in equation (14), whose value depends on Z_{st}, the country-specific stochastic trend, which is unobserved. The role of the stochastic trend in the diffusion is to account for horizontal shift in the distribution of log absolute advantage, which may result from country-specific technological progress, factor accumulation, or other sources of aggregate growth. In the estimation, we concentrate out Z_{st} by exploiting the fact that both $\hat{A}_{ist}$ and A_{ist} have generalized gamma distributions, allowing us to obtain closed-form solutions for their means, which isolates the value of the stochastic trend. To obtain an empirical estimate of Z_{st} at a given moment in time we apply equation (19), which defines the variable as the difference between the mean log value of A_{ist} and the expected value of a log gamma distributed variable (which is a function of η and ϕ). With estimated realizations for each country in each year of Z_{st} in hand, we compute realized values for $\hat{A}_{ist}$ for each country-industry in each year.

To gauge the goodness of fit of our specification, we first plot our measure of absolute advantage A_{ist}. To do so, following the earlier exercise in **Figure 2**, 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 given by 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 as well as the identified country-specific trend Z_{st}.

Figure 4 compares plots of the actual data against the diffusion implied plots for four countries in three years, 1967, 1987, 2007. **Figures A7, A8** and **A9** in the Appendix present plots for the same 28 countries in 1967, 1987 and 2007 as shown in **Figures A1, A2** and **A3** before. While **Figures A1** through **A3** depicted Pareto and log normal maximum likelihood estimates of each individual country's cross sectional distribution by year (such that the number of parameters estimated equaled the number of parameters for a distribution $\times$ number of countries $\times$ 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 4** and **Figures A7** through **A9** present the same, horizontally shifting but identically shaped, single cross-sectional distribution, as implied by the two shape relevant parameter estimates (out of the three total) that fit the global diffusion

Figure 4: **Global Diffusion Implied and Observed Cumulative Probability Distributions of Absolute Advantage for Select Countries in 1967, 1987 and 2007**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

for all country-industries and years. The country-specific trend Z_{st} terms shift the implied stationary distribution horizontally, and we cut the depicted part of that single distribution at the lower and upper bounds of the specific country's observed support in a given year to clarify the fit.

Considering that the shape of the distribution effectively depends on only two parameters for all country-industries and years, the simulated distributions fit the actual data remarkably well. There are important differences between the actual and predicted plots in only a few countries and a few years, including China in 1987, Russia in 1987 and 2007, Taiwan in 1987, and Vietnam in 1987 and 2007. Three of these four cases involve countries undergoing a transition away from central planning during the designated time period, suggesting periods of economic tumult.

There are some telling discrepancies between the actual and diffusion implied plots that are worthy of further investigation. First, for some countries the upper tail of the distribution in the actual data plots falls off more quickly than the predicted stationary distribution would imply. This suggests that for some countries comparative advantage is relatively sticky (i.e., the true value of ϕ for these countries may be larger in absolute value than that shown in **Table 2**). However, a handful of countries in East and Southeast Asia—China, Japan, Korea Rep., Malaysia, Taiwan, and Vietnam—show the opposite pattern. They exhibit less concavity in the data than in the diffusion implied distribution, revealing less stickiness in comparative advantage than the predicted stationary CDF would indicate, consistent with a ϕ that is smaller in absolute value than in **Table 2** or even positive. What remains unclear is whether these differences in fit across countries are associated with the countries or with particular industries in these countries, an issue we will explore in upcoming work.

Future empirical analysis in this paper will account for the following extensions.

1. We will use our estimates of the parameters of the generalized gamma distribution to simulate a multi-sector version of the EK model. First, we will use the generalized gamma to generate location parameters of the Fréchet distribution for firm productivity in each industry and in each country. We will then combine these location parameters with values for preference and technology parameters taken from the trade literature to simulate a global general equilibrium, which yields a gravity equation. Finally, we will add randomly generated noise to the "true" trade values and apply the gravity model to estimate exporter-industry fixed effects on the simulated data plus noise. By comparing these gravity estimates to our underlying generalized gamma draws of the location parameters, we can assess the extent to which measurement error in trade data contaminates our measurement of country export prowess.

2. We will examine alternative measures of the goodness of fit of the generalized logistic diffusion by (a)

plotting observed quantiles for absolute advantage against predicted quantiles for absolute advantage, and (b) restricting the estimation to the latter half of the sample period and using these estimates to simulate distributions for the first half of the sample period.

3. We will examine the robustness of our results to (a) using MPML-based estimates of gravity fixed effects that account for zero flows, and (b) excluding industries (mainly in electronics, electrical machinery, transportation equipment, apparel, and footwear) in which global production networks figure prominently and in which domestic value added accounts for a relatively small share of gross exports.

4. We will derive the exact discrete-time process that results from sampling from our generalized logistic diffusion at a fixed time interval Δ and compute the precise decadal evanescence rate for $\phi \neq 0$ and $\Delta = 10$ using the according generalized autoregressive parameter function of the exact discrete-time process and evaluate ρ at three percentiles of comparative advantage for the pooled sample as well as by country and sector.

5. We will re-estimate the GMM specification by explicitly allowing the absolute advantage measures A_{ist} to be aggregates of trade events between the discrete points of observation Sørensen (2011), beyond our current implementation of discrete-time trade events.

6 Conclusion

Two salient facts about comparative advantage arise from our investigation of trade flows among a large set of countries and industries over more than four decades: While at any moment of time countries exhibit hyperspecialization in only a few industries, the deviation in comparative advantage from its long-run global mean dissipates at a brisk rate, of one-quarter to one-third over a decade. This impermanence implies that the identity of the industries in which a country currently specializes changes considerably over time. Within two decades, a country's rising industries replace on average three of its top five initial industries in terms of absolute advantage.

We specify a parsimonious stochastic process for comparative advantage with only three parameters by generalizing the two-parameter logistic diffusion. The generalized logistic diffusion is consistent with both hyperspecialization in the cross section and perpetual churning in industry export ranks. We additionally allow for a country-specific stochastic trend whose removal translates absolute advantage into comparative advantage and estimate the global parameters of the generalized logistic diffusion using a recently developed GMM estimator for a well-defined mirror process. In this novel approach, we estimate the stochastic process itself, rather than the repeated cross sections, and then use the two time-invariant diffusion parameters that determine the shape of the cross-sectional distribution to assess the fit of the predicted cross-sectional distribution across countries and over time. Even though our estimator does not target the cross sections—but rather the annual diffusion—we find that the shape of the predicted stationary cross section tightly matches the shape and curvature of the observed cross-sectional distributions for the bulk of countries and years.

The exercises in this paper deliberately set aside questions about the deeper origin of comparative advantage and aim instead to characterize the empirical evolution of comparative advantage in a typical country-industry. In future research, we plan to explore natural follow-up questions.

1. We plan a systematic account of the country-industries whose evolution defies the global diffusion in the sense that their rapid success or decline over time beats the odds and lies outside a confidence bound of the likely evolution under the specified generalized logistic diffusion. Once the outside-the-odds successes and failures are accounted for, we can ask whether their subsequent performance remained outside the odds and what known market-driven forces or government interventions may account for their beating the odds. In this context, we can explore the addition of a Lévy jump process to our generalized logistic diffusion, generating a stationary distribution with no closed form, while restricting parameters so that the implied stationary distribution approximates the generalized gamma arbitrarily closely. The resulting stochastic process can potentially explain the evolution of individual country-industries more completely.

2. We plan to bring firm-level evidence on the employment and sales concentration among exporting and non-exporting firms in select countries to the project and thus complement our sector-level evidence with recent advances in firm-level theories of international trade. Countries for which we have access to firm-level data include Brazil, Germany and Sweden. Firms might withstand sector-level dissipation of comparative advantage by expanding their product scope across sectors or, alternatively, might be subject to similar rates of dissipation as their home sector. Firm-level evidence can sharpen our understanding of how the ongoing process of innovation in manufacturing industries and exploration in non-manufacturing industries contribute to hyperspecialization and industry churning.

Appendix

A Generalized Logistic Diffusion: Proof of Lemma 1

The ordinary gamma distribution arises as 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. Note that the generalized (three-parameter) gamma distribution relates 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

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

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

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

which is equivalent to the generalized gamma probability density function (15), where $\Gamma(\cdot)$ denotes the gamma function and $\hat{\theta}, \kappa, \phi$ are the three parameters of the generalized gamma distribution in our context ($a > 0$ can be arbitrarily close to zero).

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

$$\mathrm{d}X(t) = \left[\bar{\alpha} - \bar{\beta}\, X(t)\right] X(t)\, \mathrm{d}t + \bar{\sigma}\, X(t)\, \mathrm{d}W(t) \qquad \text{for} \quad X(t) > 0, \tag{A.2}$$

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):

$$f_X(x; \bar{\theta}, \kappa) = \frac{1}{\Gamma(\kappa)} \left|\frac{1}{\bar{\theta}}\right| \left(\frac{x}{\bar{\theta}}\right)^{\kappa-1} \exp\left\{-\frac{x}{\bar{\theta}}\right\} \qquad \text{for} \quad x > 0, \tag{A.3}$$

as in (A.1) with

$$\begin{aligned} \bar{\theta} &= \bar{\sigma}^2/(2\bar{\beta}) > 0, \\ \kappa &= 2\bar{\alpha}/\bar{\sigma}^2 - 1 > 0 \end{aligned} \tag{A.4}$$

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

$$\mathrm{d}X(t) = \mu_X(X(t))\, \mathrm{d}t + \sigma_X(X(t))\, \mathrm{d}W(t) \qquad \text{for} \quad X(t) > 0,$$

where

$$\mu_X(X) = (\bar{\alpha} - \bar{\beta}\, X)X \quad \text{and} \quad \sigma_X^2(X) = \bar{\sigma}^2\, X^2.$$

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

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

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

$$\mathrm{d}A(t) = \left[\alpha - \beta A(t)^{\phi}\right] A(t)\,\mathrm{d}t + \sigma A(t)\,\mathrm{d}W(t) \qquad \text{for} \quad A(t) > 0. \tag{A.5}$$

with the parameters

$$\alpha \equiv \left[\frac{1-\phi}{2}\frac{\bar{\sigma}^2}{\phi^2} + \frac{\bar{\alpha}}{\phi}\right], \qquad \beta \equiv \frac{\bar{\beta}}{\phi}, \qquad \sigma \equiv \frac{\bar{\sigma}}{\phi}. \tag{A.6}$$

The term $-\beta A(t)^{\phi}$ 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 $A(\infty) = \lim_{t\to\infty} A(t)$ is a generalized gamma distribution by (15) and by the derivations above:

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

with

$$\begin{aligned} \hat{\theta} &= \bar{\theta}^{1/\phi} = [\bar{\sigma}^2/(2\bar{\beta})]^{1/\phi} = [\phi\sigma^2/(2\beta)]^{1/\phi} > 0, \\ \kappa &= 2\bar{\alpha}/\bar{\sigma}^2 - 1 = [2\alpha/\sigma^2 - 1]/\phi > 0 \end{aligned} \tag{A.7}$$

by (A.4) and (A.6).

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 absolute 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 (16) of the generalized logistic diffusion in Lemma 1 is

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

for real parameters η, ϕ, σ. That parametrization can be related back to the parameters in (A.5) 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 (16) 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).

B Trend Identification: Proof of Proposition 1

First, consider a random variable X which 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 $\mathbb{E}[\ln \hat{A}_{is}(t) \mid s, t]$ gives

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

as stated in the text.

C Pearson-Wong Process: Proof of Proposition 2

For a random variable X with a standard logistic diffusion (the $\phi = 1$ case), the Bernoulli transformation $1/X$ maps the diffusion into the Pearson-Wong family (see e.g. Prajneshu 1980, Dennis 1989). 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 (21). 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.[39]

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.

[39] 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 $Ah = \lambda h$.

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:

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)\,\middle|\,\hat{B}_{is}(t)\right] = \exp\left\{\lambda_s t\right\} h(\hat{B}_{is}(t)). \tag{C.8}$$

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

$$\pi_{n,m} = \frac{b_{m+1}}{a_m - a_n}\pi_{n,m+1} + \frac{k_{m+2}}{a_m - a_n}\pi_{n,m+2} \tag{C.9}$$

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.8) implies a recursive condition for these conditional moments:

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

We are guaranteed that these moments exist if we restrict ourselves to the first $N < (1+1/a)/2$ moments.

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.9). These coefficients correspond to those reported in equation (22).

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

$$\Pi \cdot \mathbb{E}\left[Y_{is}(t+\Delta)\,\middle|\,\hat{B}_{is}(t)\right] = \Lambda(\Delta)\cdot\Pi\cdot Y_{is}(t) \tag{C.10}$$

with $Y_{is}(t) \equiv (1, \hat{B}_{is}(t), \ldots, \hat{B}_{is}(t)^M)'$ and the matrices $\Lambda(t) = \mathrm{diag}(e^{-a_1 t}, e^{-a_2 t}, \ldots, e^{-a_M t})$ and $\Pi = (\pi_1, \pi_2, \ldots, \pi_M)'$, where $\pi_m \equiv (\pi_{m,0}, \ldots, \pi_{m,m}, 0, \ldots, 0)'$ 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 $\Pi \cdot Y_{is}(t)$ instead of the forecast errors for $Y_{is}(t)$. Either estimator is numerically equivalent since the matrix Π is triangular by construction, and therefore invertible.

D Connection to Endogenous Growth Theory

Eaton and Kortum (1999, 2010) provide a stochastic foundation for Fréchet distributed productivity. Their fundamental unit of analysis is an idea for a new variety. An idea is a blueprint to produce a variety of good i with efficiency $\hat{q}$ (in a source country s). Efficiency is the amount of output that can be produced with a unit of input when the idea is realized, and this efficiency is common to all countries where the variety based on the idea is manufactured. Suppose an idea's efficiency $\hat{q}$ is the realization of a random variable $\hat{Q}$ drawn independently from a Pareto distribution with shape parameter θ_i and location parameter (lower bound) $\underline{\hat{q}}$.[40] Suppose further that ideas for good i arrive in continuous time at moment t according to a (non-homogeneous) Poisson process with a time-dependent rate parameter normalized to $\underline{\hat{q}}^{-\theta_i} R_{is}(t)$. In Eaton and Kortum (2010, ch. 4), the rate parameter is a deterministic function of continuous time. In future empirical implementation, we can also specify a stochastic process for the rate parameter, giving rise to a Cox process for idea generation.

In this setup, the arrival rate of ideas with an efficiency of at least $\hat{q}$ ($\hat{Q} \geq \hat{q}$) is $\hat{q}^{-\theta_i} R_{is}(t)$. If there is no forgetting, then the measure of ideas $T_{is}(t)$ expands continuously and, at a moment t, it will have reached a level

$$T_{is}(t) = \int_{-\infty}^{t} R(\tau)\, \mathrm{d}\tau.$$

As a consequence, at moment t the number of ideas about good i with efficiency $\hat{Q} \geq \hat{q}$ is distributed Poisson with parameter $\hat{q}^{-\theta_i} T_{is}(t)$. Moreover, the productivity $q = \max\{\hat{q}\}$ of the most efficient idea at moment t has an extreme value Fréchet distribution with the cumulative distribution function $F_Q(q; T_{is}(t), \theta_i) = \exp\{-T_{is}(t)\, q^{-\theta_i}\}$, where $T_{is}(t) = \underline{q}_{is}(t)^{\theta_i}$ (Eaton and Kortum 2010, ch. 4). In Section 2 we suppressed time dependency of $\underline{q}_{is}$ to simplify notation.

Similar to Grossman and Helpman (1991), we can specify a basic differential equation for the generation of new ideas:

$$\mathrm{d}T_{is}(t) = R_{is}(t) = \xi_{is}(t) \lambda_{is}(t)^{\chi} L_{is}(t), \tag{D.11}$$

where $\xi_{is}(t)$ is research productivity in country-industry is, including the efficiency of exploration in the non-manufacturing sector and the efficiency of innovation in manufacturing, $\lambda_{is}(t) = L^R_{is}(t)/L_{is}(t)$ is the fraction of employment in country-industry is devoted to research (exploration or innovation), the parameter $\chi \in (0,1)$ reflects diminishing returns to scale (whereas $chi = 1$ in Grossman and Helpman 1991) and $L_{is}(t)$ is total employment in country-industry is at moment t.

The economic value of an idea in source country s is the expected profit $\pi_{is}(t)$ from its global expected sales in industry i. Given the independence of efficiency draws, the expected profit $\pi_{is}(t)$ is equal to the total profit $\Pi_s(t)$ generated in source country s's industry i relative to the current measure of ideas $T_{is}(t)$:

$$\pi_{is}(t) = \frac{\Pi_s(t)}{T_{is}(t)} = \frac{\delta_i X_{is}(t)}{T_{is}(t)} = \frac{\delta_i}{1-\delta_i} \frac{w_s(t) L^P_{is}(t)}{T_{is}(t)},$$

where $X_{is}(t) \equiv \sum_d X_{isd}(t)$ are global sales (exports $\sum_{d' \neq s} X_{isd'}$ plus domestic sales X_{iss}) and δ_i is the

[40] The Pareto CDF is $1 - (\hat{q}/\underline{\hat{q}}_{is})^{-\theta_i}$. Eaton and Kortum (1999) speak of the "quality of an idea" when they refer to its efficiency.

fraction of industry-wide profits in industry-wide sales (for a related derivation see Eaton and Kortum 2010, ch. 7). Industry-wide expected profits vary by the type of competition. Under monopolistic competition, a CES elasticity of substitution in demand σ_i and the Pareto shape parameter of efficiency θ_i imply $\delta_i = (\sigma_i - 1)/[\theta_i \sigma_i]$ (Eaton and Kortum 2010, ch. 5). The final step follows because the wage bill of labor employed in production must be equal to the sales not paid out as profits: $w_s(t) L^P_{is}(t) = (1 - \delta_i) X_{is}(t)$.

In equilibrium, the CES demand system implies a well defined price index $P_s(t)$ for the economy as a whole, so the real value of the idea at any future date τ is $\pi_{is}(\tau)/P_s(\tau)$ and, for a fixed interest rate r, the real net present value of the idea at moment t is

$$\frac{V_{is}(t)}{P_s(t)} = \int_t^{\infty} \exp\{-r(\tau - t)\} \frac{\pi_{is}(\tau)}{P_s(\tau)} \, \mathrm{d}\tau.$$

The exact price indexes in a multi-industry and multi-country equilibrium remain to be derived (a single-industry equilibrium is derived in Eaton and Kortum 2010, ch. 5 and 6). To illustrate the optimality condition driving endogenous growth, we can consider $V_{is}(t)$ as given but we note that it will be a function of $T_{is}(t)$ in general.

Each idea has a nominal value of $V_{is}(t)$, so the total value of research output is $\xi_{is}(t)\lambda_{is}(t)^{\chi} L_{is}(t) V_{is}(t)$ at moment t, and the marginal product of engaging an additional worker in research is $\chi \xi_{is}(t)\lambda_{is}(t)^{\chi-1} V_{is}(t)$. A labor market equilibrium with some research therefore requires that

$$\chi \xi_{is}(t)\lambda_{is}(t)^{\chi-1} V_{is}(t) = w_s \quad \Longleftrightarrow \quad \lambda_{is}(t) = \left(\frac{\chi \xi_{is}(t) V_{is}(t)}{w_s} \right)^{\frac{1}{1-\chi}}.$$

The exploration of new ideas in non-manufacturing and the innovation of products in manufacturing therefore follow the differential equation

$$\mathrm{d}T_{is}(t) = \xi_{is}(t)^{\frac{1}{1-\chi}} \left(\frac{\chi V_{is}(t)}{w_s} \right)^{\frac{\chi}{1-\chi}} L_{is}(t)$$

by (D.11). The nominal value of an idea $V_{is}(t)$ is a function of $T_{is}(t)$ in general, so this is a non-degenerate differential equation. Eaton and Kortum (2010, ch. 7) derive a balanced growth path for the economy in the single-industry case. By making research productivity $\xi_{is}(t)$ stochastic, we can generate a stochastic differential equation for the measure of ideas $T_{is}(t)$ and thus the Fréchet productivity position $\underline{q}_{is}(t) = T_{is}(t)^{1/\theta_i}$.

E Classifications and Additional Evidence

In this appendix, we report country and industry classifications, as well as additional evidence to complement the reported findings in the text.

E.1 Classifications

Our empirical analysis requires a time-invariant definition of less developed countries (LDC) and industrialized countries (non-LDC). Given our data time span of more then four decades (1962-2007), we classify the 90 economies, for which we obtain exporter capability estimates, by their relative status over the entire sample period.

In our classification, there are 28 *non-LDC*: Australia, Austria, Belgium-Luxembourg, Canada, China Hong Kong SAR, Denmark, Finland, France, Germany, Greece,Ireland, Israel, Italy, Japan, Kuwait, Netherlands, New Zealand, Norway, Oman, Portugal, Saudi Arabia, Singapore,Spain, Sweden, Switzerland, Trinidad and Tobago, United Kingdom, United States.

The remaining 62 countries are *LDC*: Algeria, Argentina, Bolivia, Brazil, Bulgaria, Cameroon, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Cuba, Czech Rep., Dominican Rep., Ecuador, Egypt, El Salvador, Ethiopia, Ghana, Guatemala, Honduras, Hungary, India, Indonesia, Iran, Jamaica, Jordan, Kenya, Lebanon, Libya, Madagascar, Malaysia, Mauritius, Mexico, Morocco, Myanmar, Nicaragua, Nigeria, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Korea Rep., Romania, Russian Federation, Senegal, South Africa, Sri Lanka, Syria, Taiwan, Thailand, Tunisia, Turkey, Uganda, United Rep. of Tanzania, Uruguay, Venezuela, Vietnam, Yugoslavia, Zambia.

We split the industries in our sample by broad sector. The manufacturing sector includes all industries with an SITC one-digit code between 5 and 8. The non-manufacturing merchandise sector includes industries in the agricultural sector as well industries in the mining and extraction sectors and spans the SITC one-digit codes from 0 to 4.

E.2 Additional evidence

Table A1 shows the top two products in terms of normalized log absolute advantage $\ln A_{ist}$ for 28 of the 90 exporting countries, using 1987 and 2007 as representative years. To obtain a measure of comparative advantage, we normalize log absolute advantage by its country mean: $\ln A_{ist} - (1/I)\sum_{i'}^{I} \ln A_{i'st}$. The country normalization of log absolute advantage $\ln A_{ist}$ results in a double log difference of export capability k_{ist}—a country's log deviation from the global industry mean in export capability minus its average log deviation across all industries.

Table A2 presents estimates of the decay equation (10) for the period 1984-2007 using data from the SITC revision 2 sample. This recent sample allows us to perform regressions for log absolute advantage and the log RCA index at the two, three and four-digit level. Estimated decay rates are comparable to those in **Table 1**, which uses data for the full period 1962-2007 at the level of 135 STIC three-digit industries.

Figures A1, A2 and **A3** extend **Figure 2** 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).

To verify that the graphed cross sectional distributions in **Figures A1, A2** and **A3** are not a byproduct of specification error in estimating export capabilities from the gravity model, we repeat the plots using the revealed comparative advantage index by Balassa (1965). **Figures A4, A5** and **A6** plot, for the same 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 revealed comparative advantage $(X_{is}/\sum_{s'} X_{is'})/(\sum_{i'} X_{i's}/\sum_{i'}\sum_{s'} X_{i's'})$ in year t against that comparative advantage level for industries i. The figures also graph the fit of the revealed comparative advantage index in the cross section to a log normal distribution using maximum likelihood separately for each country in each year.

Figures A7, A8 and **A9** extend **Figure 4** in the text and plot, for 28 countries in 1967, 1987 and 2007, the observed 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. This raw data plot is identical

Table A1: **Top Two Industries by Normalized Absolute Advantage**

Country	1987		2007		Country	1987		2007	
Argentina	Maize, unmilled	5.13	Maize, unmilled	5.50	Mexico	Sulphur	3.73	Alcoholic beverages	3.97
	Animal feed	3.88	Oil seed	4.61		Crude minerals	3.26	Office machines	3.82
Australia	Wool	4.15	Cheese & curd	3.25	Peru	Metal ores & conctr.	4.24	Metal ores & conctr.	6.25
	Jute	3.83	Fresh meat	3.20		Animal feed	4.03	Coffee	4.60
Brazil	Coffee	3.34	Iron ore	5.18	Philippines	Vegetable oils & fats	3.81	Office machines	4.41
	Iron ore	3.21	Fresh meat	4.42		Pres. fruits & nuts	3.50	Electric machinery	3.51
Canada	Sulphur	4.04	Wheat, unmilled	5.13	Poland	Barley, unmilled	5.68	Furniture	3.07
	Pulp & waste paper	3.36	Sulphur	3.31		Sulphur	3.35	Glassware	2.74
China	Explosives	7.05	Sound/video recorders	4.93	Korea Rep.	Radio receivers	5.51	Television receivers	6.06
	Jute	4.24	Radio receivers	4.65		Television receivers	5.37	Telecomm. equipment	5.11
Czech Rep.	Glassware	4.05	Glassware	4.17	Romania	Furniture	3.55	Footwear	3.49
	Prep. cereal & flour	3.68	Road vehicles	3.58		Fertilizers, manuf.	2.73	Silk	3.15
Egypt	Cotton	4.52	Fertilizers, crude	4.45	Russia	Pulp & waste paper	5.16	Animal oils & fats	8.32
	Textile yarn, fabrics	2.90	Rice	3.91		Radioactive material	5.02	Fertilizers, manuf.	4.54
France	Electric machinery	3.44	Oth. transport eqpmt.	3.31	South Africa	Stone, sand & gravel	3.92	Iron & steel	4.17
	Alcoholic beverages	3.39	Alcoholic beverages	3.15		Radioactive material	3.65	Fresh fruits & nuts	3.47
Germany	Road vehicles	3.95	Road vehicles	3.10	Taiwan	Explosives	4.41	Television receivers	5.18
	General machinery	3.89	Metalworking machinery	2.70		Footwear	4.39	Office machines	5.01
Hungary	Margarine	3.19	Telecomm. equipment	4.15	Thailand	Rice	4.81	Rice	4.92
	Fresh meat	2.76	Office machines	4.08		Fresh vegetables	4.08	Natural rubber	4.50
India	Tea	4.20	Precious stones	3.86	Turkey	Fresh vegetables	3.48	Glassware	3.30
	Leather	3.90	Rice	3.61		Tobacco unmanuf.	3.41	Textile yarn, fabrics	3.20
Indonesia	Natural rubber	5.10	Natural rubber	5.26	United States	Office machines	3.96	Oth. transport eqpmt.	3.46
	Improved wood	4.74	Sound/video recorders	4.90		Oth. transport eqpmt.	3.25	Photographic supplies	2.60
Japan	Sound/video recorders	6.28	Sound/video recorders	5.90	United Kingd.	Measuring instrmnts.	3.20	Alcoholic beverages	3.26
	Road vehicles	6.08	Road vehicles	5.63		Office machines	3.15	Pharmaceutical prod.	3.12
Malaysia	Natural rubber	6.19	Radio receivers	5.78	Vietnam	Cereal meals & flour	5.34	Animal oils & fats	10.31
	Vegetable oils & fats	4.85	Sound/video recorders	5.03		Jute	5.14	Footwear	7.02

Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007.
Note: Top two industries for 28 of the 90 countries in 1987 and 2007 in terms of normalized log absolute advantage, relative to the country mean: $\ln A_{ist} - (1/I)\sum_{i'}^{I} \ln A_{i'st}$.

Table A2: DECAY REGRESSIONS FOR COMPARATIVE ADVANTAGE AT VARYING LEVELS OF INDUSTRY AGGREGATION

	Exporter capability k			Balassa RCA $\ln \hat{X}$		
SITC aggregate	4 digit	3 digit	2 digit	4 digit	3 digit	2 digit
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full sample						
Decay rate ρ	-0.365	-0.241	-0.188	-0.323	-0.288	-0.256
	(0.034)**	(0.021)**	(0.020)**	(0.014)**	(0.014)**	(0.015)**
Dissipation rate η	0.119	0.122	0.125	0.113	0.130	0.157
	(0.009)**	(0.009)**	(0.012)**	(0.004)**	(0.005)**	(0.008)**
Innovation intensity σ^2	0.766	0.451	0.333	0.693	0.523	0.377
	(0.035)**	(0.012)**	(0.008)**	(0.013)**	(0.010)**	(0.008)**
Obs.	142,664	61,280	19,815	146,644	61,577	19,815
Adj. R^2	0.142	0.142	0.160	0.121	0.129	0.132
Panel B: LDC exporters						
Decay rate ρ	-0.503	-0.326	-0.236	-0.369	-0.336	-0.296
	(0.045)**	(0.032)**	(0.028)**	(0.017)**	(0.016)**	(0.016)**
Dissipation rate η	0.113	0.122	0.121	0.099	0.115	0.138
	(0.007)**	(0.010)**	(0.013)**	(0.004)**	(0.005)**	(0.007)**
Innovation intensity σ^2	1.235	0.646	0.446	0.927	0.711	0.508
	(0.086)**	(0.027)**	(0.016)**	(0.022)**	(0.016)**	(0.012)**
Obs.	79,325	37,918	13,167	81,963	38,095	13,167
Adj. R^2	0.168	0.156	0.165	0.132	0.136	0.137
Panel C: Non-manufacturing industries						
Decay rate ρ	-0.530	-0.309	-0.251	-0.334	-0.287	-0.257
	(0.046)**	(0.027)**	(0.026)**	(0.015)**	(0.015)**	(0.015)**
Dissipation rate η	0.095	0.111	0.146	0.086	0.114	0.157
	(0.005)**	(0.008)**	(0.013)**	(0.003)**	(0.005)**	(0.008)**
Innovation intensity σ^2	1.591	0.665	0.397	0.945	0.597	0.379
	(0.118)**	(0.024)**	(0.014)**	(0.019)**	(0.013)**	(0.009)**
Obs.	37,645	17,985	7,482	40,472	18,236	7,482
Adj. R^2	0.176	0.133	0.133	0.124	0.140	0.164

Source: WTF (Feenstra et al. 2005, updated through 2008) for two-digit (61 industries), three-digit (227 industries), and four-digit (684 industries) sector definitions from 1984-2007, and CEPII.org.

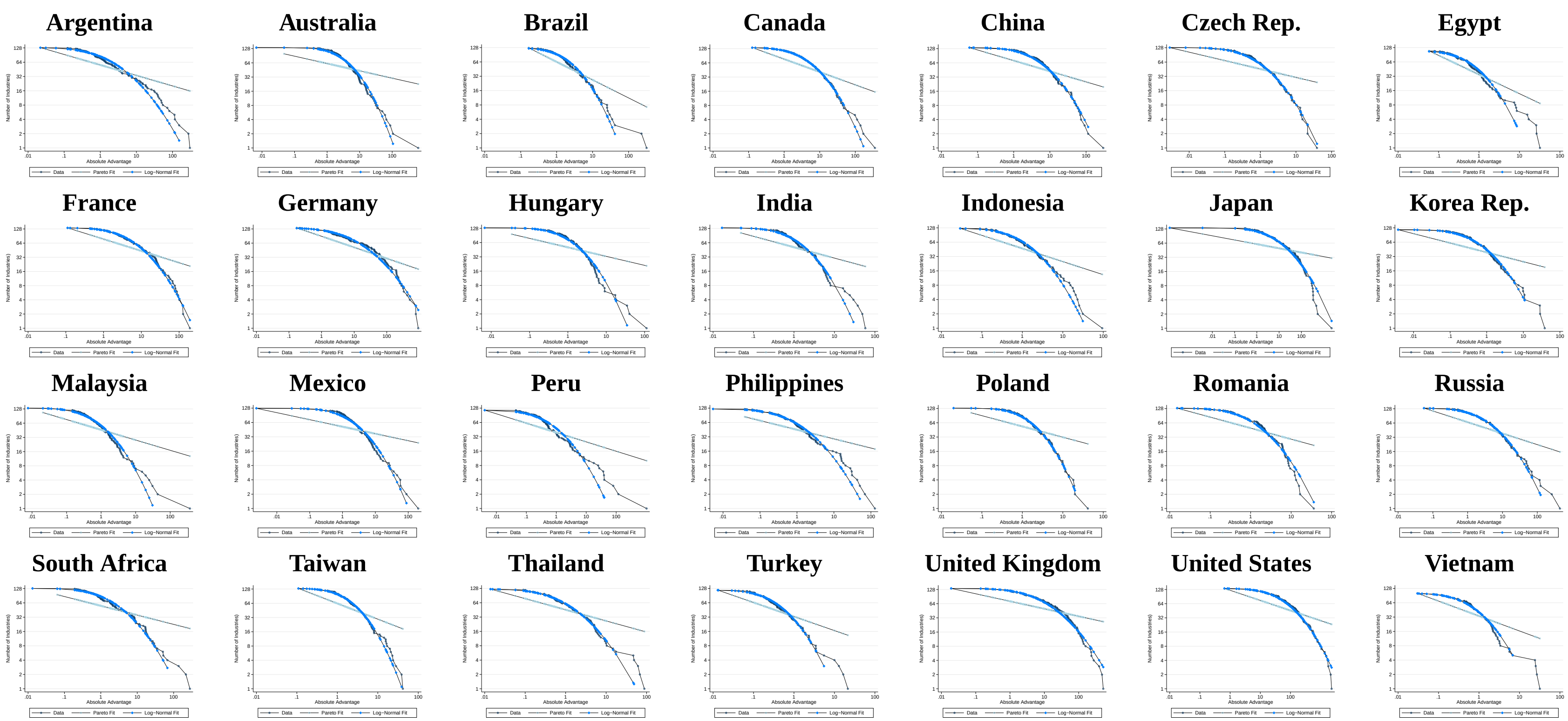
Note: Reported figures for five-year decadalized changes. Variables are OLS-estimated gravity measures of export capability k by (5) and the log Balassa index of revealed comparative advantage $\ln \hat{X}_{ist} = \ln(X_{ist}/\sum_{s'} X_{is't})/(\sum_{i'} X_{i'st}/\sum_{i'}\sum_{s'} X_{i's't})$. OLS estimation of the decadal decay rate ρ from

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

conditional on industry-year and source country-year effects δ_{it} and δ_{st} for the full pooled sample (panel A) and subsamples (panels B and C). The implied dissipation rate η and innovation intensity σ^2 are based on the decadal 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 Appendix E. Nonmanufacturing merchandise spans SITC sector codes 0-4. Standard errors (reported below coefficients) for ρ are clustered by country and for η and σ are calculated using the delta method; ** indicates significance at the 1% level.

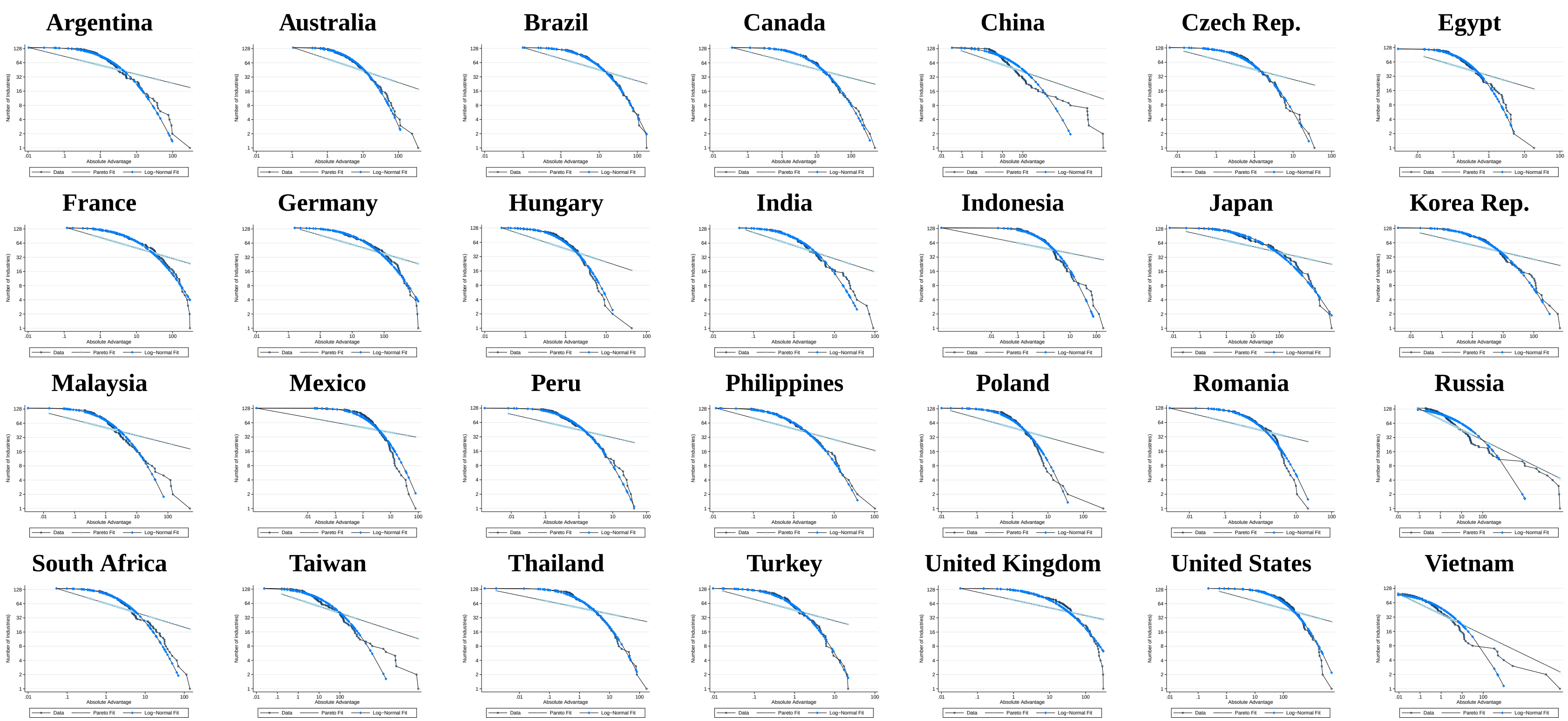
to that in **Figures A1** through **A3** and shown for the same 28 countries and years as before. In addition, **Figures A7** through **A9** now plot the implied stationary distribution based on the time series diffusion estimates in Table 2 for the full sample (column 1), using the estimates of the two shape relevant global diffusion parameters (η and ϕ), which determine the curvature of the implied single stationary distribution of comparative advantage $\hat{A}_{ist}$ (through κ and ϕ), and the recovered estimates of the unknown country-wide stochastic trends Z_{st}, which determine the horizontal position of the stationary distribution of observed absolute advantage A_{ist}.

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



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 for absolute advantage A_{ist} are based on maximum likelihood estimation by country s in year $t = 1967$.

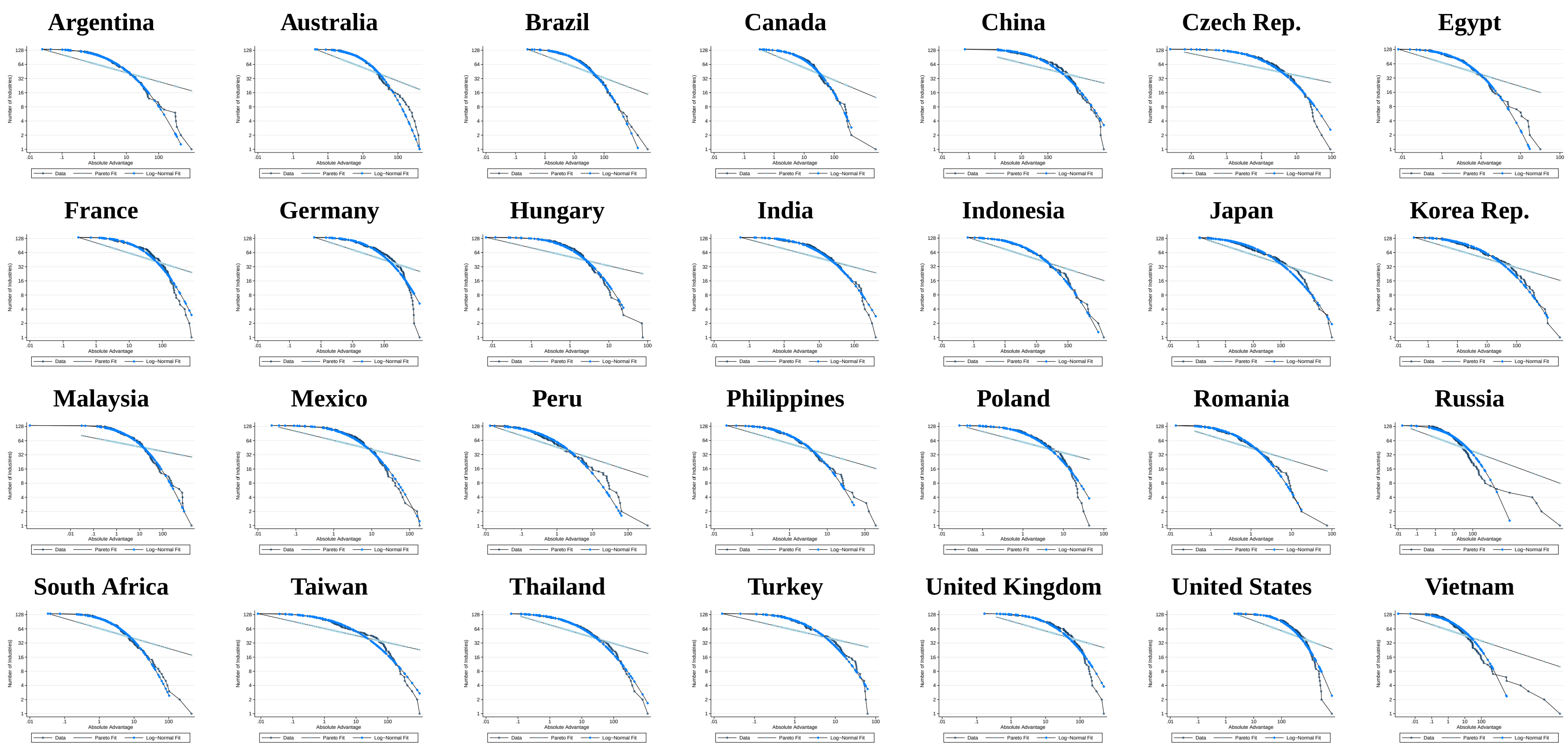
Figure A2: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 1987**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).

Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 for absolute advantage A_{ist} are based on maximum likelihood estimation by country s in year $t = 1987$.

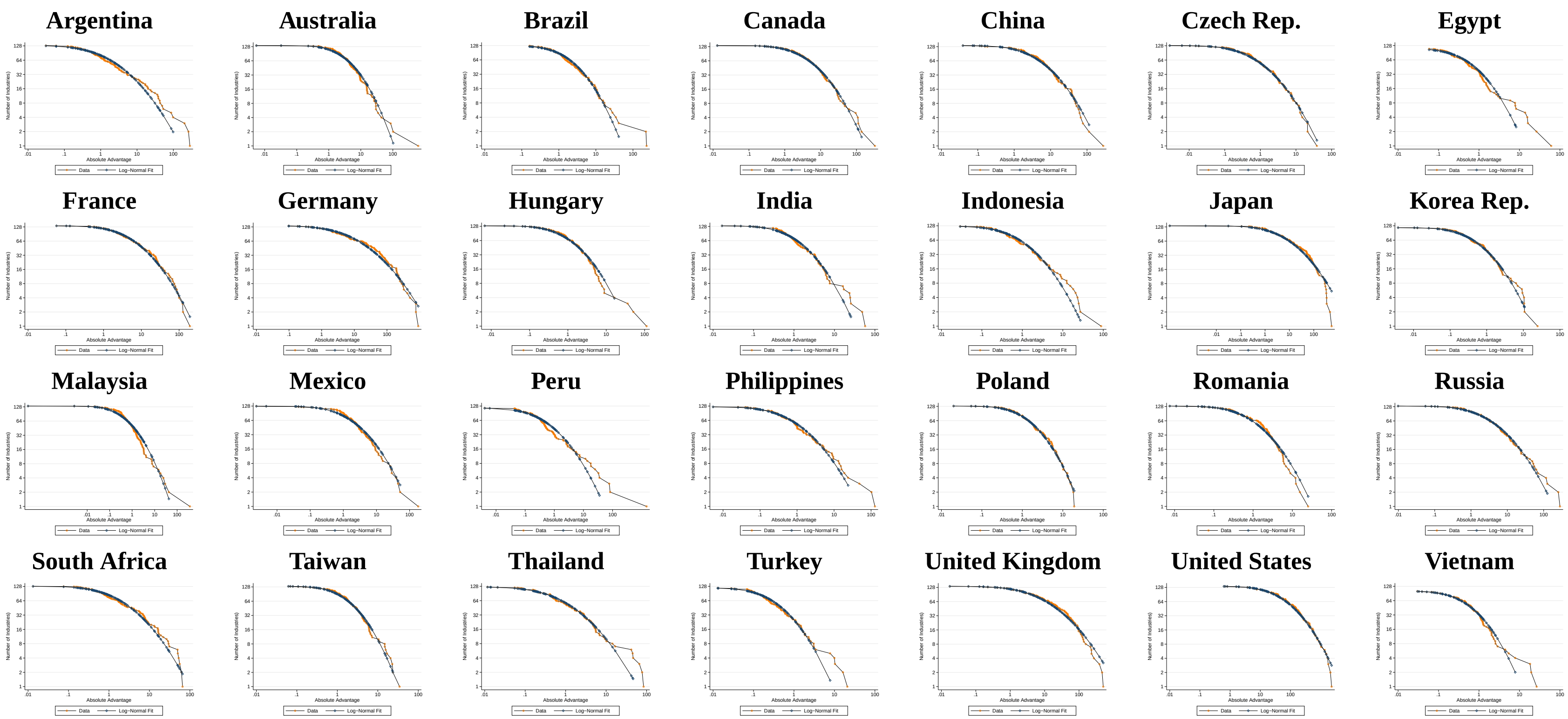
Figure A3: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 2007**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).

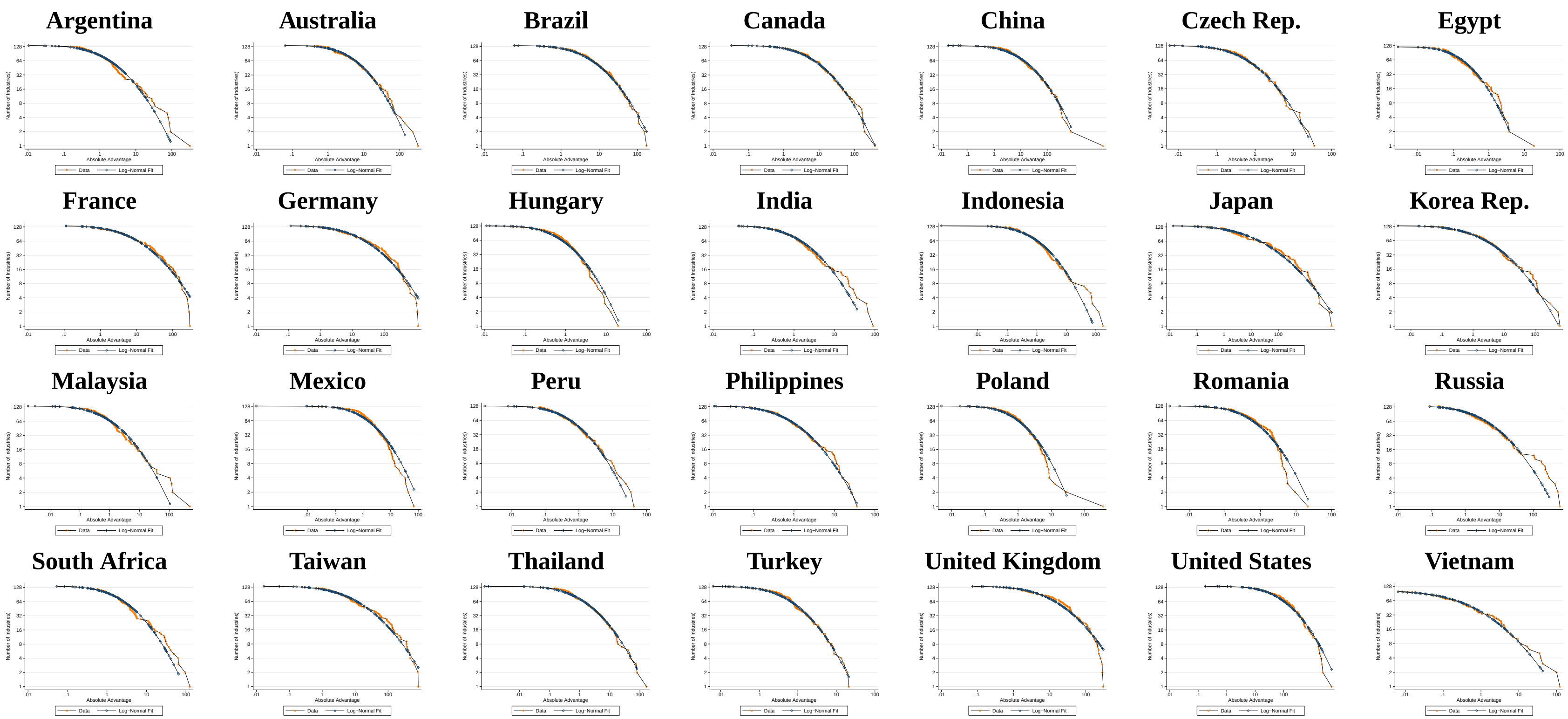
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions for absolute advantage A_{ist} are based on maximum likelihood estimation by country s in year $t = 2007$.

Figure A4: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 1967**



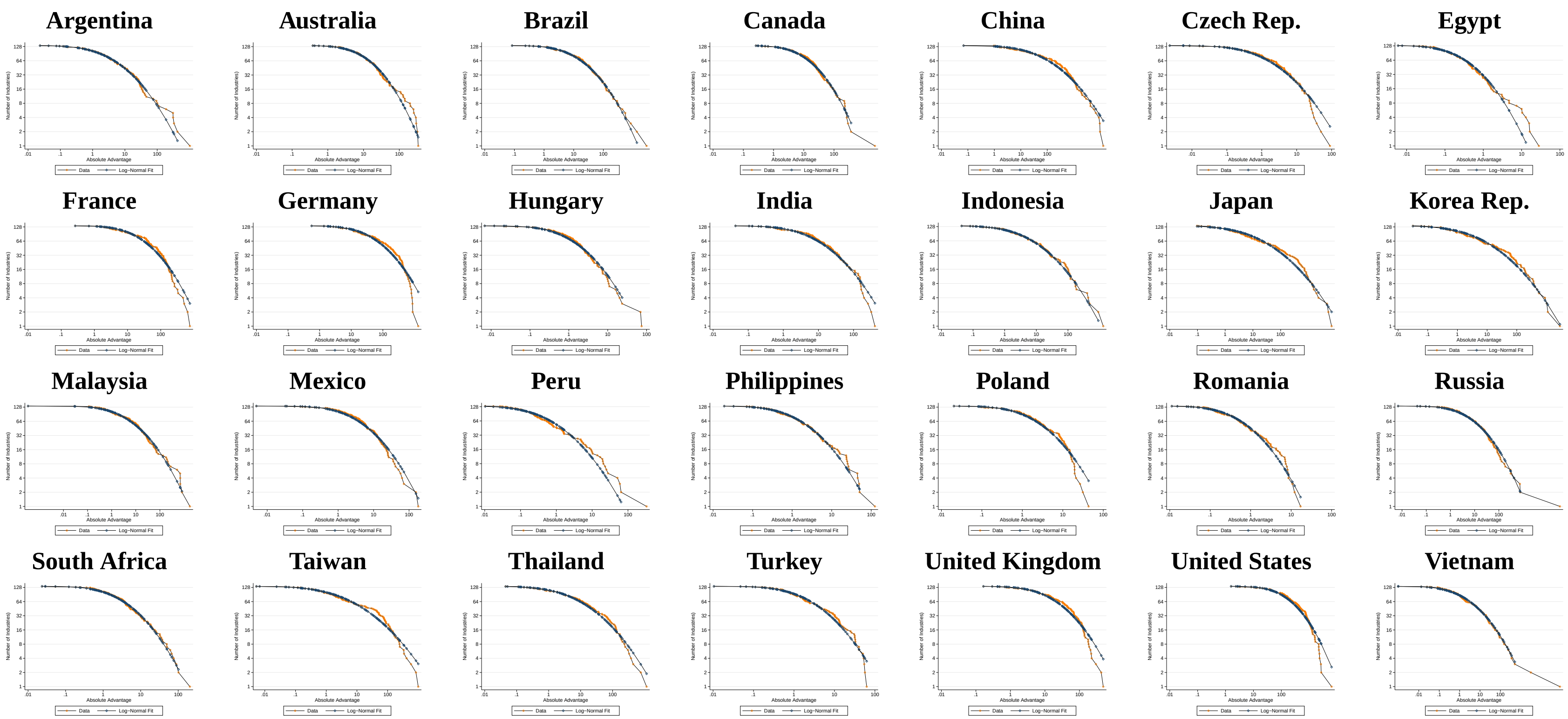
Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007.
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_{\hat{X}}(\hat{x})$ times the total number of industries $I = 135$) on the vertical axis plotted against the Balassa index of revealed comparative advantage $\hat{X} = (X_{is}/\sum_{s'} X_{is'})/(\sum_{i'} X_{i's}/\sum_{i'}\sum_{s'} X_{i's'})$ on the horizontal axis. Both axes have a log scale. The fitted log normal distribution is based on maximum likelihood estimation by country s in year $t = 1967$.

Figure A5: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 1987**



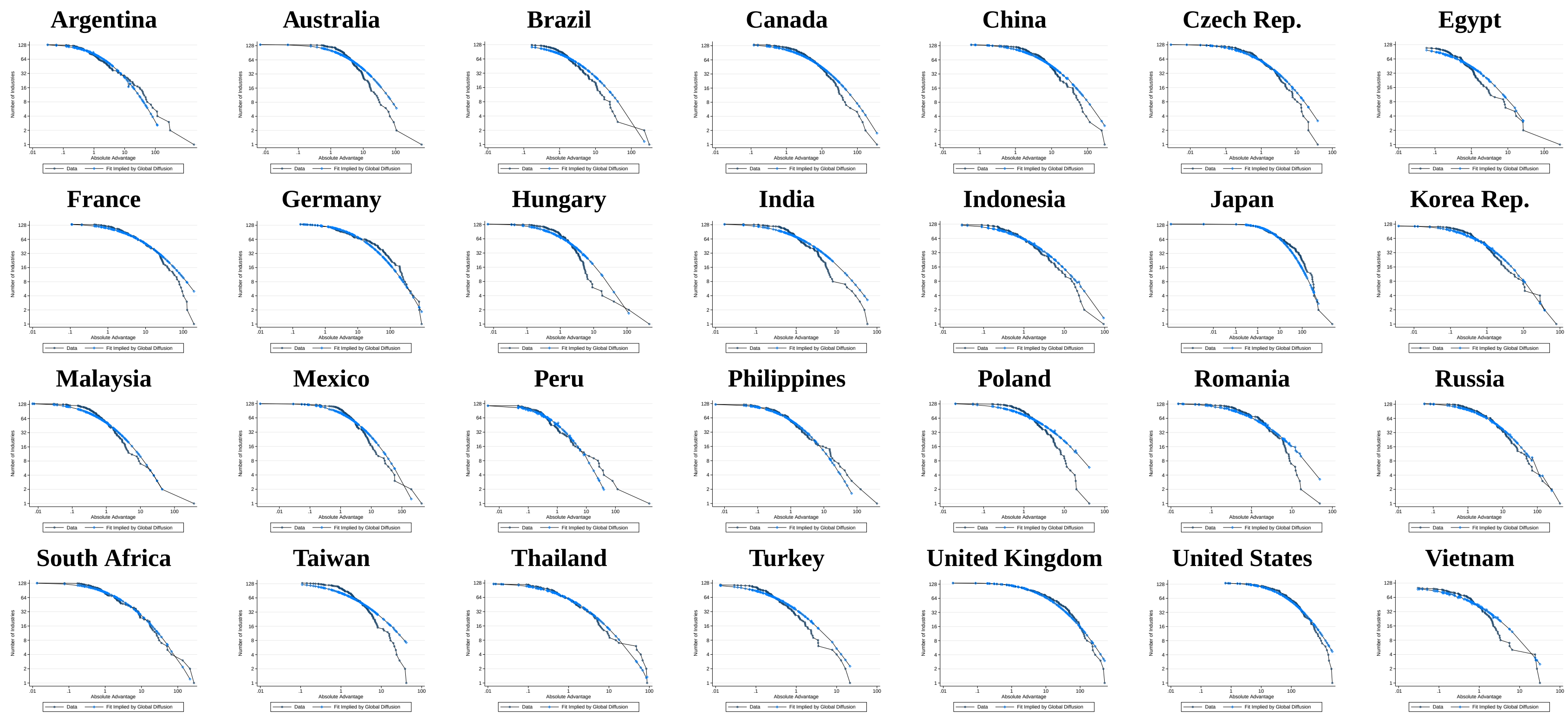
Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007.
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_{\hat{X}}(\hat{x})$ times the total number of industries $I = 135$) on the vertical axis plotted against the Balassa index of revealed comparative advantage $\hat{X} = (X_{is}/\sum_{s'} X_{is'})/(\sum_{i'} X_{i's}/\sum_{i'}\sum_{s'} X_{i's'})$ on the horizontal axis. Both axes have a log scale. The fitted log normal distribution is based on maximum likelihood estimation by country s in year $t = 1987$.

Figure A6: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 2007**



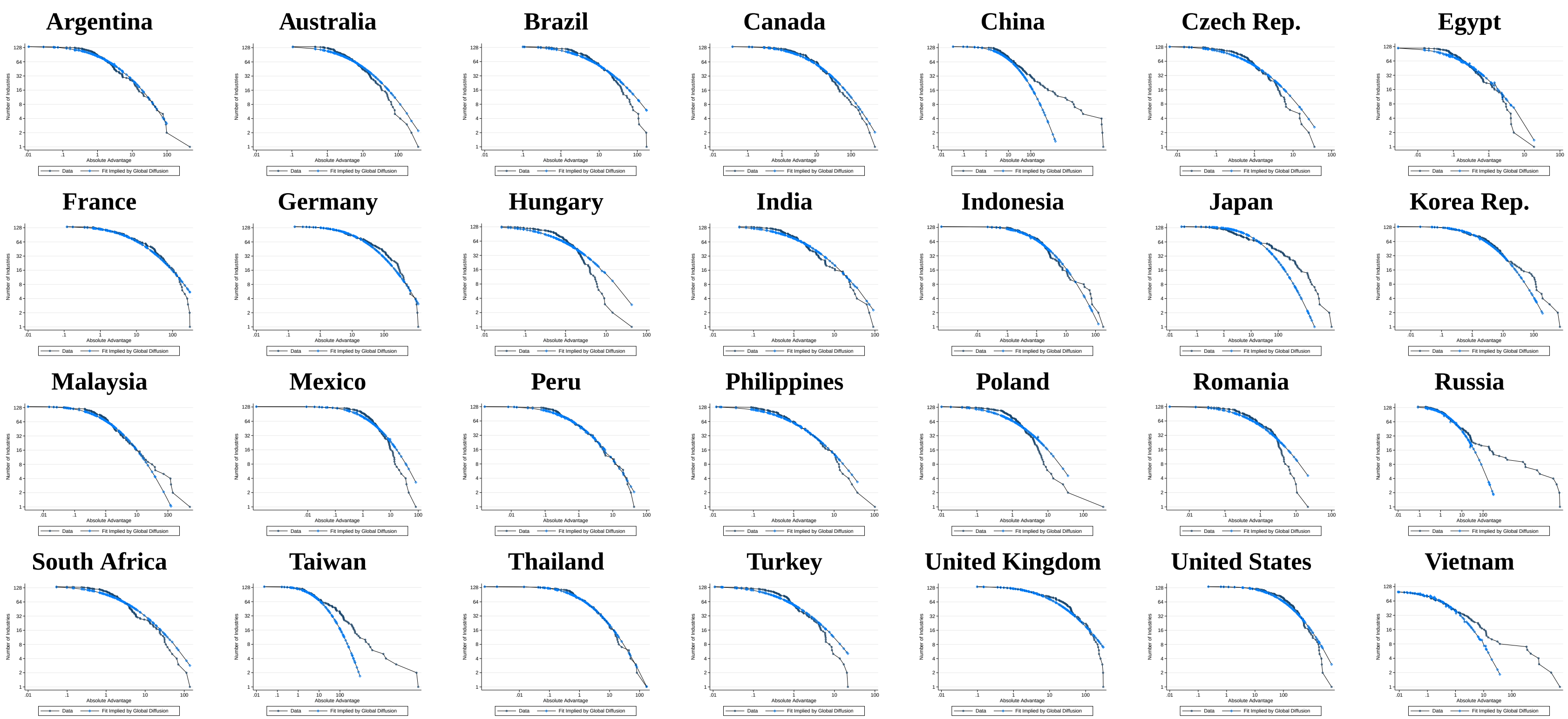
Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007.
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_{\hat{X}}(\hat{x})$ times the total number of industries $I = 135$) on the vertical axis plotted against the Balassa index of revealed comparative advantage $\hat{X} = (X_{is}/\sum_{s'} X_{is'})/(\sum_{i'} X_{i's}/\sum_{i'}\sum_{s'} X_{i's'})$ on the horizontal axis. Both axes have a log scale. The fitted log normal distribution is based on maximum likelihood estimation by country s in year $t = 2007$.

Figure A7: **Global Diffusion Implied and Observed Cumulative Probability Distributions of Absolute Advantage in 1967**



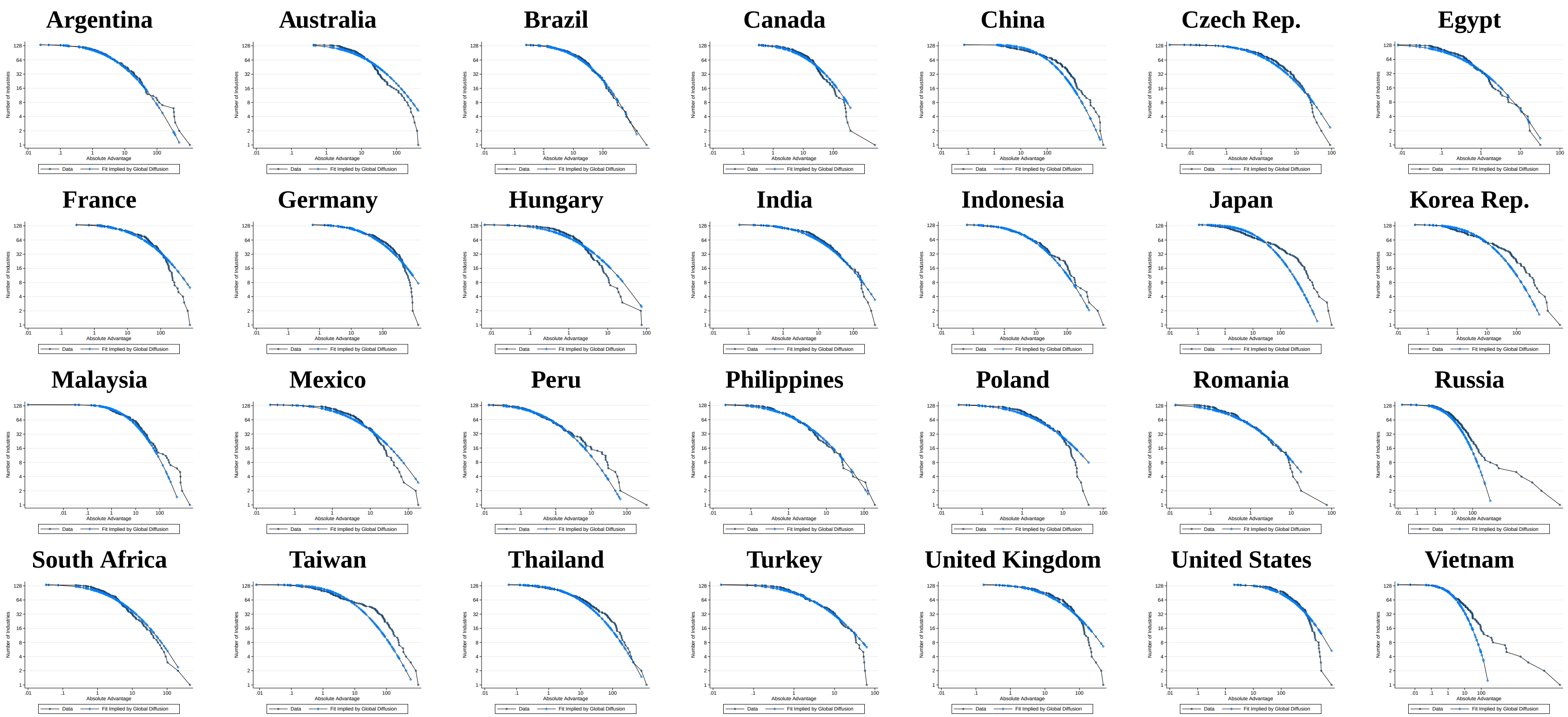
Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 $t = 1967$. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

Figure A8: **Global Diffusion Implied and Observed Cumulative Probability Distributions of Absolute Advantage in 1987**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

Figure A9: **Global Diffusion Implied and Observed Cumulative Probability Distributions of Absolute Advantage in 2007**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).

Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) 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 $t = 2007$. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

References

Abowd, John M., Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," *LEHD Technical Paper*, March 2002, *2002-06.*

Aït-Sahalia, Yacine, Lars P. Hansen, and José A. Scheinkman, "Operator Methods for Continuous-Time Markov Processes," in Yacine Aït-Sahalia and Lars P. Hansen, eds., *Handbook of Financial Econometrics*, Vol. 1 of *Handbooks in Finance*, Amsterdam: North-Holland/Elsevier, 2010, chapter 1, pp. 1–66.

Amoroso, Luigi, "Ricerche Intorno alla Curva dei Redditi," *Annali di Matematica Pura ed Applicata*, December 1925, *2* (1), 123–159.

Anderson, James E., "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, March 1979, *69* (1), 106–16.

_ , "The Gravity Model," *Annual Review of Economics*, 2011, *3* (1), 133–60.

_ **and Eric van Wincoop**, "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, March 2003, *93* (1), 170–92.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

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

Axtell, Robert L., "Zipf Distribution of U.S. Firm Sizes," *Science*, September 2001, *293* (5536), 1818–20.

Balassa, Bela, "Trade Liberalization and Revealed Comparative Advantage," *Manchester School of Economic and Social Studies*, May 1965, *33*, 99–123.

Bhagwati, Jagdish, "Free Trade: Old and New Challenges," *Economic Journal*, March 1994, *104* (423), 231–46.

Bibby, Bo M., Martin Jacobsen, and Michael Sørensen, "Estimating Functions for Discretely Sampled Diffusion-Type Models," in Yacine Aït-Sahalia and Lars P. Hansen, eds., *Handbook of Financial Econometrics*, Vol. 1 of *Handbooks in Finance*, Amsterdam: North-Holland/Elsevier, 2010, chapter 4, pp. 203–268.

Bombardini, Matilde, Giovanni Gallipoli, and German Pupato, "Skill Dispersion and Trade Flows," *American Economic Review*, August 2012, *102* (5), 2327–48.

Cabral, Luís M. B. and José Mata, "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, September 2003, *93* (4), 1075–1090.

Cadot, Olivier, Celine Carrére, and Vanessa Strauss Kahn, "Export Diversification: What's behind the Hump?," *Review of Economics and Statistics*, May 2011, *93* (2), 590–605.

Chan, K. C., G. Andrew Karolyi, Francis A. Longstaff, and Anthony B. Sanders, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, July 1992, *47* (3), 1209–27.

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

Costinot, Arnaud, "On the Origins of Comparative Advantage," *Journal of International Economics*, April 2009, *77* (2), 255–64.

__ , Dave Donaldson, and Ivana Komunjer, "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, April 2012, *79* (2), 581–608.

Crooks, Gavin E., "The Amoroso Distribution," *ArXiv*, May 2010, *1005* (3274), 1–27.

Cuñat, Alejandro and Marc J. Melitz, "Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage," *Journal of the European Economic Association*, April 2012, *10* (2), 225–54.

Davis, Donald R. and David E. Weinstein, "An Account of Global Factor Trade," *American Economic Review*, December 2001, *91* (5), 1423–53.

Deardorff, Alan V., "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?," in Jeffrey A. Frankel, ed., *The Regionalization of the World Economy*, Chicago: University of Chicago Press, January 1998, chapter 1, pp. 7–32.

Dennis, Brian, "Stochastic Differential Equations As Insect Population Models," in Lyman L. McDonald, Bryan F. J. Manly, Jeffrey A. Lockwood, and Jesse A. Logan, eds., *Estimation and Analysis of Insect Populations*, Vol. 55 of *Lecture Notes in Statistics*, New York: Springer, 1989, chapter 4, pp. 219–238. Proceedings of a Conference held in Laramie, Wyoming, January 25-29, 1988.

Easterly, William and Ariell Reshef, "African Export Successes: Surprises, Stylized Facts, and Explanations," *NBER Working Paper*, 2010, *16597.*

Eaton, Jonathan and Samuel Kortum, "International Technology Diffusion: Theory and Measurement," *International Economic Review*, August 1999, *40* (3), 537–70.

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

__ and __ , "Technology in the Global Economy: A Framework for Quantitative Analysis," March 2010. University of Chicago, unpublished manuscript.

__ , Samuel S. Kortum, and Sebastian Sotelo, "International Trade: Linking Micro and Macro," *NBER Working Paper*, 2012, *17864.*

Fadinger, Harald and Pablo Fleiss, "Trade and Sectoral Productivity," *Economic Journal*, September 2011, *121* (555), 958–89.

Fally, Thibault, "Structural Gravity and Fixed Effects," July 2012. University of California, Berkeley, unpublished manuscript.

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo, "World Trade Flows: 1962-2000," *NBER Working Paper*, January 2005, *11040.*

Finicelli, Andrea, Patrizio Pagano, and Massimo Sbracia, "Trade-Revealed TFP," *Bank of Italy Temi di Discussione (Working Paper)*, October 2009, *729.*

__ , __ , and __ , "Ricardian Selection," *Journal of International Economics*, January 2013, *89* (1), 96–109.

Forman, Julie L. and Michael Sørensen, "The Pearson Diffusions: A Class of Statistically Tractable Diffusion Processes," *Scandinavian Journal of Statistics*, September 2008, *35* (3), 438–465.

_ **and** _ , "A Transformation Approach to Modelling Multi-modal Diffusions," *Journal of Statistical Planning and Inference*, 2013, p. forthcoming (available online 7 October 2013).

Freund, Caroline L. and Martha D. Pierola, "Export Superstars," September 2013. The World Bank, unpublished manuscript.

Grossman, Gene M. and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, January 1991, *58* (1), 43–61.

Hansen, Lars P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, July 1982, *50* (4), 1029–54.

Hanson, Gordon H., "The Rise of Middle Kingdoms: Emerging Economies in Global Trade," *Journal of Economic Perspectives*, Spring 2012, *26* (2), 41–64.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik, "What You Export Matters," *Journal of Economic Growth*, March 2007, *12* (1), 1–25.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, May 2008, *123* (2), 441–87.

Imbs, Jean and Romain Wacziarg, "Stages of Diversification," *American Economic Review*, March 2003, *93* (1), 63–86.

Karlin, Samuel and Howard M. Taylor, *A second course in stochastic processes*, 1st ed., New York: Academic Press, 1981.

Kerr, William R., "Heterogeneous Technology Diffusion and Ricardian Trade Patterns," *NBER Working Paper*, November 2013, *19657.*

Kessler, Mathieu and Michael Sørensen, "Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process," *Bernoulli*, April 1999, *5* (2), 299–314.

Kotz, Samuel, Norman L. Johnson, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed., Vol. 1 of *Wiley Series in Probability and Statistics*, New York: Wiley, October 1994.

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

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

Levchenko, Andrei A., "Institutional Quality and International Trade," *Review of Economic Studies*, July 2007, *74* (3), 791–819.

_ **and Jing Zhang**, "The Evolution of Comparative Advantage: Measurement and Welfare Implications," January 2013. University of Michigan, unpublished manuscript.

Lunde, Asger and Anne F. Brix, “Estimating Stochastic Volatility Models using Prediction-based Estimating Functions,” *CREATES Research Paper*, July 2013, *23.*

Luttmer, Erzo G. J., “Selection, Growth, and the Size Distribution of Firms,” *Quarterly Journal of Economics*, August 2007, *122* (3), 1103–44.

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

Pearson, Karl, “Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material,” *Philosophical Transactions of the Royal Society of London. A*, November 1895, *186*, 343–414.

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

Quah, Danny T., “Convergence Empirics across Economies with (Some) Capital Mobility,” *Journal of Economic Growth*, March 1996, *1* (1), 95–124.

Romalis, John, “Factor Proportions and the Structure of Commodity Trade,” *American Economic Review*, March 2004, *94* (1), 67–97.

Schott, Peter K., “One Size Fits All? Heckscher-Ohlin Specialization in Global Production,” *American Economic Review*, June 2003, *93* (3), 686–708.

Shikher, Serge, “Capital, Technology, and Specialization in the Neoclassical Model,” *Journal of International Economics*, March 2011, *83* (2), 229–42.

_ , “Putting Industries into the Eaton-Kortum Model,” *Journal of International Trade and Economic Development*, December 2012, *21* (6), 807–37.

Silva, João M. C. Santos and Silvana Tenreyro, “The Log of Gravity,” *Review of Economics and Statistics*, November 2006, *88* (4), 641–58.

Sørensen, Michael, “Statistical Inference for Integrated Diffusion Processes,” in International Statistical Institute, ed., *Proceedings of the 58th World Statistics Congress 2011*, Dublin 2011, pp. 75–82.

Stacy, E. W., “A Generalization of the Gamma Distribution,” *Annals of Mathematical Statistics*, September 1962, *33* (3), 1187–1192.

Sutton, John, “Gibrat’s Legacy,” *Journal of Economic Literature*, March 1997, *35* (1), 40–59.

Trefler, Daniel, “The Case of the Missing Trade and Other Mysteries,” *American Economic Review*, December 1995, *85* (5), 1029–46.

Vasicek, Oldrich A., “An Equilibrium Characterization of the Term Structure,” *Journal of Financial Economics*, November 1977, *5* (2), 177–88.

Waugh, Michael E., “International Trade and Income Differences,” *American Economic Review*, December 2010, *100* (5), 2093–2124.

Wong, Eugene, “The Construction of a Class of Stationary Markoff Processes,” in Richard Bellman, ed., *Stochastic processes in mathematical physics and engineering*, Vol. 16 of *Proceedings of symposia in applied mathematics*, Providence, R.I.: American Mathematical Society, 1964, chapter 12, pp. 264–276.