CESifo Working Papers

www.cesifo.org/wp

The Dynamics of Comparative Advantage

Gordon H. Hanson
Nelson Lind
Marc-Andreas Muendler

CESifo Working Paper No. 5622
Category 8: Trade Policy
November 2015

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5622*

The Dynamics of Comparative Advantage

Abstract

This paper characterizes the dynamic empirical properties of country export capabilities in order to inform modelling of the long-run behavior of comparative advantage. The starting point for our analysis is two strong empirical regularities in international trade that have previously been studied incompletely and in isolation to one another. The literature has noted a tendency for countries to concentrate exports in a few sectors. We show that this concentration arises from a heavy-tailed distribution of industry export capabilities that is approximately log normal and whose shape is stable across countries, sectors, and time. Likewise, previous research has detected a tendency for mean reversion in national industry productivities. We establish that mean reversion in export capability, rather than indicative of convergence in productivities or degeneracy in comparative advantage, is instead consistent with a well behaved stochastic growth process that delivers a stationary distribution of country export advantage. In literature on the growth of cities and firms, economists have used stochastic processes to study the determinants of the long-run size distributions. Our contribution is to develop an analogous empirical framework for identifying the parameters that govern the stationary distribution of export capability. The main result of this analysis is that a generalized gamma distribution, which nests many commonly studied distributions, provides a tight fit of the data but log normality offers a reasonable approximation. Importantly, the stochastic process that generates log normality can be estimated in its discretized form by simple linear regression. Log linearity allows for an extension of our approach to multivariate diffusions, in which one can permit innovations to productivity to be transmitted intersectorally and internationally, as in recent models of trade and growth.

JEL-Codes: F140, F170, C220.

Keywords: international trade, comparative advantage, generalized logistic diffusion, estimation of diffusion process.

Gordon H. Hanson
GPS / University of California San Diego
9500 Gilman Dr MC 0519
USA – La Jolla, CA 92093-0519
gohanson@ucsd.edu

Nelson Lind
Department of Economics
University of California San Diego
9500 Gilman Dr MC 0508
USA – La Jolla, CA 92093-0508
nrlind@ucsd.edu

Marc-Andreas Muendler
Department of Economics
University of California San Diego
9500 Gilman Dr MC 0508
USA – La Jolla, CA 92093-0508
muendler@ucsd.edu

November 19, 2015

Kilian Heilmann, Ina Jäkle, Heea Jung and Chen Liu provided excellent research assistance. We thank Sebastian Sotelo for kindly sharing estimation code. We thank Davin Chor, Arnaud Costinot, Dave Donaldson, Jonathan Eaton, Keith Head, Sam Kortum and Ben Moll as well as seminar participants at NBER ITI, Carnegie Mellon U, UC Berkeley, Brown U, U Dauphine Paris, the West Coast Trade Workshop UCLA, CESifo Munich, the Cowles Foundation Yale and Princeton IES for helpful discussions and comments. We gratefully acknowledge funding from the National Science Foundation under grant SES-1427021. An Online Supplement with supporting derivations and additional empirical evidence is available at www.econ.ucsd.edu/muendler/research.

1 Introduction

Comparative advantage has made a comeback in international trade. After a long hiatus during which the Ricardian model was widely taught to students but rarely applied in research, the role of comparative advantage in explaining trade flows is again at the center of inquiry. Its resurgence is due in large part to the success of the Eaton and Kortum (2002) model (EK hereafter). A growing literature uses EK as a foundation for quantitative modelling of changes in trade policy and other trade shocks (e.g., Costinot and Rodríguez-Clare 2014, Di Giovanni et al. 2014, Caliendo and Parro 2015). In empirical work, Chor (2010) and Costinot et al. (2012) find strong support for EK in cross-section trade data. Renewed interest in comparative advantage also stems from the rapid recent growth in North-South and South-South trade, which ostensibly gives resource and technology differences between countries a prominent role in determining global commerce (Hanson 2012).

In this paper, we characterize how country export advantages evolve over time. From the gravity model of trade we extract a measure of country export capability, which we use to evaluate how export performance changes for 90 countries in 133 industries from 1962 to 2007. Distinct from Waugh (2010), Costinot et al. (2012), Levchenko and Zhang (2013), and other recent work, we do not use industry production or price data to evaluate country export prowess.[1] Instead, we rely solely on trade data, which allows us to impose less structure on the determinants of trade and to examine both manufacturing and nonmanufacturing industries at a fine degree of disaggregation and over a long time span. These features of the analysis help us uncover stable and hitherto underappreciated patterns of export dynamics.

The gravity framework is consistent with a large class of trade models (Anderson 1979, Anderson and van Wincoop 2003, Arkolakis et al. 2012). These models have in common an equilibrium relationship in which bilateral trade in a particular industry and year can be decomposed into an exporter-industry fixed effect, which measures the exporting country's *export capability* in an industry; an importer-industry fixed effect, which captures the importing country's effective demand for foreign goods in an industry; and an exporter-importer component, which accounts for bilateral trade frictions (Anderson 2011). We estimate these components for each year in our data.[2] In the EK model, the exporter-industry fixed effect embodies the location parameter of a country's productivity distribution for an industry, which fixes its sectoral efficiency in producing goods. By taking the deviation of a country's log export capability from the global mean for the industry, we obtain a measure of a country's absolute advantage in an industry. By further normalizing absolute advantage by its country-wide mean, we remove the effects of aggregate country growth. We refer to export capability after its double normalization by the global-industry and country-wide means as a measure of comparative advantage.

[1] On the gravity model and industry productivity also see Finicelli et al. (2009, 2013), Fadinger and Fleiss (2011), and Kerr (2013).

[2] We perform the estimation with and without correcting for zero bilateral trade flows. Silva and Tenreyro (2006), Helpman et al. (2008), Eaton et al. (2012), and Fally (2012) provide econometric approaches to account for zero trade.

The aim of our analysis is to identify the dynamic empirical properties of absolute and comparative advantage that any theory of their determinants must explain. Though we motivate our approach using EK, we are agnostic about the origins of export advantage. In the Krugman (1980), Heckscher-Ohlin (Deardorff 1998), Melitz (2003), and Anderson and van Wincoop (2003) models, which also yield gravity specifications, the origins of the exporter-industry fixed effect differ. However, regardless of its origin the interpretation of the exporter-industry fixed effect as a country's export capability in an industry still applies. By focusing on the dynamics of export advantage, we seek to uncover the general properties of its distribution across countries, industries, and time, whether advantage arises from the accumulation of ideas (Eaton and Kortum 1999), home-market effects (Krugman 1980), resource supplies (Trefler 1995, Davis and Weinstein 2001, Schott 2003, Romalis 2004), or the quality of institutions (Levchenko 2007, Costinot 2009, Cuñat and Melitz 2012). We verify that our results are robust to replacing our gravity-based measure of export capability with Balassa's (1965) index of revealed comparative advantage and to using data based on more disaggregated sectors.

After estimating country-industry export capabilities, our analysis proceeds in two parts. First, we document two regularities in trade behavior that motivate our modelling of trade dynamics. One is hyperspecialization in exporting.[3] In any given year, exports from a typical country are highly concentrated in a small number of industries. Across the 90 countries in our data, the median share for the top good (out of 133) in a country's total exports is 23%, for the top 3 goods is 46%, and for the top 7 goods is 64%. Consistent with strong concentration, the cross-industry distribution of absolute advantage for a country in a given year is approximately log normal, with ratios of the mean to the median of about 7. Strikingly, the log-normal shape applies both across countries that specialize in different types of goods and over time for countries at diverse stages of development.

Stability in the shape of the distribution of absolute advantage makes the second empirical regularity all the more surprising: there is continual turnover in a country's top export products. Among the goods that account for the top 5% of a country's current absolute-advantage industries, 60% were not in the top 5% two decades earlier. Churning is consistent with mean reversion in export advantage. In an OLS regression of the ten-year change in log export capability on its initial log value and industry-year and country-year fixed effects, we estimate mean reversion at the rate of about one-third per decade. Levchenko and Zhang (2013) interpret such mean reversion, which they find for 19 manufacturing industries, as evidence of international convergence in industry productivities that causes comparative advantage to degenerate.[4] Yet, the persistence of hyperspecialization in exporting suggests that a different process must be at work. Ongoing innovations to export capability somehow offset mean reversion by shifting industries along the distribution, while preserving its shape and heavy tails.

The combination of a stable cross-industry distribution for absolute advantage with churning in national

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

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

industry export rankings is characteristic of a stochastic growth model, the estimation of which occupies the second part of our analysis. As a mean-reverting AR(1) specification, our OLS decay regression is a discrete-time analogue of a continuous-time Ornstein-Uhlenbeck (OU) process.[5] An OU process is governed by two parameters, which we recover from our OLS estimates. The *dissipation rate* regulates the rate at which absolute advantage reverts to its long-run mean and determines the shape of its stationary distribution; the *innovation intensity* scales the stochastic shocks to absolute advantage and determines how frequently industries reshuffle along the distribution. Our results indicate that the dissipation rate is stable across countries and sectors, which suggests that the heavy-tailedness of export advantage is close to universal. The innovation intensity, in contrast, is more variable. It is higher for developing countries and for nonmanufacturing industries, which implies that the pace of churning in industry export ranks is idiosyncratic to countries and sectors.

If log export capability follows an OU process, the cross-industry distribution of absolute advantage for each country is log normal. The puzzle over how to reconcile hyperspecialization in exporting with churning in export ranks is then resolved. Stochastic innovations in export advantage cause a country's top industries to turn over and when paired with mean reversion imply that at any moment of time a country's exports are concentrated in a few industries. Comparative advantage is not degenerate but alive and well, if always on the move.

Although attractive for its linearity when discretized, the OU is but one of many possible stochastic processes to consider. To be as expansive as possible in our characterization of export dynamics, while still in a parametric family of Markovian stochastic processes, we next specify and estimate a *generalized logistic diffusion* (GLD) for absolute advantage. The appeal of the GLD is that it has as its stationary distribution a generalized gamma,[6] which unifies the gamma and extreme-value families and therefore nests many common distributions (Crooks 2010), including those used in influential analyses of city sizes (Gabaix and Ioannides 2004, Luttmer 2007) and firm sizes (Sutton 1997, Gabaix 1999). Relative to the OU process, the GLD adds an additional parameter to estimate—the *decay elasticity*—which allows the speed of mean reversion to differ from above versus below the mean. Slower reversion from above the mean, for instance, would indicate that absolute advantage tends to be "sticky," eroding slowly once acquired. Since the GLD nests the OU, we can easily assess log normality against the generalized gamma for the distribution of export advantage. To estimate comparative advantage as a GLD, we allow for stochastic country trends (converting absolute to comparative advantage), transform the GLD to a process for which closed-form expressions of conditional moments exist (permitting estimation by GMM), and develop a finite-sample standard error correction to account for estimation of export capability as a gravity

[5]The OU is the unique non-degenerate Markov process that has a stationary normal distribution (Karlin and Taylor 1981). It is a baseline stochastic process in the natural sciences and finance (see e.g. Vasicek 1977, Chan et al. 1992).

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

parameter (analogous to two-step estimators in Newey and McFadden (1994)).

To gauge the fit of the GLD, we take the GMM time series estimates of the three global parameters—the dissipation rate, the innovation intensity, and the decay elasticity—and predict the cross-section distribution of absolute advantage, which is not targeted in our estimation. Based on just three parameters, the predicted values match the cross-sectional distributions with considerable accuracy. While the data select the GLD over the more restrictive OU process, the difference in the performance of the two is slight. The GLD and OU yield closely similar predictions for period-to-period transition probabilities between quantiles of the distribution of export advantage. This finding is of significant practical importance for it suggests that in many applications the OU process, with its linear discrete-time analogue, will adequately characterize export dynamics. An OU process greatly simplifies estimating multivariate diffusions, which would encompass the types of intersectoral and international linkages in the transmission of knowledge that are at the core of recent theories of trade and growth (Eaton and Kortum 1999, Alvarez et al. 2013, Buera and Oberfield 2014).

Our finding that churning in export advantage applies to both manufacturing and nonmanufacturing industries suggests that "the discovery of new ideas" in models such as EK should be interpreted broadly. There may be many ways in which countries resolve uncertainty about what they are good at producing (Hausmann and Rodrik 2003). Discovery may result from R&D, such as Nokia's foray into cellular technology transforming Finland into a powerhouse for mobile telephony in the early 2000s, or from foreign direct investment, such as Intel's 1996 decision to build a chip factory in Costa Rica, which made electronics the country's largest export (Rodríguez-Clare 2001). Alternatively, discovery may arise from mineral exploration, such as Bolivia's realization in the 1980s that it held the world's largest reserves of lithium,[7] or experimentation with soil conditions that in the 1970s allowed Brazil to begin exporting soybeans (Bustos et al. 2015). Just as discovery comes in many forms, so too does its erosion. While Brazil remains a leading exporter of soybeans, the rise of smart phones has dented Finland's prominence in mobile technology, Intel's decision to close its operations in Costa Rica is abruptly shifting the country's comparative advantage, and ongoing conflicts over property rights have limited Bolivia's exports of lithium.

In Section 2 we present a theoretical motivation for our gravity specification. In Section 3 we describe the data and gravity model estimates, and document hyperspecialization in exporting and churning in top export goods. In Section 4 we introduce a stochastic process that generates a cross-sectional distribution consistent with hyperspecialization and embeds innovations consistent with observed churning, and derive a GMM estimator for this process. In Section 5 we present estimates and evaluate the fit of the model. In Section 6 we conclude.

[7] Lawrence Wright, "Lithium Dreams," *The New Yorker*, March 22, 2010.

2 Theoretical Motivation

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

2.1 Export capability, absolute advantage, and comparative advantage

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

Comparative advantage stems from the location of the industry productivity distribution, given by $\underline{q}_{is}$, which may vary by country and industry. In a country-industry with a higher $\underline{q}_{is}$, firms are more likely to have a high productivity draw, such that in this country-industry a larger fraction of firms succeeds in exporting to multiple destinations.[8] Consider the many-industry version of the EK model in Costinot et al. (2012). Exports by source country s to destination country d in industry i can be written as,

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

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

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

where $k_{is} \equiv \theta \ln(\underline{q}_{is}/w_s)$ is source country s's log *export capability* in industry i, which is a function of the

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

country-industry's efficiency ($\underline{q}_{is}$) and the country's unit production cost (w_s),[9] and

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

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

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

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

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

The normalization in (3) differences out both worldwide industry supply conditions, such as shocks to global TFP, and worldwide industry demand conditions, such as variation in the expenditure share μ_i.

Our measure of absolute advantage is somewhat unconventional. When A_{is} rises for country-industry is, we say that country s's absolute advantage has increased in industry i even though it is only strictly the case that its export capability has risen relative to the global geometric mean for i. In truth, the country's export capability in i may have gone up relative to some countries and fallen relative to others. Our motivation for using the deviation from the industry geometric mean to define absolute advantage is that in so doing we simplify the specification

[9]Export capability k_{is} depends on endogenously determined production costs w_s and therefore is not a primitive. The EK model does not yield a closed-form solution for wages, so we cannot solve for export capabilities as explicit functions of the $\underline{q}_{is}$'s. In a model with labor as the single primary factor of production, the $\underline{q}_{is}$'s are the only country and industry-specific fundamentals—other than trade costs—that determine factor prices, implying in turn that the w_s's are implicit functions of the $\underline{q}_{is}$'s.

of a stochastic process for export capability. Rather than modeling export capability itself, we model its deviation from a worldwide industry trend, which frees us from having to model a trend component that will reflect the evolution of global industry equilibrium.

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

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

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

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

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

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

[10] Eaton and Kortum (2010) allow costly research effort to affect the Poisson intensity rate and assume that there is "no forgetting" such that all previously discovered ideas are available to firms. In our simple sketch, we abstract away from research effort and treat the stock

identical predictions for the volume of bilateral trade as in equation (1). Our empirical approach is to treat the stock of ideas available to a country in an industry $T_{is}(t)$—relative to the global industry mean stock of ideas $(1/S)\sum_{\varsigma=1}^{S} T_{i\varsigma}(t)$—as following a stochastic process.[11]

2.2 Estimating the gravity model

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

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

where we have added a time subscript t. We include dummy variables to measure exporter-industry-year k_{ist} and importer-industry-year m_{idt} terms. The regressors $\mathbf{r}_{sdt}$ represent the determinants of bilateral trade costs ($\mathbf{r}_{sdt}\mathbf{b}_{it}$ adapts equation (2) by replacing $-\theta_t \ln \tau^{P}_{isdt}$), and v_{isdt} is a residual that is mean independent of $\mathbf{r}_{sdt}$. In the estimation, we exclude a constant term, include an exporter-industry-year dummy for every exporting country in each industry, and include an importer-industry-year dummy for every importing country except for one, which we select to be the United States. The variables we use to measure trade costs $\mathbf{r}_{sdt}$ in (6) are standard gravity covariates, which do not vary by industry.[12] However, we do allow the coefficient vector $\mathbf{b}_{it}$ on these variables to differ by industry and by year.[13] Absent annual measures of industry-specific trade costs for the full sample period, we model these costs via the interaction of country-level gravity variables and time-and-industry-varying coefficients.

The values that we will use for empirical analysis are the deviations of the estimated exporter-industry-year dummies from global industry means. The empirical counterpart to the definition of absolute advantage in (3)

of knowledge available to firms in a country (relative to the mean across countries) as stochastic.

[11]Buera and Oberfield (2014) microfound the innovation process in Eaton and Kortum (2010) by allowing agents to transmit ideas within and across borders through trade. A Fréchet distribution for country-industry productivity emerges as an equilibrium outcome in this environment, where the location parameter of this distribution reflects the current stock of ideas in a country. Below, we describe how the stochastic process we estimate relates to their prediction for the growth rate of the stock of ideas.

[12]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.

[13]We estimate (6) separately by industry and by year. Since in each year the regressors are the same across industries for each bilateral exporter-importer pair, there is no gain to pooling data across industries in the estimation, which helps reduce the number of parameters to be estimated in each regression.

for source country s in industry i is

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

where k_{ist}^{OLS} is the OLS estimate of k_{ist} in (6). By construction, this measure is unaffected by the choice of the omitted importer-industry-year fixed effect in the estimation.

We can place alternative distributional assumptions on the gravity equation (2). As is well known, the linear regression model (6) is inconsistent with the presence of zero trade flows, which are common in bilateral data. We recast EK to allow for zero trade by following Eaton 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. Instead of augmenting the expected log trade flow $\mathbb{E}[\ln X_{isd}]$ from gravity equation (2) with a disturbance, Eaton et al. (2012) consider the expected share of country s in the market for industry i in country d, $\mathbb{E}[X_{isd}/X_{id}]$, and write this share in terms of a multinomial logit model. This approach requires that one know total expenditure in the destination market, X_{id}, including a country's spending on its own goods. Since total spending is unobserved in our data, we invoke independence of irrelevant alternatives and specify the dependent variable as the expectation for the share of source country s in import purchases by destination d in industry i:

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

We re-estimate exporter-industry-year fixed effects and apply multinomial pseudo-maximum likelihood to (8).[14]

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

$$RCA_{ist} \equiv \frac{\sum_d X_{isdt} / \sum_\varsigma \sum_d X_{i\varsigma dt}}{\sum_j \sum_d X_{jsdt} / \sum_j \sum_\varsigma \sum_d X_{j\varsigma dt}}. \tag{9}$$

While the RCA index is ad hoc and does not correct for distortions in trade flows introduced by trade costs or proximity to market demand, it has the appealing attribute of being based solely on raw trade data. Throughout our analysis we will employ the gravity-based measure of absolute advantage (7) alongside the Balassa RCA measure (9). Reassuringly, our results for the two measures are quite similar.

[14]We thank Sebastian Sotelo for estimation code.

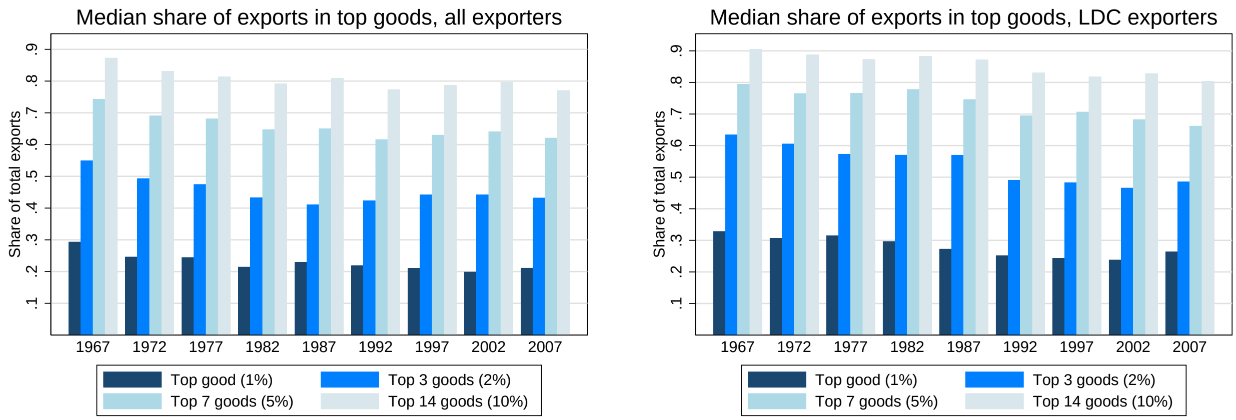
3 Data and Main Regularities

The data for our analysis are World Trade Flows from Feenstra et al. (2005),[15] which are based on SITC revision 1 industries for 1962 to 1983 and SITC revision 2 industries for 1984 and later. We create a consistent set of country aggregates in these data by maintaining as single units countries that divide or unite over the sample period.[16] 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.[17] 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 to a combination of two and three digit categories.[18] These aggregations and restrictions leave 133 industries in the data. In an extension of our main analysis, we limit the sample to SITC revision 2 data for 1984 forward, so we can check the sensitivity of our results to industry aggregation by using two-digit (60 industries) and three-digit definitions (225 industries), which bracket the industry definitions that we use for the full sample period, as well as four-digit (682) definitions.

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

Figure 1: **Concentration of Exports**

(1a) All exporters **(1b) LDC exporters**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007.
Note: Shares of industry i's export value in country s's total export value: $X_{ist}/(\sum_j X_{jst})$. For the classification of less developed countries (LDC) see Appendix E.

3.1 Hyperspecialization in exporting

We first characterize export behavior across industries for each country. For an initial take on the concentration of exports in leading products in this subsection, we tabulate the share of an industry's exports in a country's national exports across the 133 industries $X_{ist}/(\sum_j X_{jst})$. 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 3, top 7, and top 14 industries, which correspond to the top 1%, 2%, 5% and 10% of products.

For the typical country, a handful of industries dominate exports. The median export share of just the top export good is 24.6% in 1972, which declines modestly to 21.4% in 1982 and then remains stable around this level. Over the full period, the median export share of the top good averages 22.9%. For the top 3 products, the

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

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

[17]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 (in which year 42 countries had 1985 population less than 250,000, 14 had 250,000 to 500,000, and 9 had 500,000 to 1 million), reducing the sample to 116 countries (97.4% of world trade).

[18]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 sectors. From this group we drop non-standard industries: postal packages (SITC 911), special transactions (SITC 931), zoo animals and pets (SITC 941), non-monetary coins (SITC 961), and gold bars (SITC 971). We further exclude uranium (SITC 286), coal (SITC 32), petroleum (SITC 33), natural gas (SITC 341), and electrical current (SITC 351), which violate the Abowd et al. (2002) requirement of connectedness for estimating identified exporter fixed effects in most years.

median export share declines slightly from the 1960s to the 1970s and then is stable from the early 1980s onward, averaging 43.5% for 1982 to 2007. The median export shares of the top 7 and top 14 products display a similar pattern, averaging 63.1% and 78.6%, respectively, for 1982 to 2007. In **Figure 1b** we limit the sample to less developed countries (see Appendix E for the set of countries). The patterns are similar to those for all countries, though median export shares for LDCs are somewhat higher in the reported quantiles.[19]

An obvious 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 numerous countries. In 2007, for instance, the top export industry in Great Britain, France, Germany, Japan, and Mexico is road vehicles. In the same year in Korea, Malaysia, the Philippines, Taiwan, and the United States the top industry is electric machinery. We may not want to conclude from these facts that each of these countries has a comparative advantage in one of these products, simply because road vehicles and machinery are among the most traded manufactures.

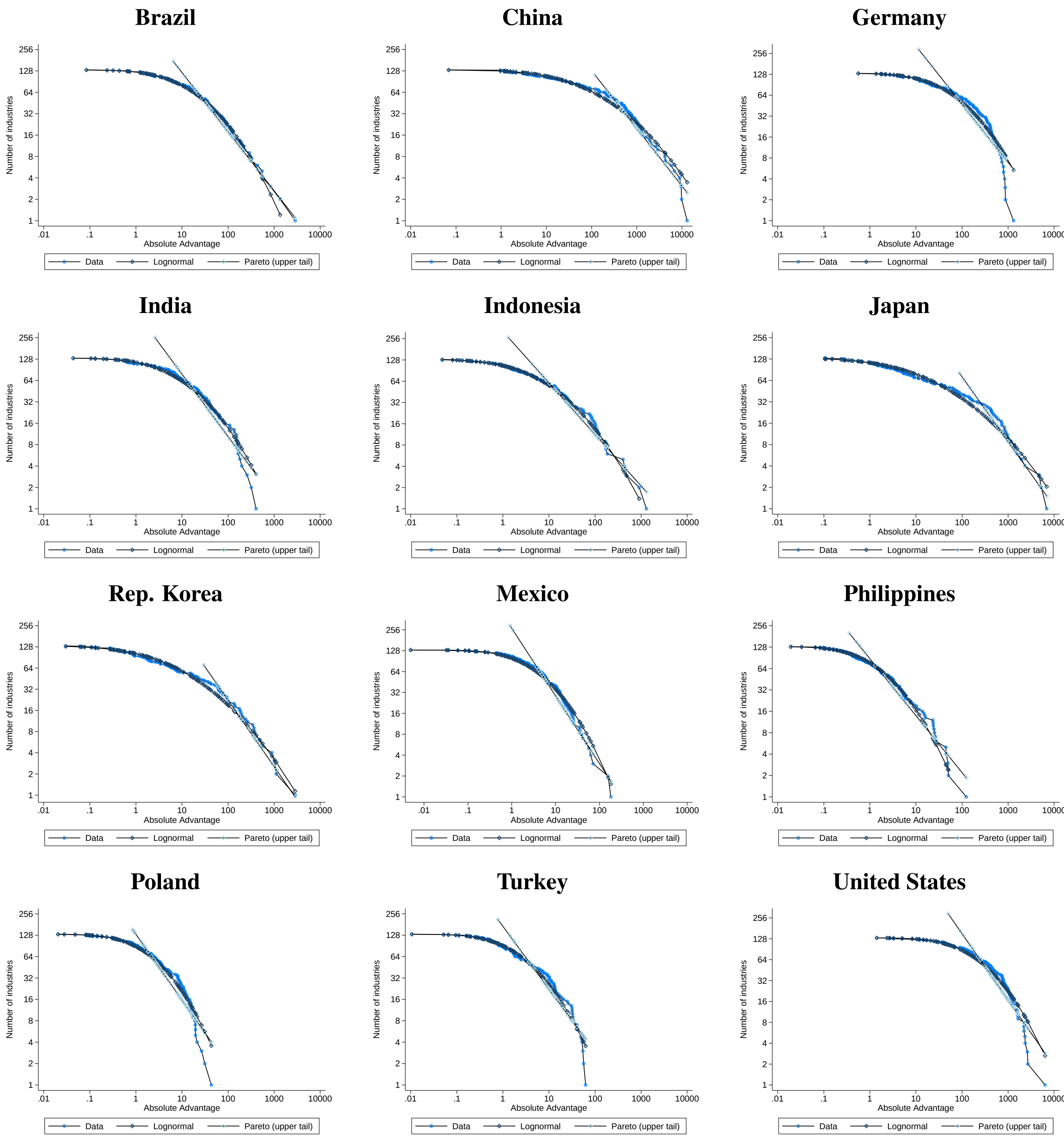
To control for variation in industry size that is associated with global preferences or technology, we turn to our measure of absolute advantage in (7) expressed in logs as $\ln A_{ist}$. To provide a sense of the identities of absolute-advantage goods and the magnitudes of their advantages, we show in Appendix **Table A1** the top two products in terms of $\ln A_{ist}$ 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 $\ln A_{ist} - (1/I) \sum_j^I \ln A_{jst}$, where I is the number of industries. The country normalization yields a double log difference—a country's log deviation from the global industry mean less its average log deviation across all industries—and therefore captures comparative advantage.

There is considerable variation across countries in the top export advantage industries. In 2007, the top industry for Argentina is maize, for Brazil is iron ore, for Germany is road vehicles, for Indonesia is rubber, for Poland is furniture, for Thailand is rice, for Turkey is glassware, and for the United States is other transport equipment (such as commercial aircraft). The magnitudes of these export advantages are enormous. Among the 90 countries in 2007, comparative advantage in the top product is over 400 log points in 76 of the cases and over 300 log points in 88 of the cases.[20]

[19] In analyses of developing-country trade, Easterly and Reshef (2010) document the tendency of a small number of destination markets to dominate national exports by industry and Freund and Pierola (2013) describe the prominent export role of country's very largest firms.

[20] To verify that our measure of export advantage does not peg obscure industries as top sectors, we plot $\ln A_{ist}$ against the log of the share of the industry in national exports $\ln(X_{ist}/(\sum_j X_{jst}))$. The graphs are in the Online Supplement (Figure S1). In all years, there is a strongly positive correlation between log absolute advantage and the log industry share of national exports. This correlation is 0.77 in 1967, 0.78 in 1987, and 0.83 in 2007. (For comparison, the correlation between $\ln A_{ist}$ and the log Balassa RCA index is these same years is 0.69, 0.70, and 0.68, respectively.)

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



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity 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 $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 2007$ (Pareto fit to upper five percentiles only).

To characterize the full distribution of absolute advantage across industries for a country, we next plot the log number of source country s industries that have at least a given level of absolute advantage in year t against the corresponding log level of absolute advantage $\ln A_{ist}$ 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 countries $\times$ number of years). We choose the Pareto and the log normal as comparison cases because these are standard options in the literature on the distribution of city and firm sizes (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 in a log-log plot. Relevant to our later analysis, both the Pareto and the log normal distributions are special cases of the generalized gamma distribution that we will specify. To verify that the graphed cross-section distributions are not a byproduct of specification error in estimating export capabilities from the gravity model, we repeat the plots using the Balassa RCA index in 1987 and 2007, with similar results. And to verify that the patterns we uncover are not a consequence of arbitrary industry aggregations we construct plots also at the 4-digit level based on SITC revision 2 data in 1987 and 2007, again with similar results.[21]

The cumulative distribution plots clarify that the empirical distribution of absolute advantage is decidedly not Pareto. The log normal, in contrast, fits the data closely. The concavity of the cumulative distribution plots drawn for the data indicate that gains in absolute advantage fall off progressively more rapidly as one moves up the rank order of absolute advantage, a feature absent from the scale-invariant Pareto but characteristic of the log normal. This concavity could indicate limits on industry export size associated with resource depletion, congestion effects, or general diminishing returns. Though the log normal well approximates the shape of the distribution for absolute advantage, there are 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 (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.

Overall, countries tend to have a strong export advantage in only a few industries, where this pattern is stable both across countries and over time. Before examining the time series of export advantage in more detail, we consider whether a log normal distribution of absolute advantage across industries could be an incidental consequence of gravity estimation. The exporter-industry fixed effects are estimated sample parameters, which

[21]Each of these additional sets of results is available in the Online Supplement: Figures S2 and S3 for the Balassa measure as well as Figures S4 and S5) for the four-digit industry definitions under SITC revision 2.

by the Central Limit Theorem converge to being normally distributed around their respective population moments as the sample size becomes large. However, normality of this log export capability estimator does not imply that the cross-sectional distribution of absolute advantage becomes log normal. The Central Limit Theorem does not relate to the distribution of the population moments. If no other stochastic element but the residual noise from gravity estimation generated log normality in absolute advantage, then the cross-sectional distribution of absolute advantage between industries in a country would be degenerate around a single mean.

The empirics are decidedly in favor of non-degeneracy for the distribution of absolute advantage. First, **Figure 2** and its counterparts (**Figures A1, A2** and **A3** in the Appendix) document that industries within a country differ markedly in terms of their mean export capability. In turn, the top export advantage industries in Appendix **Table A1**, far from being random selections, appear closely related to country resource abundance (e.g., Brazil's specialization in iron ore, Indonesia's specialization in rubber) or technological prowess (e.g., Germany's specialization in road vehicles; Turkey's specialization in glassware). Second, the distribution of Balassa revealed comparative advantage is also approximately log normal, indicating that alternative measures of comparative advantage elicit similar distributional patterns. Third, the dynamics of absolute advantage, which we study next, reveal a pace of mean reversion in $\ln A_{ist}$ that appears to be too slow to be consistent with iid sampling errors being the sole explanation for the variable's normal shape. The distribution of absolute advantage therefore appears to reflect meaningful economic variation in export capability across industries within countries.

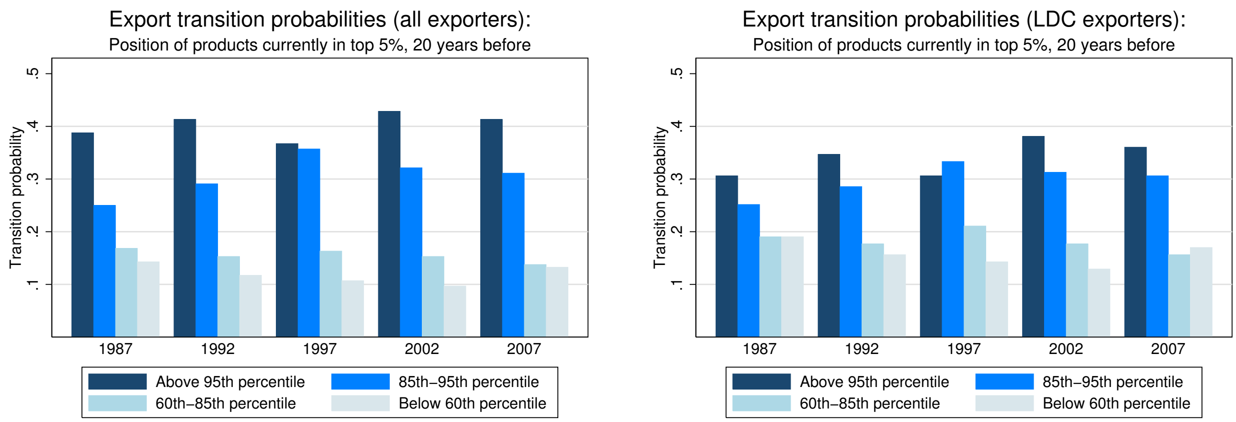
3.2 Churning in export advantage

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 and the shape, though not the position, of the cumulative distribution function remains similar across countries and years. Yet, this stability masks considerable turnover in industry rankings of absolute advantage. Initial evidence of churning is evident in Appendix **Table A1**. Between 1987 and 2007, Canada's top good switches from sulfur to wheat, China's from explosives (fireworks) to TVs and telecommunications equipment, Egypt's from cotton to crude fertilizers, India's from tea to precious stones, and Poland's from barley to furniture. Of the 90 total exporters, 68 have a change in the top comparative-advantage industry between 1987 and 2007. Moreover, most new top products in 2007 were not the number two product in 1987 but from lower down the ranking. Churning thus appears to be both pervasive and disruptive.

To characterize turnover in industry export advantage more completely, in **Figure 3** we calculate the fraction of top products in a given year that were also top products some time ago. For each country in each year, we identify where in the distribution the top 5% of absolute-advantage products (in terms of A_{ist}) were 20 years before, with the categories being top 5%, next 10%, next 25% or bottom 60%. We then average across outcomes

Figure 3: **Absolute Advantage Transition Probabilities**

(3a) All exporters **(3b) LDC exporters**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 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.

for the 90 export countries. The fraction of top 5% products in a given year that were also top 5% products two decades earlier ranges from a high of 42.9% in 2002 to a low of 36.7% in 1997. Averaging over all years, the share is 40.2%, indicating a 60% chance that a good in the top 5% in terms of absolute advantage today was not in the top 5% two decades before. On average, 30.6% of new top products come from the 85th to 95th percentiles, 15.5% come from the 60th to 85th percentiles, and 11.9% come from the bottom six deciles. Outcomes are similar when we limit the sample to developing economies.

Turnover in top export goods suggests that over time export advantage dissipates—countries' strong sectors weaken and their weak sectors strengthen. We test for mean reversion in export capability by specifying the following AR(1) process,

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

where k^{OLS}_{ist} is the OLS estimate of log export capability from gravity equation (6). In (10), the dependent variable is the ten-year change in export capability and the predictors are the initial value of export capability and dummies for the industry-year δ_{it} and for the country-year δ_{st}. We choose a long time difference for export capability—a full decade—to help isolate systematic variation in country export advantages. Controlling for industry-year fixed effects converts export capability into a measure of absolute advantage; controlling further for country-year fixed effects allows us to evaluate the dynamics of comparative advantage. The coefficient ρ captures the fraction of comparative advantage that decays over ten years. The specification in (10) is similar to the

Table 1: DECAY REGRESSIONS FOR COMPARATIVE ADVANTAGE

	Full sample		LDC exporters		Nonmanufacturing	
	Exp. cap. k	$\ln RCA$	Exp. cap. k	$\ln RCA$	Exp. cap. k	$\ln RCA$
	(1)	(2)	(3)	(4)	(5)	(6)
Decay Regression Coefficients						
Decay rate ρ	-0.355	-0.303	-0.459	-0.342	-0.457	-0.293
	(0.002)***	(0.01)***	(0.002)***	(0.013)***	(0.003)***	(0.012)***
Var. of residual s^2	2.104	2.318	2.424	2.849	2.522	2.561
	(0.024)***	(0.006)***	(0.025)***	(0.009)***	(0.039)***	(0.009)***
Implied Ornstein-Uhlenbeck (OU) Parameters						
Dissipation rate η	0.277	0.222	0.292	0.199	0.280	0.195
	(0.003)***	(0.006)***	(0.003)***	(0.006)***	(0.005)***	(0.006)***
Intensity of innovations σ	0.562	0.570	0.649	0.648	0.661	0.596
	(0.003)***	(0.005)***	(0.004)***	(0.009)***	(0.006)***	(0.006)***
Observations	324,978	324,983	202,010	202,014	153,768	153,773
Adjusted R^2 (within)	0.227	0.216	0.271	0.224	0.271	0.214
Years t	36	36	36	36	36	36
Industries i	133	133	133	133	68	68
Source countries s	90	90	62	62	90	90

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org.

Note: Reported figures for ten-year changes. Variables are OLS-estimated gravity measures of export capability k from (6) and the log Balassa index of revealed comparative advantage $\ln RCA_{ist} = \ln(X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_j X_{jst}/\sum_j \sum_{\varsigma} X_{j\varsigma t})$. OLS estimation of the ten-year decay rate ρ from

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

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

productivity convergence regressions reported in Levchenko and Zhang (2013), except that we use trade data to calculate country advantage in an industry, examine industries at a more disaggregate level, and include both manufacturing and nonmanufacturing sectors in the analysis. Because we estimate log export capability k_{ist}^{OLS} from the first-stage gravity estimation in (6), we need to correct the standard errors in (10) for the presence of generated variables. To do so, we apply the generated-variable correction discussed in Appendix D.[22]

Table 1 presents coefficient estimates for equation (10). The first two columns report results for all countries and industries, first for log export capability and next for the log Balassa RCA index. Subsequent pairs of columns show results separately for developing countries and nonmanufacturing industries. Estimates for ρ are uniformly negative and precisely estimated, consistent with mean reversion in export advantage. We soundly reject the

[22]This correction is for GMM. See the Online Supplement S.1 for a discussion of the OLS correction as a special case of the GMM correction.

hypothesis that there is no decay (H_0: $\rho = 0$) and also the hypothesis that there is instantaneous dissipation (H_0: $\rho = -1$). Estimates in columns 1 and 2 are similar in value, equal to -0.36 when using log export capability and -0.30 when using log RCA. These magnitudes indicate that over the period of a decade the typical country-industry sees approximately one-third of its comparative advantage (or disadvantage) erode. In columns 3 and 4, we present comparable results for the subsample of developing countries. Decay rates for this group are larger than the worldwide averages in columns 1 and 2, indicating that in less developed economies mean reversion in comparative advantage is more rapid. In columns 5 and 6, we present results for nonmanufacturing industries (in all countries). For export capability, though not for Balassa RCA, decay rates are larger in absolute value for the nonmanufacturing sector (agriculture, mining, and other primary commodities).

To account for zero trade flows, we re-estimate exporter-industry-year fixed effects under the distributional assumptions of Eaton et al. (2012), as described in Section 2, and use multinomial pseudo-maximum likelihood (MPML) on (8). With the resulting gravity-based export capability measures at hand, we re-estimate the decay regression (10). Results are reported in Appendix **Table A3**. MPML-estimated export capability exhibits a somewhat stronger rate of decay ρ, in the full sample (-0.43 for MPML gravity but only -0.36 for OLS gravity) as well as in the LDC (-0.50 instead of -0.46) and nonmanufacturing (-0.49 instead of -0.46) subsamples, and a roughly one-third larger residual variance.

As an additional robustness check, we re-estimate (10) for the period 1984-2007 using data from the SITC revision 2 sample, reported in Appendix **Table A5**. Estimated decay rates are comparable to those in **Table 1**. At the two-digit level (60 industries), the ten-year decay rate for absolute advantage using the full sample is -0.26, at the three-digit level (224 industries) it is -0.38; the two decay rate estimates for two- and three-digit industries bracket the earlier estimate of -0.36 for our preferred industry aggregate between the two- and three-digit levels. At the four-digit level (682 industries) we estimate a decay rate of -0.51. When using log RCA, decay rates vary less across aggregation levels, ranging from -0.31 at the two-digit level to -0.33 at the three-digit level and -0.34 at the four-digit level. The qualitative similarity in decay rates across definitions of export advantage and levels of industry aggregation suggest that our results are neither the byproduct of sampling error nor the consequence of arbitrary industry definitions.

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

3.3 Comparative advantage as a stochastic process

On its own, reversion to the mean in log export capability is uninformative about the dynamics of its distribution.[23] While mean reversion is consistent with a stationary cross-sectional distribution, it is also consistent with a non-ergodic distribution or a degenerate comparative advantage that collapses at a long-term mean of one (log comparative advantage of zero). Degeneracy in comparative advantage is the interpretation that Levchenko and Zhang (2013) give to their finding of cross-country convergence in industry productivities. Yet, the combination of mean reversion in **Table 1** and temporal stability of the cumulative distribution plots in **Figure 2** is suggestive of a balance between random innovations to export capability and the dissipation of these capabilities. Such balance is characteristic of a stochastic process that generates a stationary cross-section distribution.

To explore the dynamics of comparative advantage, we limit ourselves to the family of stochastic processes known as *diffusions*. Diffusions are Markov processes for which all realizations of the random variable are continuous functions of time and past realizations. As a first exercise, we exploit the fact that the decay regression in (10) is consistent with the discretized version of a commonly studied diffusion, the Ornstein-Uhlenbeck (OU) process. Consider log comparative advantage $\ln \hat{A}_{is}(t)$—export capability normalized by industry-year and country-year means.[24] Suppose that, when expressed in continuous time, comparative advantage $\hat{A}_{is}(t)$ follows an OU process given by

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

where $W^{\hat{A}}_{is}(t)$ is a Wiener process that induces stochastic innovations in comparative advantage.[25] The parameter η regulates the rate at which comparative advantage reverts to its global long-run mean and the parameter σ scales

[23] See, e.g., Quah's (1996) critique of using cross-country regressions to test for convergence in rates of economic growth.

[24] In Section 4, we will specify log generalized comparative advantage in continuous time as

$$\ln \hat{A}_{is}(t) \equiv \ln A_{is}(t) - \ln Z_s(t),$$

where $\ln A_{is}(t)$ is log absolute advantage from (7) and $\ln Z_s(t)$ is an unobserved country-wide stochastic trend that we will ultimately estimate. For now, we simply define $\ln Z_s(t) \equiv (1/I) \sum_j \ln A_{js}(t)$.

[25] To relate equation (11) to trade theory, our specification for the evolution of export advantage is analogous to the equation of motion for a country's stock of ideas in the dynamic EK model of Buera and Oberfield (2014). In their model, each producer in source country s draws a productivity from a Pareto distribution, where this productivity combines multiplicatively with ideas learned from other firms, either within the same country or in different countries. Learning—or exposure to ideas—occurs at an exogenous rate $\alpha_s(t)$ and the transmissibility of ideas from one producer to another depends on the parameter β. In equilibrium, the distribution of productivity across suppliers within a country is Fréchet, with location parameter equal to a country's current stock of ideas. The OU process in (11) emerges from the equation of motion for the stock of ideas in Buera and Oberfield (2014, eq. (4)) as the limiting case with the transmissibility parameter $\beta \to 1$, provided that the learning rate $\alpha_s(t)$ is subject to random shocks and producers in a country only learn from suppliers within the same country. In Section 6, we discuss how equation (11) could be extended to allow for learning across national borders.

time and therefore the Brownian innovations $\mathrm{d}W_{is}^{\hat{A}}(t)$.[26]

Comparative advantage reflects a double normalization of export capability, so it is natural to consider a global mean of one, implying a global mean of zero for $\ln \hat{A}_{is}(t)$. The OU case is the unique non-degenerate Markov process that has a stationary normal distribution (Karlin and Taylor 1981). An OU process of log comparative advantage $\ln \hat{A}_{is}(t)$ therefore implies that $\hat{A}_{is}(t)$ has a stationary log normal distribution. In other words, if we were to plot comparative advantage $\hat{A}_{is}(t)$ in a manner similar to **Figure 2**, we would find a log normal shape if and only if the underlying Markov process of $\ln \hat{A}_{is}(t)$ is an OU process.

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

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

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

which implies for the reduced-form decay parameter

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

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

The reduced-form decay coefficient ρ in (12) is a function both of the dissipation rate η and the intensity of innovations σ and may differ across samples because either or both of these parameters vary. That is, ρ may be

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

large either because the number of standard deviation adjustments toward the mean at each unit of time is large or because the size of these standard deviation adjustments is large. This distinction is important because ρ may vary even if the shape of the distribution of comparative advantage does not change.[27]

From OLS estimation of (12), we can obtain estimates of η and σ^2 using the solutions,

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

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

Table 1 shows estimates of η and σ^2 implied by the decay regression results, with standard errors obtained using the multivariate delta method.[28] Estimates of η based on log export capability, at 0.28 in column 1 of **Table 1**, are somewhat larger than those based on the log RCA index, at 0.22 in column 2 of **Table 1**, implying that the distribution of export capability will be more concave to the origin. But estimates for each indicate strong concavity, consistent with the visual evidence in **Figure 2** (as well as **Figures A1, A2** and **A3** in the Appendix) for log absolute advantage (and Figures S2 and S3 for Balassa RCA in the Online Supplement). Patterns of interest emerge when we compare η and σ^2 across subsamples.

First, compare the estimate for ρ in the subsample of developing economies in column 3 of **Table 1** to that for the full sample of countries in column 1. The larger estimate of ρ in the former sample (-0.46 in column 3 versus -0.36 in column 1) implies that *reduced-form* mean reversion is relatively rapid in developing countries. However, this result is silent about how the shape of the distribution of comparative advantage varies across nations. The absence of a statistically significant difference in the estimated dissipation rate η between the developing-country sample ($\eta = 0.29$) and the full-country sample ($\eta = 0.28$) indicates that comparative advantage is similarly heavy-tailed in the two groups. The larger reduced-form decay rate ρ for developing countries results from their having a larger intensity of innovations ($\sigma = 0.65$ in column 3 versus $\sigma = 0.56$ in column 1, where this difference is statistically significant). In other words, a one-standard-deviation shock to comparative advantage in a developing country dissipates at roughly the same rate as in an industrialized country. But because the magnitude of this shock is larger for the developing country, its observed rate of decay will be faster, too (otherwise the country's export capabilities would not have a stationary cross-sectional distribution).

Second, compare nonmanufacturing industries in column 5 to the full sample of industries in column 1.

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

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

Whereas the average nonmanufacturing industry differs from the average overall industry in the reduced-form decay rate ρ (-0.46 in column 5 versus -0.36 in column 1), it shows no such difference in the estimated dissipation rate η (0.28 in both columns 1 and 5). This implies that comparative advantage has comparably heavy tails among manufacturing and nonmanufacturing industries. However, the intensity of innovations σ is larger for nonmanufacturing industries (0.66 in column 5 versus 0.56 in column 1), due perhaps to higher output volatility associated with occasional major resource discoveries. These nuances regarding the implied shape of and the convergence speed toward the cross-sectional distribution of comparative advantage are not at all apparent when one considers the reduced-form decay rate ρ alone.

Finally, we compare results for two-, three- and four-digit industries in Appendix **Table A5** for the subperiod 1984-2007. Whereas reduced-form decay rates ρ increase in magnitude by a factor of 1.4 as one goes from the two- to the three-digit level or the three- to the four-digit level (from -0.26 in column 1 to -0.38 in column 3 and -0.51 in column 5), dissipation rates η move in the opposite direction (from 0.31 in column 1 to 0.30 in column 3 and 0.26 in column 5). The difference in reduced-form decay rates ρ is driven largely by a higher intensity of innovations σ among the more narrowly defined industries at the three- and four-digit levels. Intuitively, the magnitude of shocks to comparative advantage is larger in the more disaggregated three- and four-digit product groupings.

The diffusion model in (11) and its discrete-time analogue in (12) reveal a deep connection between hyper-specialization in exporting and churning in industry export ranks. Random innovations in absolute advantage cause industries to alternate positions in the cross-sectional distribution of comparative advantage for a country at a rate of innovation precisely fast enough so that the deterministic dissipation of absolute advantage creates a stable, heavy-tailed distribution of export prowess. Having established the plausibility of comparative advantage following a stochastic growth process, we turn next to a more rigorous analysis of the properties of this process.

4 The Diffusion of Comparative Advantage

We now search in a more general setting for a parsimonious stochastic process that characterizes the dynamics of export advantage. In **Figure 2**, the cross-sectional distributions of absolute advantage drift steadily rightward for each country, implying that absolute advantage is non-stationary. However, the cross-sectional distributions preserve their shape over time. We therefore consider absolute advantage as a proportionally scaled outcome of an underlying stationary and ergodic variable: comparative advantage. We specify *generalized comparative advantage* in continuous time as

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

where $A_{is}(t)$ is observed absolute advantage and $Z_s(t)$ is an unobserved country-wide stochastic trend. This measure satisfies the properties of comparative advantage in (5), which compares country and industry pairs.

To find a well-defined stochastic process that is consistent with the churning of absolute advantage over time and with heavy tails in the cross section, we implement a generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$, which then has a generalized gamma as its stationary distribution. We denote comparative advantage in the cross section with $\hat{A}_{is}$ (not a function of time) and take it to have a time-invariant stationary distribution given constant parameters of the stochastic process. Absolute advantage $A_{is}(t)$, in contrast, has a trend-scaled generalized gamma as its cross-sectional distribution, with stable shape but moving position as in **Figure 2**.

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

4.1 Generalized logistic diffusion

The regularities in Section 3.1 indicate that the log normal distribution is a plausible benchmark for the cross section of absolute advantage. But the graphs in **Figure 2** (as well as **Figures A1** through **A3** in the Appendix) also indicate that for many countries, the number of industries drops off more quickly or more slowly in the upper tail than the log normal distribution can capture. We require a distribution that generates kurtosis and that is not simply a function of the lower-order moments, as would be the case in the two-parameter log normal. The generalized gamma distribution, which unifies the gamma and extreme-value distributions as well as many others (Crooks 2010), offers a candidate family.[29] There are a number of alternative generalizations to the ordinary gamma distribution (see e.g. Kotz et al. 1994, Ch. 17, Section 8.7). Our implementation of the generalized gamma uses three parameters, as in Stacy (1962).[30]

In a cross section of the data, after arbitrarily much time has passed, the generalized gamma probability

[29] 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.

[30] In the early 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.

density function for a realization $\hat{a}_{is}$ of the random variable comparative advantage $\hat{A}_{is}$ is given by:

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

where $\Gamma(\cdot)$ denotes the gamma function and $(\hat{\theta},\kappa,\phi)$ are real parameters with $\hat{\theta},\kappa > 0$.[31] The generalized gamma nests the ordinary gamma distribution for $\phi = 1$ and the log normal or Pareto distributions when ϕ tends to zero.[32] The parameter restriction $\phi = 1$ clarifies that the generalized gamma distribution results when one takes an ordinary gamma distributed variable and raises it to a finite power $1/\phi$. The exponentiated random variable is then generalized gamma distributed—a result that also points to a candidate stochastic process that has a stationary generalized gamma distribution. The ordinary *logistic diffusion*, a widely used stochastic process, generates an ordinary gamma as its stationary distribution (Leigh 1968). By extension, the *generalized* logistic diffusion has a *generalized* gamma as its stationary distribution.[33]

Lemma 1. *The generalized logistic diffusion*

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

for real parameters (η,σ,ϕ) *has a stationary distribution that is generalized gamma with a probability density* $f_{\hat{A}}(\hat{a}_{is}|\hat{\theta},\kappa,\phi)$ *given by* (15) *and the real parameters*

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

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

Proof. See Appendix A. □

The term $(\sigma^2/2)[1-\eta\{\hat{A}_{is}(t)^{\phi}-1\}/\phi]$ in (16) is a deterministic drift that regulates the relative change in comparative advantage $\mathrm{d}\hat{A}_{is}(t)/\hat{A}_{is}(t)$. The variable $W^{\hat{A}}_{is}(t)$ is the Wiener process. The generalized logistic

[31] 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.

[32] 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).

[33] Returning to the connection between our estimation and the dynamic EK model in Buera and Oberfield (2014)—also see footnotes 11 and 25—the specification in (16) is equivalent to their equation of motion for the stock of ideas (Buera and Oberfield 2014, equation (4)) under the assumptions that producers only learn from suppliers within their national borders and the learning rate $\alpha_s(t)$ is constantly growing across industries, countries, and over time but subject to idiosyncratic shocks. The parameter ϕ in (16) is equivalent to the value $\beta - 1$ in their model, where β captures the transmissibility of ideas between producers. Our finding, discussed in Section 5, that ϕ is small and negative implies that the value of β in the Buera and Oberfield model is large (but just below 1, as they require).

diffusion nests the Ornstein-Uhlenbeck process ($\phi \to 0$ and η finite), leading to a log normal distribution in the cross section. In the estimation, we will impose the constraint that $\eta > 0$.[34]

The deterministic drift involves both constant parameters (η, σ, ϕ) and a level-dependent component $\hat{A}_{is}(t)^{\phi}$, where ϕ is the elasticity of the mean reversion with respect to the current level of absolute advantage. We call ϕ the *elasticity of decay*. The ordinary logistic diffusion has a unitary elasticity of decay ($\phi = 1$). In our benchmark case of the OU process ($\phi \to 0$), the relative change in absolute advantage is neutral with respect to the current level. If $\phi > 0$, then the level-dependent drift component $\hat{A}_{is}(t)^{\phi}$ leads to a faster than neutral mean reversion from above than from below the mean, indicating that the loss of absolute advantage in a currently strong industry tends to occur more rapidly than the buildup of absolute disadvantage in a currently weak industry. Conversely, if $\phi < 0$ then mean reversion tends to occur more slowly from above than below the long-run mean, indicating that absolute advantage above average is sticky. Only in the level neutral case of $\phi \to 0$ is the rate of mean reversion from above and below the mean the same.

The parameters η and σ in (16) inherit their interpretations from the OU process in (11) as the rate of dissipation and the intensity of innovations, respectively. The intensity of innovations σ again plays a dual role: on the one hand magnifying volatility by scaling up the Wiener innovations and on the other hand regulating how fast time elapses in the deterministic part of the diffusion. This dual role guarantees that the diffusion will have a non-degenerate stationary distribution. Scaling the deterministic part of the diffusion by $\sigma^2/2$ ensures that dissipation occurs at the right speed to offset the unbounded random walk that the Wiener process would otherwise induce for each country-industry. Under the generalized logistic diffusion, the dissipation rate η and decay elasticity ϕ jointly determine the heavy tail of the cross-sectional distribution of comparative advantage, with the intensity of innovations σ determining the speed of convergence to this distribution but having no effect on its shape.

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

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

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

[34] If η were negative, comparative advantage would collapse over time for $\phi < 0$ or explode for $\phi \geq 0$. We do not constrain η to be finite.

4.2 The cross-sectional distributions of comparative and absolute advantage

If comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion by (16), then the stationary distribution of comparative advantage is a generalized gamma distribution with density (15) and parameters $\hat{\theta} = \left(\phi^2/\eta\right)^{1/\phi} > 0$ and $\kappa = 1/\hat{\theta}^{\phi} > 0$ by Lemma 1. From this stationary distribution of comparative advantage $\hat{A}_{is}$, we can infer the cross-sectional distribution of absolute advantage $A_{is}(t)$. Note that, by definition (14), absolute advantage is not necessarily stationary because the stochastic trend $Z_s(t)$ may not be stationary.

Absolute advantage is related to comparative advantage through a country-wide stochastic trend by definition (14). Plugging this definition into (15), we can infer that the probability density of absolute advantage must be proportional to

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

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

Lemma 2. *If comparative advantage $\hat{A}_{is}(t)$ follows a generalized logistic diffusion*

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

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

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

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

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

for the strictly positive parameters

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

Proof. Derivations above establish that the cross-sectional distributions are generalized gamma. The cumulative distribution functions follow from Kotz et al. (1994, Ch. 17, Section 8.7). □

Returning to the graphs in **Figure 2**, Lemma 2 clarifies that a country-wide stochastic trend $Z_s(t)$ shifts log absolute advantage horizontally and that the shape related parameter κ is not country specific if comparative advantage follows a diffusion with a common set of three deep parameters $(\hat{\theta}, \kappa, \phi)$ worldwide.

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

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

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

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

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

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

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

4.3 Implementation

The generalized logistic diffusion model (16) has no known closed-form transition density when $\phi \neq 0$. We therefore cannot evaluate the likelihood of the observed data and cannot perform maximum likelihood estimation. However, an attractive feature of the generalized logistic diffusion is that it can be transformed into a diffusion that belongs to the Pearson-Wong family, for which closed-form solutions of the conditional moments exist.[35] We construct a consistent GMM estimator based on the conditional moments of a transformation of comparative advantage, using results from Forman and Sørensen (2008).

Our model depends implicitly on the unobserved stochastic trend $Z_s(t)$. We use a closed form expression for the mean of a log-gamma distribution to identify and concentrate out this trend. For a given country and year, the cross-section of the data across industries has a generalized gamma distribution. The mean of the log of this distribution can be calculated explicitly as a function of the model parameters, enabling us to identify the trend

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

from the relation that $\mathbb{E}_{st}[\ln \hat{A}_{is}(t)] = \mathbb{E}_{st}[\ln A_{is}(t)] - \ln Z_s(t)$ by definition (14). We adopt the convention that the expectations operator $\mathbb{E}_{st}[\cdot]$ denotes the conditional expectation for source country s at time t. This result is summarized in the following proposition:

Proposition 1. *If comparative advantage $\hat{A}_{is}(t)$ follows the generalized logistic diffusion* (16) *with real parameters η, σ, ϕ ($\eta > 0$), then the country specific stochastic trend $Z_s(t)$ is recovered from the first moment of the logarithm of absolute advantage as:*

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

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

Proof. See Appendix B. □

This proposition implies that for any GMM estimator, and at any iterative estimation step, we can obtain detrended data based on the sample analog of equation (19):

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

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

Next, we perform a change of variable to recast our model as a Pearson-Wong diffusion, which 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 on time series data.[36] The following proposition presents an invertible transformation of comparative advantage that makes estimation possible.

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

- *The transformed variable*

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

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

follows the diffusion

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

and belongs to the Pearson-Wong family.

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

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

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

Proof. See Appendix C. □

Consider time series observations for $\hat{B}_{is}(t)$ at times $t_1, \ldots, t_T$.[37] By (22) in Proposition 2, we can calculate a closed form for the conditional moments of the transformed diffusion at time t_τ conditional on the information set at time $t_{\tau-1}$. We then compute the forecast error based on using these conditional moments to forecast the m-th power of $\hat{B}_{is}(t_\tau)$ with time $t_{\tau-1}$ information. These forecast errors are uncorrelated with any function of past $\hat{B}_{is}(t_{\tau-1})$. We can therefore construct a GMM criterion for estimation. Denote the forecast error with

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

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

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

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

[37] Taking advantage of our continuous time more, our estimation does not require evenly spaced observations. For estimation based a given horizon of h (say 5 years), we include in our sample every possible pair of observations such that no observation is used in more than two pairs, the gap between observations is at least h years, and each time series gap is made as small as possible.

instruments for each of the M moments that we include in our GMM criterion.[38] In estimation, we use $K = 2$ instruments and $M = 2$ conditional moments, providing us with $K \cdot M = 4$ equations and overidentifying the three parameters (η, σ, ϕ).

Let T_{is} denote the number of time series observations available in industry i and country s. Given sample size of $N = \sum_i \sum_s T_{is}$, our GMM estimator solves the minimization problem

$$(\eta^*, \sigma^*, \phi^*) = \arg \min_{(\eta,\sigma,\phi)} \left(\frac{1}{N} \sum_i \sum_s \sum_\tau g_{is\tau}(\eta, \sigma, \phi) \right)' \mathbf{W} \left(\frac{1}{N} \sum_i \sum_s \sum_\tau g_{is\tau}(\eta, \sigma, \phi) \right) \tag{23}$$

for a given weighting matrix $\mathbf{W}$. Being overidentified, we adopt a two-step estimator. On the first step we compute an identity weighting matrix, which provides us with a consistent initial estimate. On the second step we update the weighting matrix to an estimate of the optimal weighting matrix by setting the inverse weighting matrix to $\mathbf{W}^{-1} = (1/N) \sum_i \sum_s \sum_\tau g_{is\tau}(\eta, \sigma, \phi) g_{is\tau}(\eta, \sigma, \phi)'$, which is calculated at the parameter value from the first step. Forman and Sørensen (2008) establish asymptotics for a single time series as $T \to \infty$.[39] For estimation, we impose the constraints that $\eta > 0$ and $\sigma^2 > 0$ by reparametrizing the model in terms of $\ln \eta > -\infty$ and $2 \ln \sigma > -\infty$. We evaluate the objective function (23) at values of (η, σ, ϕ) by detrending the data at each iteration to obtain $\hat{A}_{is}^{\text{GMM}}(t)$ from equation (20), transforming these variables into their mirror variables $\hat{B}_{is}^{\text{GMM}}(t) = [\hat{A}_{is}^{\text{GMM}}(t)^{-\phi} - 1]/\phi$, and using equation (22) to compute forecast errors. Then we calculate the GMM criterion function for each industry and country pair by multiplying these forecast errors by instruments constructed from $\hat{B}_{is}^{\text{GMM}}(t)$, and finally sum over industries and countries to arrive at the value of the GMM objective.

Standard errors of our estimates need to account for the preceding estimation of our absolute advantage $\ln A_{is}(t)$ measures. Newey and McFadden (1994) present a two-step estimation method for GMM, which accounts for the presence of generated (second-stage) variables that are predicted (from a first stage). In contrast to that generated-regressor correction, our absolute advantage $\ln A_{is}(t)$ measures are not predicted variables but parameter estimates from a gravity equation: $\ln A_{is}(t)$ is a normalized version of the estimated exporter-sector-year fixed effect in equation (6). The Newey-McFadden results require a constant first stage parameter, but the number of parameters we estimate in our first stage increases with our first stage sample size. Moreover, the moments in GMM time series estimation here (just as the variables in OLS decay estimation in Section 3.2 above)

[38] We work with a suboptimal 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 a global numerical optimization algorithm. For these computational concerns we sacrifice efficiency and use suboptimal instruments.

[39] Our estimator would also fit into the standard GMM framework of Hansen (1982), which establishes consistency and asymptotic normality of our second stage estimator as $IS \to \infty$. To account for the two-step nature of our estimator, we use an asymptotic approximation where each dimension of our panel data gets large simultaneously (see Appendix D).

involve pairs of parameter estimates from different points in time—$\ln A_{is}(t)$ and $\ln A_{is}(t+\Delta)$—and thus require additional treatments of induced covariation in the estimation. In Appendix D, we extend Newey and McFadden (1994) to our specific finite-sample context, which leads to an alternative two-step estimation method that we employ for the computation of standard errors. We use the multivariate delta method to calculate standard errors for transformed functions of the estimated parameters.

5 Estimates

Following the GMM procedure described in Section 4.3, we proceed to estimate the dissipation rate η, innovation intensity σ, and decay elasticity ϕ in the diffusion of comparative advantage, subject to an estimated country-specific stochastic trend $Z_s(t)$. The trend allows absolute advantage to be non-stationary but because it is common to all industries in a country has no bearing on comparative advantage. Thus, we aim to describe the global evolution of comparative advantage using just three parameters, which must apply to all industries in all countries and in all time periods. This approach contrasts sharply with our descriptive exercise in **Figure 2**, where we fit the cumulative distribution to the log normal and Pareto distributions separately for each country and each year. Specifying the GLD allows us further to test the strong distributional assumptions implicit in the OLS estimation of the discretized OU process in **Table 1**.

5.1 GMM results for the Generalized Logistic Diffusion

Table 2 presents our baseline GMM estimation results using moments on five-year or single-year intervals. The parameter η measures the rate at which shocks dissipate in continuous time so that log comparative advantage reverts towards the global mean. In combination with the decay elasticity ϕ, the dissipation rate η controls both the magnitude of the long-run mean and the curvature of the cross-sectional distribution. Across specifications in **Table 2**, η takes a value of about one-quarter for gravity estimates of export capability (odd-numbered columns) and about one-fifth for the Balassa RCA index (even-numbered columns). To gain intuition about η, suppose the intensity of innovations of the Wiener process is unity ($\sigma = 1$) and comparative advantage is sticky neither above nor below the mean ($\phi = 0$). Then a value of η equal to 0.25 means that it will take 5.5 years for half of the initial shock to log comparative advantage to dissipate and 18.4 years for 90 percent of the initial shock to dissipate.[40] Alternatively, if η equals 0.20 it will take 6.9 years for half of the initial shock to decay.

The sign of ϕ captures asymmetry in mean reversion for comparative advantage. The estimate of ϕ is negative

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

Table 2: GMM ESTIMATES OF COMPARATIVE ADVANTAGE DIFFUSION

	5-year transitions						1-yr. trans.	
	Full sample		LDC exp.		Non-manuf.		Full sample	
	$\ln A$	$\ln RCA$	$\ln A$	$\ln RCA$	$\ln A$	$\ln RCA$	$\ln A$	$\ln RCA$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Generalized Logistic Diffusion Parameters								
Dissipation rate η	0.256	0.212	0.270	0.194	0.250	0.185	0.256	0.211
	(0.005)***	(0.003)***	(0.006)***	(0.004)***	(0.006)***	(0.006)***	(0.004)***	(0.005)***
Intensity of innovations σ	0.745	0.713	0.842	0.789	0.875	0.712	1.434	1.068
	(0.01)***	(0.024)***	(0.015)***	(0.025)***	(0.018)***	(0.028)***	(0.026)***	(0.035)***
Elasticity of decay ϕ	-0.040	0.006	-0.066	-0.011	-0.033	-0.026	-0.040	-0.065
	(0.016)**	(0.025)	(0.022)***	(0.025)	(0.017)*	(0.024)	(0.015)***	(0.023)***
Implied Parameters								
Log gen. gamma scale $\ln \hat{\theta}$	127.080	-1424.400	62.229	652.090	161.920	213.770	128.010	59.850
	(71.847)*	(7244.308)	(30.520)**	(1818.956)	(116.263)	(267.438)	(69.392)*	(31.538)*
Log gen. gamma shape $\ln \kappa$	5.079	8.657	4.120	7.337	5.412	5.596	5.089	3.902
	(0.817)***	(8.282)	(0.661)***	(4.367)*	(1.056)***	(1.821)***	(0.783)***	(0.682)***
Mean/median ratio	8.108	10.259	8.064	13.913	8.337	17.724	8.104	15.645
Observations	392,850	392,860	250,300	250,300	190,630	190,630	439,810	439,820
Industry-source obs. $I \times S$	11,542	11,542	7,853	7,853	5,845	5,845	11,854	11,854
Root mean sq. forecast error	1.852	1.76	2.018	1.931	1.969	1.903	1.735	1.693
Min. GMM obj. ($\times$ 1,000)	3.06e-13	6.79e-12	7.14e-13	2.16e-11	6.70e-13	2.47e-11	4.76e-14	3.36e-12

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org.
Note: GMM estimation at the five-year (and one-year) horizon for the generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$,

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

using absolute advantage measures $A_{is}(t) = \hat{A}_{is}(t) Z_s(t)$, and the Balassa index of revealed comparative advantage $RCA_{ist} = (X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_j X_{jst}/\sum_j \sum_{\varsigma} X_{j\varsigma t})$. Parameters η, σ, ϕ are estimated under the constraints $\ln \eta, \ln \sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (21), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (18). Less developed countries (LDC) as listed in Appendix E. The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

and statistically significantly different from zero in all specifications for gravity-based comparative advantage, but for the Balassa RCA index ϕ is imprecisely estimated at the five-year time interval. Thus, for gravity-based comparative advantage 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. Industries that randomly churn into the upper tail of the cross section will tend to retain their comparative advantage for longer than those below the mean, affording high-advantage industries with opportunities to reach higher levels of comparative advantage as additional innovations arrive. However, the value of ϕ is not far from zero, suggesting that in practice deviations of comparative advantage from log normality may be modest. We explore this possibility in more detail in Section 5.4.

The parameter σ measures both the volatility of the Wiener innovations to comparative advantage and the speed of convergence on the deterministic dissipation. This dual role binds the parameter estimate of σ to a level such that a non-degenerate stationary distribution of comparative advantage exists. The intensity of innovations therefore does not play a role in determining the cross-sectional distribution. That job is performed by κ and $\hat{\theta}$, which depend exclusively on η and ϕ, so that we are describing the shape and scale of the cross-sectional distribution with just two parameters. Similar to the earlier decay regressions in Section 3.2, our estimates for η and σ are stable across subsamples. In the LDC subsample (column 3) η is slightly higher, while in the manufacturing subsample (column 5) it is slightly reduced. The estimate of ϕ is also robust. In the LDC subsample, the negative value of ϕ gains in absolute magnitude (reflecting increased stickiness of comparative advantage from above) but this estimate is not significantly different from the point estimate for the full sample. In the nonmanufacturing subsample, ϕ is again not statistically significantly different from the full-sample estimate.

The parameter estimates for η and ϕ together imply that the distributions of absolute and comparative advantage have considerable mass in the upper tail. The mean exceeds the median by a factor of more than eight for gravity-based comparative advantage, both among developing and industrialized countries. This concentration of gravity-based comparative advantage is slightly more pronounced in nonmanufacturing industries (column 5). When we use the Balassa RCA index, the mean/median ratio increases to 10 (column 2) in the full sample and exceeds 17 in the nonmanufacturing sample (column 6). One interpretation of the greater concentration in revealed comparative advantage relative to our gravity-based measure is that geography reinforces comparative advantage by making countries appear overspecialized in the goods in which their underlying capability is strong.

To fully describe the time-series dynamics, we go to the shortest time horizon our data allow and estimate the GLD at one-year intervals. Similar to attenuation bias driving estimates of persistence to zero in auto-regression models, one might expect measurement error in the GLD to deliver larger values of σ at shorter horizons. In the limit when σ becomes arbitrarily large, the GLD would exhibit no persistence, converging to an iid process.

Columns 7 and 8 of **Table 2** report GMM estimates of the GLD using moments at the one-year horizon. The estimates for ϕ and η are statistically no different from the estimates at five-year intervals but, consistent with the interpretation that measurement error biases estimation towards faster mean reversion, the estimate of σ nearly doubles for the gravity-based measure of comparative advantage and increases by a factor of 1.5 for RCA.

To compare estimates to the previous OLS decay regression, and to further assess the sensitivity of σ to the choice of interval length, we next fit the GLD to moments based on ten-year intervals. **Table A7** in the Appendix shows the results. We obtain estimates comparable to those at the five-year horizon, with the dissipation rate η still around one-quarter for gravity-based comparative advantage and with a somewhat lower value of σ. There is a less pronounced drop in σ as we go from five- to ten-year intervals, compared to the more marked drop when moving from the one- to the five-year horizon, which suggests that measurement error in the GLD is less of a concern when expanding the time horizon beyond five years.

By relating GLD parameter estimates back to those from the decay regressions in Section 3.2, we can assess the importance of relaxing the assumption of log normality for comparative advantage. Because the GLD nests the Ornstein-Uhlenbeck process as a limiting case with $\phi \to 0$ (and finite η), the dissipation rate η and the intensity of innovations σ carry over one-for-one between the two models. We can therefore compare GLD parameter estimates of η and σ at the decade horizon in **Table A7** to the ten-year decay estimates from specification (11) reported in **Table 1**. GLD and OLS estimates of η are very similar (0.27 versus 0.28 for gravity-based comparative advantage and 0.23 versus 0.22 for the RCA index), as are estimates for the innovation intensity σ. Although we reject the OU process in favor the GLD in five of the eight samples, the similarity of the estimates in the two models suggest that imposing log normality introduces little in the way of specification bias.

As an additional robustness check, we repeat the GMM procedure using MPML-based estimates of gravity fixed effects that account for zero trade flows. Appendix **Table A4** reports GMM estimates of the GLD at the five-year and one-year horizons for MPML-based export capability. A comparison to our benchmark estimates in **Table 2** shows little change in coefficient estimates for η, σ, or ϕ. Finally, we re-estimate the GLD on data for the post-1984 period with SITC revision 2 industries at the two-, three- and four-digit levels. We report the results in Appendix **Table A6**. The findings are largely in line with those in **Table 2**. Estimates of the dissipation rate η are slightly larger for the post-1984 period than for the full sample period but become smaller as we move from broader to narrower industry classifications. Estimates of the decay elasticity ϕ are negative but only at the two-digit level is ϕ statistically significantly different from zero.

5.2 Model fit I: Matching the empirical cross-section distribution

We have given the GMM estimator a heavy task: to fit the export dynamics across 90 countries for 46 years using only three time invariant parameters (η, σ, ϕ), conditional on stochastic country-wide growth trends. The moments we use in GMM estimation are based on five-year intervals (in our benchmark specification) that reflect the time-series behavior of country-industry exports. In other words, our estimator targets the diffusion of comparative advantage but not its stationary cross-sectional distribution. We can therefore use the stationary generalized gamma distribution implied by the GLD process to assess how well our model captures the heavy tails of export advantage observed in the repeated cross-section data. For this comparison, we use the benchmark estimates from **Table** 2 at five-year intervals in column 1.

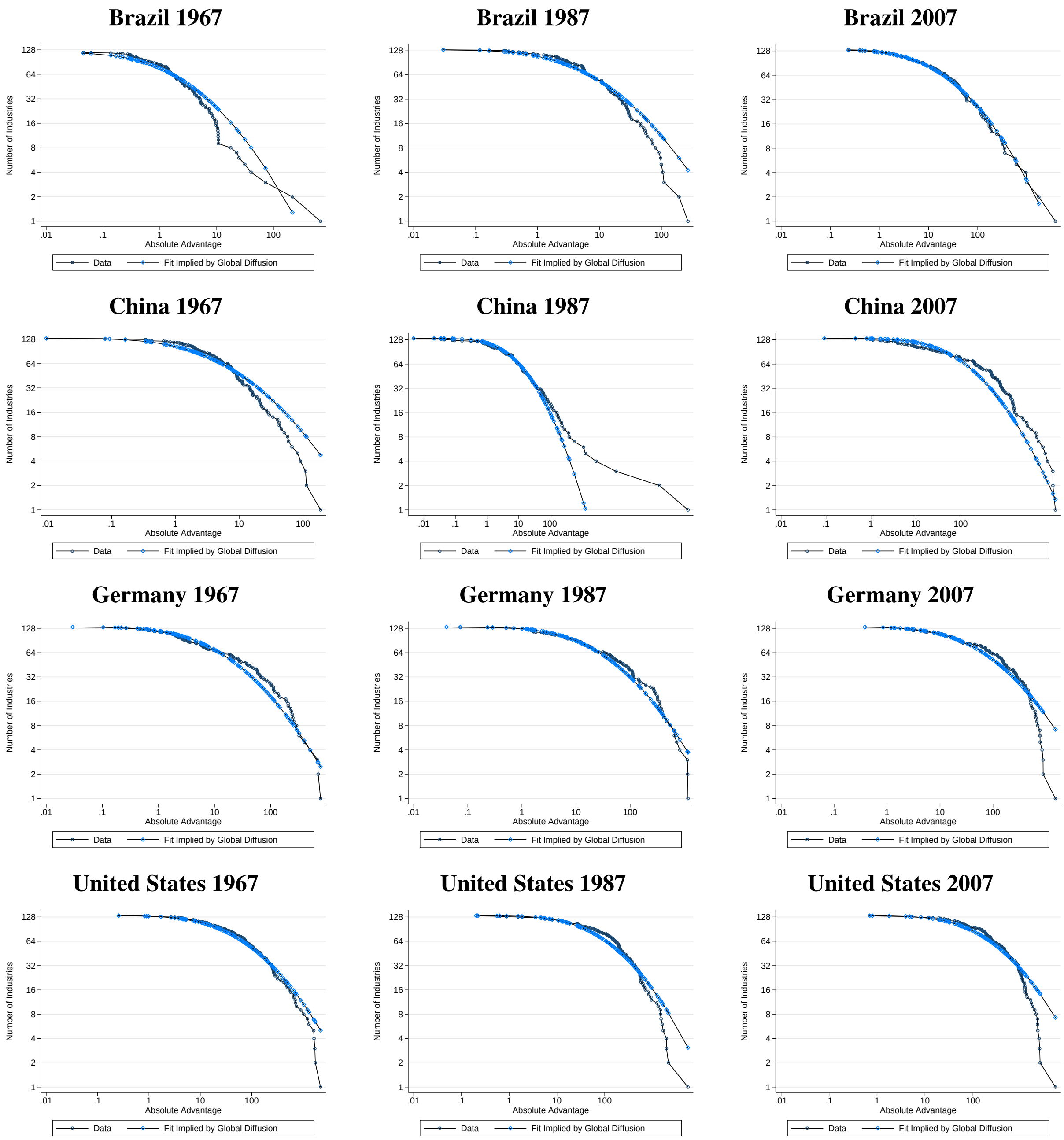
For each country in each year, we project the cross-sectional distribution of comparative advantage implied by the parameters estimated from the diffusion and compare it to the empirical distribution. To implement this validation exercise, we need a measure of $\hat{A}_{ist}$ in (14), the value of which depends on the unobserved country-specific stochastic trend Z_{st}. This trend accounts for the observed horizontal shifts in distribution of log absolute advantage over time, which may result from country-wide technological progress, factor accumulation, or other sources of aggregate growth. In the estimation, we concentrate out Z_{st} by (19), which allows us to estimate its realization for each country in each year. Combining observed absolute advantage A_{ist} with the stochastic-trend estimate allows us to compute realized values of comparative advantage $\hat{A}_{ist}$.

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

Figure 4 compares plots of the actual data against the GLD-implied distributions for four countries in three years, 1967, 1987, 2007. **Figures A4, A5** and **A6** in the Appendix present plots in these years for the 28 countries that are also shown in **Figures A1, A2** and **A3**.[41] While **Figures A1** through **A3** depict Pareto and log normal maximum likelihood estimates of each individual country's cross-sectional distribution by year (number

[41] Because the country-specific trend Z_{st} shifts the implied stationary distribution horizontally, to clarify fit 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.

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



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 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 $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

of parameters estimated = number of countries × number of years), our exercise now is vastly more parsimonious and based on a fit of the time-series evolution, not the observed cross sections. **Figure 4** and **Figures A4** through **A6** show that the empirical distributions and the GLD-implied distributions have the same concave shape and horizontally shifting position. Considering that the shape of the distribution effectively depends on only two parameters for all country-industries and years, the GLD-predicted distributions are remarkably accurate. There are important differences between the actual and predicted plots in only a few countries and a few years, including China in 1987, Russia in 1987 and 2007, Taiwan in 1987, and Vietnam in 1987 and 2007. Three of these four cases involve countries undergoing a transition away from central planning during the designated time period, suggesting periods of economic disruption.

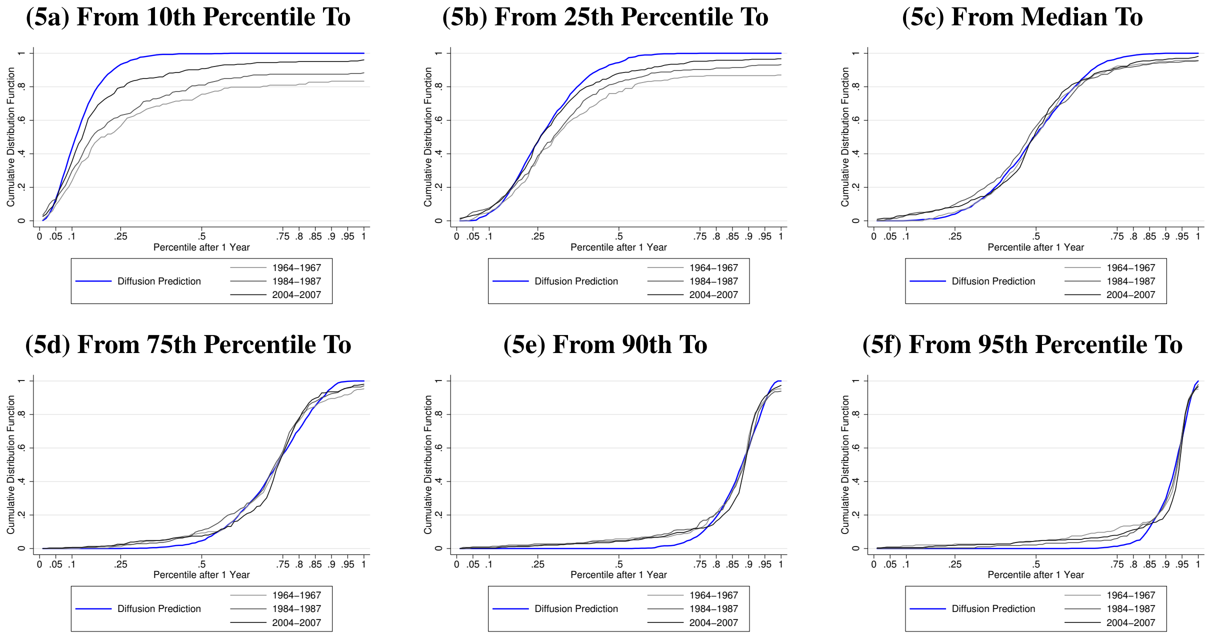
There are other, minor discrepancies between the empirical distributions and the GLD-implied distributions that merit further attention. In 2007 in a handful of countries in East and Southeast Asia—China, Japan, Rep. Korea, Malaysia, Taiwan, and Vietnam—the empirical distributions exhibit less concavity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

5.3 Model fit II: Matching dynamic transition probabilities

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

Figure 5 shows empirical and model-predicted conditional cumulative distribution functions for annual transitions of comparative advantage. We pick select percentiles in the start year: the 10th and 25th percentile, the median, the 75th, 90th and 95th percentile. The left-most upper panel in **Figure 5**, for example, considers industries that were at the tenth percentile of the cross-section distribution of comparative advantage in the start year; panel **Figure 5c** shows industries that were at the median of the distribution in the start year. Each curve in a panel then plots the conditional CDF for the transitions from the given percentile in the start year to any percentile of the cross section one year later. By design, data that are re-sampled under an iid distribution would show up at a 45-degree line, while complete persistence of comparative advantage would make the CDF a step

Figure 5: **Diffusion Predicted Annual Transitions**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity measures of absolute advantage (7).
Note: Predicted cumulative distribution function of comparative advantage $\hat{A}_{is,t+1}$ after one year, given the percentile (10th, 25th, median, 75th, 90th, 95th) of current comparative advantage $\hat{A}_{is,t}$. Predictions based on simulations using estimates from Table 2 (column 1). Observed cumulative distribution function from mean annual transitions during the periods 1964-1967, 1984-1987, and 2004-2007.

function.

To characterize the data, we use three windows of annual transitions: the mean annual transitions during the years 1964-67 around the beginning of our sample period, the mean annual transitions during the years 1984-87 around the middle of our sample, and the mean annual transitions during the years 2004-07 towards the end of the sample. These transitions are shown in grey. Our GLD estimation constrains parameters to be constant over time, so the model predicted transition probabilities give rise to a time-invariant CDF shown in blue.

The five-year GLD performs well in capturing the annual dynamics of comparative advantage for most industries. As **Figure 5** shows, the model-predicted conditional CDF's tightly fit their empirical counterparts for industries at the median and higher percentiles in the start year. It is only in the lower tail, in particular around the 10th percentile, that the fit of the GLD model becomes less close, though the model predictions are more comparable to the data in later than in earlier periods. Country-industries in the bottom tail have low trade volumes, especially in the early sample period, meaning that estimates of the empirical transition probabilities in the lower tail are not necessarily precisely estimated and may fluctuate more over time. **Figure 5** indicates that the dynamic fit becomes relatively close for percentiles at around the 25th percentile. The discrepancies in the lowest tail notwithstanding, for industries with moderate to high trade values, which account for the bulk of global trade, the model succeeds in matching empirical transition probabilities.

5.4 Model fit III: The GLD versus the OU process

The GLD model also affords us with a framework to rigorously assess how well a simple Ornstein-Uhlenbeck process approximates trade dynamics. As a final exercise we estimate the GLD under the constraint $\phi = 0$, which yields the Ornstein-Uhlenbeck process and a stationary log normal distribution in the cross section.

Table 3 presents GMM estimates for the GLD under the constraint that $\phi = 0$, paired with the benchmark GMM estimates from **Table 2** that use gravity-based measures of absolute advantage. We can now formally compare the OU coefficient estimates from the OLS decay regression in Section 3.2 (**Table 1**) to those from the GMM estimation (we informally attempted such a comparison in Section 5.1 before). As suggested by the earlier results, the constrained coefficient estimates of η and σ are close to the unconstrained benchmark estimates across subsamples and for different time horizons. This parameter stability implies that the special case of the OU process captures the broad persistence and overall variability of comparative advantage. Still, the extension to a GLD does help explain the degree of export concentration documented in **Section 3.1**. The estimated mean-median ratio increases from $6.2 - 7.0$ under the constrained estimation of the OU process to $8.1 - 8.3$ under the unconstrained case. As noted above, the negative and significant decay elasticity ϕ implies that the data reject the constrained model in favor of the unconstrained model (as least for gravity-based measures of absolute advantage).

In a statistical horse race between the unconstrained GLD and the OU process, the former clearly wins because we soundly reject that $\phi = 0$ in **Table 3**. Yet, estimating the unconstrained GLD is substantially more burdensome than estimating the simple OU process, which has a convenient linear form when discretized. Moreover, for finer industry aggregates at the three- and four-digit level, the unconstrained GLD no longer clearly wins as it becomes statistically impossible to reject the null hypothesis that $\phi = 0$ (see Appendix **Table A6**).

For both empirical and theoretical modelling, it is important to understand how much is lost by imposing the constraint that comparative advantage has a stationary log normal distribution. To address this economic question, we compare the implied dynamics of the unconstrained GLD and the OU process. Following **Figure 5**, we simulate trajectories of the GLD, once from unconstrained estimates and once from constrained estimates, using coefficients from columns 1 and 2 in **Table 3**. The simulations predict the theoretical transition probabilities over the one-year horizon across percentiles of the cross-section distribution.

Figure 6 shows the empirical cumulative distribution functions for annual transitions of comparative advantage over the full sample period 1962-2007 (in grey) and compares the empirical distribution to the two model-

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

	5-year transitions						1-yr. trans.	
	Full sample		LDC exp.		Non-manuf.		Full sample	
		$\phi = 0$		$\phi = 0$		$\phi = 0$		$\phi = 0$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Generalized Logistic Diffusion Parameters								
Dissipation rate η	0.256	0.264	0.270	0.274	0.250	0.256	0.256	0.266
	(0.005)***	(0.003)***	(0.006)***	(0.003)***	(0.006)***	(0.004)***	(0.004)***	(0.003)***
Intensity of innovations σ	0.745	0.739	0.842	0.836	0.875	0.868	1.434	1.419
	(0.01)***	(0.009)***	(0.015)***	(0.012)***	(0.018)***	(0.015)***	(0.026)***	(0.025)***
Elasticity of decay ϕ	-0.04		-0.066		-0.033		-0.04	
	(0.016)**		(0.022)***		(0.017)*		(0.015)***	
Implied Parameters								
Log gen. gamma scale $\ln \hat{\theta}$	127.080		62.229		161.920		128.010	
	(71.847)*		(30.520)**		(116.263)		(69.392)*	
Log gen. gamma shape $\ln \kappa$	5.079		4.120		5.412		5.089	
	(0.817)***		(0.661)***		(1.056)***		(0.783)***	
Mean/median ratio	8.108	6.668	8.064	6.214	8.337	7.040	8.104	6.539
Observations	392,850	392,850	250,300	250,300	190,630	190,630	439,810	439,810
Industry-source obs. $I \times S$	11,542	11,542	7,853	7,853	5,845	5,845	11,854	11,854
Root mean sq. forecast error	1.852	1.731	2.018	1.827	1.969	1.869	1.735	1.614
Min. GMM obj. (× 1,000)	3.06e-13	2.48e-12	7.14e-13	1.95e-11	6.70e-13	8.24e-12	4.76e-14	1.47e-12

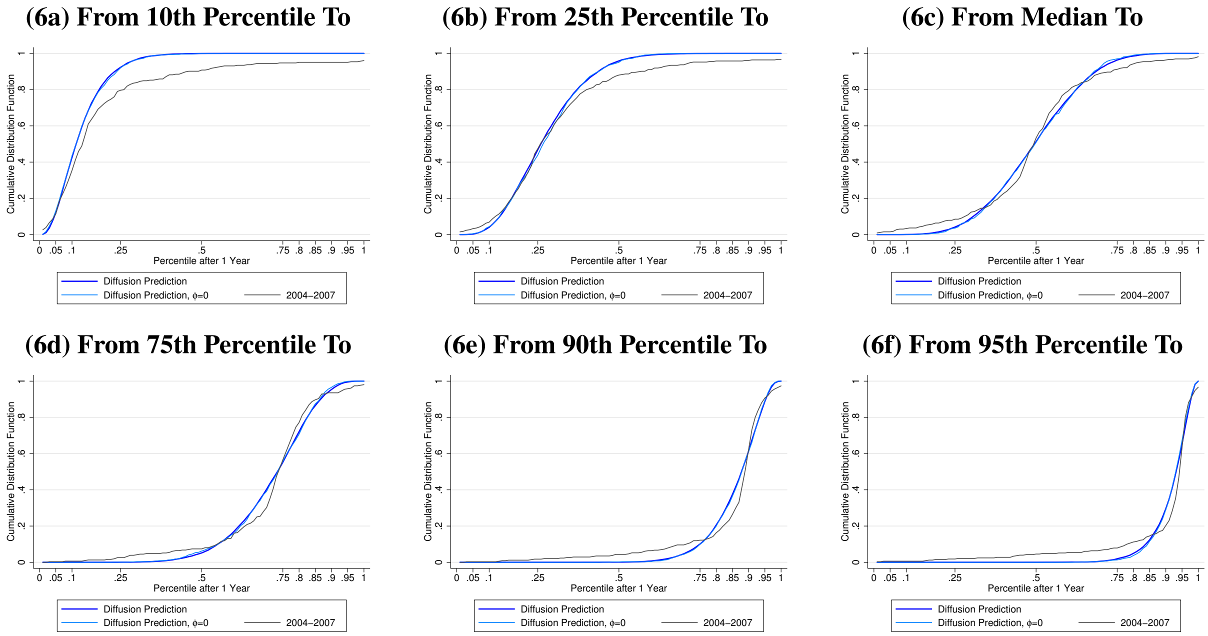
Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity measures of absolute advantage (7).

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

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

using absolute advantage measures $A_{is}(t) = \hat{A}_{is}(t) Z_s(t)$, unrestricted and restricted to $\phi = 0$. Parameters η, σ, ϕ are estimated under the constraints $\ln \eta, \ln \sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (21), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (18). Less developed countries (LDC) as listed in Appendix E. The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

Figure 6: **Diffusion Predicted Annual Transitions, Constrained and Unconstrained ϕ**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007; gravity measures of absolute advantage (7).
Note: Predicted cumulative distribution function of comparative advantage $\hat{A}_{is,t+1}$ after one year, given the percentile (10th, 25th, median, 75th, 90th, 95th) of current comparative advantage $\hat{A}_{is,t}$. Predictions based on simulations using estimates from Table 2 (column 1) and Table 3 (column 2, $\phi = 0$). Observed cumulative distribution function from mean annual transitions during the period 2006-2009.

predicted cumulative distribution functions (light and dark blue), where the fit of the unconstrained GLD model (dark blue) is the same as depicted in **Figure 5** above. Each panel in **Figure 6** considers industries that were at a given percentile of the cross-section distribution of comparative advantage in the start year: the left-most upper panel in **Figure 6**, for instance, considers industries that were at the tenth percentile in the start year; the right-most lower panel shows industries that were at the 95th percentile in the start year. Each curve in a panel shows the conditional CDF for the transitions from the given percentile in the start year to any percentile of the cross section one year later. For all start-year percentiles, the model-predicted transitions hardly differ between the constrained specification (light blue) and the unconstrained specification (dark blue). When alternating between the two models, the shapes of the model-predicted conditional CDF's are very similar, even in the upper tail. In the lower tail, where the GLD produces the least tight dynamic fit, the constrained OU specification performs no worse than the unconstrained GLD. We conclude that while the GLD extension is important for capturing hyper-specialization in the upper tail of the cross-section distribution precisely, the simple OU process approximates the empirical dynamics in a manner that is similar to the GLD extension.

6 Conclusion

The traditional Ricardian trade model has long presented a conundrum to economists. Although it offers a simple and intuitive characterization of comparative advantage, it yields knife-edge predictions for country specialization patterns that fit the data poorly. Eaton and Kortum (2002) have reinvigorated the Ricardian framework. By treating the capability of firms from a country in a sector as probabilistic rather than deterministic, they derive realistically complex country specialization patterns and provide a robust framework for quantitative analysis. The primitives in the EK model are the parameters of the distribution for industry productivity, which pin down country export capabilities. Comparative advantage arises from these capabilities varying across countries. Our goal in this paper is to characterize the dynamic empirical properties of export capability in order to inform modelling of the deep origins of comparative advantage.

The starting point for our analysis is two strong empirical regularities in trade that economists have studied incompletely and in isolation. Many papers have noted the tendency for countries to concentrate their exports in a relatively small number of sectors. Our first contribution is to show that this concentration arises from a heavy-tailed distribution of industry export capability that is approximately log normal and whose shape is stable across countries, sectors, and time. Likewise, the trade literature has detected in various forms a tendency for mean reversion in national industry productivities. Our second contribution is to establish that mean reversion in export capability, rather than indicative of convergence in productivities and degeneracy in comparative advantage, is instead consistent with a disciplined stochastic growth process, whose properties are also common across borders and sectors. In literature on the growth of cities and the growth of firms, economists have used stochastic processes to study the determinants of the long-run distribution of sizes. Our third contribution is to develop an analogous empirical framework for identifying the parameters that govern the stationary distribution of export capability. The main result of this analysis is that, while a generalized gamma distribution provides the best fit, log normality offers a reasonable approximation. The stochastic process that generates log normality can be estimated in its discretized form by simple linear regression.

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

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.

Alvarez, Fernando E., Francisco J. Buera, and Robert E. Jr. Lucas, "Idea Flows, Economic Growth, and Trade," *NBER Working Paper*, 2013, *19667.*

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.

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.

Buera, Francisco J. and Ezra Oberfield, "The Global Diffusion of Ideas," May 2014. Federal Reserve Bank of Chicago, unpublished manuscript.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *Universitat Pompeu Fabra Economics Working Paper*, June 2015, *1403.*

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.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, January 2015, *82* (1), 1–44.

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

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

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

_ **and Andrés Rodríguez-Clare**, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Elhanan Helpman, Kenneth Rogoff, and Gita Gopinath, eds., *Handbook of International Economics*, Vol. 4, Amsterdam: Elsevier, 2014, chapter 4, pp. 197–261.

_ **, 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.

Di Giovanni, Julian, Andrei A. Levchenko, and Jing Zhang, "The Global Welfare Impact of China: Trade Integration and Technological Change," *American Economic Journal: Macroeconomics*, July 2014, *6* (3), 153–83.

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

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

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

Hausmann, Ricardo and Dani Rodrik, "Economic Development as Self-Discovery," *Journal of Development Economics*, December 2003, *72* (2), 603–33.

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

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

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

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

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

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

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

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

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

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

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

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

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

Newey, Whitney K. and Daniel L. McFadden, "Large Sample Estimation and Hypothesis Testing," in Robert F. Engle and Daniel L. McFadden, eds., *Handbook of Econometrics*, Vol. 4, Amsterdam: Elsevier Science, 1994, chapter 23, pp. 2111–2245.

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.

Rodríguez-Clare, Andrés, "Costa Rica's Development Strategy Based on Human Capital and Technology: How It Got There, the Impact of Intel, and Lessons for Other Countries," *Journal of Human Development*, July 2001, *2* (2), 311–24.

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.

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

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

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

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

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

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

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

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

Appendix

A Generalized Logistic Diffusion: Proof of Lemma 1

The ordinary gamma distribution arises as the stationary distribution of the stochastic logistic equation (Leigh 1968). We generalize this ordinary logistic diffusion to yield a generalized gamma distribution as the stationary distribution in the cross section. Note that the generalized (three-parameter) gamma distribution relates to the ordinary (two-parameter) gamma distribution through a power transformation. Take an ordinary gamma distributed random variable X with two parameters $\bar{\theta}, \kappa > 0$ and the density function

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

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

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

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

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

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

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

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

as in (A.1) with

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

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

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

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

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

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

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

with the parameters

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

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

We infer that the stationary distribution of $A(\infty) = \lim_{t\to\infty} A(t)$ is a generalized gamma distribution by (15) and by the derivations above:

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

with

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

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

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

Our preferred parametrization (16) of the generalized logistic diffusion in Lemma 1 is

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

for real parameters η, σ, ϕ. That parametrization can be related back to the parameters in (A.5) by setting $\alpha = (\sigma^2/2) + \beta$ and $\beta = \eta\sigma^2/(2\phi)$. In this simplified formulation, the no-collapse and no-explosion conditions are satisfied for the single restriction that $\eta > 0$. The reformulation in (16) also clarifies that one can view our generalization of the drift term $[\hat{A}_{is}(t)^{\phi} - 1]/\phi$ as a conventional Box-Cox transformation of $\hat{A}_{is}(t)$ to model the level dependence.

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

B Trend Identification: Proof of Proposition 1

First, consider a random variable X which has a gamma distribution with scale parameter θ and shape parameter κ. For any power $n \in \mathbb{N}$ we have

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

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

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

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

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

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

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

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

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

as stated in the text.

C Pearson-Wong Process: Proof of Proposition 2

For a random variable X with a standard logistic diffusion (the $\phi = 1$ case), the Bernoulli transformation $1/X$ maps the diffusion into the Pearson-Wong family (see e.g. Prajneshu 1980, Dennis 1989). We follow up on that transformation with an additional Box-Cox transformation and apply $\hat{B}_{is}(t) = [\hat{A}_{is}(t)^{-\phi} - 1]/\phi$ to comparative advantage, as stated in (21). Define $W_{is}^{\hat{B}}(t) \equiv -W_{is}^{\hat{A}}(t)$. Then $\hat{A}_{is}^{-\phi} = \phi\hat{B}_{is}(t) + 1$ and, by Itō's lemma,

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

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

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

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

To construct a GMM estimator based on this Pearson-Wong representation, we apply results in Forman and Sørensen (2008) to construct closed form expressions for the conditional moments of the transformed data and then use these moment conditions for estimation. This technique relies on the convenient structure of the Pearson-Wong class and a general result in Kessler and Sørensen (1999) on calculating conditional moments of diffusion processes using the eigenfunctions and eigenvalues of the diffusion's infinitesimal generator.[42]

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

Given an eigenfunction and eigenvalue pair (h_s, λ_s) of the infinitesimal generator of $\hat{B}_{is}(t)$, we can follow

[42] For a diffusion

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

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

Kessler and Sørensen (1999) and calculate the conditional moment of the eigenfunction:

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

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

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

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

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

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

We now need to match coefficients on terms.

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

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

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

Using the normalization that $\pi_{n,n} = 1$, equation (C.8) implies a recursive condition for these conditional moments:

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

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

To arrive at the result in the second part of Proposition 2, set the parameters as $q_s = \sigma^2(\eta-\phi^2)/2$, $\bar{X}_s = \phi/(\eta-\phi^2)$, $a_s = \phi^2/(\eta-\phi^2)$, $b_s = 2\phi/(\eta-\phi^2)$, and $c_s = 1/(\eta-\phi^2)$. From these parameters, we can construct eigenvalues and their associated eigenfunctions using the recursive condition (C.9). These coefficients correspond to those reported in equation (22).

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

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

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

$\mathbf{\Pi}$ is triangular by construction, and therefore invertible.

D Correction for Generated Variables in GMM Estimation

D.1 Sampling variation in estimated absolute and comparative advantage

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

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

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

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

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

We assume that the two-way least squares dummy variable estimator for each industry time pair it is consistent and asymptotically normal for an individual industry i shipping from source country s to destination d at time t,[43] and state this assumption formally.

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

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

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

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

with $\sigma^2_{it} \equiv \mathbb{E}_{it}v^2_{isdt}$,

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

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

In finite samples, uncertainty as captured by $\mathbf{\Sigma}_{it}$ can introduce sampling variation in second-stage estimation because $\mathbf{k}^{\text{OLS}}_{i\cdot t}$ is a generated variable. To perform an according finite sample correction, we use

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

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

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

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

$$a_{ist} \equiv \ln A_{ist} = k_{ist} - \frac{1}{|\mathcal{S}_{it}|} \sum_{\varsigma \in \mathcal{S}_{it}} k_{i\varsigma t} \quad \text{and} \quad \hat{a}_{ist} \equiv \ln \hat{A}_{ist} = a_{ist} - \frac{1}{|\mathcal{I}_{st}|} \sum_{j \in \mathcal{I}_{st}} a_{jst}. \tag{D.11}$$

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

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

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

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

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

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

The two matrices $\mathbf{A}_t^{\text{OLS}}$ and $\hat{\mathbf{A}}_t^{\text{OLS}}$ are linearly related to the matrix containing our estimates of export capability $\mathbf{K}_t^{\text{OLS}}$. From equation (D.11), the matrix $\mathbf{A}_t^{\text{OLS}}$ is related to $\mathbf{K}_t^{\text{OLS}}$ and $\mathbf{H}_t$ by

$$\textbf{vec}(\mathbf{A}_t^{\text{OLS}}) = \textbf{Trans}(I,S) \underbrace{\begin{pmatrix} \mathbf{I}_S - \frac{[\mathbf{H}_t]_{1\cdot}'[\mathbf{H}_t]_{1\cdot}}{[\mathbf{H}_t]_{1\cdot}[\mathbf{H}_t]_{1\cdot}'} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I}_S - \frac{[\mathbf{H}_t]_{I\cdot}'[\mathbf{H}_t]_{I\cdot}}{[\mathbf{H}_t]_{I\cdot}[\mathbf{H}_t]_{I\cdot}'} \end{pmatrix}}_{\equiv \mathbf{Z}_{IS}(\mathbf{H}_t)} \textbf{vec}[(\mathbf{K}_t^{\text{OLS}})']. \tag{D.12}$$

Here $\textbf{vec}(\cdot)$ stacks the columns of a matrix into a vector and $\textbf{Trans}(I,S)$ is a vectorized-transpose permutation matrix.[44] The function $\mathbf{Z}_{IS}(\mathbf{H}_t)$ maps the matrix $\mathbf{H}_t$ into a block diagonal $IS \times IS$ matrix, which removes the global industry average across countries. The matrix of comparative advantage estimates is then:

$$\textbf{vec}(\hat{\mathbf{A}}_t^{\text{OLS}}) = \underbrace{\begin{pmatrix} \mathbf{I}_I - \frac{[\mathbf{H}_t']_{1\cdot}'[\mathbf{H}_t']_{1\cdot}}{[\mathbf{H}_t']_{1\cdot}[\mathbf{H}_t']_{1\cdot}'} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I}_I - \frac{[\mathbf{H}_t']_{S\cdot}'[\mathbf{H}_t']_{S\cdot}}{[\mathbf{H}_t']_{S\cdot}[\mathbf{H}_t']_{S\cdot}'} \end{pmatrix}}_{\equiv \mathbf{Z}_{SI}(\mathbf{H}_t')} \textbf{vec}(\mathbf{A}_t^{\text{OLS}}) \equiv \mathbf{Z}_{SI}(\mathbf{H}_t')\textbf{vec}(\mathbf{A}_t^{\text{OLS}}). \tag{D.13}$$

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

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

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

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

average across industries.

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

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

We then have the following result.

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

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

and

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

with

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

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

Proof. Assumptions 1 and 2 along with $\bar{D}_{it} \to D \to \infty$ for all (it) implies that $\sqrt{D}(\mathbf{vec}[(\mathbf{K}^{\text{OLS}}_t)'] - \mathbf{vec}[\mathbf{K}'_t]) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^*_t)$. The results then follow from equation (D.12) and equation (D.13). □

D.2 Second-stage generated variable correction

We estimate two time series models which both can be implemented as GMM estimators. For brevity, we focus on GLD estimation here. (We present the case of OLS estimation of the decay regression in Online Supplement S.1, which simply uses a different GMM criterion and absolute advantage as data instead of comparative advantage.) GLD estimation is based on a conditional moment of the form:

$$\mathbf{0} = \mathbb{E}_{is,t-\Delta}\mathbf{g}\left(\boldsymbol{\theta}, \hat{a}_{ist}, \hat{a}_{is,t-\Delta}\right), \tag{D.14}$$

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

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

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

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

Our finite sample analog for second-stage estimation is:

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

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

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

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

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

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

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

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

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

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

Proposition 3. *Suppose that*

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

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

3. *for any given positive definite matrix* W *and for each* N, *there is a unique minimizer of* $Q_N(\theta;\mathbf{W})$ *given by* $\hat{\theta}_N$;
4. *both* $\mathbb{E}_{it}k_{ist}$ *and* $\mathbb{E}_{st}k_{ist}$ *exist and are finite.*

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

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

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

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

where

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

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

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

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

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

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

Then,

$$\mathbf{0} = \tilde{\mathbf{\Lambda}}'_N\mathbf{W}\frac{1}{N}\sum_{i=1}^{I}\sum_{t=2}^{T}\sum_{s\in\mathcal{S}^P_{it}}\frac{N}{I|\mathcal{S}^P_{it}|(T-1)}\left[\mathbf{G}^0_{ist} + \tilde{\mathbf{G}}^1_{ist}(\hat{\theta}_N - \theta_0) + \tilde{\mathbf{G}}^2_{ist}(\hat{a}^{\text{OLS}}_{ist} - \hat{a}_{ist}) + \tilde{\mathbf{G}}^3_{ist}(\hat{a}^{\text{OLS}}_{is\tau^P_{ist}} - \hat{a}_{is\tau^P_{ist}})\right]$$

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

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

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
[\mathbf{G}_t]_{\cdot j} = \mathbf{1}\left\{s(j)\in\mathcal{S}_{i(j),t}^P\right\}&\frac{N}{I\left|\mathcal{S}_{i(j),t}^P\right|(T-1)}\mathbf{G}_{i(j),s(j),t}^2\\
&+ \mathbf{1}\left\{s(j)\in\mathcal{S}_{i(j),\tau_{i(j),s(j),t}^F}^F\right\}\frac{N}{I\left|\mathcal{S}_{i(j),\tau_{i(j),s(j),t}^F}^F\right|(T-1)}\mathbf{G}_{i(j),s(j),\tau_{i(j),s(j),t}^F}^3
\end{aligned}$$

for

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

We can now re-write the sum as

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

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

For an adaption of the GMM generated-variable correction to second-stage OLS estimation, see the Online Supplement S.1.

E Classifications

In this appendix, we report country and industry 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 export 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, Rep. Korea, 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 nonmanufacturing 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.

F Additional Evidence

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

F.1 Top products

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,

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

Country	1987		2007		Country	1987		2007	
Argentina	Maize, unmilled	4.22	Maize, unmilled	5.48	Mexico	Sulphur	3.63	Alcoholic beverages	4.00
	Animal feed	3.88	Oil seed	4.59		Other crude minerals	3.22	Office machines	3.85
Australia	Wool	3.74	Cheese and curd	3.23	Peru	Metal ores & concntr.	4.21	Metal ores & concntr.	6.33
	Jute	3.78	Fresh meat	3.18		Animal feed	4.00	Coffee	4.69
Brazil	Iron ore	3.61	Iron ore	5.18	Philippines	Vegetable oils & fats	3.85	Office machines	4.48
	Coffee	3.39	Fresh meat	4.47		Preserved fruits & nuts	3.54	Electric machinery	3.59
Canada	Sulphur	3.99	Wheat, unmilled	5.16	Poland	Barley, unmilled	5.34	Furniture	2.19
	Iron ore	3.57	Sulphur	3.34		Sulphur	3.26	Glassware	2.77
China	Explosives	7.33	Sound/video recorders	4.99	Rep. Korea	Radio receivers	5.57	Television receivers	6.06
	Maize, unmilled	6.89	Radio receivers	4.71		Television receivers	5.43	Telecomm. equipmt.	5.11
Czech Rep.	Glassware	4.06	Glassware	4.26	Romania	Furniture	3.55	Footwear	3.50
	Prep. cereal & flour	3.69	Road vehicles	3.67		Fertilizers, manuf.	2.73	Silk	3.16
Egypt	Cotton	4.46	Fertilizers, crude	4.34	Russia	Maize, unmilled	5.63	Animal oils & fats	8.11
	Textile yarn, fabrics	2.84	Rice	3.79		Pulp & waste paper	5.04	Fertilizers, manuf.	4.34
France	Electric machinery	3.52	Other transport eqpmt.	3.42	South Africa	Stone, sand & gravel	3.90	Iron & steel	4.17
	Alcoholic beverages	3.47	Alcoholic beverages	3.26		Radioactive material	3.62	Fresh fruits & nuts	3.47
Germany	Road vehicles	4.08	Metalworking mach.	2.78	Taiwan	Explosives	4.74	Television receivers	5.24
	General machinery	4.02	Meters & counters	0.75		Footwear	4.45	Office machines	5.06
Hungary	Margarine	3.21	Telecomm. equipmt.	4.21	Thailand	Rice	4.92	Rice	4.99
	Fresh meat	2.79	Office machines	4.14		Fresh vegetables	4.18	Natural rubber	4.57
India	Tea	4.23	Precious stones	3.89	Turkey	Fresh vegetables	3.45	Glassware	3.35
	Leather	3.92	Rice	3.65		Tobacco, unmanuf.	3.38	Textile yarn, fabrics	3.25
Indonesia	Natural rubber	5.02	Natural rubber	5.24	United States	Office machines	4.00	Other transport eqpmt.	3.49
	Improved wood	4.66	Sound/video recorders	4.88		Other transport eqpmt.	3.29	Photographic supplies	2.63
Japan	Sound/video recorders	6.37	Sound/video recorders	5.97	United Kingd.	Measuring instrmnts.	3.22	Alcoholic beverages	3.30
	Road vehicles	6.17	Road vehicles	5.70		Office machines	3.17	Pharmaceutical prod.	3.16
Malaysia	Natural rubber	6.19	Radio receivers	5.77	Vietnam	Maize, unmilled	7.63	Animal oils & fats	10.26
	Vegetable oils & fats	4.85	Sound/video recorders	5.01		Jute	5.16	Footwear	6.97

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007.
Note: Top two industries for 28 of the 90 countries in 1987 and 2007 in terms of normalized log absolute advantage, relative to the country mean: $\ln A_{ist} - (1/I)\sum_{i'}^{I} \ln A_{i'st}$.

Table A2: **Top Two Industries by Balassa Comparative Advantage**

Country	1987		2007		Country	1987		2007	
Argentina	Cereals, unmilled	3.65	Animal feed	3.74	Mexico	Stone, sand & gravel	2.14	Television receivers	2.20
	Dyeing extracts	3.20	Oil seed	3.34		Sulphur	2.12	Fresh vegetables	1.45
Australia	Wool	3.74	Uranium	4.57	Peru	Metal ores & concntr.	3.79	Animal oils & fats	4.07
	Uranium	3.58	Wool	4.04		Ores, precious metals	3.19	Metal ores & concntr.	3.83
Brazil	Iron ore	3.34	Iron ore	5.18	Philippines	Vegetable oils & fats	3.81	Office machines	4.41
	Preserved fruits/nuts	2.64	Tobacco, unmanuf.	3.21		Pres. fruits & nuts	3.50	Electric machinery	3.51
Canada	Sulphur	2.24	Cinemat. film, exposed	2.64	Poland	Sulphur	3.78	Smoked fish	2.19
	Pulp & waste paper	1.90	Sulphur	2.45		Preserved meat	2.66	Wood manuf.	1.71
China	Silk	3.77	Silk	1.97	Rep. Korea	Travel goods	1.88	Optical instrmnts.	2.15
	Jute	3.30	Travel goods	1.43		Footwear	1.78	Synthetic fibres	1.49
Czech Rep.	Glassware	2.03	Television receivers	1.61	Romania	Jute	3.28	Leather manuf.	3.03
	Metalworking mach.	1.74	Glassware	1.33		Fertilizers, manuf.	2.71	Silk	2.03
Egypt	Cotton	4.47	Fertilizers, crude	3.70	Russia	Ferrous scrap metal	2.58	Fertilizers, manuf.	3.00
	Vegetable fibres	2.57	Vegetable fibres	3.63		Raw furskins	5.02	Radioactive material	2.72
France	Alcoholic beverages	1.70	Vegetable fibres	2.21	South Africa	Uranium	3.92	Ores, precious metals	3.09
	Radioactive material	1.49	Alcoholic beverages	1.72		Ores, precious metals	3.02	Natural abrasives	2.71
Germany	Synthetic dye	0.95	Other man-made fibres	0.99	Taiwan	Travel goods	2.09	Optical instrmnts.	2.27
	Dyeing extracts	0.77	Meters & counters	0.75		Footwear	1.96	Synthetic fibres	1.51
Hungary	Preserved meat	2.22	Television receivers	1.63	Thailand	Rice	3.93	Natural rubber	3.24
	Crude animal materials	2.17	Maize, unmilled	1.58		Cereal meals & flour	3.31	Rice	3.03
India	Tea	3.81	Iron ore	2.68	Turkey	Tobacco unmanuf.	3.17	Tobacco unmanuf.	2.28
	Spices	3.10	Precious stones	2.65		Crude minerals	2.45	Lime, cement	2.11
Indonesia	Improved wood	3.90	Natural rubber	3.57	United States	Maize, unmilled	1.42	Maize, unmilled	1.64
	Natural rubber	3.88	Vegetable oils & fats	3.03		Oil seed	1.39	Cotton	1.56
Japan	Sound/video recorders	1.76	Photographic supplies	1.36	United Kingd.	Cinemat. film, exposed	1.47	Alcoholic beverages	1.27
	Photographic eqpmnt.	1.33	Photographic eqpmnt.	1.23		Precious stones	1.41	Rags	1.07
Malaysia	Natural rubber	3.90	Proc. animal/plant oils	2.71	Vietnam	Fresh shellfish	4.59	Rice	3.74
	Proc. animal/plant oils	3.56	Vegetable oils & fats	2.36		Spices	3.42	Natural rubber	3.12

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 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 log revealed comparative advantage, using the log Balassa (1965) index $\ln RCA_{ist} = \ln(X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_j X_{jst}/\sum_j \sum_{\varsigma} X_{j\varsigma t})$.

we normalize log absolute advantage by its country mean: $\ln A_{ist} - (1/I)\sum_j^I \ln A_{jst}$. 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 less its average log deviation across all industries. For comparison, **Table A2** presents the top two products in terms of the Balassa RCA index.

F.2 Cumulative probability distribution of log absolute advantage

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).

F.3 Gravity-based comparative advantage from multinomial pseudo-maximum likelihood estimation

We re-estimate exporter-industry-year fixed effects under the distributional assumptions of Eaton et al. (2012), as described in Section 2, and use multinomial pseudo-maximum likelihood (MPML) on (8). With the resulting gravity-based export capability measures at hand, we re-estimate the decay regression (10) at then-year intervals and the GLD model (23) using GMM at five-year intervals.

For the decay regression, **Table A3** restates the benchmark results (from **Table 1**) in columns 1, 3 and 5 and contrasts them with the respective results for MPML-estimated export capability in columns 2, 4 and 6. MPML estimated absolute advantage exhibits both a somewhat stronger rate of decay ρ in the full sample as well as the LDC and nonmanufacturing subsamples and a larger residual variance. When translated into the rate of dissipation η and the intensity of innovations σ, the difference in coefficient estimates loads onto both η, which drops from slightly above one-quarter for OLS-estimated absolute advantage to slightly below one-quarter for MPML-estimated absolute advantage, and σ, which increases.

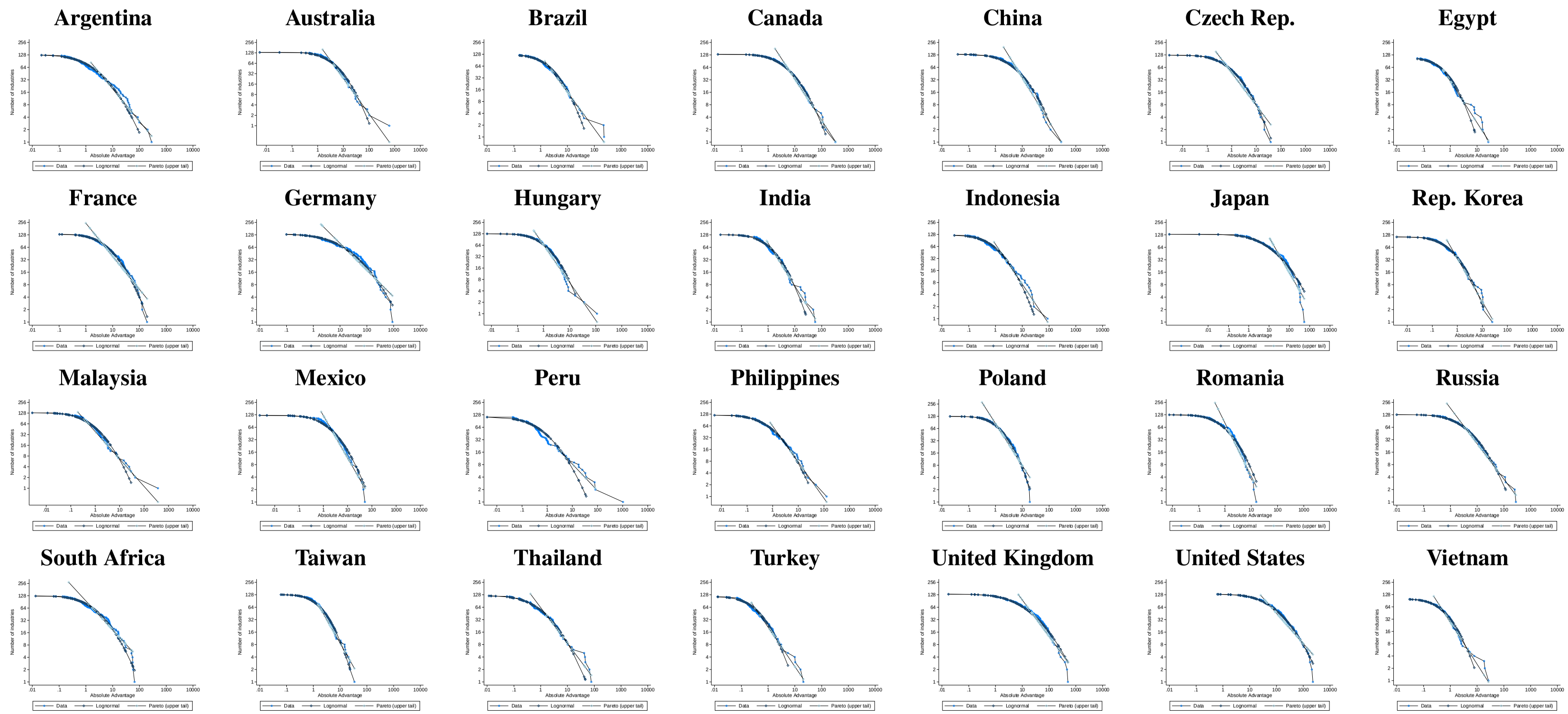
For GLD estimation with GMM at the five-year and one-year horizons, **Table A4** restates the benchmark results (from **Table 2**) in columns 1, 3, 4 and 7 and contrasts them with the respective results for MPML-estimated absolute advantage in columns 2, 4, 6 and 8.

F.4 Comparative advantage at varying industry aggregates

As a robustness check, we restrict the sample to the period 1984-2007 with industry aggregates from the SITC revision 2 classification. Data in this late period allow us to construct varying industry aggregate, from the four-digit STIC revision 2 level to the two-digit level. We first obtain gravity-based estimates of log absolute advantage from OLS (6) at the refined industry aggregates. Following our benchmark specifications in the text, we then estimate the decay regression (10) at ten-year intervals and the GLD model (23) using GMM at five-year intervals.

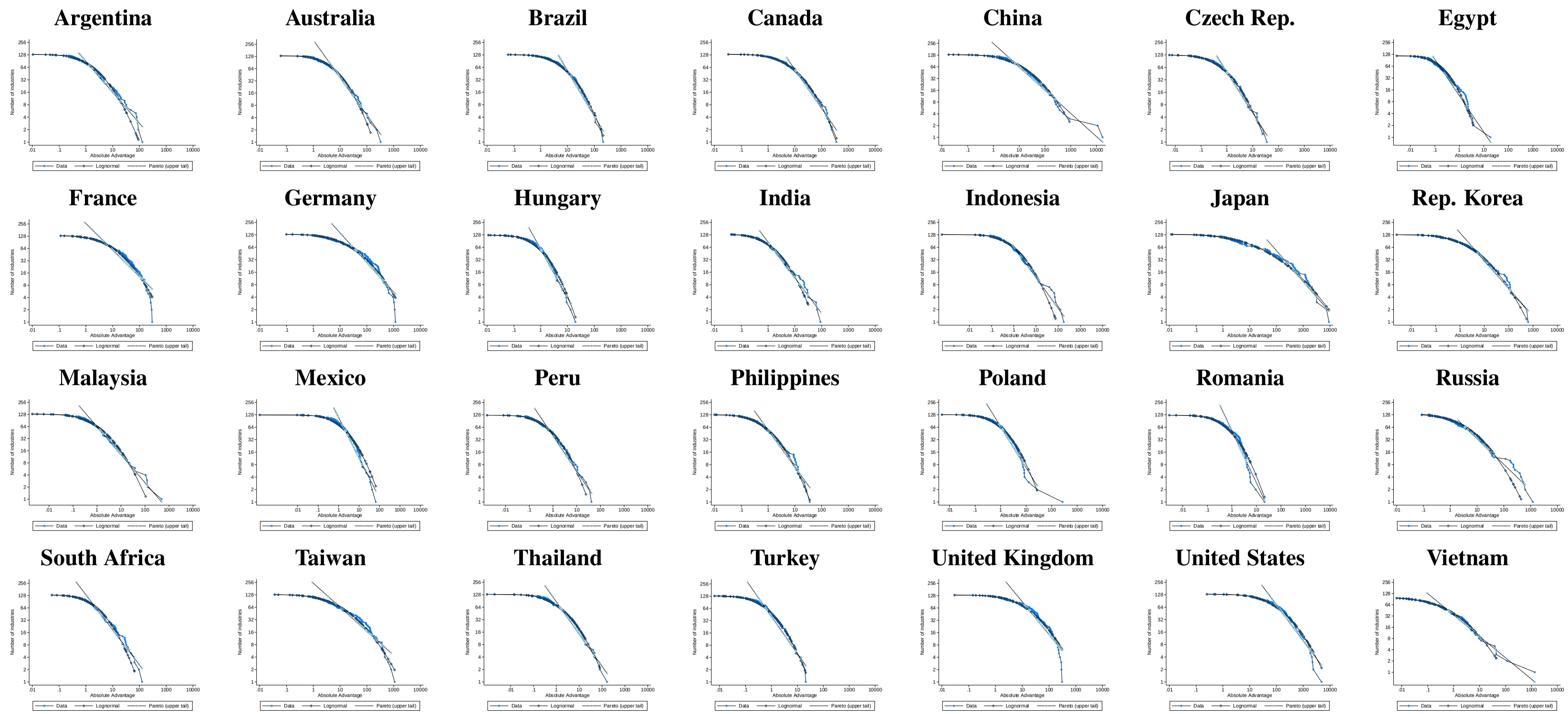
For the decay regression, **Table A5** shows that estimated decay rates are comparable to those in **Table 1** for our benchmark sector aggregates at the SITC 2-3 digit level (133 industries). At the two-digit level (60 industries), the ten-year decay rate for absolute advantage using all countries and industries is -0.26, at the three-digit level (224 industries) it is -038, and at the four-digit level (682 industries) it is -0.51. When using log RCA, decay rates vary less across aggregation levels, ranging from -0.31 at the two-digit level to -0.34 at the four-digit level. The qualitative similarity in decay rates across definitions of export advantage and levels of industry aggregation suggest that our results are neither the byproduct of sampling error nor the consequence of industry definitions.

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



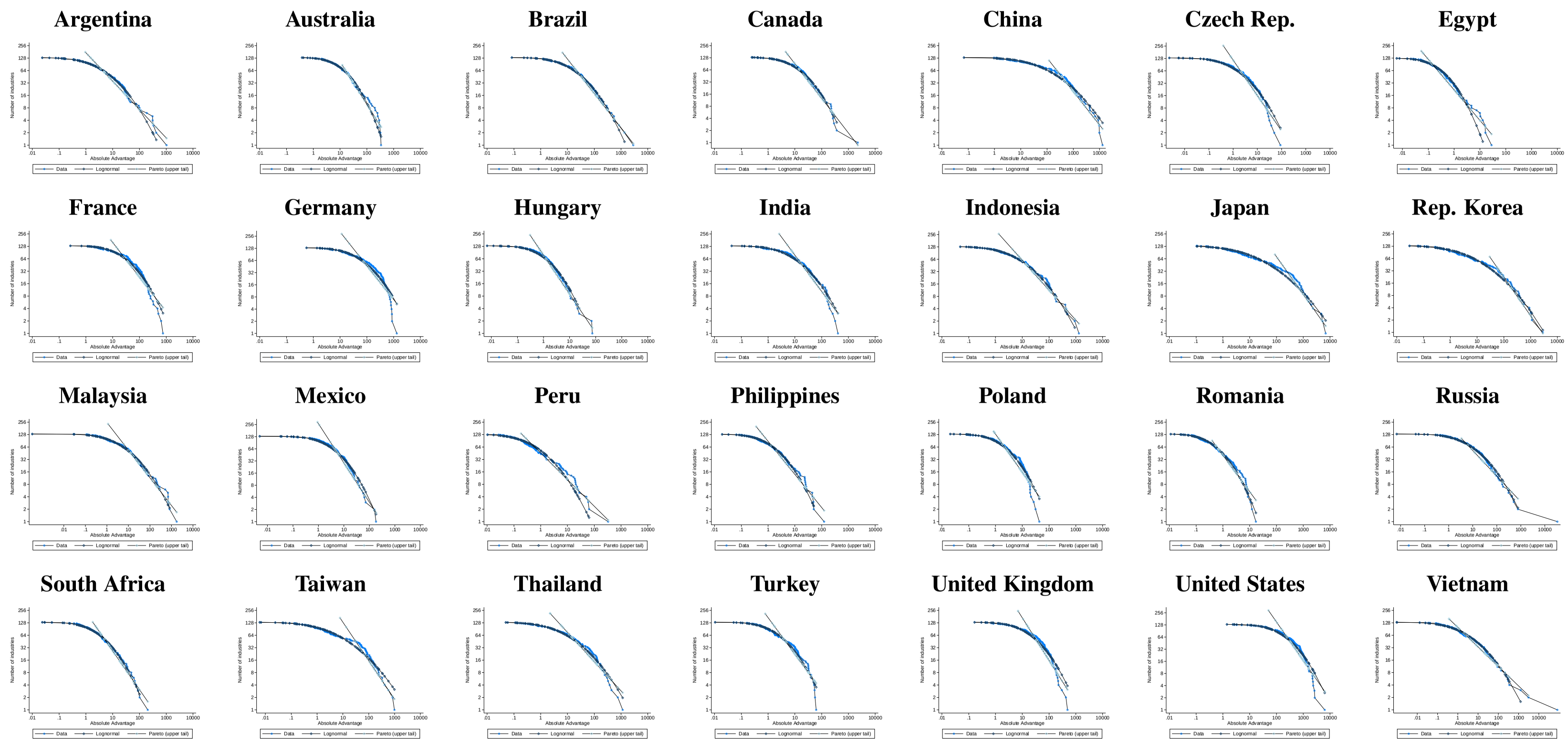
Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity measures of absolute advantage (7).
Note: The graphs show the frequency of industries (the cumulative probability $1 - F_A(a)$ times the total number of industries $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 1967$ (Pareto fit to upper five percentiles only).

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



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity 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 $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 1987$ (Pareto fit to upper five percentiles only).

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



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity 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 $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The fitted Pareto and log normal distributions are based on maximum likelihood estimation by country s in year $t = 2007$ (Pareto fit to upper five percentiles only).

Table A3: DECAY REGRESSIONS FOR COMPARATIVE ADVANTAGE, MPML GRAVITY ESTIMATES

	Full sample		LDC exporters		Nonmanufacturing	
	OLS k	MPML k	OLS k	MPML k	OLS k	MPML k
	(1)	(2)	(3)	(4)	(5)	(6)
Decay Regression Coefficients						
Decay rate ρ	-0.355	-0.430	-0.459	-0.496	-0.457	-0.487
	(0.002)***	(0.0002)***	(0.002)***	(0.0003)***	(0.003)***	(0.0004)***
Var. of residual s^2	2.104	2.829	2.424	3.297	2.522	3.499
	(0.024)***	(0.018)***	(0.025)***	(0.022)***	(0.039)***	(0.032)***
Implied OU Parameters						
Dissipation rate η	0.277	0.239	0.292	0.226	0.280	0.211
	(0.003)***	(0.002)***	(0.003)***	(0.002)***	(0.005)***	(0.002)***
Intensity of innovations σ	0.562	0.687	0.649	0.778	0.661	0.796
	(0.003)***	(0.002)***	(0.004)***	(0.003)***	(0.006)***	(0.004)***
Observations	324,978	324,978	202,010	202,010	153,768	153,768
Adjusted R^2 (within)	0.227	0.277	0.271	0.303	0.271	0.298
Years t	36	36	36	36	36	36
Industries i	133	133	133	133	68	68
Source countries s	90	90	62	62	90	90

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org.
Note: Reported figures for ten-year changes. Variables are OLS-estimated gravity measures of export capability k from (6) in columns 1, 3 and 5 (as previously reported in Table 1), and MPML-estimated gravity measures of export capability k from (8) following Eaton et al. (2012) in columns 2, 4 and 6. OLS estimation of the ten-year decay rate ρ from

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

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

For the GLD model under the GMM procedure, Table A6 confirms that results remain largely in line with those in **Table 2** before, for the benchmark aggregates at the SITC 2-3 digit level (133 industries). Estimates of the dissipation rate η are slightly larger during the post-1984 period than over the full sample period and, similar to the decay regressions, become larger as we move from broader to finer classifications of industry disaggregation. Estimates of the elasticity of dissipation ϕ are negative in all cases except one—when we measure export prowess using log absolute advantage (based on the gravity fixed effects) at the four-digit SITC level. As mentioned in Section 3.1, with nearly 700 four-digit SITC revision 2 industries we frequently have few destination markets per exporter-industry with which to estimate the gravity fixed effects, contributing to noise in the estimated exporter-industry coefficients.

F.5 GMM estimates of comparative advantage diffusion at ten-year horizon

We repeat GMM estimation of the generalized logistic diffusion of comparative advantage at the ten-year horizon. **Table A7** shows that estimated coefficients at the ten-year horizon are comparable to those for our benchmark

Table A4: GMM Estimates of Comparative Advantage Diffusion, MPML Gravity Estimates

	5-year transitions						1-yr. trans.	
	Full sample		LDC exp.		Non-manuf.		Full sample	
	$\ln A$	MPML $\ln A$	$\ln A$	MPML $\ln A$	$\ln A$	MPML $\ln A$	$\ln A$	MPML $\ln A$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Generalized Logistic Diffusion Parameters								
Dissipation rate η	0.256	0.215	0.270	0.207	0.250	0.187	0.256	0.206
	(0.005)***	(0.002)***	(0.006)***	(0.002)***	(0.006)***	(0.002)***	(0.004)***	(0.002)***
Intensity of innovations σ	0.745	0.883	0.842	0.984	0.875	1.028	1.434	1.606
	(0.01)***	(0.004)***	(0.015)***	(0.005)***	(0.018)***	(0.006)***	(0.026)***	(0.009)***
Elasticity of decay ϕ	-0.040	-0.020	-0.066	-0.029	-0.033	-0.008	-0.040	-0.018
	(0.016)**	(0.004)***	(0.022)***	(0.005)***	(0.017)*	(0.005)*	(0.015)***	(0.004)***
Implied Parameters								
Log gen. gamma scale $\ln \hat{\theta}$	127.080	315.900	62.229	188.150	161.920	985.810	128.010	349.740
	(71.847)*	(88.416)***	(30.520)**	(44.560)***	(116.263)	(693.289)	(69.392)*	(93.572)***
Log gen. gamma shape $\ln \kappa$	5.079	6.294	4.120	5.492	5.412	7.963	5.089	6.417
	(0.817)***	(0.427)***	(0.661)***	(0.348)***	(1.056)***	(1.126)***	(0.783)***	(0.412)***
Mean/median ratio	8.108	11.247	8.064	13.066	8.337	15.123	8.104	12.406
Observations	392,850	392,850	250,300	250,300	190,630	190,630	439,810	439,810
Industry-source obs. $I \times S$	11,542	11,542	7,853	7,853	5,845	5,845	11,854	11,854
Root mean sq. forecast error	1.852	1.985	2.018	2.131	1.969	2.139	1.735	1.804
Min. GMM obj. ($\times$ 1,000)	3.06e-13	3.17e-13	7.14e-13	7.26e-13	6.70e-13	7.36e-13	4.76e-14	1.39e-13

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org.
Note: GMM estimation at the five-year and one-year horizons for 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 absolute advantage measures $A_{is}(t) = \hat{A}_{is}(t) Z_s(t)$, and the Balassa index of revealed comparative advantage $RCA_{ist} = (X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_j X_{jst}/\sum_j \sum_{\varsigma} X_{j\varsigma t})$. Variables are OLS-estimated gravity measures of export capability k from (6) in columns 1, 3, 5 and 7, and MPML-estimated gravity measures of export capability k from (8) following Eaton et al. (2012) in columns 2, 4, 6 and 8. Parameters η, σ, ϕ are estimated under the constraints $\ln \eta, \ln \sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (21), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (18). Less developed countries (LDC) as listed in Appendix E. The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

Table A5: DECAY REGRESSIONS FOR COMPARATIVE ADVANTAGE, VARYING INDUSTRY AGGREGATES

	2-digit Industries		3-digit Industries		4-digit Industries	
	Exp. cap. k	$\ln RCA$	Exp. cap. k	$\ln RCA$	Exp. cap. k	$\ln RCA$
	(1)	(2)	(3)	(4)	(5)	(6)
Decay Regression Coefficients						
Decay rate ρ	-0.262	-0.307	-0.375	-0.326	-0.512	-0.335
	(0.003)***	(0.017)***	(0.002)***	(0.01)***	(0.002)***	(0.006)***
Var. of residual s^2	1.472	1.678	2.021	2.270	2.941	2.672
	(0.027)***	(0.009)***	(0.028)***	(0.007)***	(0.115)***	(0.005)***
Implied OU Parameters						
Dissipation rate η	0.309	0.310	0.301	0.241	0.259	0.209
	(0.006)***	(0.014)***	(0.004)***	(0.006)***	(0.01)***	(0.003)***
Intensity of innovations σ	0.443	0.486	0.558	0.573	0.744	0.625
	(0.004)***	(0.008)***	(0.004)***	(0.006)***	(0.015)***	(0.004)***
Observations	70,609	70,609	230,583	230,584	566,225	566,494
Adjusted R^2 (within)	0.241	0.233	0.266	0.224	0.311	0.213
Years t	14	14	14	14	14	14
Industries i	60	60	225	225	682	682
Source countries s	90	90	90	90	90	90

Source: WTF (Feenstra et al. 2005, updated through 2008) for 60 time-consistent industries at the 2-digit SITC level, 225 industries at 3 digits and 682 industries at 4 digits in 90 countries from 1984-2007 and CEPII.org.
Note: Reported figures for ten-year changes. Variables are OLS-estimated gravity measures of export capability k from (6). OLS estimation of the ten-year decay rate ρ from

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

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

estimation at the five-year horizon. To facilitate the comparison, columns 7 and 8 of **Table A7** here restate the full-sample estimates at the five-year horizon from **Table 2** (columns 1 and 2). The qualitative similarity in global diffusion coefficients at varying intervals for the estimation moments suggest that our results tightly characterize the dynamics of comparative advantage.

F.6 Diffusion predicted and observed cumulative probability distributions of absolute advantage

Figures A4, A5 and **A6** present plots for the same 28 countries in 1967, 1987 and 2007 as shown before (in **Figures A1, A2** and **A3**). **Figures A4** through **A6** contrast graphs of the actual data with the diffusion implied predictions.

Table A6: GMM Estimates of Comparative Advantage Diffusion, Varying Industry Aggregates

	2-digit Industries		3-digit Industries		4-digit Industries	
		$\phi = 0$		$\phi = 0$		$\phi = 0$
	(1)	(2)	(3)	(4)	(5)	(6)
Estimated Generalized Logistic Diffusion Parameters						
Dissipation rate η	0.304	0.319	0.289	0.289	0.250	0.245
	(0.011)***	(0.006)***	(0.012)***	(0.033)***	(0.094)***	(0.002)***
Intensity of innovations σ	0.559	0.562	0.719	0.719	0.922	0.921
	(0.01)***	(0.007)***	(0.035)***	(0.072)***	(0.204)***	(0.005)***
Elasticity of decay ϕ	-0.056		-0.004		-0.005	
	(0.021)***		(0.126)		(0.041)	
Implied Parameters						
Log gen. gamma scale $\ln \hat{\theta}$	82.094		2,715.3		1,752.5	
	(45.495)*		(112,227.4)		(16,736.2)	
Log gen. gamma shape $\ln \kappa$	4.581		9.971		9.127	
	(0.786)***		(68.726)		(15.535)	
Mean/median ratio	6.028	4.799	5.710	5.648	7.542	7.677
Observations	96,989	96,989	323,140	323,140	733,820	733,820
Industry-source obs. $I \times S$	5,332	5,332	19,187	19,187	52,240	52,240
Root mean sq. forecast error	1.643	1.489	1.694	1.686	1.969	1.955
Min. GMM obj. ($\times$ 1,000)	1.79e-12	8.56e-11	9.41e-13	9.72e-13	4.35e-13	4.32e-13

Source: WTF (Feenstra et al. 2005, updated through 2008)for 60 time-consistent industries at the 2-digit SITC level, 225 industries at 3 digits and 682 industries at 4 digits in 90 countries from 1984-2007 and CEPII.org.

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

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

using absolute advantage measures $A_{is}(t) = \hat{A}_{is}(t) Z_s(t)$, unrestricted and restricted to $\phi = 0$. Parameters η, σ, ϕ are estimated under the constraints $\ln \eta, \ln \sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (21), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (18). Less developed countries (LDC) as listed in Appendix E. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

Table A7: GMM Estimates of Comparative Advantage Diffusion at Ten-Year Intervals

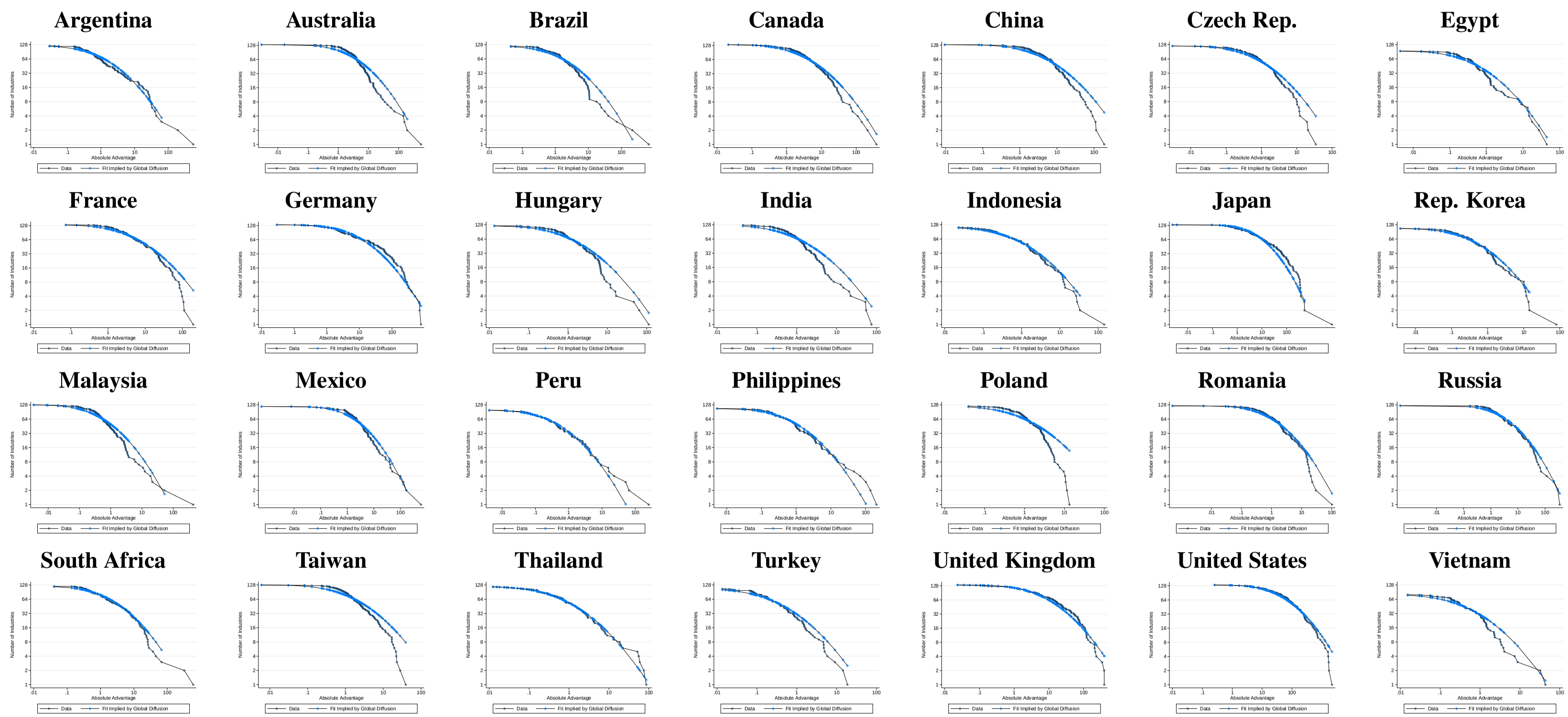
	10-year transitions						5-yr. trans.	
	Full sample		LDC exp.		Non-manuf.		Full sample	
	$\ln A$	$\ln RCA$	$\ln A$	$\ln RCA$	$\ln A$	$\ln RCA$	$\ln A$	$\ln RCA$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Generalized Logistic Diffusion Parameters								
Dissipation rate η	0.265	0.225	0.287	0.202	0.269	0.196	0.256	0.212
	(0.004)***	(0.006)***	(0.007)***	(0.004)***	(0.005)***	(0.007)***	(0.005)***	(0.003)***
Intensity of innovations σ	0.574	0.604	0.661	0.672	0.668	0.619	0.745	0.713
	(0.007)***	(0.038)***	(0.02)***	(0.028)***	(0.012)***	(0.033)***	(0.01)***	(0.024)***
Elasticity of decay ϕ	-0.028	0.027	0.003	0.016	-0.028	0.004	-0.040	0.006
	(0.014)**	(0.046)	(0.051)	(0.031)	(0.015)*	(0.033)	(0.016)**	(0.025)
Implied Parameters								
Log gen. gamma scale $\ln\hat{\theta}$	212.700	-216.920	-3716.100	-404.590	204.890	-2081.500	127.080	-1424.400
	(144.489)	(510.607)	(79239.510)	(990.621)	(147.523)	(18573.640)	(71.847)*	(7244.308)
Log gen. gamma shape $\ln\kappa$	5.858	5.763	10.491	6.626	5.813	9.209	5.079	8.657
	(1.017)***	(3.515)	(35.878)	(3.779)*	(1.074)***	(14.709)	(0.817)***	(8.282)
Mean/median ratio	7.190	8.387	5.668	10.970	7.008	12.604	8.108	10.259
Observations	335,820	335,820	211,640	211,640	161,940	161,940	392,850	392,860
Industry-source obs. $I \times S$	11,213	11,213	7,556	7,556	5,588	5,588	11,542	11,542
Root mean sq. forecast error	1.877	1.887	1.893	2.049	1.995	2.001	1.852	1.760
Min. GMM obj. ($\times$ 1,000)	2.89e-12	2.02e-11	1.07e-11	6.24e-11	1.44e-11	8.75e-11	3.06e-13	6.79e-12

Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org.
Note: GMM estimation at the ten-year (and five-year) horizon for the generalized logistic diffusion of comparative advantage $\hat{A}_{is}(t)$,

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

using absolute advantage measures $A_{is}(t) = \hat{A}_{is}(t)Z_s(t)$, and the Balassa index of revealed comparative advantage $RCA_{ist} = (X_{ist}/\sum_{\varsigma} X_{i\varsigma t})/(\sum_j X_{jst}/\sum_j\sum_{\varsigma} X_{j\varsigma t})$. Parameters η, σ, ϕ are estimated under the constraints $\ln\eta, \ln\sigma^2 > -\infty$ for the mirror Pearson (1895) diffusion of (21), while concentrating out country-specific trends $Z_s(t)$. The implied parameters are inferred as $\hat{\theta} = (\phi^2/\eta)^{1/\phi}$, $\kappa = 1/\hat{\theta}^{\phi}$and the mean/median ratio is given by (18). Less developed countries (LDC) as listed in Appendix E. The manufacturing sector spans SITC one-digit codes 5-8, the nonmanufacturing merchandise sector codes 0-4. Robust errors in parentheses (corrected for generated-regressor variation of export capability k): * marks significance at ten, ** at five, and *** at one-percent level. Standard errors of transformed and implied parameters are computed using the multivariate delta method.

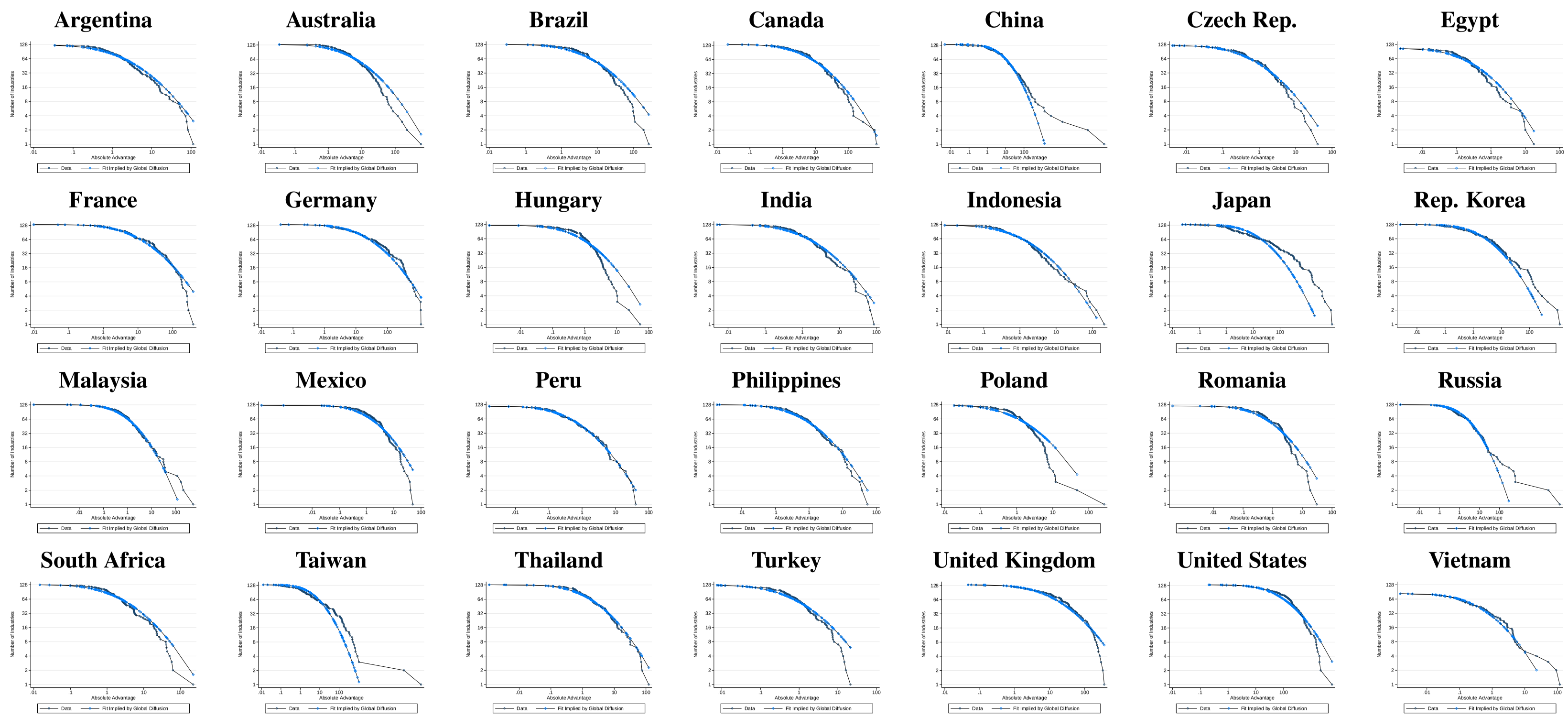
Figure A4: **Diffusion Predicted and Observed Cumulative Probability Distributions of Absolute Advantage in 1967**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity 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 $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

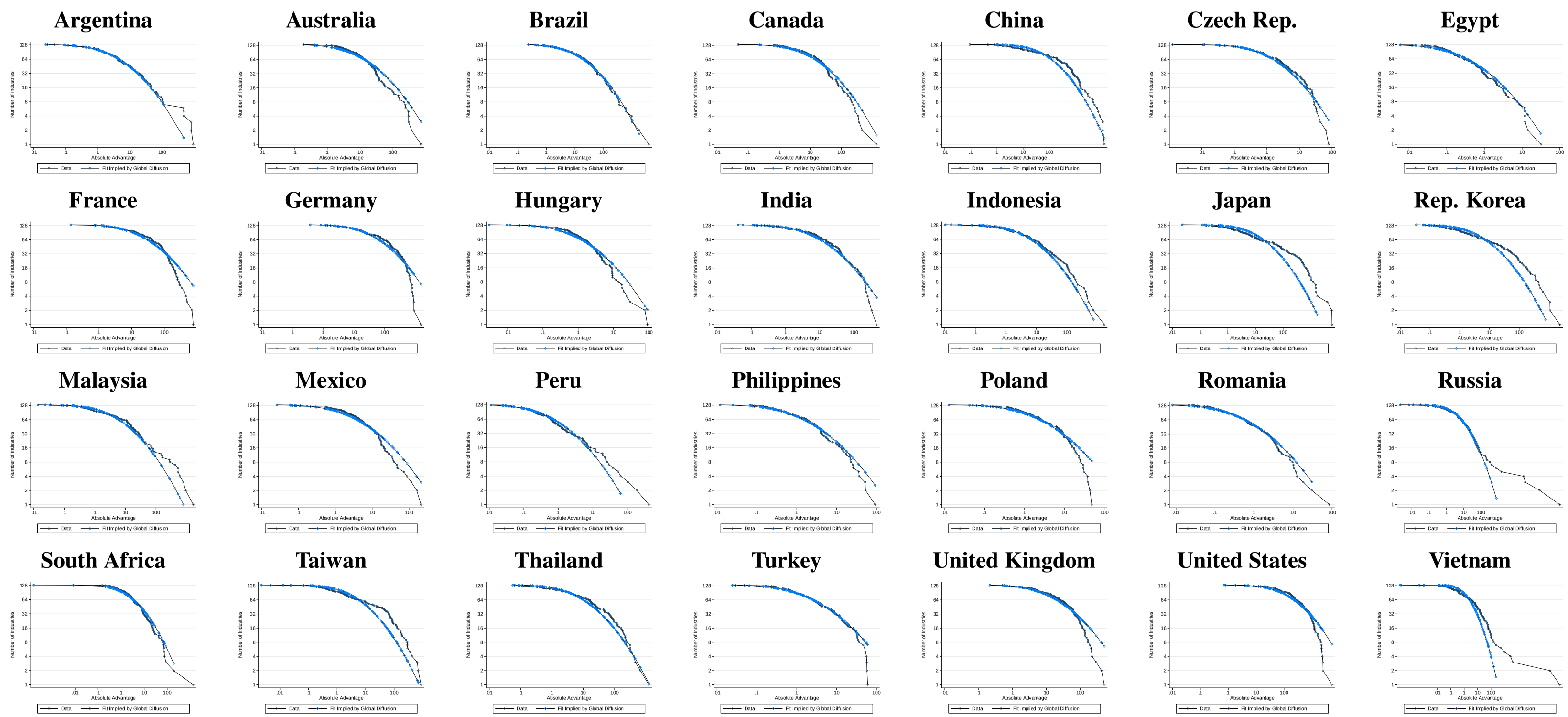
Figure A5: **Diffusion Predicted and Observed Cumulative Probability Distributions of Absolute Advantage in 1987**



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity 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 $I = 130$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis, for the year $t = 1987$. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.

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



Source: WTF (Feenstra et al. 2005, updated through 2008) for 133 time-consistent industries in 90 countries from 1962-2007 and CEPII.org; gravity 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 $I = 133$) on the vertical axis plotted against the level of absolute advantage a (such that $A_{ist} \geq a$) on the horizontal axis. Both axes have a log scale. The predicted frequencies are based on the GMM estimates of the comparative advantage diffusion (17) in Table 2 (parameters η and phi in column 1) and the inferred country-specific stochastic trend component $\ln Z_{st}$ from (19), which horizontally shifts the distributions but does not affect their shape.