Global Innovation and Knowledge Diffusion*

Nelson Lind
Emory University

Natalia Ramondo
UCSD and NBER

PRELIMINARY AND INCOMPLETE
(DO NOT CIRCULATE)

February 15, 2019

Abstract

This paper develops a model of economic growth and trade in which countries innovate ideas that diffuse across the globe. This model dynamically generates max-stable multivariate Fréchet productivity distributions and implies a mixed-CES import demand system. This demand system allows for rich substitution patterns in trade flows that arise from spatial correlation in technology. In the special case of a pure innovation model where countries do not share ideas, productivities are independent across space, and the demand system is CES. As a consequence, departures from CES reflect how knowledge diffusion generates technological similarity. In the general case with diffusion, high innovation countries tend to have dissimilar technology and their goods are less substitutable. These theoretical results provide a direct connection between estimable substitution patterns and the underlying dynamics of innovation and knowledge diffusion.

JEL Codes: F1. Key Words: innovation; knowledge diffusion; international trade; generalized extreme value; Fréchet distribution; growth accounting.

*TBD.

1 Introduction

How do the dynamics of innovation and knowledge diffusion shape similarities and differences in technology across countries? Since heterogeneity in technology determines patterns of comparative advantage and specialization, trade flows reflect the creation and spread of ideas. In particular, countries with correlated productivity tend to have larger trade elasticities. As a consequence, estimates of import demand systems reflect the underlying dynamics of innovation and knowledge diffusion.

This paper develops a new quantitative model of global innovation, knowledge diffusion, and trade that dynamically generates productivity with a max-stable multivariate Fréchet distribution. This distribution allows for rich patterns of correlation in productivity across production locations, and leads to a mixed-CES import demand system. Our model links spatial correlation to the underlying evolution of the state of knowledge. As a result, we can estimate the dynamics of innovation and knowledge diffusion using data on trade flows and trade costs.

The static Ricardian trade model in Lind and Ramondo (2018b)—which also incorporates flexible spatial correlation patterns—can be interpreted as the outcome of this dynamic model at a fixed point in time. Moreover, the growth model in Eaton and Kortum (2001), which generates the Ricardian trade model in Eaton and Kortum (2002) (EK), corresponds to the special case of our framework when there is no knowledge diffusion—which implies independent Fréchet productivity. Intuitively, if ideas are never shared between countries, then there can be no correlation in technology. More generally, our framework examines how innovation and diffusion interact to generate rich patterns of dependence between countries.

Our analysis begins by considering a general model of trade over continuous time whose import demand system satisfies the property of gross substitutes. We establish that this model is observationally equivalent to a model with max-stable Fréchet productivity. Next, we extend the characterization result in Lind and Ramondo (2018b)—which links max-stable Fréchet productivity to underlying Poisson processes—to incorporate time. Specifically, productivity follows a max-stable Fréchet process over both discrete space (countries) and continuous time if and only if productivity arises from the adoption of technological innovations with a particular Poisson structure.

Next, we reinterpret this representation as a model of innovation and knowledge diffusion. Innovations are discovered over time by individual countries according to a Poisson innovation process. Conditional on an innovation's discovery location and time, other countries learn about the idea according to a second Poisson diffusion process. This setup leads to a mixed-CES import demand system, which is well known to approximate any demand system satisfying gross substitutes (Fosgerau et al., 2013). This result comes from the combined forces of innovation—which pushes the demand system toward the case of independence—and diffusion—which generates correlation between countries with similar access to ideas. This link between mixed-CES demand and underlying innovation and diffusion processes gives us a framework to infer knowledge dynamics from data on trade flows and trade costs over time.[1]

This paper builds on the literature that generates Fréchet productivity from Poisson processes. The basic idea was introduced in Kortum (1997)—if the production technology is determined by the best "idea", or blueprint, and if ideas become available according to a Poisson process, then after a sufficiently long period of time productivity can be approximated by an extreme value distribution.[2] However, this literature—such as Eaton and Kortum (1999), Eaton and Kortum (2001), and Buera and Oberfield (2016)—restricts to the case of independence (see the survey in Lind and Ramondo, 2018a). As a consequence, the resulting models cannot capture how the offsetting forces of innovation and diffusion shape correlation in productivity and substitution patterns in trade flows. We contribute to this literature by generalizing this class of growth models to generate spatial correlation in productivity.

Because this generalization maps directly to the class of GEV models studied in Lind and Ramondo (2018b), it also links the large body of quantitative models inspired by EK to models of innovation and diffusion.[3] This paper implies that the

[1]In particular, mixed Logit (CES) models are identified off of local variation in characteristics (Fox et al., 2012), and, due to gross substitutes, our model fits into the class of models with invertible demand systems and the non-parametric identification results of Adao et al. (2017) apply.

[2] There is a relatively large literature that models innovation and diffusion of technologies as stochastic processes in closed economy setups, starting by the early work of Jovanovic and McDonald (1994) and Jovanovic and Rob (1989), followed by Eeckhout and Jovanovic (2002), Alvarez et al. (2008), Lucas (2009), Luttmer (2012), Lucas and Moll (2014), Perla and Tonetti (2014), and Benhabib et al. (2017), among others.

[3]Specifically, the GEV class accommodates many disaggregate Ricardian models of trade, such as multi-sector models (Costinot et al., 2012; Costinot and Rodrìguez-Clare, 2014; Levchenko and Zhang, 2014; DiGiovanni et al., 2014; Caliendo and Parro, 2015; Ossa, 2015; Levchenko and Zhang, 2016; French, 2016; Lashkaripour and Lugovskyy, 2017), multinational production models (Ra-

fundamentals in these models—such as those estimated in Hanson et al. (2015) and Levchenko and Zhang (2016)—link directly to underlying innovation and diffusion processes. We contribute to this large quantitative literature by showing how estimates of the substitution patterns in these models—based solely on trade flow and trade cost data—can be used to learn about and to discipline models of the underlying determinants of economic growth.[4]

2 A Model of Ricardian Trade

Consider a global economy consisting of N countries that produce and trade in a continuum of product varieties $v \in [0,1]$. Countries act as both origins where goods get produced and destinations where goods get delivered. We will use n to denote a generic country, o when a country is the origin where a good gets produced, and d when a country is the destination where a good gets delivered. Time is continuous and indexed by $t \in \mathbb{R}$.

Each country is populated by consumers and firms. Consumers supply a unit of time each period inelastically while firms hire their labor to produce goods.

In each period, consumers have identical CES preferences with elasticity of substitution $\sigma > -1$, $C_d(t) = \left(\int_0^1 C_d(t;v)^{\frac{\sigma}{\sigma+1}} \mathrm{d}v\right)^{\frac{\sigma+1}{\sigma}}$. Expenditure on variety v is $X_d(t;v) \equiv P_d(t;v)C_d(t;v) = (P_t(t;v)/P_d(t))^{-\sigma} X_d(t)$ where $P_d(t;v)$ is the price of the variety, $P_d(t) = \left(\int_0^1 P_d(t;v)^{-\sigma} \mathrm{d}v\right)^{-\frac{1}{\sigma}}$ is the price level in country d, and $X_d(t) = P_d(t)C_d(t)$ is total expenditure.

2.1 Productivity and Import Demand Systems

Our goal is to use trade flow data to learn how productivity changes over time. We begin by establishing equivalence results that allow us to focus on the special case of max-stable multivariate Fréchet productivity distributions. These equivalence results are in terms of the import demand system implied by various assumptions on production. We start by considering the general demand system generated un-

mondo and Rodríguez-Clare, 2013; Alviarez, 2018), global value chain models (Antràs and de Gortari, 2017), and models of trade with domestic geography (Fajgelbaum and Redding, 2014; Ramondo et al., 2016; Redding, 2016).

[4]For direct efforts to measure technology adoption and knowledge diffusion see Comin and Hobijn (2004) and Comin and Hobijn (2010). Also see Comin and Mestieri (2014) for a review.

der stochastic productivity, and then specialize to max-stable multivariate Fréchet productivity distributions that exactly match this general case.

We assume that the production function for varieties exhibits constant returns to scale in labor and depends on the origin country where the good is produced, the destination market where the good gets delivered, and time.[5] Output in o when delivering to destination d at time t is

$$Y_{od}(t;v) = A_{od}(t;v)L_{od}(t;v),$$

where $L_{od}(t;v)$ is labor used, and $A_{od}(t;v)$ is *productivity*.

The unit cost of producing in o and delivering to d is $c_{od}(t;v) = W_o(t)/A_{od}(t;v)$ where $W_o(t)$ is the nominal wage in country o. Under perfect competition, potential import prices match units costs so that $P_{od}(t;v) = c_{od}(t;v)$.

Consumers in destination d import from whichever origin offers the lowest price. The price of variety v in d is then

$$P_d(t;v) = \min_{o=1,\ldots,N} P_{od}(t;v).$$

We can then characterize the price level and expenditure as follows.

First, we make an assumption of the distribution of productivity.

Assumption 1 (I.I.D. Productivity)**.** *The distribution of productivity is iid over varieties with $\mathbb{E}A_{od}(v)^\sigma < \infty$.*

Under this assumption, we can characterize the aggregate price level using moments of the productivity distribution.

Lemma 1 (Price Aggregation)**.** *Under Assumption 1, the price level in d is*

$$P_d(t) = G^d(T_{1d}(t)W_1(t)^{-\sigma}, \ldots, T_{Nd}(t)W_N(t)^{-\sigma}; t)^{-\frac{1}{\sigma}}$$

where $T_{od}(t) \equiv \mathbb{E}[A_{od}(v;t)^\sigma]$ and

$$G^d(x_1, \ldots, x_N; t) \equiv \mathbb{E}\left[\max_{o=1,\ldots,N} \frac{A_{od}(v;t)^\sigma}{T_{od}(t)} x_o\right].$$

[5] This structure means that production incorporates the delivery technology.

Proof. The proof follows directly from the definition of the price index and cost-minimization by consumers. See Appendix A. □

The variable $T_{od}(t)$ is a *productivity index* and the function G^d is an *aggregation function*. The aggregation function alongside the productivity indices together define how wages across origin countries get reflected in destination market price levels. Higher wages or lower productivity indices (e.g. higher trade costs) increase the price level in a destination market.

We can use these quantities (which simply reflect the distribution of productivity) to compute expenditure shares. Expenditure by d on goods from o is

$$X_{od}(t) \equiv \mathbb{E}\left[\left(\frac{P_{od}(v;t)}{P_d(t)}\right)^{-\sigma} \mathbf{1}\{P_{od}(v;t) = P_d(v;t)\}\right] X_d(t).$$

Note that

$$\frac{\partial}{\partial x_o} G^d(x_1, \ldots, x_N; t) = \sum_{o=1}^{N} \mathbb{E}\left[\frac{A_{od}(v;t)^\sigma}{T_{od}(t)} \mathbf{1}\left\{\frac{A_{od}(v;t)^\sigma}{T_{od}(t)} x_o = \max_{o'=1,\ldots,N} \frac{A_{o'd}(v;t)^\sigma}{T_{o'd}(t)} x_{o'}\right\}\right].$$

and so we can use Shepard's lemma to characterize expenditure shares based on these productivity indices and the aggregation function.

Lemma 2 (Expenditure Shares). *Under Assumptions 1, the share of expenditure on country o goods equals the elasticity of the price level to the wage in o:*

$$\pi_{od}(t) \equiv \frac{X_{od}(t)}{X_d(t)} = \frac{\partial \ln P_d(t)}{\partial \ln W_o(t)} = \frac{T_{od}(t) W_o(t)^{-\sigma} G_o^d(T_{1d}(t)W_1(t)^{-\sigma}, \ldots, T_{Nd}(t)W_N(t)^{-\sigma})}{G^d(T_{1d}(t)W_1(t)^{-\sigma}, \ldots, T_{Nd}(t)W_N(t)^{-\sigma})}$$

where $G_o^d(x_1, \ldots, x_N; t) \equiv \partial G^d(x_1, \ldots, x_N; t)/\partial x_o$.

Proof. The price level is the minimum cost of one unit of aggregate consumption in d. As a result the map $(W_1(t), \ldots, W_N(t)) \mapsto P_d(t)C_d(t)$ is the expenditure function to achieve consumption of $C_d(t)$. By Shephard's Lemma, the share of expenditure on o is the elasticity of expenditure to $W_o(t)$ for fixed $C_d(t)$, which is just the elasticity of $P_d(t)$ to $W_o(t)$. □

This result gives a closed form characterization of expenditure shares in this general Ricardian model in terms of productivity indices, wages, and the aggregation function. The map $(W_1(t), \ldots, W_N(t)) \mapsto \pi_{od}(t)$ is the *import demand system* for destination d at time t.

We now establish that whenever goods are substitutes ($\sigma > 0$), this general case is observationally equivalent to a model where productivity has a max-stable multivariate Fréchet distribution across origins.

Proposition 1 (Observational Equivalence to θ-Fréchet Productivity)**.** *Suppose that $\sigma > 0$. Then for any productivity distribution with implied import demand system $\{\pi_{od}(\cdot, t)\}_{o=1}^N$, there exists a multivariate θ-Fréchet productivity distribution with $\theta = \sigma$ such that if preferences are CES with elasticity of substitution below σ then the implied import demand system $\{\pi_{od}^{GEV}(\cdot, t)\}_{o=1}^N$ satisfies*

$$\pi_{od}(W_1, \ldots, W_N, t) = \pi_{od}^{GEV}(W_1, \ldots, W_N, t)$$

for all $(W_1, \ldots, W_N) \in \mathbb{R}_+^N$.

Proof. The aggregation function is a correlation function so the import demand system is a GEV import demand system as defined in Lind and Ramondo (2018b). Therefore, if preferences are CES with a sufficiently low elasticity of substitution and the joint distribution of productivity is iid over varieties and θ-Fréchet with $\theta = \sigma$, scale parameters of $T_{od}(t)$, and correlation function $G^d(\cdot, t)$, then expenditure shares are identical to $\pi_{od}(t)$. □

This result implies that so long as the import demand system satisfies the property of gross substitutes, then it can be explained using a Ricardian model with CES preferences and θ-Fréchet productivity.

2.2 Global Innovation Representation

This section further relates the general Ricardian model to models of innovation. In particular, Theorem 1 in Lind and Ramondo (2018b) establishes that static Ricardian models with max-stable multivariate Fréchet distributions can always be related to an underlying structure of technology where countries adopt global innovations. We extend this result to case with continuous time. Specifically, the following two assumptions define the *global innovation representation* for productivity. They are necessary and sufficient for productivity to be a max-stable Fréchet process across countries and time.

Assumption 2 (Innovation Decomposition). *There exists a measurable space of attributes* $(\Omega, \mathcal{F})$ *and for each* $v \in [0,1]$ *an infinite but countable set of global innovations,* $i = 1, 2, \ldots$*, with global productivity* $Z_i(v) \geq 0$ *and attributes* $\omega_i(v) \in \Omega$ *such that*

$$A_{od}(r, t; v) = \max_{i=1,2,\ldots} Z_i(v) A_{od}(t, \omega_i(v)). \tag{1}$$

for some measurable function $\omega \mapsto A_{od}(t, \omega)$ *for each* $o, d = 1, \ldots, N$*, and* $t \in \mathbb{R}$*.*

Global innovations represent physical techniques (i.e., blueprints) for producing a good that any firm can potential use in production. The *core productivity*, $Z_i(v)$, of the innovation represents its fundamental efficiency as it is constant across origins, destinations, and time. Meanwhile, *attributes*, $\omega_i(v)$, represent anything specific to the innovation relevant for productivity across origins, destinations, and time. For instance, one attribute of an innovation might be the time at which it was first discovered, and the country in which it was discovered. The function $A_{od}(t, \omega)$ determines how spatial factors, time, and attributes combine to determine productivity. We refer to the variable $A_{iod}(t; v) \equiv A_{od}(t, \omega_i(v))$ as the *applicability* of idea i to production in o for delivery to d at time t. This assumption states that at each point in time productivity arises from adoption of whichever global innovation is most efficient given core productivities, and applicability.

Our next assumption states that innovations follow a Poisson process over core productivities and characteristics.

Assumption 3 (Poisson Innovations). *There exists* $\theta > 0$ *and a* σ*-finite measure* μ *such that* $\int A_{od}(r, t, \omega)^\theta d\mu(\omega) < \infty$ *and the collection* $\{Z_i(v), \omega_i(v)\}_{i=1,2,\ldots}$ *consists of the points of a Poisson process with intensity measure* $\theta z^{-\theta-1} dz d\mu(\omega)$*, i.i.d. over* $v \in [0,1]$*.*

This assumption defines the stochastic properties of core productivities and attributes. It specifies that core productivities are independent of attributes. As a consequence, any random variables we define as measurable functions of $\omega_i(v)$ are independent of core productivities.[6]

The following theorem is a generalization of the characterization result in Lind and Ramondo (2018b) to arbitrary θ-Fréchet processes for productivity over space and time.

[6] Any random variable defined as a measurable function of $\omega_i(v)$ has its stochastic properties derive from the measure space $(\Omega, \mathcal{F}, \mu)$. We denote the probability measure of the underlying probability space on which core productivities and these random variables are jointly defined by $\mathbb{P}$.

Theorem 1 (Global Innovation Representation)**.** *Productivity across origins, destinations, and time, $\{A_{od}(r,t;v)\}_{o,d=1,\dots,N,t\in\mathbb{R}}$, is a θ-Fréchet process if and only if assumptions 2 and 3 hold. In this case, for any fixed d and t, the joint distribution of productivity over origins is*

$$\mathbb{P}[A_{1d}(t;v)\le a_1,\dots,A_{Nd}(t;v)\le a_N]=\exp\left[-G^d(T_{1d}(t)a_1^{-\theta},\dots,T_{Nd}(t)a_N^{-\theta},t\right]$$

for scale parameters $T_{od}(t)\equiv\int A_{od}(t,\omega)^\theta d\mu(\omega)$ and correlation function

$$G^d(x_1,\dots,x_K,t)\equiv\int\max_{o=1,\dots,N}\frac{A_{od}(t,\omega)^\theta}{T_{od}(t)}x_o d\mu(\omega).$$

Proof. Sufficiency follows from showing that the finite dimensional distributions of the productivity are θ-Fréchet by applying Campbell's theorem. Necessity follows from Theorem 1 in Kabluchko (2009), which states that any θ-Fréchet process has a spectral representation, which we call the global innovation representation. See Appendix B. □

This theorem tells us that any model where productivity is θ-Fréchet is equivalent to a model incorporating the latent adoption of innovations. Moreover, the fundamental properties of these underlying innovations are time invariant and any productivity dynamics come from changes in knowledge over time. Combined with Proposition 1, we can always relate any model with CES preferences and stochastic productivity to an observationally equivalent model where firms adopt global innovations based on the applicability of those innovations.

3 A Model of Innovation and Knowledge Diffusion

We now explore the consequences of this global innovation representation under the additional assumption that innovations (ideas) are discovered and learned over time by firms. Our assumptions will lead to a mixed-CES demand system, which is well known to approximate any GEV demand system. As a consequence, the resulting model can closely match any changing substitution patterns that we may estimate from trade flow data.

3.1 Innovation and Diffusion

We now add additional assumptions to the Ricardian model developed in Section 2 that can be interpreted as firms discovering innovations and learning about innovations over time. Prior to a firm knowing an idea—by either discovering it or learning it—it is unavailable for production. Once an innovation is discovered, other firms can learn about the idea over time and begin to use it in production.

The following assumption formalizes this setup by adding structure to the applicability function, $A_{od}(t,\omega)$, and the attributes of ideas, $\omega_i(v)$.

Assumption 4 (Applicability and Attributes Under Innovation and Diffusion)**.** *There exists measurable functions $A_{od} : [0,1] \to \mathbb{R}_+$ for each $o, d = 1, \ldots, N$ such that*

$$A_{od}(t, \omega_i(v)) = A_{od}(\chi_i(v)) \max_{j=0,1,\ldots} U_{ij}(v)\mathbf{1}\{t_{ij}(v) \leq t\}$$

where attributes take the form $\omega_i(v) = \{\chi_i(v), \{U_{ij}(v), n_{ij}(v), t_{ij}(v)\}_{j=0,1,\ldots}\}$ with $t_{ij}(v) < t_{ij'}(v)$ if and only if $j < j'$.

Each innovation has some characteristics $\chi_i(v)$ which influence applicability symmetrically over time and may influence applicability differently across origin countries and destination markets. The dynamics of applicability arise from when innovations become available to firms.

The first firm with access to the idea is in $n_{i0}(v)$, the *innovation location*. The innovator's understanding of the idea is $U_{i0}(v)$. The *innovation time* is $t_{i0}(v)$. We refer to the collection of core productivities, characteristics, initial understandings, innovation locations, and innovation times, $\{Z_i(v), \chi_i(v), U_{i0}(v), n_{i0}(v), t_{i0}(v)\}_{i=1,2,\ldots}$, as the *innovation process*.

The j'th firm to learn the idea after its initial discovery is in country $n_{ij}(v)$. They have understanding $U_{ij}(v)$, and learn the idea at time $t_{ij}(v)$. All learning occurs after the initial discovery because $t_{ij}(v) > t_{i0}(v)$ for all $j = 1, 2, \ldots$. We refer to $\{U_{ij}(v), n_{ij}(v), t_{ij}(v)\}_{j=1,2,\ldots}$ as the *diffusion process* for idea i.

Assumption 4 implies that if a firm has not either discovered or learned an innovation by time t, then their applicability is zero. That is, firms can only apply innovations that they know. In turn, the firm with the best understanding of the idea within each origin determines country-level applicability.

The following assumption specifies that the diffusion process is Poisson conditional on the initial discovery of the idea.

Assumption 5 (Poisson Diffusion)**.** *For each $v \in [0,1]$ and $i = 1, 2, \ldots$, the diffusion process is Poisson conditional on $\chi_i(v)$, $U_{i0}(v)$, $n_{i0}(v)$, and $t_{i0}(v)$ with intensity for country n of $\mathbf{1}\{t > t_{i0}(v)\}\vartheta(\chi_i(v))u^{-\vartheta(\chi_i(v))-1}\lambda_n^D(t \mid \chi_i(v), n_{i0}(v))dudt$.*

To interpret this assumption, denote the number of firms in n who have learned the idea with understanding above u as of time t by

$$J_{in}(u, t; v) \equiv \sum_{j=1,2,\ldots} \mathbf{1}\{U_{ij}(v) > u, n_{ij}(v) = n, t_{ij}(v) \leq t\}.$$

This assumption states that this random variable is Poisson and has expected value conditional on the idea's characteristics and discovery location and time of

$$\begin{aligned}\mathbb{E}[J_{in}(u, t; v) \mid \chi_i(v), n_{i0}(v), t_{i0}(v)] &= \int_{t_{i0}(v)}^{t} \int_{u}^{\infty} \vartheta(\chi_i(v))\tilde{u}^{-\vartheta(\chi_i(v))-1}\lambda_n^D(\tilde{t} \mid \chi_i(v), n_{i0}(v))\mathrm{d}\tilde{u}\mathrm{d}\tilde{t} \\ &= u^{-\vartheta(\chi_i(v))}\Lambda_n^D(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v))\end{aligned}$$

where $\Lambda_n^D(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v)) \equiv \int_{t_{i0}(v)}^{t} \lambda_n^D(\tilde{t} \mid \chi_i(v), n_{i0}(v))\mathrm{d}t$. The variable $\vartheta(\chi_i(v))$ is the *similarity of understanding* of the idea. It determines how likely it is for firms to have a high understanding of the idea. When similarity is high, firms are less likely to have an understanding above u, meaning that understandings are concentrated at a low level. The variable $\lambda_n^D(t \mid \chi_i(v), n_{i0}(v))$ is the rate of learning in n at time t about the idea and $\Lambda_n^D(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v))$ is accumulated learning. The learning rate may depend on the idea's characteristics (reflecting the possibility that certain kinds of ideas are harder to learn) as well as the location where it was discovered (reflecting spatial factors influencing diffusion of ideas).

As an intermediate result, we first characterize the distribution of the best learned understanding of each idea within each country. The best learned understanding in n is

$$U_{in}^*(v) \equiv \max_{j=1,2,\ldots} U_{ij}(v)\mathbf{1}\{n_{ij}(v) = n, t_{ij}(v) \leq t\}.$$

Note that this variable doesn't include the initial understanding of the idea, only those understandings learned by firms after the initial discovery. The following lemma establishes that this random variable is independent across countries and Fréchet.

Lemma 3 (Distribution of Best Learned Understanding)**.** *Under assumption 5, the best learned understanding of each idea within country n at time t is Fréchet across varieties and independent across countries conditional on $\chi_i(v)$, $n_{i0}(v)$, and $t_{i0}(v)$ with shape $\vartheta(\chi_i(v))$ and scale $\Lambda_n^D(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v))$.*

Proof. See Appendix C. □

Alongside the next assumption, this result will ensure we get a closed form characterization for the aggregate correlation function.

We next asume that the innovation process is Poisson.

Assumption 6 (Poisson Innovation)**.** *For each $v \in [0,1]$, the innovation process is Poisson with intensity for country n of $\theta z^{-\theta-1} f(u, \vartheta(\chi_i(v)))\lambda_n^I(\chi, t) dz d\chi du dt$ where $f(u, \vartheta) = \vartheta u^{-\vartheta-1} e^{-u^{-\vartheta}}$ is the density of a Fréchet random variable with unit scale and shape ϑ.*

Define the total number of ideas for variety v with productivity of at least z, characteristic in $[0, \chi]$, and initial understand below u that were discovered by firms in n prior to t by

$$I_n(z, u, \chi, t; v) \equiv \sum_{i=1,2,\ldots} \mathbf{1}\{Z_i(v), \chi_i(v) \in [0, \chi], U_{i0}(v) \leq u, n_{i0}(v) = n, t_{i0}(v) \leq t\}.$$

This assumption states that this random variable is Poisson with expected value

$$\begin{aligned}\mathbb{E}[I_n(z, u, \chi, t; v)] &= \int_{-\infty}^{t} \int_0^u \int_0^\chi \int_z^\infty \theta \tilde{z}^{-\theta-1} f(\tilde{u}, \vartheta(\chi_i(v)))\lambda_n^I(\chi, t) \mathrm{d}\tilde{z} \mathrm{d}\tilde{\chi} \mathrm{d}\tilde{u} \mathrm{d}\tilde{t} \\ &= z^{-\theta} \int_0^\chi e^{-u^{-\vartheta(\tilde{\chi})}} \Lambda_n^I(\tilde{\chi}, t) \mathrm{d}\tilde{\chi}\end{aligned}$$

where $\Lambda_n^I(\chi, t) \equiv \int_{-\infty}^t \lambda_n^I(\chi, t) \mathrm{d}\tilde{t}$. The variable $\lambda_n^I(\chi, t)$ governs how quickly innovations with characteristic χ arrive to country n over time and the distribution of these innovations across characteristics. In particular, conditional on an idea being known at time t with core productivity of at least $\underline{z}$, the likelihood that an idea is known with core productivity of z and characteristic χ is

$$\frac{\theta z^{-\theta-1} \Lambda_n^I(\chi, t)}{\underline{z}^{-\theta} \int_0^\chi \Lambda_n^I(\tilde{\chi}, t) \mathrm{d}\tilde{\chi}} = \theta z^{-\theta-1} \underline{z}^\theta \frac{\Lambda_n^I(\chi, t)}{\int_0^1 \Lambda_n^I(\tilde{\chi}, t) \mathrm{d}\tilde{\chi}}.$$

That is, among those ideas with $Z_i(v) > \underline{z}$ core productivities are Pareto with lower bound $\underline{z}$ and shape θ, and accumulated innovation determines the distribution of ideas across characteristics. Conditional on having characteristic χ, the likelihood of an initial understanding equal to u is Fréchet with unit scale and shape $\vartheta(\chi)$.

Combining Assumption 6 with Lemma 3, we get the following result for the distribution of productivity across countries.

Proposition 2 (Country-Level Productivity Distribution)**.** *Under assumptions 2, 4, 5, and 6 country-level productivity is multivariate θ-Fréchet. For each destination d and time t, the joint distribution across o is*

$$\mathbb{P}\left[A_{1d}(t;v) \leq a_1, \ldots, A_{Nd}(t;v) \leq a_N\right] = \exp\left[-G^d(T_{1d}(t)a_1^{-\theta}, \ldots, T_{Nd}(t)a_N^{-\theta}, t)\right]$$

where the scale of productivity in o when delivering to d at time t is

$$T_{od}(t) \equiv \int_{-\infty}^{t} \int_{0}^{1} \sum_{n=1}^{N} T_{od}(t \mid \chi, n, \tilde{t}) \lambda_n^I(\chi, \tilde{t}) d\chi d\tilde{t}$$

for $T_{od}(t \mid \chi, n, \tilde{t}) \equiv \Gamma(\rho(\chi)) A_{od}(\chi)^{\theta} [\mathbf{1}\{n = o\} + \Lambda_o^D(t \mid \chi, n, \tilde{t})]^{1-\rho(\chi)}$ and $\rho(\chi) \equiv 1 - \theta/\vartheta(\chi)$, and the correlation function is

$$G^d(x_1, \ldots, x_N, t) \equiv \int_{-\infty}^{t} \int_{0}^{1} \sum_{n=1}^{N} \left[\sum_{o=1}^{N} \left(\alpha_{od}(t \mid \chi, n, \tilde{t}) x_o\right)^{\frac{1}{1-\rho(\chi)}}\right]^{1-\rho(\chi)} \lambda_n^I(\chi, \tilde{t}) d\chi d\tilde{t}.$$

for $\alpha_{od}(t \mid \chi, n, \tilde{t}) \equiv T_{od}(t \mid \chi, n, \tilde{t}) / T_{od}(t)$.

Proof. See Appendix D □

To get some intuition for this result, first consider a pure innovation model where ideas never diffuse. In this case $\Lambda_o^D(t \mid \chi, n, \tilde{t}) = 0$, and ideas can only be applied in the country where they were discovered. As a consequence $T_{od}(t \mid \chi, n, \tilde{t}) = 0$ for $o \neq n$ and $T_{od}(t \mid \chi, n, \tilde{t}) = \Gamma(\rho(\chi)) A_{od}(\chi)^{\theta}$ otherwise. The scale of productivity then simplifies to

$$T_{od}(t) \equiv \int_{-\infty}^{t} \int_{0}^{1} \Gamma(\rho(\chi)) A_{od}(\chi)^{\theta} \lambda_o^I(\chi, \tilde{t}) \mathrm{d}\chi \mathrm{d}\tilde{t}$$

That is, only innovations in o generate increases in productivity in o. As a consequence, the aggregation weights are zero for any country other than the place of innovation: $\alpha_{od}(t, \chi, n, t_0) = 0$ if $o \neq n$. This isolation of ideas to the country where they get innovated implies that the correlation function reduces to the case of independent productivity:

$$G^d(x_1, \ldots, x_N, t) \equiv \int_0^1 \int_{-\infty}^t \sum_{o=1}^N \alpha_{od}(t, \chi, o, \tilde{t}) x_o \mathrm{d}\chi \mathrm{d}\tilde{t} = \sum_{o=1}^N x_o.$$

Independence in productivity implies a CES demand system as in EK. As a consequence, this model interprets any deviation from a CES demand system as reflecting the diffusion of ideas over space.

4 Measuring Knowledge Dynamics

4.1 Identification

Our ability to measure the dynamics of knowledge comes from the fact that the random coefficients in mixed logit models are identified off of local variation in characteristics (Fox et al., 2012). This model also fits into the class of models with invertible demand systems (due to gross-substitutability), and so the non-parametric identification results of Adao et al. (2017) (ACD) also apply. They establish that the demand system is non-parametrically identified using only variation in aggregate trade costs. As a consequence, we can use observable aggregate trade flows and aggregate trade costs to estimate the import demand system and infer knowledge dynamics. In the case of a single cross-section, we can focus on a steady state of the model and identify steady state learning flows and innovation rates.

TBA

4.2 Estimation

We follow ACD and proceed parametrically. We assume that learning depends on bilateral characteristics of countries—specifically, their proximity as measured by distance.

TBA

4.3 Growth Accounting

TBA

5 Conclusion

TBA

References

Adao, R., A. Costinot, and D. Donaldson (2017). Nonparametric counterfactual predictions in neoclassical models of international trade. *The American Economic Review 107*(3), 633–689.

Alvarez, F. E., F. J. Buera, and R. E. Lucas (2008). Models of idea flows. *NBER Working Paper Series* (14135).

Alviarez, V. (2018). Multinational production and comparative advantage. *Manuscript UBC*.

Antràs, P. and A. de Gortari (2017). On the geography of global value chains. *Mimeo, Harvard University*.

Benhabib, J., J. Perla, and C. Tonetti (2017). Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier. *Unpublished*.

Buera, F. J. and E. Oberfield (2016). The global diffusion of ideas. NBER Working Paper Series *21844*.

Caliendo, L. and F. Parro (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies 82*(1), 1–44.

Comin, D. and B. Hobijn (2004). Cross-country technology adoption: making the theories face the facts. *Journal of Monetary Economics 51*, 39Ð83.

Comin, D. and B. Hobijn (2010). An exploration of technology diffusion. *American Economic Review 100*, 2031Ð2059.

Comin, D. and M. Mestieri (2014). Technology diffusion: Measurement, causes, and consequences. In *Handbook of economic growth*, Volume 2, pp. 565–622. Elsevier.

Costinot, A., D. Donaldson, and I. Komunjer (2012). What goods do countries trade? a quantitative exploration of ricardo's ideas. *Review of Economic Studies 79*, 581–608.

Costinot, A. and A. Rodrìguez-Clare (2014). Trade theory with numbers: Quantifying the consequences of globalization. Technical Report 4.

DiGiovanni, J., A. Levchenko, and J. Zhang (2014). The global welfare impact of china: Trade integration and technological change. *American Economic Journal: Macroeconomics 6*(3), 153–183.

Eaton, J. and S. Kortum (1999). International technology diffusion: Theory and measurement. *International Economic Review 40*(3), 537–570.

Eaton, J. and S. Kortum (2001). Technology, trade, and growth: A unified framework. *European Economic Review 45*, 742–755.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica*, 1741–1779.

Eeckhout, J. and B. Jovanovic (2002). Knowledge spillovers and inequality. *The American Economic Review 92*(5), 1290–1307.

Fajgelbaum, P. and S. Redding (2014). External integration, structural transformation and economic development: Evidence from argentina 1870-1914. *NBER Working Paper 20217*.

Fosgerau, M., D. McFadden, and M. Bierlaire (2013). Choice probability generating functions. *Journal of Choice Modeling 8*, 1–18.

Fox, J. T., K. il Kim, S. P. Ryan, and P. Bajari (2012). The random coefficients logit model is identified. *Journal of Econometrics 166*(2), 204–212.

French, S. (2016). The composition of trade flows and the aggregate effects of trade barriers. *Journal of International Economics 98*, 114–137.

Hanson, G. H., N. Lind, and M.-A. Muendler (2015). The dynamics of comparative advantage. Technical report, National bureau of economic research.

Jovanovic, B. and G. M. McDonald (1994). Competitive diffusion. *Journal of Political Economy 102*, 24–52.

Jovanovic, B. and R. Rob (1989). The growth and diffusion of knowledge. *Review of Economic Studies 56*, 569–582.

Kabluchko, Z. (2009). Spectral representations of sum-and max-stable processes. *Extremes 12*(4), 401.

Kortum, S. (1997). Research, patenting, and technological change. *Econometrica 65*(6).

Lashkaripour, A. and V. Lugovskyy (2017). Industry-level scale economies: from micro-estimation to macro-implications. *Mimeo, Indiana University*.

Levchenko, A. and J. Zhang (2014). Ricardian productivity differences and the gains from trade. *European Economic Review 65*, 45–65.

Levchenko, A. and J. Zhang (2016). The evolution of comparative advantage: Measurement and welfare implications. *Journal of Monetary Economics 78*, 96–111.

Lind, N. and N. Ramondo (2018a). Innovation, knowledge diffusion, and globalization. *NBER WP 25071*.

Lind, N. and N. Ramondo (2018b). Trade with correlation. *NBER Working Paper Series* (24380).

Lucas, R. E. (2009). Ideas and growth. *Economica 76*(301), 1–19.

Lucas, R. E. and B. Moll (2014). Knowledge growth and the allocation of time. *Journal of Political Economy 122*(1), 1–51.

Luttmer, E. (2012). Eventually, noise and imitation implies balanced growth. *Federal Reserve Bank of Minneapolis Working Paper* (699).

Ossa, R. (2015). Why trade matters after all. *Journal of International Economics 97*(2), 266–277.

Perla, J. and C. Tonetti (2014). Equilibrium imitation and growth. *Journal of Political Economy 122*, 512–76.

Ramondo, N. and A. Rodríguez-Clare (2013). Trade, multinational production, and the gains from openness. *Journal of Political Economy 121*(2), 273–322.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016). Trade, domestic frictions, and scale effects. *American Economic Review 106*(10), 3159–84.

Redding, S. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167.

A Proof of Lemma 1

Proof.

$$
\begin{aligned}
P_d(t) &= \left(\int_0^1 P_d(t;v)^{-\sigma}\mathrm{d}v\right)^{-\frac{1}{\sigma}} = \mathbb{E}\left[P_d(t;v)^{-\sigma}\right]^{-\frac{1}{\sigma}} \\
&= \mathbb{E}\left[\left(\min_{o=1,\ldots,N}\frac{W_o(t)}{A_{od}(t)}\right)^{-\sigma}\right]^{-\frac{1}{\sigma}} = \mathbb{E}\left[\max_{o=1,\ldots,N} A_{od}(t)^{\sigma}W_o(t)^{-\sigma}\right]^{-\frac{1}{\sigma}} \\
&= G^d(T_{1d}(t)W_1(t)^{-\sigma},\ldots,T_{Nd}(t)W_N(t)^{-\sigma};t)^{-\frac{1}{\sigma}}
\end{aligned}
$$

□

B Proof of Theorem 1

Proof. Sufficiency follows from showing that the finite dimensional distributions of the productivity are θ-Fréchet. This result follows from Campbell's theorem. Fix K origin-destination-time tuples $\{o_k, d_k, t_k\}_{k=1}^K$. Then

$$
\begin{aligned}
&\mathbb{P}\left[A_{o_k d_k}(t_k;v) \le a_k, \forall k = 1,\ldots,K\right] \\
&\mathbb{P}\left[\max_{i=1,2,\ldots} Z_i(v)A_{o_k d_k}(t_k,\omega_i(v)) \le a_k, \forall k = 1,\ldots,K\right] \\
&\mathbb{P}\left[Z_i(v)A_{o_k d_k}(t_k,\omega_i(v)) \le a_k, \forall k = 1,\ldots,K, \forall i = 1,2,\ldots\right] \\
&\mathbb{P}\left[Z_i(v) \le \min_{k=1,\ldots,K} a_k/A_{o_k d_k}(t_k,\omega_i(v)), \forall i = 1,2,\ldots\right] \\
&\mathbb{P}\left[Z_i(v) > \min_{k=1,\ldots,K} a_k/A_{o_k d_k}(t_k,\omega_i(v)), \text{ for no } i = 1,2,\ldots\right]
\end{aligned}
$$

This last expression is a void probability for the Poisson process. We can compute it by applying Campbell's theorem.

$$\begin{aligned}
&\mathbb{P}\left[Z_i(v) > \min_{k=1,\dots,K} a_k/A_{o_k d_k}(t_k, \omega_i(v)),\ \text{for no } i = 1, 2, \dots\right] \\
&= \exp\left[-\int\int_{\min_{k=1,\dots,K} a_k/A_{o_k d_k}(t_k,\omega)}^{\infty} \theta z^{-\theta-1} \mathrm{d}z \mathrm{d}\mu(\omega)\right] \\
&= \exp\left[-\int \max_{k=1,\dots,K} A_{o_k d_k}(t_k, \omega)^{\theta} a_k^{-\theta} \mathrm{d}\mu(\omega)\right] \\
&= \exp\left[-\int \max_{k=1,\dots,K} \frac{A_{o_k d_k}(t_k, \omega)^{\theta}}{T_{o_k d_k}(t_k)} T_{o_k d_k}(t_k) a_k^{-\theta} \mathrm{d}\mu(\omega)\right]
\end{aligned}$$

for $T_{o_k d_k}(t_k) \equiv \int A_{o_k d_k}(t_k, \omega)^{\theta} \mathrm{d}\mu(\omega)$. This final expression is the joint distribution of a multivariate θ-Fréchet random variable with scale parameters $T_{o_k d_k}(t_k)$ for each $k = 1, \dots, K$ and correlation function $(x_1, \dots, x_K) \mapsto \int \max_{k=1,\dots,K} \frac{A_{o_k d_k}(t_k,\omega)^{\theta}}{T_{o_k d_k}(t_k)} x_k \mathrm{d}\mu(\omega)$. Since this result holds for arbitrary collections $\{o_k, d_k, t_k\}_{k=1}^{K}$ for any finite K, all of the finite dimensional distributions of productivity are θ-Fréchet, and productivity is a θ-Fréchet process.

Necessity follows from Theorem 1 in Kabluchko (2009), which states that any θ-Fréchet process has a spectral representation. Let $A_{od}(t)$ be a θ-Fréchet process on $\{1, \dots, N\} \times \{1, \dots, N\} \times \mathbb{R}$. Then there exists a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$, spectral functions $\{A_{od}(t, \omega)\}_{o,d=1,\dots,N,t\in\mathbb{R}}$ with $\int A_{od}(t, \omega)\mathrm{d}\omega < \infty$, and a Poisson process $\{Z_i, \omega_i\}_{i=1,2,\dots}$ with intensity $\theta z^{-\theta-1} \mathrm{d}z \mathrm{d}\mu(\omega)$ such that $A_{od}(t) = \max_{i=1,2,\dots} Z_i A_{od}(t, \omega_i)$. Taking $A_{od}(t; v)$ across $v \in [0, 1]$ to be *iid* copies of this process completes the proof. □

C Proof of Lemma 3

Proof. The joint distribution of the best learned understanding conditional on $\chi_i(v)$, $n_{i0}(v)$, and $t_{i0}(v)$ is

$$
\begin{aligned}
&\mathbb{P}\left[U^*_{in}(v) \leq u_n, \forall n = 1, \ldots, N \mid \chi_i(v), n_{i0}(v), t_{i0}(v)\right] \\
&= \mathbb{P}\left[\max_{j=1,2,\ldots} U_{ij}(v)\mathbf{1}\{n_{ij}(v) = n, t_{ij}(v) \leq t\} \leq u_n, \forall n = 1, \ldots, N \mid \chi_i(v), n_{i0}(v), t_{i0}(v)\right] \\
&= \mathbb{P}\left[\max_{n=1,\ldots,N} \mathbf{1}\{n_{ij}(v) = n, t_{ij}(v) \leq t\} u_n^{-1} \leq U_{ij}(v)^{-1}, \forall j = 1, 2, \cdots \mid \chi_i(v), n_{i0}(v), t_{i0}(v)\right] \\
&= \mathbb{P}\left[\mathbf{1}\{t_{ij}(v) \leq t\} u_{n_{ij}(v)}^{-1} \leq U_{ij}(v)^{-1}, \forall j = 1, 2, \cdots \mid \chi_i(v), n_{i0}(v), t_{i0}(v)\right] \\
&= \mathbb{P}\left[\mathbf{1}\{t_{ij}(v) \leq t\} u_{n_{ij}(v)}^{-1} > U_{ij}(v)^{-1}, \text{ for no } j = 1, 2, \cdots \mid \chi_i(v), n_{i0}(v), t_{i0}(v)\right]
\end{aligned}
$$

Applying Campbell's theorem,

$$
\begin{aligned}
&\mathbb{P}\left[U^*_{in}(v) \leq u_n, \forall n = 1, \ldots, N \mid \chi_i(v), n_{i0}(v), t_{i0}(v)\right] \\
&= \exp\left[-\int_{t_{i0}(v)}^{t} \sum_{n=1}^{N} \int_{u_n}^{\infty} \vartheta(\chi_i(v)\tilde{u}^{-\vartheta(\chi_i(v))-1} \lambda_n^D(\tilde{t} \mid \chi_i(v), n_{i0}(v)) \mathrm{d}\tilde{u}\mathrm{d}\tilde{t}\right] \\
&= \exp\left[-\sum_{n=1}^{N} \Lambda_n^D(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v)) u_n^{-\vartheta(\chi_i(v))}\right]
\end{aligned}
$$

This distribution is the joint distribution of independent Fréchet random variables with scales equal to $\Lambda_n^D(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v))$ and common shape of $\vartheta(\chi_i(v))$. □

D Proof of Proposition 2

Proof. Productivity is

$$
A_{od}(t; v) = \max_{i=1,2,\ldots} Z_i(v) A_{iod}(t; v).
$$

where applicability is

$$
\begin{aligned}
A_{iod}(t; v) &\equiv A_{od}(\chi_i(v)) \max_{j=0,1,\ldots} U_{ij}(v)\mathbf{1}\{n_{ij}(v) = o, t_{ij}(v) \leq t\} \\
&= A_{od}(\chi_i(v)) \max\left\{U_{i0}(v)\mathbf{1}\{n_{i0}(v) = o, t_{i0}(v) \leq t\}, U^*_{io}(t; v)\right\}
\end{aligned}
$$

By Lemma 3, $U^*_{io}(t;v) \mid \chi_i(v), n_{i0}(v), t_{i0}(v)$ is independent Fréchet over o with scale $\Lambda^D_n(t \mid \chi_i(v), n_{i0}(v), t_{i0}(v))$ and shape $\vartheta(\chi_i(v))$. Under Assumption 6, $U_{i0}(v) \mid \chi_i(v), n_{i0}(v), t_{i0}(v)$ is iid and distribution Fréchet with unit scale and shape $\vartheta(\chi_i(v))$.

Then the collection $\{Z_i(v), \chi_i(v), U_{i0}(v), n_{i0}(v), t_{i0}(v), \{U^*_{i0}(t;v)\}_{o=1}^N\}_{i=1,2,\ldots}$ is a marked Poisson process. By the marking theorem, it can be taken as a Poisson process on the product space $\mathbb{R}_+ \times [0,1] \times \mathbb{R}_+ \times \{1,\ldots,N\} \times \mathbb{R} \times \mathbb{R}^N_+$ with intensity for country n of

$$\theta z^{-\theta-1} \frac{\partial}{\partial u} e^{-u^{-\vartheta(\chi)}} \lambda^I_n(\chi,t) \prod_{o=1}^N \frac{\partial}{\partial u_o} e^{-\Lambda^D_n(t|\chi_i(v),n_{i0}(v),t_{i0}(v)) u_o^{-\vartheta(\chi)}} \mathrm{d}z\mathrm{d}\chi\mathrm{d}u\mathrm{d}t\mathrm{d}u_1\ldots\mathrm{d}u_N$$

Then the joint distribution of productivity is

$$\begin{aligned}
&\mathbb{P}[A_{od}(t;v) \le a_o, \forall o = 1,\ldots,N] \\
&= \mathbb{P}\left[\max_{i=1,2,\ldots} Z_i(v) A_{iod}(t;v) \le a_o, \forall o = 1,\ldots,N\right] \\
&= \mathbb{P}\left[\max_{o=1,\ldots,N} A_{iod}(t;v) a_o^{-1} \le Z_i(v)^{-1}, \forall i = 1,2,\ldots\right] \\
&= \mathbb{P}\Bigg[\max_{o=1,\ldots,N} A_{od}(\chi_i(v)) \max\left\{U_{i0}(v)\mathbf{1}\{n_{i0}(v) = o, t_{i0}(v) \le t\}, U^*_{io}(t;v)\right\} a_o^{-1} > Z_i(v)^{-1}, \\
&\qquad\qquad \text{for no } i = 1,2,\ldots\Bigg]
\end{aligned}$$

By Campbell's theorem

$$\begin{aligned}
&\mathbb{P}[A_{od}(t;v) \le a_o, \forall o = 1,\ldots,N] \\
&= \exp\Bigg[-\int_0^\infty \cdots \int_0^\infty \int_{-\infty}^t \sum_{n=1}^N \int_0^\infty \int_0^1 \int_{1/\max_{o=1}^N A_{od}(\chi)\max\{u\mathbf{1}\{n=o\},u_o\}a_o^{-1}}^\infty \\
&\qquad \times \theta z^{-\theta-1} \frac{\partial}{\partial u} e^{-u^{-\vartheta(\chi)}} \lambda^I_n(\chi,\tilde{t}) \prod_{o=1}^N \frac{\partial}{\partial u_o} e^{-\Lambda^D_n(t|\chi,n,\tilde{t}) u_o^{-\vartheta(\chi)}} \mathrm{d}z\mathrm{d}\chi\mathrm{d}u\mathrm{d}\tilde{t}\mathrm{d}u_1\ldots\mathrm{d}u_N\Bigg] \\
&= \exp\Bigg[-\int_{-\infty}^t \int_0^1 \sum_{n=1}^N \int_0^\infty \cdots \int_0^\infty \int_0^\infty \left[\max_{o=1,\ldots,N} A_{od}(\chi) \max\left\{u\mathbf{1}\{n=o\}, u_o\right\} a_o^{-1}\right]^\theta \\
&\qquad \times \frac{\partial}{\partial u} e^{-u^{-\vartheta(\chi)}} \prod_{o=1}^N \frac{\partial}{\partial u_o} e^{-\Lambda^D_n(t|\chi,n,\tilde{t}) u_o^{-\vartheta(\chi)}} \mathrm{d}u\mathrm{d}u_1\ldots\mathrm{d}u_N \lambda^I_n(\chi,\tilde{t})\mathrm{d}\chi\mathrm{d}\tilde{t}\Bigg]
\end{aligned}$$

Note that $\frac{\partial}{\partial u} e^{-u^{-\vartheta(\chi)}} \prod_{o=1}^N \frac{\partial}{\partial u_o} e^{-\Lambda_n^D(t|\chi,n,\tilde{t}) u_o^{-\vartheta(\chi)}}$ is the joint density of $N+1$ independent $\vartheta(\chi)$-Fréchet random variables and under this density the distribution of $\{\max\{u\mathbf{1}\{n=o\}, u_o\}\}_{o=1}^N$ is also $\vartheta(\chi)$-Fréchet, by max-stability, and for each o has scale $\mathbf{1}\{n = o\} + \Lambda_o^D(t \mid \chi, n, \tilde{t})$. As a consequence, the distribution of $\max_{o=1,\ldots,N} A_{od}(\chi) \max\{u\mathbf{1}\{n=o\}, u_o\} a_o^{-1}$ under this density is $\vartheta(\chi)$-Fréchet with scale of $\sum_{o=1}^N A_{od}(\chi)^{\vartheta(\chi)}[\mathbf{1}\{n=o\} + \Lambda_o^D(t \mid \chi, n, \tilde{t})] a_o^{-\vartheta(\chi)}$. As a result,

$$
\begin{aligned}
&\mathbb{P}[A_{od}(t;v) \le a_o, \forall o = 1, \ldots, N] \\
&= \exp\left[-\int_{-\infty}^t \int_0^1 \sum_{n=1}^N \int_0^\infty x^\theta \frac{\partial}{\partial u} e^{-\sum_{o=1}^N A_{od}(\chi)^{\vartheta(\chi)}[\mathbf{1}\{n=o\}+\Lambda_o^D(t|\chi,n,\tilde{t})] a_o^{-\vartheta(\chi)} x^{-\vartheta(\chi)}} \mathrm{d}x \lambda_n^I(\chi,\tilde{t}) \mathrm{d}\chi \mathrm{d}\tilde{t} \right] \\
&= \exp\Bigg[-\int_{-\infty}^t \int_0^1 \sum_{n=1}^N \Gamma(1-\theta/\vartheta(\chi)) \left[\sum_{o=1}^N A_{od}(\chi)^{\vartheta(\chi)}[\mathbf{1}\{n=o\} + \Lambda_o^D(t \mid \chi, n, \tilde{t})] a_o^{-\vartheta(\chi)} \right]^{\frac{\theta}{\vartheta(\chi)}} \\
&\qquad\qquad \times \lambda_n^I(\chi,\tilde{t}) \mathrm{d}\chi \mathrm{d}\tilde{t} \Bigg] \\
&= \exp\left[-\int_{-\infty}^t \int_0^1 \sum_{n=1}^N \left[\sum_{o=1}^N \left(T_{od}(t \mid \chi, n, \tilde{t}) a_o^{-\theta} \right)^{\frac{1}{1-\rho(\chi)}} \right]^{1-\rho(\chi)} \lambda_n^I(\chi,\tilde{t}) \mathrm{d}\chi \mathrm{d}\tilde{t} \right]
\end{aligned}
$$

where $T_{od}(t \mid \chi, n, \tilde{t}) \equiv \Gamma(\rho(\chi)) A_{od}(\chi)^\theta [\mathbf{1}\{n=o\} + \Lambda_o^D(t \mid \chi, n, \tilde{t})]^{1-\rho(\chi)}$ and $\rho(\chi) \equiv 1 - \theta/\vartheta(\chi)$. This result states that productivity is distributed θ-Fréchet with scale parameters of

$$
T_{od}(t) \equiv \int_{-\infty}^t \int_0^1 \sum_{n=1}^N T_{od}(t \mid \chi, n, \tilde{t}) \lambda_n^I(\chi,\tilde{t}) \mathrm{d}\chi \mathrm{d}\tilde{t}
$$

and correlation function

$$
G^d(x_1, \ldots, x_N, t) \equiv \int_{-\infty}^t \int_0^1 \sum_{n=1}^N \left[\sum_{o=1}^N \left(\alpha_{od}(t \mid \chi, n, \tilde{t}) x_o \right)^{\frac{1}{1-\rho(\chi)}} \right]^{1-\rho(\chi)} \lambda_n^I(\chi,\tilde{t}) \mathrm{d}\chi \mathrm{d}\tilde{t}.
$$

where $\alpha_{od}(t \mid \chi, n, \tilde{t}) \equiv T_{od}(t \mid \chi, n, \tilde{t}) / T_{od}(t)$. $\square$