The Dynamics of Comparative Advantage: Hyperspecialization and Evanescence*

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

November 30, 2013

Abstract

We explore the long run evolution of comparative advantage and employ the gravity model of trade to extract a measure of export capability, purified of geographic and demand side confounds, for 135 industries in 90 countries from 1962 to 2007. We use the resulting measure of a country-industry's export capability by year to document two striking empirical regularities in comparative advantage. One is hyperspecialization in exporting: In the typical country, export success is highly concentrated in a handful of industries. Hyperspecialization is consistent with a heavy upper tail in the distribution of export capabilities across industries within a country, which we find is well approximated by a generalized gamma distribution whose shape remains relatively stable over time. The second empirical regularity is a high rate of turnover in a country's top export industries. The evanescence of top exports reflects a high rate of decay in a typical country's export capability, which we estimate to be on the order of 35% to 55% per decade. To reconcile persistent hyperspecialization in exporting with evanescence in export capability, we specify a generalized logistic diffusion for comparative advantage which has a generalized gamma as a stationary distribution. Our results provide an empirical roadmap for dynamic theoretical models of the determinants of comparative advantage.

*We thank Ina Jäkle and Heea Jung for excellent research assistance.
†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 is making 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 return 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 an expanded 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). At least superficially, exports by emerging economies—with China and Mexico specializing in labor-intensive manufactures, Brazil and Indonesia in agricultural commodities, Nigeria and Russia in oil and gas, and Peru and South Africa in minerals—suggest that resource and technology differences between countries contribute to driving trade.

In this paper, we explore the evolution of comparative advantage over the last four and a half decades. Using the gravity model of trade, we extract a measure of countries' industry export capabilities which we then use to evaluate how comparative advantage 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 comparative advantage. Instead, we rely on trade data only, which allows us to impose minimal theoretical structure on the determinants of trade, examine industries at a fine degree of disaggregation, and include both manufacturing and non-manufacturing sectors in our analysis. These features turn out to be extremely helpful in identifying the stable and heretofore unknown 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 can be decomposed into three components (Anderson 2011): an *exporter-industry component*, which captures the exporting country's average export capability in the industry; an *importer-industry component*, which captures the importing country's effective demand for goods in the industry; and an *exporter-importer component*, which captures bilateral trade costs between the 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 component is the product of a country's over-

[1] Dornbusch et al. (1977), on which EK builds, generalizes Ricardian trade theory to many industries but does not easily extend beyond two countries (Collins 1985). Shikher (2011, 2012) expands 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.

all efficiency in producing goods and its unit production costs. In the Krugman (1980), Heckscher-Ohlin (Deardorff 1998), Melitz (2003), Anderson and van Wincoop (2003) models, which also yield gravity specifications, the form of the exporter-industry component varies 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 country's absolute advantage in an industry, and by estimating a country-specific normalization in addition we obtain a measure of a country-industry's 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 confront. Though we motivate our approach using EK, we remain agnostic about the origins of a country's export prowess. 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 searching for the covariates of export capability, as in Chor (2010), we seek the stable features of its distribution across countries, industries, and time.

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 deep underlying patterns in comparative 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 top good (out of 135 total) in a country's total exports is 21%, for the top 3 goods (top 2% of export goods) is 45%, and for the top 7 goods (top 5% of export goods) is 64%. Consistent with this strong concentration, the cross-industry distribution of absolute advantage in a given year appears to be roughly log normal, reminiscent of the distributions of firm size (e.g., Cabral and Mata 2003, Luttmer 2007) and city size (e.g., Eeckhout 2004, Gabaix and Ioannides 2004). Strikingly, this approximation applies to countries specializing in discrepant types of goods and at diverse stages of their economic development.

Stability in the shape of the distribution of absolute advantage makes the second empirical regularity regarding exports all the more surprising: There is a high rate of 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 previously. Such evanescence is consistent with fast mean reversion in export superiority, which we confirm by regressing the change in a country-industry's export capability on its initial value, obtaining decadal decay rates on the order of 33% to 57%. These regressions

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

control for industry-time and country-time fixed effects, so they are properly interpreted as summarizing decay rates in 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 consequence that the dominance of a country in its top export products may be short lived.

In the second stage of our analysis, we seek to reconcile hyperspecialization in exports with fast mean reversion in comparative advantage by modelling absolute advantage as a stochastic process. We specify a generalized logistic diffusion for absolute advantage that allows for Brownian innovations (thus accounting for surges in a country's relative export prowess), for a country-wide stochastic trend (thus flexibly transforming absolute into comparative advantage), as well as for deterministic mean reversion in absolute advantage (thus rendering surges impermanent). The generalized logistic diffusion has the generalized gamma as a stationary distribution (Kotz et al. 1994).[4] The generalized gamma unifies the gamma and extreme-value families (Crooks 2010) and therefore flexibly nests many common distributions (including the log normal, exponential, Pareto, Fréchet, and Weibull). To gauge the fit of the model, we take the three global parameters estimated from the time series and project the cross-sectional distribution of absolute advantage for each country in each year. Based on just these three parameters (and controlling for the country-wide stochastic trend), the simulated values match the cross-sectional distributions, country-by-country and period-by-period, with remarkable 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 evanescent, with the identity of top industries churning steadily.

We then allow model parameters to vary by groups of countries and by broad industry. The three parameters of the generalized gamma govern the rate at which the process reverts to the global long-run mean of comparative advantage (the dissipation of comparative advantage), the rate at which industries are reshuffled within the distribution (the intensity of innovations in comparative advantage), and the degree of asymmetry in mean reversion from above versus below the mean (the stickiness of comparative advantage). We examine how these parameters differ between developed versus developing economies and for manufacturing versus nonmanufacturing industries. The differences we uncover are informative about possible origins of comparative advantage.

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 data for a single year, Chor (2010) explores whether the interaction of industry factor intensity with national characteristics can explain cross-industry

[4] Cabral and Mata (2003) use the generalized gamma distribution to study firm-size distributions. The finance literature considers a wide family of stochastic asset prices processes with linear drift and power diffusion terms (see, e.g., Chan et al. 1992, on interest rate movements) but, to our knowledge, does typically neither nest 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).

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

Our approach differs from the literature in two important 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 EK model parameters) and are thus able to examine all merchandise sectors (including nonmanufacturing) at the finest level of industry disaggregation possible. We gain from this approach a perspective on hyperspecialization in exporting and evanescence 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 comparative advantage in the cross section, which our estimator does not target, and use the fit of the cross-section projections as a check on the goodness of fit.

Section 2 of the paper presents a simple theoretical motivation for our gravity specification. Section 3 describes the data and our estimates of country export capabilities, and documents the two key empirical regularities regarding comparative advantage (hyperspecialization in exporting and evanescence in countries' top export goods). Section 4 describes a stochastic process that has a stationary distribution consistent with hyperspecialization and a drift consistent with turnover, and introduces a generalized method of moments (GMM) estimator to identify the fundamental parameters. Section 5 presents the estimates, documents the close fit of the diffusion estimates by comparing the diffusion-implied cross-sectional distribution of absolute advantage by country and year to the actual cross-sectional distribution, and discusses the potential implications for dynamic explanations of trade. Section 6 concludes.

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

[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).

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's firm that manufactures a variety in industry i is determined by a random draw from a Fréchet distribution with the cumulative distribution function (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 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. 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.

To formalize this reasoning, 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

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

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. Because EK does not yield a closed-form solution for wages, we cannot solve 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 (as well as the industry-specific μ_i's and θ_i's).

In principle, our concept of export capability k_{is} can be related to a deeper *origin of comparative advantage*. For instance, the country-industry specific Féchet position parameter $T_{is} \equiv (\underline{q}_{is})^{\theta_i}$ can be modelled as the outcome of an exploration and innovation process in our version of the EK model. We sketch this connection in Appendix D.

Any trade model that has a gravity structure will have an associated exporter-industry fixed effect and therefore 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 industry marginal production costs raised to the power of the trade-cost elasticity. 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 commonality across these models is that export capability is related to a country's productive potential in an industry, be it associated with resource endowments, 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). To illustrate the idea, 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 $\left(\tau_{isd}\right)^{\theta}$ removes the distortion that geography exerts on export capability when trade flows are realized.[8] When compared to any country s', country s has a comparative advantage in industry i relative to industry i' if the following condition holds,

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

Importantly, 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 inviable (and cumbersome) to conduct double comparisons of a country-industry is to all other industries and all other countries. We therefore need an adequate summary measure. Our gravity-based correction of trade flows for geographic frictions gives rise to a natural construction of such a measure.

2.2 Estimating the gravity model

A measure of comparative advantage should be uncontaminated by geographic happenstance or idiosyncratic demand conditions in nearby countries. We extract a measure of export capability from the gravity model of trade that is free from the confounding effects of geography and foreign demand shocks. The presence of unobserved trade costs introduces 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 ordinary least squares (OLS) estimation. 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

[8] This adjustment ignores any impact of trade costs on the factor prices that determine unit production costs.

(4) are standard gravity covariates, which do not vary by industry.[9] We allow the coefficients b_{it} on these variables to differ by industry and by year.[10] For we lack annual measures of industry-specific trade costs for all years in our sample period, we model these costs via the interaction of country-level gravity variables and time-and-industry-varying coefficients.

We estimate *export capability* using exporter-industry-year dummy variables. There is therefore a disconnect between the estimated coefficients and the underlying theoretical values they represent. 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 but one, which in all cases is 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 importer-industry-year fixed effect for the United States, and k_{ist} is the underlying log export capability. By construction, the importer-industry-year dummy for the excluded importing country loads onto the estimated exporter-industry-year dummies and prevents isolation of log exporter capability k_{ist}. 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-year dummies from the global industry means:

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

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 total factor productivity (TFP) growth in the industry, worldwide demand change for the industry, or movements in the industry's global producer price index. The normalization allows us to conduct analysis of comparative advantage with trade data exclusively.

[9] 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.

[10] 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.

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

$$A_{ist} \equiv \exp\{\hat{k}_{ist}\} = \frac{\exp\left\{k_{ist}^{\mathrm{OLS}}\right\}}{\exp\left\{\frac{1}{S}\sum_{s'=1}^{S} k_{ist}^{\mathrm{OLS}}\right\}} = \frac{\exp\left\{k_{ist}\right\}}{\exp\left\{\frac{1}{S}\sum_{s'=1}^{S} k_{ist}\right\}}. \tag{7}$$

By construction, this measure of absolute advantage is unaffected by the choice of the omitted importer-industry-year fixed effect ($m_{i\mathrm{US}\,t}$ in our implementation). 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 reason for using the deviation from the geometric mean in defining absolute advantage is twofold. One is that our statistic removes the global industry component of estimated export capability, thus 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 modelling 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 include global industry TFP growth, global industry demand shifts, and global producer price index changes. We establish the main regularities regarding the cross section and the dynamics of exporter performance using absolute advantage A_{ist} in the upcoming data Section 3. In Section 4, we will 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 (especially at the sectoral level). One can 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:

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

which we convert into a regression equation by allowing for unobserved trade costs. We re-estimate exporter-industry-year fixed effects by applying multinomial pseudo-maximum likelihood to (8).

3 Data and Main Regularities

The data for our analysis are World Trade Flows (WTF) from Feenstra et al. (2005),[11] which are based on SITC revision 1 industries for 1962 to 1983 and SITC revision 2 industries for 1984 and later.[12] We create a consistent set of country aggregates in these data by maintaining as single units countries that split up over the sample period.[13] To further maintain consistency in the countries present, we restrict the sample to nations that trade in all years and that exceed a minimal size threshold, which leaves 116 country units.[14] 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.[15] For much of the analysis we additionally exclude uranium, oil, and natural gas. These aggregations and restrictions leave 135 industries in the data.

A further set of country restrictions are required to estimate importer and exporter fixed effects. To estimate exporter dummies that are 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 over the sample period among the original 116 country units. We also need that exporters ship to overlapping groups of importing countries. As Abowd et al. (2002) have shown, this type of connectedness assures that all exporter fixed effects are separately identified from importer fixed effects.[16] This restriction

[11]We use a version of these data that have been extended to 2008 by Robert Feenstra and Gregory Wright.

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

[13]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.

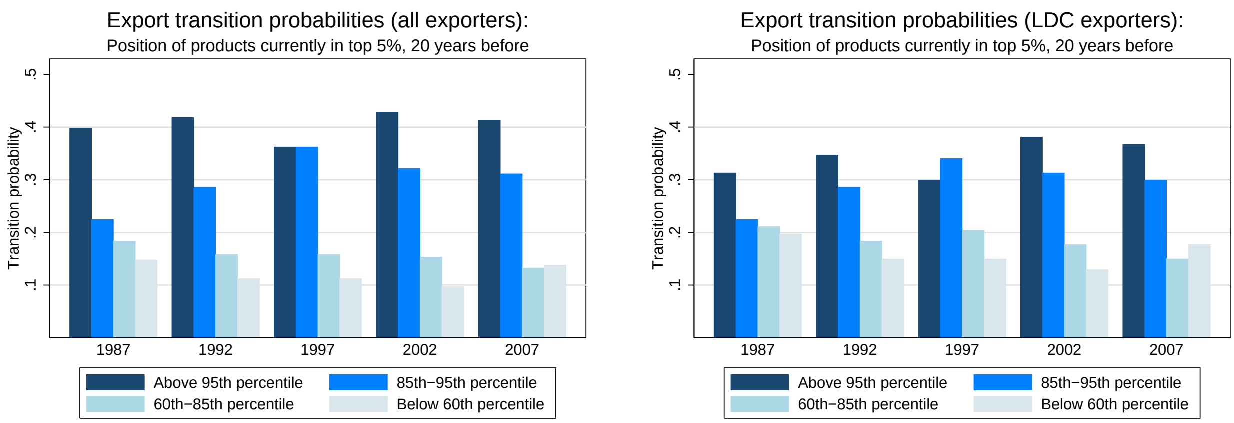
[14]This reporting restriction leaves 141 importers (97.7% of world trade) and 139 exporters (98.2% of world trade) and is roughly equivalent to dropping small countries from the sample. For consistency in terms of country size, we drop countries with fewer than 1 million inhabitants in 1985 (42 countries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

[15]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 further 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.

[16]Countries that export to mutually exclusive sets of destinations would not allow us to separately identify the exporter fixed

Figure 1: **Concentration of Exports**

(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: 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.

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. We now investigate the properties of exporter success and absolute advantage.

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 percentage of a country-industry's exports $X_{ist}/(\sum_{i'} X_{i'st})$ in the country's total exports across the 135 industries, which include all merchandise except oil, gas, and uranium. 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.

The message of this exercise is that for the typical country a handful of industries dominate exports.[17] The median export share of just the top export good is 24% in 1972, which declines modestly over time to

effect from the importer fixed effects.

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

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 (LDC; for our classification see Appendix E). The patterns are quite similar to those for all countries, though median export shares for developing countries are modestly higher in the reported quantiles.

One natural 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. And in the same year in Korea Rep., Malaysia, the Philippines, Taiwan, and the United States the top industry is electric machinery. One certainly would not want to conclude from this fact that all of these countries have 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 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. We display these values in two forms. 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 Korea Rep. it is TVs, in Poland it is furniture, in Thailand it is rice, in Turkey it is glassware, in the United Kingdom it is alcoholic beverages, and in the United States it is other transport equipment (e.g., aircraft). The implied magnitudes of the advantages are enormous. Among the 28 countries in 2007, comparative advantage in the top product—i.e., the double log difference—is over 400 log points in 15 of the cases.[18]

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. Each is a special case of the generalized gamma distribution.

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 raw 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 decent 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.

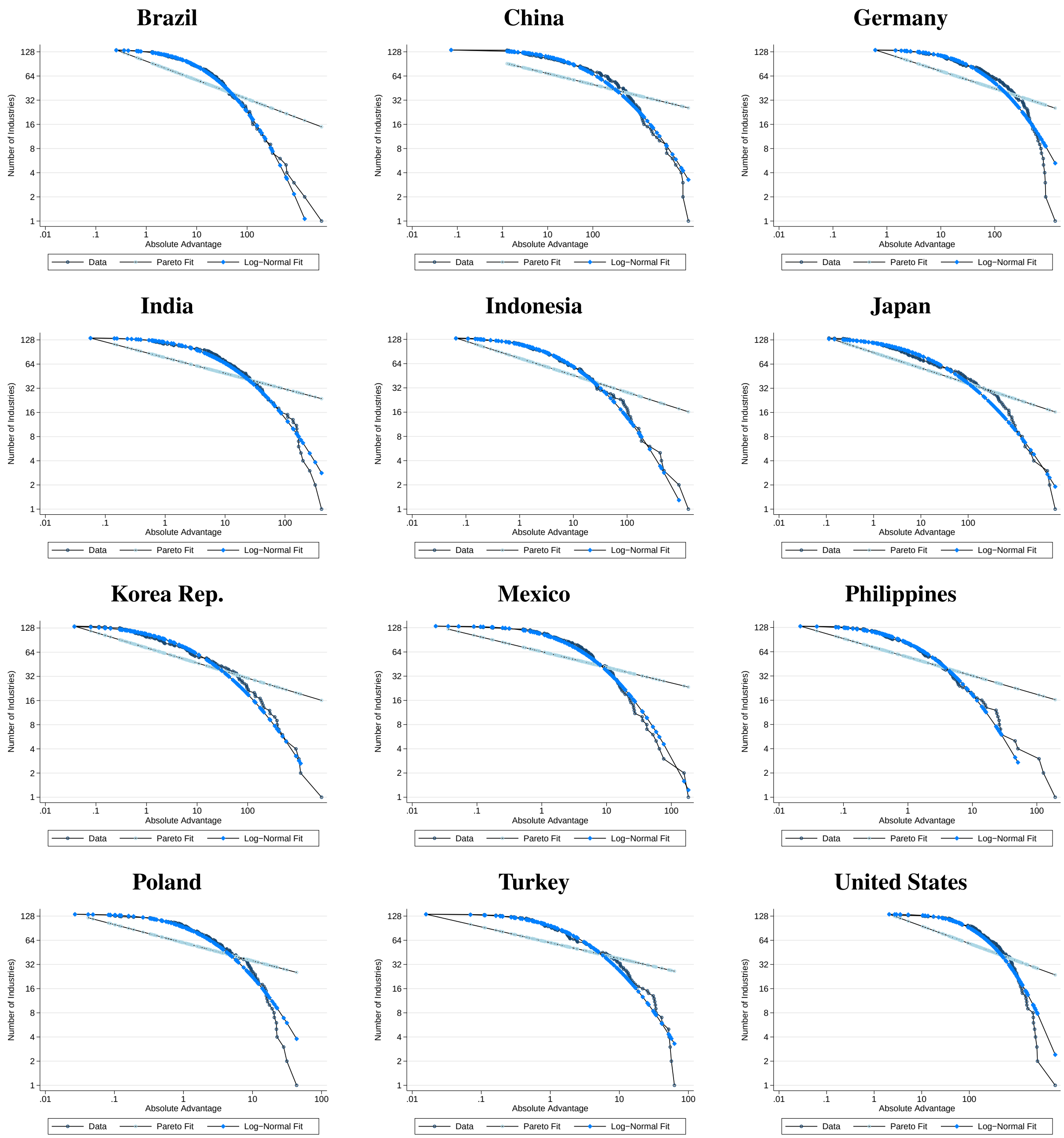
Overall, we see that at any moment in time countries have a strong export advantage in just a few industries, where this general pattern is quite stable both across countries and over time.

3.2 The evanescence 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

[18]These countries are Argentina, Brazil, Canada, China, the Czech Republic, Hungary, Indonesia, Japan, Korea Rep., Malaysia, the Philippines, South Africa, Taiwan, Thailand, and Vietnam.

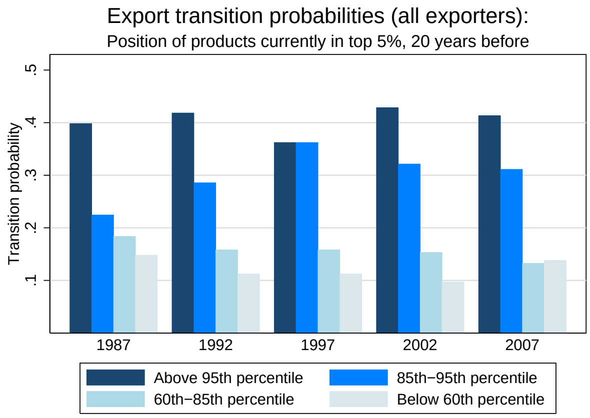
Figure 2: **Cumulative Probability Distribution of Absolute Advantage for Select Countries in 2007**



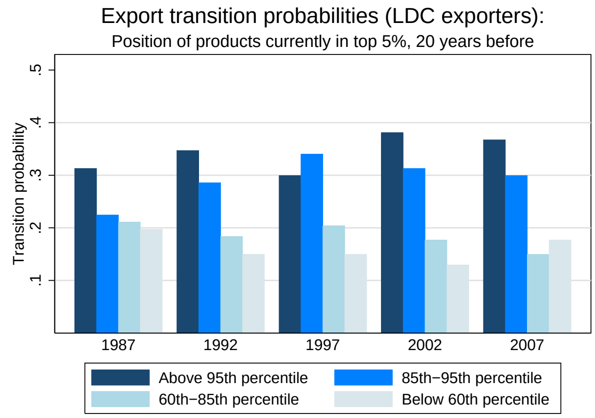
Source: WTF (Feenstra et al. 2005, updatedthrough 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number 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 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.

distribution of absolute advantage, which we investigate next. Initial evidence of churning is evident in Appendix **Table A1**. Of the 28 countries shown, 21 exhibit a change in the top comparative-advantage industry between 1987 and 2007. Canada's top good switches from sulphur to wheat, China's from explosives to telecommunications equipment, Egpyt's from cotton to crude fertilizers, India's from tea to precious stones, Malaysia's from rubber to radios, the Philippines' from vegetable oils to office machines, and Romania's from furniture to footwear. Of the 21 countries with a change in its top comparative-advantage good, in only three cases was the new top product in 2007 the number two product in 1987, with the rest coming from lower down in the distribution. Churning is thus 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. In particular, we calculate for each country in each year where in the distribution the top 5% of absolute-advantage products (in terms of A_{ist}) 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. We see that 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 fraction 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 past. 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.

Turnover in top export goods suggests that absolute advantage dissipates over time. To evaluate this impermanence, we estimate decay regressions for absolute advantage of the form,

$$\ln A_{is,t+10} - \ln A_{ist} = \rho \ln A_{ist} + \delta_{st} + \varepsilon_{ist}. \tag{9}$$

In (9), the dependent variable is the ten-year change in log absolute advantage $\ln A_{ist}$ and the predictors are the initial value of absolute advantage and dummy variables for the country-year δ_{st}.[19] Because absolute advantage is itself a deviation from the global industry mean and the inclusion of country-year dummies imply a further level of differencing from the country-year mean, the regression in (9) evaluates the dynamic evolution of comparative advantage. Where $\rho > 0$, growth in comparative advantage builds on past success; whereas mean reversion implies that $\rho < 0$. The regression in (9) is a crude version of a diffusion observed in discrete time. With normally distributed innovations ε_{ist}, the diffusion would be an Ornstein-Uhlenbeck process, which would imply a log normal distribution for absolute advantage in the cross section. This simple regression gives us an initial impression of the evanescence of export prowess.

Table 1 presents coefficient estimates for equation (9). We perform regressions separately decade by decade to evaluate whether the coefficient ρ changes over time. The first column reports results for all countries and industries, whereas the subsequent columns impose sample restrictions. Estimates for ρ are uniformly negative and precisely estimated, with values ranging from -0.33 for 1997-2007 to -0.57 for 1987-1997. These magnitudes indicate that over the period of a decade the typical country-industry sees 33% to 57% of it comparative advantage erode. Random shocks lift some industries up and push others down. In the second two columns, we divide the sample between developed (non-LDC) economies and developing (LDC) economies. Coefficient estimates for ρ are systematically larger in absolute value for the LDC sample, with decadal decay rates of 42% to 67% compared to 21% to 35% for developed economies. Comparative advantage thus appears to erode more quickly in developed economies. The final two columns divide the sample between manufacturing industries and non-manufacturing industries, where the latter includes agricultural products, minerals, and other commodities. Decay rates are larger for non-manufacturing industries—indicating more rapid dissipation of comparative advantage—though differences in ρ between the two groups of industries are not statistically significant in all decades.

If the white noise in the decay regression in (9) were negligible, then our estimate of the evanescence parameter ρ would imply that, over time, the log of absolute advantage in all industries in all countries would

[19]In implementing equation (9), we choose the alternative but equivalent approach of regressing the decadal change in export capability k_{ist} on its initial value, industry-year dummies, and country-year dummies. Because the inclusion of industry-year dummies effectively converts k_{ist} to $\ln A_{ist}$, this regression is equivalent to that in (9), with appropriately adjusted standard errors.

Table 1: PREDICTION OF 10-YEAR CHANGE IN LOG ABSOLUTE ADVANTAGE BY INITIAL LOG LEVEL

		Exporter countries		Sectors	
$\ln A_{is,t+10} - \ln A_{ist}$	All	LDC	Non-LDC	Manuf.	Nonmanuf.
	(1)	(2)	(3)	(4)	(5)
		Change from 1967-1977			
$\ln A_{ist}$	-0.366	-0.492	-0.301	-0.440	-0.431
	(0.026)**	(0.024)**	(0.042)**	(0.033)**	(0.027)**
Observations	9,271	5,811	3,460	4,665	4,606
Adjusted R^2	0.792	0.754	0.888	0.811	0.804
		Change from 1977-1987			
$\ln A_{ist}$	-0.478	-0.596	-0.333	-0.499	-0.565
	(0.028)**	(0.024)**	(0.068)**	(0.042)**	(0.027)**
Observations	10,411	6,777	3,634	5,281	5,130
Adjusted R^2	0.743	0.704	0.871	0.657	0.791
		Change from 1987-1997			
$\ln A_{ist}$	-0.571	-0.674	-0.345	-0.477	-0.691
	(0.023)**	(0.020)**	(0.048)**	(0.030)**	(0.021)**
Observations	11,071	7,385	3,686	5,547	5,524
Adjusted R^2	0.777	0.762	0.878	0.762	0.820
		Change from 1997-2007			
$\ln A_{ist}$	-0.331	-0.416	-0.207	-0.291	-0.460
	(0.032)**	(0.039)**	(0.059)**	(0.037)**	(0.032)**
Observations	11,520	7,792	3,728	5,656	5,864
Adjusted R^2	0.688	0.648	0.836	0.740	0.703

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 export capability (5).
Note: OLS estimation of the decadal decay

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

conditional on source country effects δ_{st}, using annual export capability measures k_{ist} and conditioning on an additional industry effect δ_{it} for the full pooled sample (column 1) and subsamples (columns 2-5). 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.

revert to a single degenerate log level of zero (conditional on the country fixed effect that transforms absolute advantage into comparative advantage in the regression). Note that an absolute advantage below one results in $\ln A_{ist} < 0$ so that $\rho < 0$ reverts the below-zero industries to the long-run mean of zero (absolute advantage of one) from below, just as it symmetrically reverts the above-zero industries to the long-run mean of zero from above. The estimated evanescence therefore implies a degeneracy in comparative advantage, unless the white noise follows a particular process that pulls individual industries away from the global degenerate long-run mean at precisely the right rate. In the next sections, we turn to a joint specification of both a generalized decay regression to match evanescence and a cross-sectional distribution that is not degenerate but persistently hyperspecialized.

4 The Diffusion of Comparative Advantage

Guided by the two key empirical regularities—hyperspecialization in export activities and churning in the identity of these activities—we search for a parsimonious stochastic process that can characterize the dynamics of comparative advantage. We consider absolute advantage in continuous time $A_{is}(t)$ and adopt the convention to denote absolute advantage in the cross section with $A_{is} \equiv \lim_{t\to\infty} A_{is}(t) = A_{is}(\infty)$. We explore a family of well-defined stochastic processes that are consistent with evanescence in absolute advantage $A_{is}(t)$ over time and with heavy tails of A_{is} in the cross section.

For parsimony and to guarantee existence of a closed-form stationary distribution, we restrict ourselves to a diffusion of absolute advantage. A diffusion is a Markovian process for which all realizations of the random variable are continuous functions of time and past realizations. We implement a generalized logistic diffusion, which has a generalized gamma with three parameters as its stationary distribution. The attractive feature of the generalized gamma is that it nests many distributions as special cases, making the diffusion we employ flexible in nature. We construct a GMM estimator by turning to an invertible mirror diffusion of the generalized logistic diffusion. Our estimator uses an exhaustive set of conditional moments of the mirror diffusion to accommodate 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 stationary distribution fits the actual cross-section data.

4.1 Stationary distribution

The regularities in Section 3.1 on hyperspecialization indicate that the log normal distribution is a plausible benchmark distribution for the cross section of absolute advantage.[20] 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 therefore seek a more flexible 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 offers a candidate family; it unifies the gamma and extreme-value distributions as well as many other distributions (Crooks 2010).[21] Our implementation of the generalized gamma uses three parameters, as in Stacy (1962).[22]

In a cross section of the data, after arbitrarily much time has passed, the proposed relevant generalized gamma probability density function for a realization a_{is} of our random variable absolute advantage A_{is} is given by:

$$f_A(a_{is};\theta_s,\kappa,\phi)=\frac{1}{\Gamma(\kappa)}\left|\frac{\phi}{\theta_s}\right|\left(\frac{a_{is}}{\theta_s}\right)^{\phi\kappa-1}\exp\left\{-\left(\frac{a_{is}}{\theta_s}\right)^{\phi}\right\}\qquad\text{for}\quad a_{is}>0, \tag{10}$$

where $\Gamma(\cdot)$ denotes the gamma function and (θ_s,κ,ϕ) are real parameters with $\theta_s,\kappa>0$.[23] Note that we permit one parameter, θ_s, to vary by country, which allows for a country-specific horizontal shift in the distribution of absolute advantage, consistent with national differences in aggregate TFP. 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.[24]

The parameter restriction $\phi=1$ (indicating an ordinary gamma) is particularly noteworthy. It clarifies

[20] 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.

[21] 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.

[22] In the original Amoroso (1925) formulation, the 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.

[23] 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.

[24] 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).

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. This result points to a candidate stochastic process that has a stationary generalized gamma distribution. The ordinary *logistic diffusion*, a widely used stochastic process (in biology it is termed the competitive Lotka-Volterra model), generates an ordinary gamma as its stationary distribution (Leigh 1968). By extension, the *generalized* logistic diffusion has a *generalized* gamma as its stationary distribution.

Lemma 1. *The generalized logistic diffusion*

$$\frac{\mathrm{d}A_{is}(t)}{A_{is}(t)} = \left[\alpha - \beta_s A_{is}(t)^{\phi}\right] \mathrm{d}t + \sigma \, \mathrm{d}W_{is}^{A}(t), \qquad A_{is}(t) > 0, \tag{11}$$

for real parameters $\alpha, \beta_s, \sigma, \phi$ has a stationary distribution that is generalized gamma with a probability density $f_A(a_{is}; \theta_s, \kappa, \phi)$ given by (10)*, for $A_{is} = \lim_{t\to\infty} A_{is}(t) = A_{is}(\infty)$ and the real parameters*

$$\theta_s = [\phi\sigma^2/(2\beta_s)]^{1/\phi} > 0 \quad \textit{and} \quad \kappa = [2\alpha/\sigma^2 - 1]/\phi > 0.$$

Proof. See Appendix A. □

The generalized logistic nests the Lotka-Volterra model ($\phi = 1$)—leading to an ordinary gamma distribution—and the widely used Ornstein-Uhlenbeck process ($\phi \to 0$)—leading to a log normal distribution.[25]

The term $\left[\alpha - \beta_s A_{is}(t)^{\phi}\right]$ in the generalized logistic diffusion (11) is a deterministic drift that regulates the relative change in absolute advantage $\mathrm{d}A_{is}(t)/A_{is}(t)$; $W_{is}^{A}(t)$ is the Wiener process (standard Brownian motion) that induces stochastic changes in absolute advantage. The drift has two components: a constant drift parameter α, and a level-dependent drift component $\beta_s A_{is}(t)^{\phi}$ where ϕ is the elasticity of the mean reversion (decay) 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 decay ($\phi = 1$). In our benchmark

[25] To see how (11) relates to commonly specified diffusions, consider an arbitrary diffusion of the change in absolue advantage $\mathrm{d}A/A$ over time t:

$$\frac{\mathrm{d}A_{is}(t)}{A_{is}(t)} = \mu\left(A_{is}(t), t\right) \mathrm{d}t + \sigma\left(A_{is}(t), t\right) \mathrm{d}W_{is}^{A}(t),$$

where $\mu(A_{is}, t)$ and $\sigma(A_{is}, t)$ are infinitesimal parameters that are functions of the current level of absolute advantage and of time. There are two potential sources of uncertainty that affect the change in absolute advantage: potentially stochastic components in the infinitesimal parameters $\mu(\cdot)$ and $\sigma(\cdot)$ and the stationary innovation $\mathrm{d}W_{is}^{A}(t)$. (The Wiener innovation is stationary but the level of absolute advantage would follow a random walk if $\mu(\cdot) = 0$ and $\sigma(\cdot)$ is constant.) In the literatures on firm size (Sutton 1997) and city size distributions (Gabaix 1999), a common assumption is that $\mu(\cdot) = \tilde{\mu}(t)$ is stochastic and independent of the random variable's current level, while $\sigma(\cdot) = \sigma/A_{is}(t)$ for a constant σ. This diffusion is a Kesten (1973) process in continuous time and the limiting stationary distribution in the cross section approximates Zipf's law in expectation (a Pareto distribution with an expected shape parameter of one). The graphs in **Figure 2** clearly do not support a Pareto in the cross section of absolute advantage A_{is}, which motivates our search for alternative specifications of $\mu(\cdot)$ and $\sigma(\cdot)$.

case of $\phi \to 0$, the relative change in absolute advantage is neutral with respect to the current level. If $\phi > 0$ then the level-dependent drift component $\beta_s A_{is}(t)^{\phi}$ 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 $\phi < 0$ (requiring $\beta_s < 0$ for existence of the stationary distribution) 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 existence of a non-degenerate stationary distribution with $\theta_s, \kappa > 0$ circumscribes how the parameters of the diffusion α, β_s, σ and ϕ relate to each other. A strictly positive θ_s implies that $\text{sign}(\beta_s) = \text{sign}(\phi)$. 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 $f_A(\cdot)$ 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. Finally, in the benchmark case with $\phi \to 0$, we need $1/\alpha \to 0$ for the limiting distribution to be log normal.

Lemma 1 provides a useful initial specification in our context, tying the observed cross-section distribution of absolute advantage back to a clearly defined stochastic process. However, fitting the stationary generalized gamma distribution to the cross section of absolute advantage for individual countries at varying points in time—which we perform to gauge the validity of our approach—would in the presence of time trends in absolute advantage (due, e.g., to cross-country differences in aggregate TFP growth) require the parameters (θ_s, κ, ϕ) in (10) to vary with time.[26] In point of fact, the distribution plots in **Figures A1** through **A3** suggest that there is marked rightward shift in the distribution of absolute advantage over the years in many countries, consistent with such time trends. We therefore need a standardization of absolute advantage that results in a stationary variable. One candidate stationary variable is, of course, a conventional measure of comparative advantage such as Balassa's revealed comparative advantage because it divides absolute advantage by its country-wide average and thus removes a specific type of country-wide trend. Instead of limiting ourselves to a narrowly imposed form, we allow for more generality by introducing a country-specific stochastic trend in absolute advantage.

[26]Results showing this outcome in our data are available from the authors upon request.

4.2 Diffusion

Adding a trend component to absolute advantage, we specify *generalized comparative advantage* as

$$\hat{A}_{is}(t) \equiv \frac{A_{is}(t)}{Z_s(t)}, \tag{12}$$

where $Z_s(t)$ is an unknown country-wide stochastic trend. The generalized comparative advantage measure preserves the properties of the crude comparative advantage statistic (3) that compares country pairs and industry pairs individually:

$$\frac{\tilde{X}_{isdt}/\tilde{X}_{is'dt}}{\tilde{X}_{i'sdt}/\tilde{X}_{i's'dt}} = \frac{K_{ist}/K_{is't}}{K_{i'st}/K_{i's't}} = \frac{A_{ist}/A_{is't}}{A_{i'st}/A_{i's't}} = \frac{\hat{A}_{ist}/\hat{A}_{is't}}{\hat{A}_{i'st}/\hat{A}_{i's't}}.$$

To derive the diffusion for generalized comparative advantage, note that by (12) $\hat{A}_{is}$ is just a scaled version of absolute advantage, so that comparative advantage must also be generalized gamma by the properties of the generalized gamma distribution (which we formally show in Lemma 2 below). We therefore consider a generalized logistic diffusion of comparative advantage $\hat{A}_{is}$, which has the generalized gamma as its stationary distribution by Lemma 1. We write this diffusion as

$$\frac{\mathrm{d}\hat{A}_{is}(t)}{\hat{A}_{is}(t)} = \frac{\sigma^2}{2}\left[1 - \eta\,\frac{\hat{A}_{is}(t)^{\phi} - 1}{\phi}\right]\mathrm{d}t + \sigma\,\mathrm{d}W_{is}^{\hat{A}}(t), \tag{13}$$

with real parameters η, σ, ϕ. This parametrization is consistent with log normality in the limit as $\phi \to 0$ but rules out a Pareto distribution in the limit, consistent with our evidence.[27] In this simplified formulation, the no-collapse and no-explosion conditions are satisfied for the single restriction that $\eta > 0$, which we assume holds (and impose as a restriction in the estimation).[28] Note that in our specification of the diffusion of comparative advantage in (13) we do not allow any parameter to be country-specific but we do allow $Z_s(t)$ to be an unknown country-wide stochastic trend. Later in the analysis, we allow parameters to vary by broad country group and broad sector.

How do we interpret the three parameters, η, σ, and ϕ in (13)? The parameter η regulates the rate of convergence at which comparative advantage reverts to its global long-run mean of around one. If there

[27] Parametrization (13) of the generalized logistic diffusion can be related back to the parameters in (11) by setting $\alpha = (\sigma^2/2) + \beta_s$ and $\beta_s = \eta\sigma^2/(2\phi)$. As $\phi \to 0$, both α and β_s tend to infinity; if β_s 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 $\beta_s \to \infty$ as $\phi \to 0$, which our parametrization accommodates. 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).

[28] The reformulation in (13) 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.

were no stochastic innovations to comparative advantage (i.e., no Wiener process), then the differential equation (13) governing comparative advantage would be a deterministic (generalized logistic) equation. In the absence of stochastic innovations, comparative advantage would dissipate and gradually revert to the long-run steady state level $[(\eta+\phi)/\eta]^{1/\phi}$. [29] Our estimates will imply that this steady-state value is close to one (because ϕ is close to zero). In short, η regulates how fast comparative advantage would collapse to a degenerate level of around one in all industries and all countries if there were no stochastic innovations. We therefore call the parameter η the *rate of dissipation* of comparative advantage.

There are stochastic shocks to comparative advantage, however, and the Wiener process does not allow comparative advantage to degenerate. Instead, comparative advantage exhibits a robust, and non-degenerate, stationary distribution in the cross section period after period, as seen in Section 3.1. In our statistical explanation, the Wiener process pulls individual industries away from the single global long-run mean. The parameter that regulates the existence of a non-degenerate stationary distribution is the *intensity of innovations* σ. The parameter plays a dual role. One the one hand, σ magnifies volatility by scaling up the Wiener innovations, $\mathrm{d}W_{is}^{\hat{A}}(t)$. On the other hand, σ regulates how fast time elapses in the deterministic part of our generalized logistic diffusion (i.e., σ scales $\mathrm{d}t$ in (13)). It is precisely this dual role that 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. We therefore call σ the intensity of innovations in comparative advantage because it plays the dual role of a volatility scalar on the Wiener process and of a speed of convergence scalar on the deterministic decay.

As discussed before, ϕ is the level elasticity of dissipation and determines whether there is slower reversion to the long-run mean from above, as occurs if $\phi < 0$, versus from below, as occurs if $\phi > 0$. We thus interpret this parameter as describing the stickiness of comparative advantage, with the benchmark log normal case ($\phi = 0$) exhibiting no asymmetry in stickiness on either side of the long-run mean.

For subsequent derivations, it is convenient to restate the generalized logistic diffusion (13) 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 by (13) and Itō's lemma. To understand the properties of this diffusion, we relate it back to

[29] Set $\mathrm{d}\hat{A}_{is}(t)$ to zero in (13) and ignore the Wiener innovation, $\mathrm{d}W_{is}^{\hat{A}}(t)$. Then the singleton steady-state value follows from $1-\eta(\hat{A}^{\phi}-1)/\phi = 0$.

the empirical regularities documented in Section 3. We show how the diffusion is connected to evanescence in the decay regression (9), and we link the diffusion to hyperspecialization consistent with the observed concavity in the cross sectional distribution of absolute advantage.

4.2.1 Evanescence revisited

In the limit when the drift becomes level neutral ($\phi \to 0$), the diffusion of $\ln \hat{A}_{is}(t)$ turns into an Ornstein-Uhlenbeck process with

$$\mathrm{d} \ln \hat{A}_{is}(t) = -\frac{\eta\sigma^2}{2} \ln \hat{A}_{is}(t)\, \mathrm{d}t + \sigma\, \mathrm{d}W_{is}^{\hat{A}}(t) \qquad \text{as } \phi \to 0$$

because $\lim_{\phi\to 0}[\hat{A}_{is}(t)^{\phi} - 1]/\phi = \ln \hat{A}_{is}(t)$. The Ornstein-Uhlenbeck process is a continuous-time analogue to a mean reverting AR(1) process in discrete time and a baseline stochastic process in the natural sciences and finance (see e.g. Vasicek 1977, Chan et al. 1992). In our case, the Ornstein-Uhlenbeck process reverts to a mean of zero (that is an expected $\ln \hat{A}_{is}(t)$ level of zero where a country's industry has neither a comparative advantage nor a comparative disadvantage). The parameter σ captures the volatility caused by the Wiener process, and $\eta\sigma^2/2 > 0$ is the rate at which the shocks dissipate and log comparative advantage reverts towards the zero long-run mean.

In Section 3, we documented evanescence with a decay regression of exporter capability. The specification of the decay regression (9) is in fact the limiting case of (13) for a level-neutral drift ($\phi \to 0$). To see this, relate log absolute advantage $\ln A_{is}(t)$ to log comparative advantage with $\ln A_{is}(t) = \ln \hat{A}_{is}(t) + \ln Z_s(t)$ at any moment in time by (12). Consider an arbitrary interval of time Δ (e.g., a decade) and write out the first difference:

$$\ln A_{is}(t+\Delta) - \ln A_{is}(t) = [\ln \hat{A}_{is}(t+\Delta) - \ln \hat{A}_{is}(t)] + [\ln Z_s(t+\Delta) - \ln Z_s(t)].$$

The discrete-time process that results from sampling from a Ornstein-Uhlenbeck process at a fixed time interval Δ is known to be 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).[30] Applying this result to the first-difference equation above, we obtain our decay regression in (9):

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

[30] 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 level-neutral drift ($\phi \to 0$) and the disturbance $\varepsilon_{ist}(t, t+\Delta) \sim \mathcal{N}(0, [1 - \exp\{-\eta\sigma^2\Delta\}]/\eta)$.

for the evanescence 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 $(1 - \exp\{-\eta\sigma^2\Delta\})/\eta$.[31]

The Ornstein-Uhlenbeck process is the unique non-degenerate Markovian process that has a stationary normal distribution (Karlin and Taylor 1981, ch. 15, proposition 5.1). The Ornstein-Uhlenbeck 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**, then we would find a log normal shape if and only if the underlying Markovian process of log comparative advantage $\ln \hat{A}_{is}(t)$ is an Ornstein-Uhlenbeck process. In **Figure 2**, we only observe absolute advantage, however, so it remains to connect the two stationary distributions of comparative and absolute advantage.

4.2.2 The stationary distributions of comparative and absolute advantage

If comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion by (13), then the stationary distribution of comparative advantage is a generalized gamma distribution with density

$$f_{\hat{A}}(\hat{a}_{is}; \hat{\theta}, \kappa, \phi) = \frac{1}{\Gamma(\kappa)} \left|\frac{\phi}{\hat{\theta}}\right| \left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi\kappa - 1} \exp\left\{-\left(\frac{\hat{a}_{is}}{\hat{\theta}}\right)^{\phi}\right\} \qquad \text{for} \quad \hat{a}_{is} > 0, \tag{14}$$

and parameters

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

by Lemma 1, setting $\alpha = (\sigma^2/2) + \beta_s$ and $\beta_s = \eta\sigma^2/(2\phi)$ in the equivalent of (10) for comparative advantage. From this stationary distribution of comparative advantage $\hat{A}_{is}(t)$, we can infer the stationary distribution of absolute advantage $A_{is}(t)$.

[31] As we explained in footnote 19, we implement the decay regression (9) using log exporter capability $k_{is}(t)$ and present that specification here. Relate exporter capability $k_{is}(t)$ to log comparative advantage with $k_{is}(t) = \ln \hat{A}_{is}(t) + \ln Z_s(t) + \bar{k}_i(t)$ at any moment in time by inserting (12) into (7), where $\bar{k}_i(t) \equiv \sum_{s'=1}^{S} k_{is'}(t)/N$. For an interval of time Δ, the first difference is:

$$k_{is}(t{+}\Delta) - k_{is}(t) = [\ln \hat{A}_{is}(t{+}\Delta) - \ln \hat{A}_{is}(t)] + [\bar{k}_i(t{+}\Delta) - \bar{k}_i(t)] + [\ln Z_s(t{+}\Delta) - \ln Z_s(t)].$$

Given the Gaussian first-order autoregressive process of log comparative advantage in discrete time with autoregressive parameter $\exp\{-\eta\sigma^2\Delta/2\}$ and innovation variance $(1 - \exp\{-\eta\sigma^2\Delta\})/\eta$ we have:

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

for $\rho \equiv -(1 - \exp\{-\eta\sigma^2\Delta/2\}) < 0$ and the unobserved fixed effects $\delta_i(t) \equiv \bar{k}_i(t{+}\Delta) - (1{+}\rho)\bar{k}_i(t)$ and $\delta_s(t) \equiv \ln Z_s(t{+}\Delta) - (1{+}\rho)\ln Z_s(t)$, where the residual $\varepsilon_{ist}(t, t{+}\Delta)$ is normally distributed with mean zero and variance $(1 - \exp\{-\eta\sigma^2\Delta\})/\eta$.

Absolute advantage is related to comparative advantage through a country-wide stochastic trend with $A_{is}(t) = \hat{A}_{is}(t)Z_s(t)$ by definition (12). Plugging this definition into (14), 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*

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

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

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

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 stationary distribution of absolute advantage $A_{is}(t)$ is generalized gamma with the cumulative distribution function

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

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 stationary 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 $\ln a$ 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 $\hat{\theta}, \kappa, \phi$:

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

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

$$\textit{Mean/median ratio} = \frac{\Gamma(\kappa + 1/\phi)/\Gamma(\kappa)}{(G^{-1}[.5;\kappa])^{1/\phi}}. \tag{16}$$

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 (13) 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 (1895) family, for which closed-form solutions of the conditional moments exist. 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 (12). 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* (13) *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:*

$$Z_s(t) = \exp\left\{\mathbb{E}_{st}[\ln A_{is}(t)] - \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\} \tag{17}$$

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

Proof. See Appendix B. □

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 (17):

$$\hat{A}_{is}^{GMM}(t) = \exp\left\{\ln A_{is}(t) - \frac{1}{I}\sum_{j=1}^{I} \ln A_{js}(t) + \frac{\ln(\phi^2/\eta) + \Gamma'(\eta/\phi^2)/\Gamma(\eta/\phi^2)}{\phi}\right\} \tag{18}$$

Detrending absolute advantage to arrive at an estimate of comparative advantage completes the first step in implementing model (13).

Next, we perform a change of variable to recast our model as a Pearson (1895) diffusion. Rewriting our model as a member of the Pearson (1895) 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.[32] The following proposition presents an invertible transformation of comparative advantage that facilitates estimation.

Proposition 2. *If comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion* (13) *with real parameters η, σ, ϕ ($\eta > 0$), then:*

[32] 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).

1. *The transformed variable*

$$\hat{B}_{is}(t) = [\hat{A}_{is}(t)^{-\phi} - 1]/\phi \tag{19}$$

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 (1895) family.

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

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^n \,\middle|\, \hat{B}_{is}(t) = b\right] = \exp\left\{-a_n\Delta\right\}\sum_{m=0}^{n}\pi_{n,m}b^m - \sum_{m=0}^{n-1}\pi_{n,m}\mathbb{E}\left[\hat{B}_{is}(t+\Delta)^m \,\middle|\, \hat{B}_{is}(t) = b\right] \tag{20}$$

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

Proof. See Appendix C. □

Transformation (19) 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 (20) 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. Because these forecast errors must be uncorrelated with any function of past $\hat{B}_{is}(t_{\tau-1})$, we can 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.[33]

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

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

for a given weighting matrix W.

We evaluate this objective function at values of ϕ, η, and σ^2 by de-trending the data to obtain $\hat{A}_{is}^{GMM}(t)$ from equation (18), transforming these variables into their mirror variables $\hat{B}_{is}^{GMM}(t) = [\hat{A}_{is}^{GMM}(t)^{-\phi} - 1]/\phi$, and using equation (20) 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. Although Forman and Sørensen (2008) establish asymptotics as $T \to \infty$, our framework fits into the standard GMM framework of Hansen (1982), which establishes consistency and asymptotic normality of our estimator as the product $IS \to \infty$. We impose the constraints that $\eta > 0$ and $\sigma^2 > 0$ by re-parameterizing 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.

[33]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.

Table 2: GMM ESTIMATES OF COMPARATIVE ADVANTAGE DIFFUSION

	All	Exporter countries: LDC	Exporter countries: Non-LDC	Sectors: Manuf.	Sectors: Nonmanuf.
	(1)	(2)	(3)	(4)	(5)
Estimated Generalized Logistic Diffusion Parameters					
Dissipation rate η	0.265	0.265	0.264	0.416	0.246
	(0.004)**	(0.004)**	(0.008)**	(0.005)**	(0.004)**
Intensity of innovations σ	1.396	1.587	0.971	1.157	1.625
	(0.039)**	(0.045)**	(0.073)**	(0.011)**	(0.042)**
Level elasticity of dissipation ϕ	-.034	-.027	-.036	-.064	-.028
	(0.004)**	(0.005)**	(0.01)**	(0.012)**	(0.004)**
Implied Parameters					
Log generalized gamma shape $\ln \kappa$	160.720	223.840	146.370	72.643	201.560
	(25.325)**	(52.487)**	(57.776)*	(20.257)**	(42.274)**
Log generalized gamma scale $\ln \hat{\theta}$	5.443	5.933	5.302	4.628	5.722
	(0.236)**	(0.355)**	(0.585)**	(0.395)**	(0.316)**
Mean/median ratio	7.375	7.181	7.533	3.678	8.457
Observations	459,680	296,060	163,620	230,890	228,790
Root mean sq. forecast error	1.250	1.381	.913	1.022	1.409
Minimized GMM objective ($\times$ 1,000)	0.411	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 $\hat{A}_{is}(t)$,

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

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 2-5) and fitting the deep parameters η, σ, ϕ under the restrictions $\ln \eta, \ln \sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (19), 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 (16). 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.

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. **Table 2** presents the estimation results.

The magnitude of the estimate of η, which captures the dissipation of comparative advantage, is 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 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 close to zero, suggesting that in practice deviations in comparative advantage from log normality may be modest.

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 in all economies, both among developing and industrialized countries. This considerable concentration is mainly driven by industries in the nonmanufacturing merchandise sector, which exhibit a mean/median ration of more than eight, whereas the ratio is less than four for industries in the manufacturing sector.

The parameters themselves give no indication of the fit of the model. To evaluate fit, we exploit the fact that our GMM estimation targeted exclusively the diffusion of comparative advantage and not its cross sectional dimension. Thus, the cross section 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 section distribution of comparative advantage implied by the parameters estimated from the diffusion.

To implement our validation exercise, we need a measure of $\hat{A}_{is}(t)$ in equation (12), whose value depends on $Z_s(t)$, 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, for instance, from country-specific technological progress. In the estimation, we concentrate out $Z_s(t)$ by exploiting the fact that both $\hat{A}_{is}(t)$ and $A_{is}(t)$ have generalized gamma distributions, allowing us to obtain closed-form solutions for their means, which isolates the value of the trend. To obtain an empirical estimate of $Z_s(t)$ at a given moment in time we apply equation (17), which defines the variable as the difference between the mean log value of $A_{is}(t)$ and the expected value of a log gamma distributed variable (a function of η and ϕ). With estimated realizations for each country in each year Z_{st} in hand, we compute realized values for $\hat{A}_{ist}$ for each country-industry in each year.

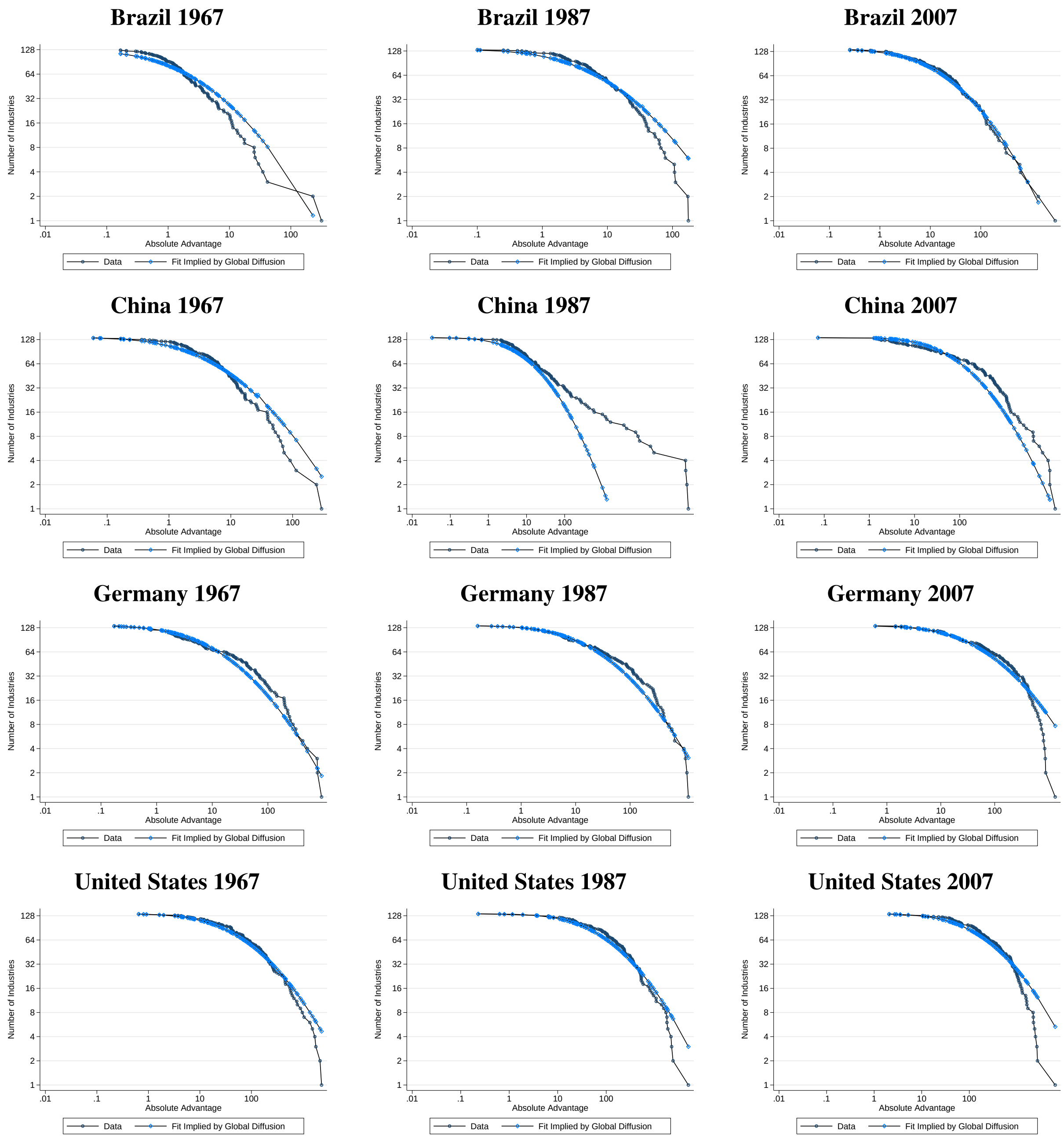
To gauge the goodness of fit of our specification, we first plot our measure of absolute advantage A_{ist}. To do so, following the earlier exercise in **Figure 2**, we rank order the data and plot for each country-industry observation the level of absolute advantage (in log units) against the log number 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 A4, A5** and **A6** 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 equalled the number of parameters for a distribution $\times$ number of countries $\times$ number of years), our exercise now is considerably more parsimonious and based on a fit of the time series evolution, not the observed cross sections. **Figure 4** and **Figures A4** through **A6** present the same, horizontally shifting but identically shaped, single cross sectional distribution, as implied by just two shape relevant parameter estimates (out of three total) that fit the global diffusion 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 (e.g., China in 1987, Russia in 1987 and 2007, Taiwan in 1987, and Vietnam in 1987 and 2007), with the common feature that most of these countries undergo a transition away from central planning during the sample period.

There are some telling minor 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 (China, Japan, Korea

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, updatedthrough 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 (15) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_s(t)$ from (17), which horizontally shifts the distributions but does not affect their shape.

Rep., Malaysia, Taiwan, 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 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 (as established in Forman and Sørensen 2008), beyond our current implementation of discrete-time trade events.

2. We examine an alternative measure of the goodness of fit by plotting observed quantiles for absolute advantage against predicted quantiles for absolute advantage.

3. We restrict the estimation to the latter half of the sample period and use these estimates to simulate distributions for the first half of the sample period, as yet another method for evaluating fit.

4. We allow the diffusion to differ between groups of countries (developing countries, developing countries) or between broad sectors (manufacturing, nonmanufacturing) in a single estimation step.

5. We simulate the model to evaluate predictions for the frequency of turnover in top comparative advantage industries and the predicted likelihood for industries of movements between quantiles of the distribution.

6. We 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.

6 Conclusion

Two salient facts about comparative advantage arise from an investigation of trade flows among a large set of countries and time-consistent industries over more than four decades: Countries exhibit hyperspecialization in only a few industries at any moment in time, but the deviation in comparative advantage from its long-run mean of around one dissipates fast, at a rate of one-third to one-half over a decade. This evanescence implies that the identity of the industries in which a country hyperspecializes at any moment changes considerably over time. Within two decades, a country's newcomer industries replace on average three out of five of a country's top five industries in terms of absolute advantage.

We specify a parsimonious global stochastic process for comparative advantage with only three parameters that is consistent with both perennial hyperspecialization in the cross sections and perpetual evanescence. We allow for a country-specific stochastic trend whose removal translates absolute advantage into comparative advantage and estimate the three worldwide 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 (out of three) time-invariant global diffusion parameters that exclusively matter for the shape of the cross sectional distribution to assess the fit of the predicted cross sectional distribution across countries and over four decades. Even though our estimator explicitly does not target the cross sections, but only the annual diffusion over time, we find the shape of the predicted single stationary cross section from just two time-invariant global parameters to tightly match the shape and curvature of the individual observed cross sectional distributions for the bulk of countries over four decades. By construction, our adjustment for the estimated country-level trend does not affect the shape and curvature, which account for hyperspecialization, but only the country-wide level of absolute advantage and thus the absolute position. These results provide an empirical roadmap for dynamic theoretical models of the determinants of comparative advantage.

Our exercise in this paper deliberately sets aside questions about the deeper origins of apparent comparative advantage on the production side and aims to characterize the typical evolution of comparative advantage in a country's industry. In future research, we plan to explore natural follow-up questions.

1. We plan to specify a stochastic gravity equation, allowing for underlying diffusions of technology, endowments, home market effects and institutions that drive the time varying exporter fixed effect as well as allowing for per-capita income and income dispersion for non-homothetic preferences that drive the time varying importer fixed effect. Given the typically short time series of proxies for endowments, institutions and income dispersion, their parameters may have to be fit from the cross sectional stationary distributions in select years. We expect the specification of a stochastic gravity equation to reveal the extent to which "deep" production and "deep" demand forces drive global trade flows.

2. 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 at present, generating a stationary distribution with no closed form, while we can restrict 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 adequately.

3. We plan a decomposition of export values into local intermediate inputs, foreign intermediate inputs and local value added using global input-output accounts. The result will be an interrelated system of diffusions that we can estimate. The estimates will document to what extent the formation of global supply chains drives the evolution of a country-industry's apparent comparative advantage, which can now depend systematically on export capability and thus apparent comparative advantage in other countries from where a given country-industry sources its intermediate inputs. A hypothesis to test and evaluate is whether, and to what extent, apparent comparative advantage may depend less and less on local resources over time, but increasingly on the efficient utilization of resources elsewhere as global vertical specialization progresses industry by industry.

4. 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 the sector-level evidence on evanescence with recent advances in firm-level theories of international trade. Countries for which firm-level data have become available to the authors in earlier research projects include, for example, Brazil, Germany and Sweden. Firms might withstand sector-level evanescence by branching out their product scope across sectors or might be subject to similar evanescence as sectors. Firm-level evidence can sharpen our understanding for the causes of evanescence as well as the product- or process-driven innovations in exploration among nonmanufacturing merchandise industries and in innovation among manufacturing sector industries.

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 $\theta = \bar{\theta}^{1/\phi}$ because

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

which is equivalent to the generalized gamma probability density function (10), where $\Gamma(\cdot)$ denotes the gamma function and θ, κ, ϕ 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

$$\bar{\theta} = \bar{\sigma}^2/(2\bar{\beta}) > 0, \qquad \text{(A.4)}$$
$$\kappa = 2\bar{\alpha}/\bar{\sigma}^2 - 1 > 0$$

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*:

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

with the parameters

$$\alpha \equiv \left[\frac{1-\phi}{2}\frac{\bar{\sigma}^2}{\phi^2} + \frac{\bar{\alpha}}{\phi}\right], \qquad \beta \equiv \frac{\bar{\beta}}{\phi}, \qquad \sigma \equiv \frac{\bar{\sigma}}{\phi}. \qquad \text{(A.5)}$$

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 (10) and by the derivations above:

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

with

$$\begin{aligned}\theta &= \bar{\theta}^{1/\phi} = [\bar{\sigma}^2/(2\bar{\beta})]^{1/\phi} = [\phi\sigma^2/(2\beta)]^{1/\phi} > 0, \\ \kappa &= 2\bar{\alpha}/\bar{\sigma}^2 - 1 = [2\alpha/\sigma^2 - 1]/\phi > 0 \end{aligned} \tag{A.6}$$

by (A.4) and (A.5).

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 (12) 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 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 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 (19). 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 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 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 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.[34]

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

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

$$\mathbb{E}\left[\hat{B}_{is}(t+\Delta)\,\middle|\,\hat{B}_{is}(t)\right] = \exp\{\lambda_s t\}\, h(\hat{B}_{is}(t)). \tag{C.7}$$

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 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

[34]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 $A(f)(x) = \mu_X(x)\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{2}\sigma_X(x)^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}$. An eigenfunction with associated eigenvalue $\lambda \neq 0$ is any function h in the domain of A satisfying $Ah = \lambda h$.

backward recursion from this terminal condition using the equation

$$\pi_{n,m} = \frac{b_{m+1}}{a_m - a_n}\pi_{n,m+1} + \frac{k_{m+2}}{a_m - a_n}\pi_{n,m+2} \tag{C.8}$$

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.7) 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.8). These coefficients correspond to those reported in equation (20).

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

$$\Pi \cdot \mathbb{E}\left[Y_{is}(t+\Delta)\,\middle|\,\hat{B}_{is}(t)\right] = \Lambda(\Delta)\cdot\Pi\cdot Y_{is}(t) \tag{C.9}$$

with $Y_{is}(t) \equiv (1, \hat{B}_{is}(t), \ldots, \hat{B}_{is}(t)^M)'$ and the matrices $\Lambda(t) = \text{diag}(e^{-a_1 t}, e^{-a_2 t}, \ldots, e^{-a_M t})$ 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) $\underline{\hat{q}}$.[35] 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:

$$\mathrm{d}T_{is}(t) = R_{is}(t) = \xi_{is}(t) \lambda_{is}(t)^{\chi} L_{is}(t), \tag{D.10}$$

where $\xi_{is}(t)$ is research productivity in country-industry is, including the efficiency of exploration in the nonmanufacturing 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

[35] 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 nonmanufacturing 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.10). 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.

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).

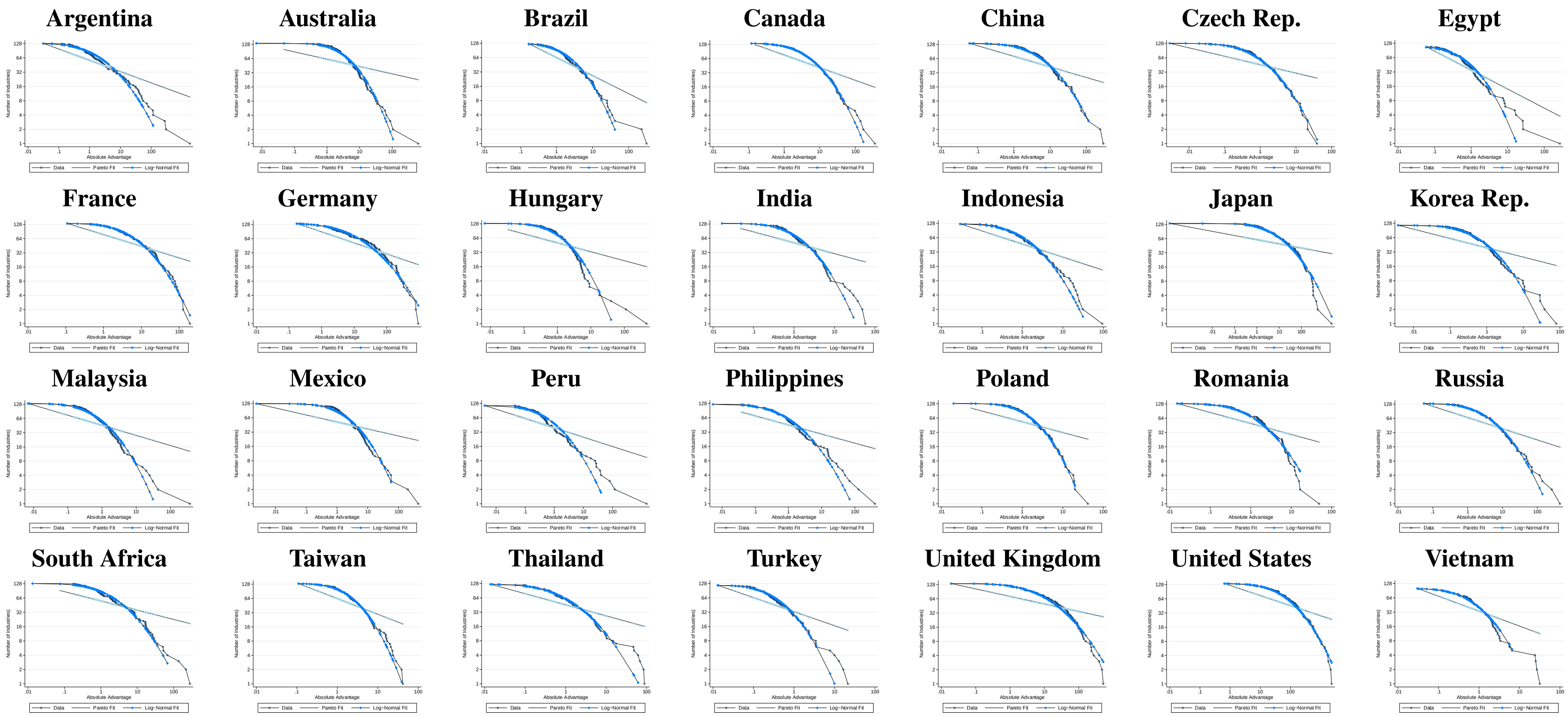
Figures A4, A5 and **A6** extend **Figure 4** in the text and plot, for 28 countries in 1967, 1987 and 2007, the observed log number of a source country s's industries that have at least a given level of absolute advantage in year t against that log absolute advantage level $\ln A_{ist}$ for industries i. This raw data plot is identical to that in **Figures A1** through **A3** and shown for the same 28 countries and years as before. In addition, **Figures A4** through **A6** 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 $Z_s(t)$, which determine the horizontal position of the stationary distribution of observed absolute advantage A_{ist}.

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

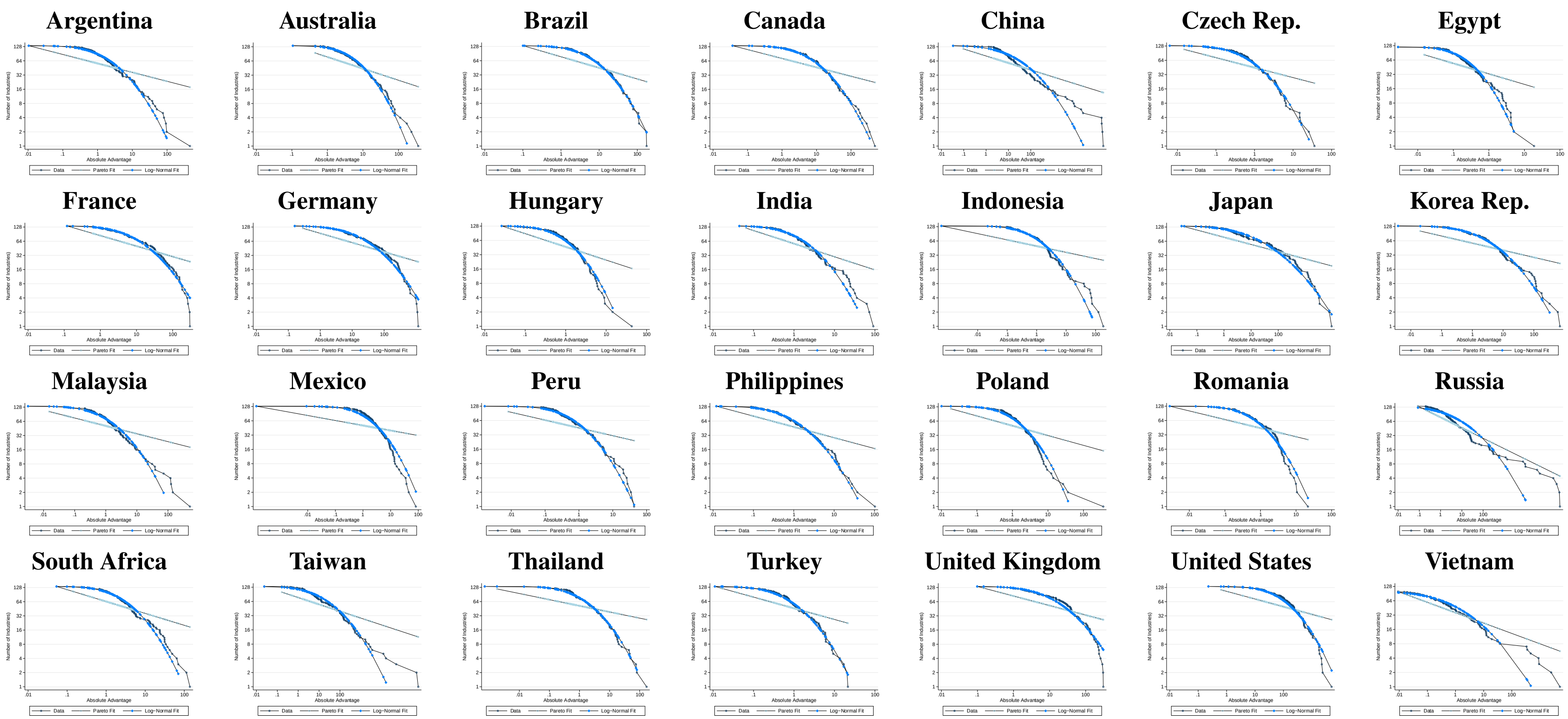
Source: WTF (Feenstra et al. 2005, updated through 2008) for 135 time-consistent industries in 90 countries from 1962-2007.
Note: Top two industries for 28 of the 90 countries in 1987 and 2007 in terms of normalized log absolute advantage, relative to the country mean: $\ln A_{ist} - (1/I)\sum_{i'}^{I}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

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



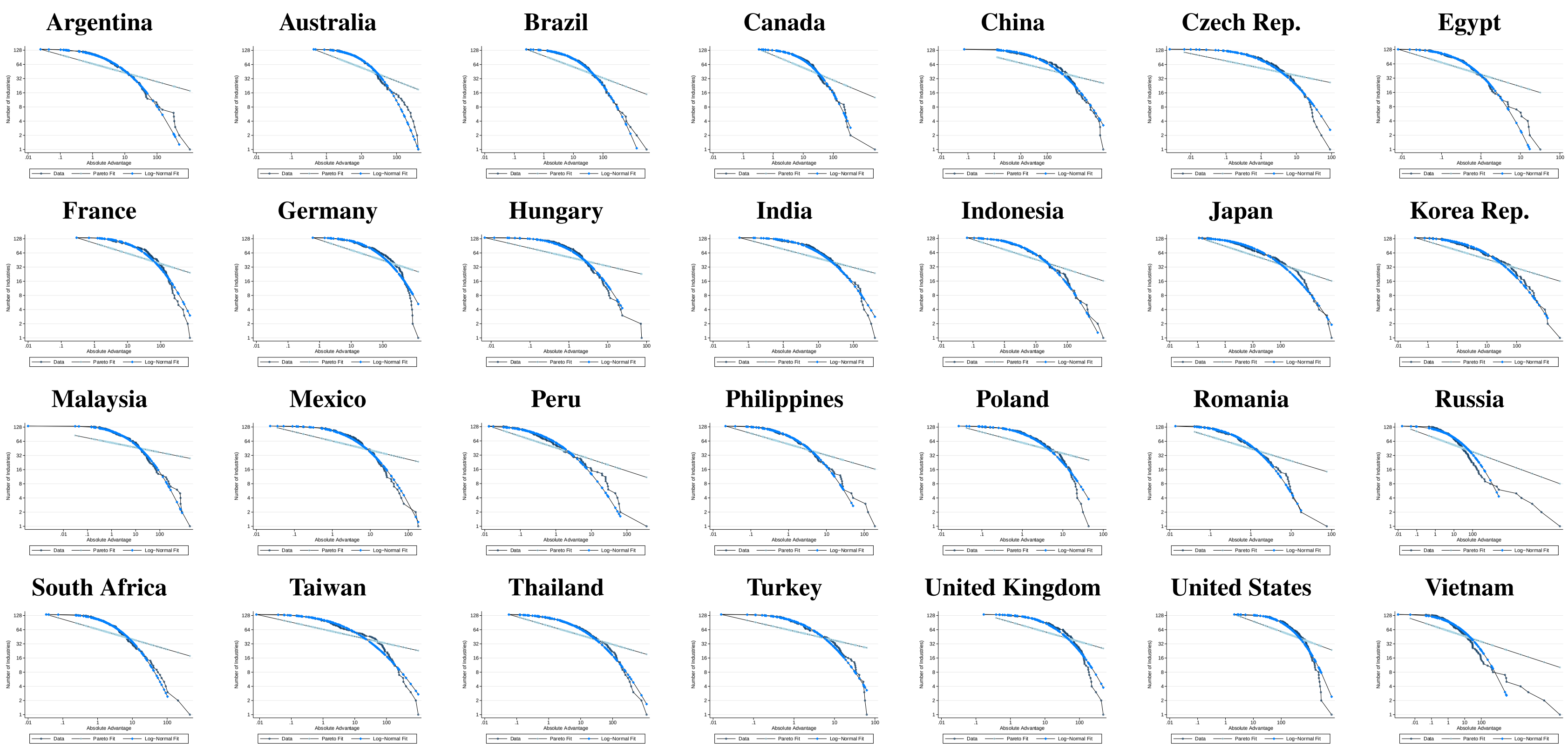
Source: WTF (Feenstra et al. 2005, updatedthrough 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number 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, updatedthrough 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 $t = 1987$.

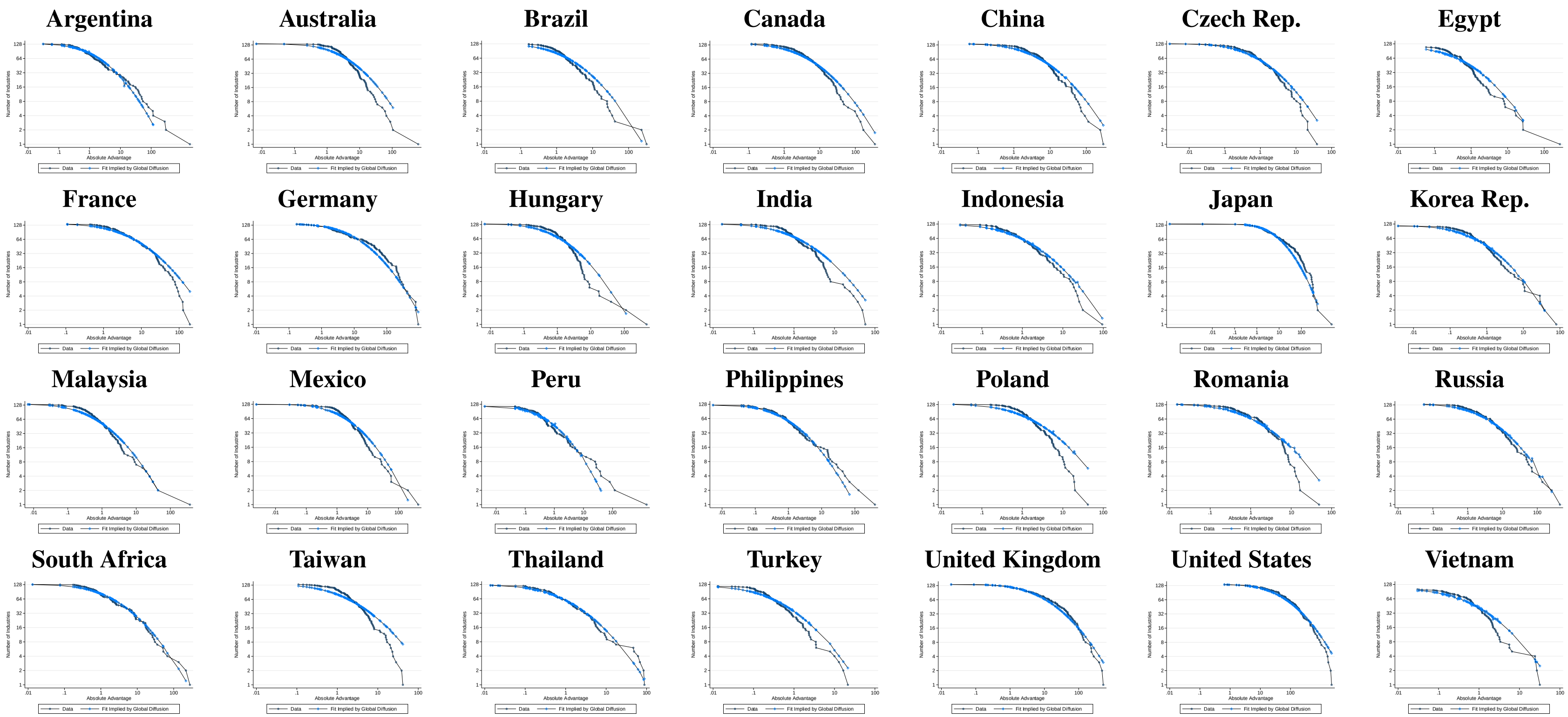
Figure A3: **Cumulative Probability Distribution of Absolute Advantage for 28 Countries in 2007**



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

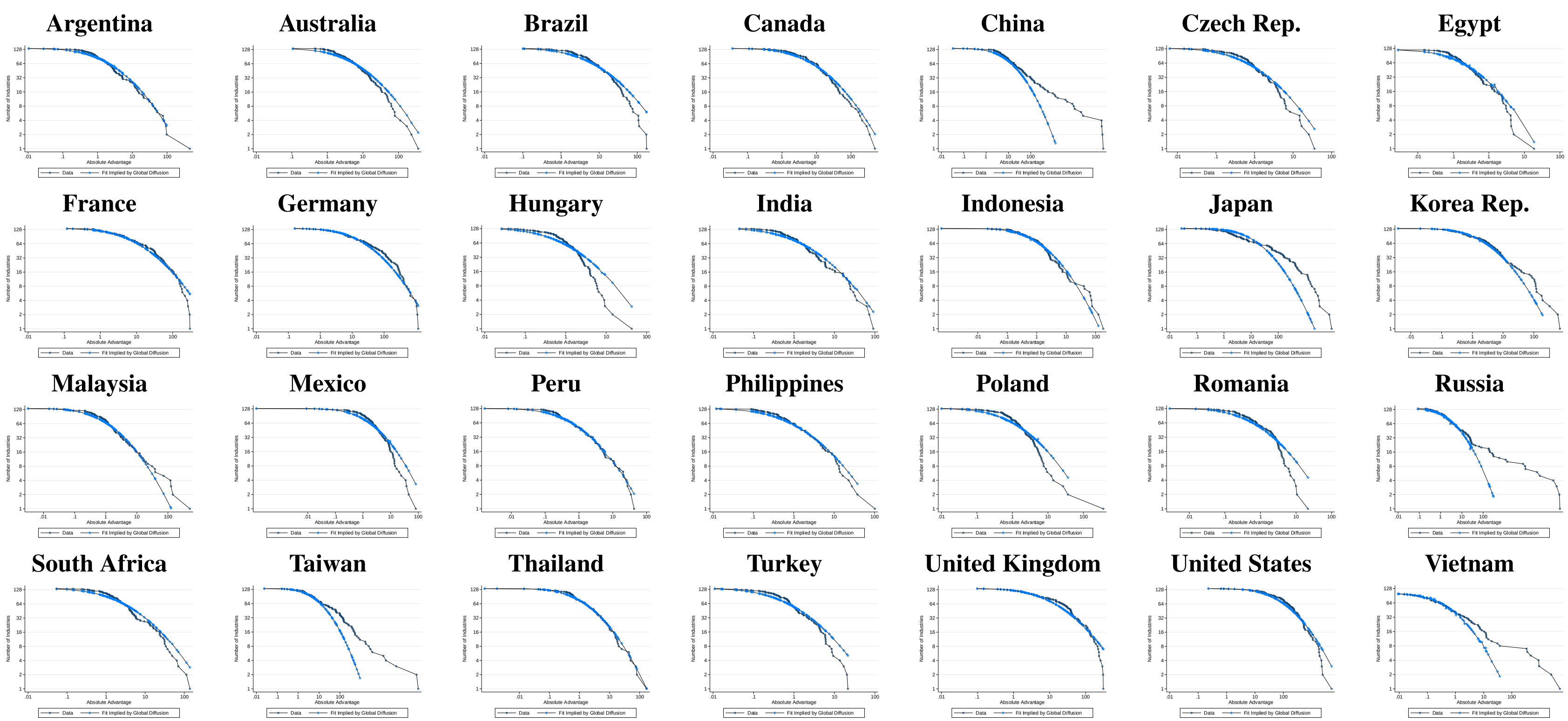
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 135$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted 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: **Global Diffusion Implied and Observed Cumulative Probability Distributions of Absolute Advantage in 1967**



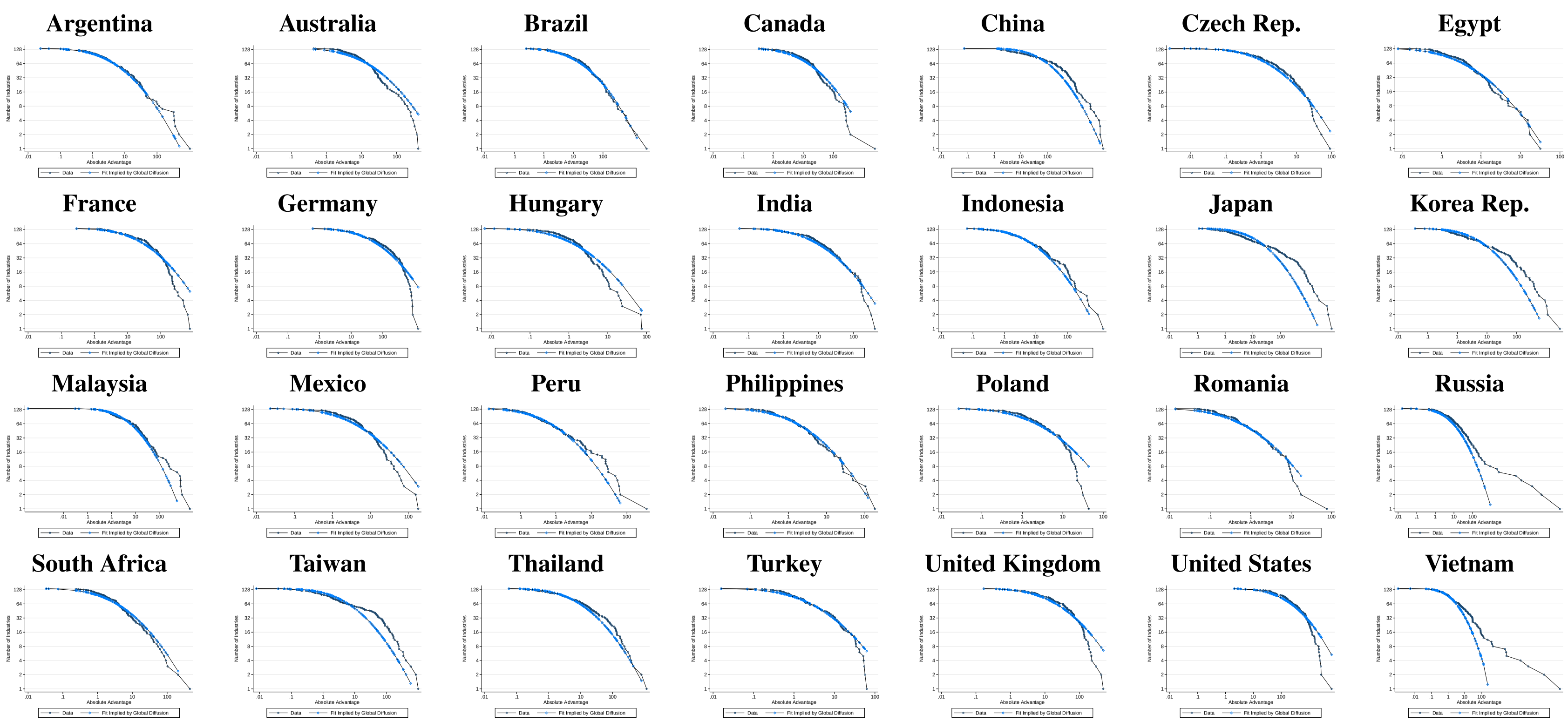
Source: WTF (Feenstra et al. 2005, updatedthrough 2008) for 135 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity-based measures of absolute advantage (7).
Note: The graphs show the observed and predicted frequency of industries (the cumulative probability $1 - F_A(a)$ 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 (15) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_s(t)$ from (17), which horizontally shifts the distributions but does not affect their shape.

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



Source: WTF (Feenstra et al. 2005, updatedthrough 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 (15) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_s(t)$ from (17), which horizontally shifts the distributions but does not affect their shape.

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



Source: WTF (Feenstra et al. 2005, updatedthrough 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 (15) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_s(t)$ from (17), 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.

Collins, Susan M., "Technical Progress in a Three-Country Ricardian Model with a Continuum of Goods," *Journal of International Economics*, August 1985, *19* (1-2), 171–79.

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.

Dornbusch, Rudiger, Stanley Fischer, and Paul A. Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, December 1977, *67* (5), 823–39.

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.*

Eeckhout, Jan, "Gibrat's Law for (All) Cities," *American Economic Review*, December 2004, *94* (5), 1429–51.

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 Colorado, Boulder, 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.

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

_ and Yannis M. Ioannides, "The Evolution of City Size Distributions," in J. Vernon Henderson and Jacques Francois Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4 of *Handbooks in Economics*, Amsterdam, San Diego and Oxford: Elsevier, 2004, chapter 53, pp. 2341–78.

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.

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.

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

Kesten, Harry, "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Mathematica*, December 1973, *131* (1), 207–248.

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.

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, Joao 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.