From Discrete Choice to Endogenous Distributions: Selection and Heterogeneity

Sergey Nigai*

January, 2026

Abstract

This paper develops a tractable framework in which heterogeneity, captured by the distribution of agent characteristics within an alternative (such as a sector, location, or occupation), emerges endogenously in equilibrium through selection rather than being imposed exogenously. I characterize conditions under which widely used extreme-value and power-law distributions arise as equilibrium outcomes of payoff-based discrete choice with multivariate shocks. The framework delivers a one-to-one analytical mapping between equilibrium allocation shares and the distribution of agent characteristics within each alternative. As a result, heterogeneity is endogenously determined in equilibrium through selection. Under counterfactual shocks, this mechanism reshapes heterogeneity and generates adjustment margins absent from standard models. Applications to trade, economic geography, and occupational sorting show how endogenous distributions generate richer responses in canonical general-equilibrium settings with minimal structural modifications.

Keywords: *Endogenous Distributions, Endogenous Heterogeneity, International trade, Fréchet distribution, Dagum distribution, Heavy-tailed distributions, Economic geography, Selection*

JEL-codes: F10, F20, R12, C6.

*University of Colorado Boulder and CESIfo. I thank Costas Arkolakis, Hamid Firooz, Taylor Jaworski, Wolfgang Keller and participants of EIIT 2025 for comments and suggestions. I am grateful to Carlos Martins-Filho for insightful discussions on multivariate distributions and copula dependence.
The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

Heavy-tailed distributions are widely used in quantitative economics to model heterogeneity in fundamental characteristics of economic agents, such as firm productivity and worker skill. In most frameworks, however, the distribution of heterogeneity among agents operating within a given economic alternative, such as a sector, location, or occupation, is assumed to be exogenous and therefore invariant to economic conditions. This paper develops a tractable framework in which heterogeneity is an equilibrium outcome of payoff-based discrete choice that adjusts endogenously to economic shocks.

The key insight is that, under suitable optimization and stochastic structures, the allocation of agents across discrete alternatives and the heterogeneity of agents active within a given alternative are two sides of the same structural relationship. I establish this equivalence for two widely used heavy-tailed families, extreme-value and power-law, in settings where agents draw shocks from multivariate distributions and select among alternatives. The resulting mappings operate in both directions and extend to general equilibrium, while preserving analytical transparency. As a result, the framework endogenizes heavy-tailed heterogeneity in a wide class of heterogeneous-agent models through discrete-choice allocation problems.

To fix ideas on the usefulness of the results, consider heterogeneous-firm models of international trade in which firm heterogeneity takes the form of productivity draws from extreme-value or power-law distributions. In these settings, the framework developed here implies that the productivity distribution of firms operating in a given country is not an exogenous primitive, but an equilibrium outcome shaped by firms' location choices. As a result, the mass of active firms and their productivity distribution are jointly determined in equilibrium. This approach to counterfactual analysis therefore differs from standard approaches by allowing both firm allocation and within-country heterogeneity to adjust endogenously, rather than assuming fixed productivity distributions.

More generally, the framework provides a unified way to endogenize distributions of heterogeneity across economic environments by deriving them from payoff-based discrete choice. Changes in fundamentals translate into systematic adjustments in selection patterns and the resulting distributions of active agents, generating richer

equilibrium responses than in models where distributions are imposed exogenously. Because the framework accommodates correlated shocks, it applies broadly across settings with heterogeneous agents and discrete choices. Applications to international trade, spatial models, and occupational sorting illustrate how distributions that are typically treated as fixed inputs can instead be recovered as endogenous equilibrium objects within existing general-equilibrium structures.

Section 3 introduces a payoff-based selection environment and specifies conditions under which both allocation shares and within-alternative heterogeneity are endogenously determined in equilibrium rather than imposed exogenously. For Fréchet shocks with logistic dependence, equilibrium choice shares admit closed-form expressions, and the within-alternative distribution remains Fréchet, with its scale parameter determined by the corresponding share. For Dagum shocks with Clayton dependence, an analogous characterization holds for power-law heterogeneity. In both cases, the framework delivers explicit mappings between equilibrium allocation shares and within-alternative distributions, accommodates dependence across alternatives, and preserves full analytical tractability.

Section 4 applies the results in three settings. In trade, the framework delivers endogenous firm location choice within the Eaton and Kortum (2002) and Melitz (2003) environments without altering their underlying consumption, production, or trade structures. In economic geography, it extends the Allen and Arkolakis (2014) model to incorporate heterogeneous labor productivity both across and within locations while retaining explicit congestion and agglomeration effects. In a Roy model of occupational choice, as in Hsieh et al. (2019), it yields closed-form sorting patterns with correlated abilities across occupations. Collectively, these applications illustrate how endogenizing heterogeneity within alternatives introduces new, analytically tractable adjustment margins into standard general-equilibrium frameworks, yielding richer counterfactual predictions.

2 Related Literature

This paper relates to a large literature that uses distribution-based approaches to model heterogeneity in economics. In this work, heterogeneity is typically represented using extreme-value or power-law families, which provide parsimonious and tractable representations of empirically relevant characteristics. A common feature

of these approaches is that the distribution of heterogeneity within a given economic environment is imposed exogenously and treated as a primitive of the model. The framework developed here speaks directly to this literature by providing a structural interpretation of heavy-tailed distributions.

First, the paper relates to the literature that employs extreme-value distributions to model heterogeneity. The Fréchet distribution, introduced into international trade by Eaton and Kortum (2002), has become a workhorse for modeling firm heterogeneity (e.g., Bernard et al., 2003; Costinot et al., 2011; Caliendo and Parro, 2015), multinational production (e.g., Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2017), migration decisions, and assignment outcomes (e.g., Lagakos and Waugh, 2013; Redding, 2016; Fajgelbaum et al., 2018; Burstein et al., 2019; Hsieh et al., 2019; Tombe and Zhu, 2019). In these models, Fréchet-distributed heterogeneity is typically treated as a primitive. The results in this paper provide a structural foundation for interpreting such outcomes as equilibrium distributions generated by discrete selection across alternatives. This perspective makes it possible to introduce new margins – such as endogenous location choice by firms – into standard frameworks like Eaton and Kortum (2002) with minimal modification to their underlying structure.

Second, the results also encompass models based on power-law distributions, building on a long tradition of power-law modeling in economics (e.g., Simon and Bonini, 1958; Axtell, 2001; Gabaix, 1999, 2009). The Pareto distribution, combined with the firm-heterogeneity framework of Melitz (2003), has become standard for analyzing firm selection and market outcomes (e.g., Chaney, 2008; Arkolakis et al., 2012; di Giovanni and Levchenko, 2012),[1] selection into trade and export modes (Helpman et al., 2004; Arkolakis et al., 2008; Arkolakis, 2010; Eaton et al., 2011), and the distributional effects of globalization (Helpman et al., 2010; Nigai and Yang, 2024).[2] The framework developed here provides a structural link between discrete selection and heterogeneity when the latter follows a power law, allowing distributions that are typically imposed exogenously to be recovered as endogenous equilibrium outcomes.

Third, because the goal of this paper is to endogenize distributional heterogeneity while preserving analytical tractability, it builds on the random-utility tradition in

[1] For alternative distributions capturing firm-level heterogeneity, see Nigai (2017).

[2] Beyond trade, Pareto distributions are widely used in macroeconomic analyses of firm dynamics and income distribution (e.g., Luttmer, 2007; Gabaix, 2011), as well as in spatial and regional economics (Rossi-Hansberg and Wright, 2007).

discrete-choice theory. I show that both extreme-value and power-law distributions can be obtained as conditional outcome distributions arising from discrete-choice problems in which agents maximize economic payoffs. The analysis is therefore rooted in the framework of McFadden (1974) for conditional logit models with independent alternatives and its extensions to correlated choices through generalized extreme value and nested logit models (McFadden, 1978). These typically focus on choice probabilities and leave the equilibrium distribution of outcomes implicit, while this paper brings these distributions to the foreground. Discrete-choice models have been widely applied to study worker and firm sorting across locations and sectors (e.g., Grogger and Hanson, 2011; Allen and Arkolakis, 2014; Ahlfeldt et al., 2015; Holmes and Sieg, 2015; Redding and Rossi-Hansberg, 2017; Monte et al., 2018; Bryan and Morten, 2019; Fajgelbaum and Gaubert, 2020), as well as labor reallocation responses to trade shocks (Artuç et al., 2010; Dix-Carneiro, 2014; Caliendo et al., 2019; Galle et al., 2023), and they integrate naturally into spatial and differentiated-product settings (Anderson et al., 1992). This paper complements these approaches by making the distributional implications of selection explicit and endogenous.

Finally, this paper relates to frameworks that employ multivariate probability distributions within the extreme-value and power-law families. Existing work exploits special properties of these distributions, such as max-stability, to preserve analytical tractability. For example, Lind and Ramondo (2023a) and Lind and Ramondo (2023b) use multivariate Fréchet distributions with correlation to capture shared technology draws across sectors, where max-stability plays a central role in delivering closed-form results. The objective of the present paper is complementary. Rather than taking the joint distribution as given, it provides microfoundations for potentially correlated heavy-tailed distributions by linking their structure to underlying payoff-based discrete-choice behavior. In this sense, while the analysis also employs multivariate Fréchet distributions in the extreme-value case, it shows that the resulting conditional distributions arise endogenously as equilibrium outcomes of the sorting process itself. As a consequence, changes driven by economic shocks reshape the distribution of heterogeneity within each economic environment, yielding richer counterfactual responses.

A similar perspective applies to multivariate Pareto frameworks with Pareto tails and Pareto maxima, developed by Arkolakis et al. (2018) and Arkolakis et al. (2025),

which are particularly well suited for capturing extreme realizations such as product- or destination-level performance. While these approaches treat the joint distribution as a maintained object throughout counterfactual analysis, the framework developed here endogenizes it. This provides a theoretical foundation for the emergence of correlated heavy-tailed distributions and introduces additional margins of adjustment in comparative statics.

In sum, this paper links discrete selection and heterogeneity within a single, internally consistent, and analytically tractable structure. The approach is useful for heterogeneous-agent frameworks that would benefit from endogenizing distributional heterogeneity, as well as for models based on discrete-choice mechanisms that would benefit from an explicit characterization of conditional outcome distributions. As a result, it extends naturally to a wide range of settings in trade, economic geography, labor, industrial organization, and macroeconomics in which both selection and heterogeneity play a central role.

3 Selection and Heterogeneity

The central idea of this paper is that heterogeneity can be viewed as an equilibrium outcome of a selection problem. Before turning to the formal results, it is useful to build intuition through a simple example. Consider an economy with two alternatives, A and B, which can be interpreted as different locations, sectors, or occupations. Agents draw idiosyncratic characteristics ε, which may be interpreted as firm productivity or worker skill, and select between these alternatives.

Suppose that the marginal distributions of ε associated with alternatives A and B take the Fréchet form:

$$F_A(\varepsilon) = \exp(-T_A\varepsilon^{-\theta}) \quad \text{and} \quad F_B(\varepsilon) = \exp(-T_B\varepsilon^{-\theta}).$$

For the purpose of exposition, I begin by assuming independence of shocks across alternatives, as in Eaton and Kortum (2002), and postpone the discussion of correlated environments.

Suppose that $T_B > T_A$, so that the expected value of ε is higher in alternative B. Economically, this can be interpreted as B being associated with more productive firms or more skilled workers on average. While the mapping from distributions $F_A(\varepsilon)$

and $F_B(\varepsilon)$ to equilibrium outcomes is well understood in many economic settings, such as comparative advantage and international trade (Eaton and Kortum, 2002), much less attention has been paid to how these distributions themselves arise and why they differ across alternatives.

The key point emphasized in this paper is that selection into alternatives A and B, driven by economic forces, may explain the differences between the distributions $F_A(\varepsilon)$ and $F_B(\varepsilon)$. They can be interpreted as conditional distributions of ε among agents who choose each alternative. To see this, suppose that agents draw a vector of idiosyncratic characteristics $\boldsymbol{\varepsilon} = (\varepsilon_A, \varepsilon_B)$ from a joint distribution with cumulative distribution function:

$$F(\boldsymbol{\varepsilon}) = \exp\left(-\left[\varepsilon_A^{-\theta} + \varepsilon_B^{-\theta}\right]\right),$$

which corresponds to a multivariate Fréchet distribution with identical independent marginals. While this example still assumes independence for simplicity, the full framework accommodates correlation across alternatives.

Upon drawing $\boldsymbol{\varepsilon} = (\varepsilon_A, \varepsilon_B)$, agents choose between A and B by maximizing payoffs of the form:

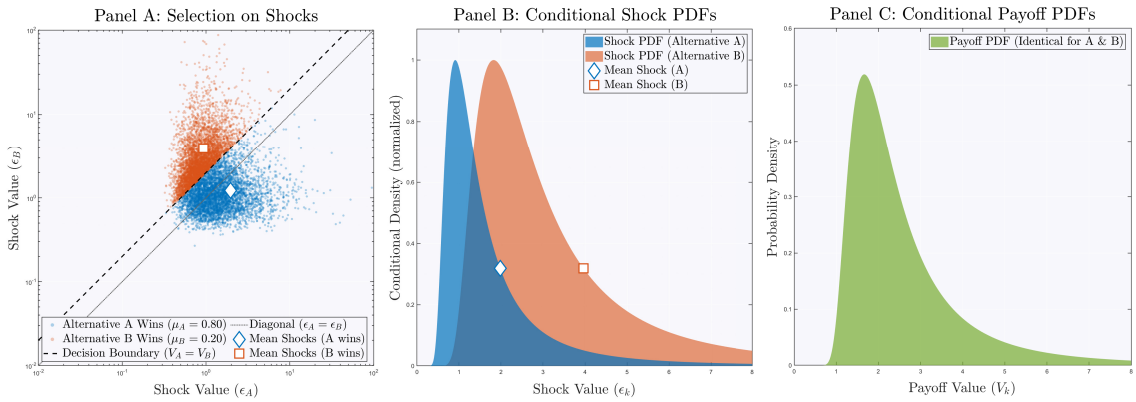
$$V_k = W_k \varepsilon_k^{\omega}, \qquad W_k > 0,\ \omega > 0,$$

where W_k denotes an alternative-specific payoff shifter. In the context of firms, W_k can be interpreted as profits available in alternative k, while in the context of workers it can be interpreted as wages. The parameter ω governs the responsiveness of agents' choices to differences in idiosyncratic characteristics across alternatives.

Let $W_A > W_B$, so that A is the "favorite" alternative and B is the "underdog". Figure 1 provides a visual illustration of how selection shapes observed heterogeneity in this setting. The three panels trace the mechanism from the joint distribution of shocks and the selection rule (Panel A), to the resulting within-alternative distributions $F_A(\varepsilon)$ and $F_B(\varepsilon)$ (Panel B), and finally to the equilibrium implications summarized by the distribution of realized payoffs (Panel C).

Panel A illustrates the selection process in the two-dimensional space of idiosyncratic shocks. Each point corresponds to an agent's draw $(\varepsilon_A, \varepsilon_B)$ and is colored according to the alternative selected. The dashed black line depicts the decision boundary defined by payoff indifference, $V_A = V_B$. Because the favorite alternative A has a higher

payoff shifter, $W_A > W_B$, this boundary is steeper than the main diagonal (shown by the dotted gray line, where $\varepsilon_A = \varepsilon_B$). Since selection operates on payoffs rather than on raw shocks, agents who select the underdog B must draw sufficiently large ε_B to offset both the competing shock ε_A and the intrinsic advantage of alternative A. As a result, the conditional mean of ε_B is higher among agents who choose B. Let μ_A and μ_B denote the equilibrium shares of agents selecting A and B, respectively; in this example, the selection outcome is characterized by $\mu_A > \mu_B$.



Notes: The figure illustrates the result with two alternatives, A and B, with choice shares $\mu_A = 0.8$ and $\mu_B = 0.2$. The asymmetry is driven by different payoff shifters ($W_A > W_B$) and $\omega = 0.8$, while shocks are drawn from identical i.i.d. Fréchet distributions ($\theta = 2$). Panel A shows simulated draws in $(\varepsilon_A, \varepsilon_B)$ space, the decision boundary ($V_A = V_B$), the diagonal ($\varepsilon_A = \varepsilon_B$), and the mean winning shocks (markers). Panel B plots normalized conditional shock PDFs. Panel C plots the conditional payoff PDFs.

Figure 1: Simulated Draws from a Bivariate Fréchet Model

Panel B shows that the differences in heterogeneity across alternatives A and B can be understood as outcomes of the selection process. The distributions $F_A(\varepsilon)$ and $F_B(\varepsilon)$, introduced earlier as exogenous objects, are in fact the conditional distributions of winning shocks in each alternative. Crucially, differences between these distributions are generated by differences in economic payoffs, even when the joint distribution of shocks $F(\boldsymbol{\varepsilon})$ is symmetric. The conditional distribution for the favorite (shown in blue) is relatively concentrated near the origin, whereas the conditional distribution for the underdog (shown in red) is shifted to the right. Because the underdog is intrinsically disadvantaged, it succeeds only when it receives an exceptionally large shock, while the favorite can prevail with more moderate draws. Building on this intuition, the results in this paper show formally that the equilibrium choice shares μ_A and μ_B determine the parameters of the conditional distributions, such as T_A and

T_B, which are therefore endogenous and respond to economic shocks captured by changes in W_A and W_B, respectively.

The final panel illustrates another crucial property of the approach developed in this paper, namely the general-equilibrium consistency of the conditional distributions $F_A(\varepsilon)$ and $F_B(\varepsilon)$. Although the conditional distributions of shocks differ across alternatives, as shown in Panel B, the implied conditional distributions of payoffs must coincide in equilibrium. Otherwise, agents would have incentives to reallocate. In this example, the higher average shock required for the underdog to win is exactly offset by its lower deterministic payoff shifter, W_B. Selection on shocks therefore compensates for differences in deterministic payoffs, yielding an identical equilibrium distribution of realized payoffs across alternatives. This mechanism lies at the core of the results in this paper: endogenizing distributions through selection remains analytically tractable because complex, alternative-specific selection on shocks collapses into a simple, choice-invariant distribution of equilibrium payoffs.

This simple bivariate example illustrates the core logic behind endogenizing heterogeneity: choice shares and alternative-specific distributions are jointly determined by the same stochastic structure and adjust together in equilibrium. The example highlights three features that carry over to the general framework. First, equilibrium choice probabilities admit closed-form expressions. Second, the conditional distributions of shocks have parameters that are pinned down by these choice shares. Third, equilibrium requires expected payoffs to be equalized across alternatives. I now extend this intuition to the general case.

I begin with the case that endogenizes extreme-value heterogeneity under a potentially dependent shock structure. I show that Fréchet distributions arise as conditional equilibrium outcomes of a payoff-based discrete-choice model in which the underlying shocks follow a multivariate Fréchet distribution with logistic (Gumbel–Hougaard) dependence. This dependence structure is the same as in Lind and Ramondo (2023a). However, while their work uses correlated Fréchet shocks to capture shared technology draws across sectors and to study their implications for patterns of international trade, the focus here is different. The results in this paper establish complementary and novel economic properties of this distributional family by showing that marginal Fréchet distributions can be rationalized as endogenous equilibrium outcomes of payoff-based selection.

Proposition 1 (Endogenizing Fréchet-Based Heterogeneity)**.** Consider a Fréchet distribution of the form

$$F_k(\varepsilon) = \exp\left(-T_k \varepsilon^{-\theta}\right), \qquad T_k > 0, \theta > 0,$$

interpreted as the distribution of heterogeneity among agents in alternative k.

There exists a payoff-based discrete-choice environment in which this distribution arises endogenously in equilibrium. In particular, suppose agents choose among alternatives $k \in \mathcal{K}$ by solving

$$k^* = \arg\max_{k \in \mathcal{K}} \{V_k\}, \qquad V_k = W_k \varepsilon_k^{\omega}, \quad W_k > 0,\ \omega > 0,$$

where V_k is total payoff and the vector of shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ has joint cumulative distribution function

$$F(\boldsymbol{\varepsilon}) = \exp\left(- \left(\sum_{k=1}^{K} \left[H_k \varepsilon_k^{-\theta} \right]^{\rho} \right)^{1/\rho} \right),$$

where $H_k > 0$ and $\rho \geq 1$. The parameter ρ governs the dependence structure across alternatives: $\rho = 1$ corresponds to independent shocks, while larger values of ρ allow for positive dependence. Let $\mu_k \equiv \mathbb{P}(k^* = k)$ denote the equilibrium choice share of alternative k. Then the following results hold.

(i) The equilibrium choice shares are given by

$$\mu_k = \frac{\left(H_k W_k^{\theta/\omega} \right)^{\rho}}{\sum_{n=1}^{K} \left(H_n W_n^{\theta/\omega} \right)^{\rho}}.$$

(ii) The conditional distribution of ε_k given selection coincides with $F_k(\varepsilon)$:

$$F_k(\varepsilon) = \exp\left(-H_k \, \mu_k^{-1/\rho} \varepsilon^{-\theta} \right) \qquad \text{such that} \qquad T_k = H_k \, \mu_k^{-1/\rho}.$$

Hence, the scale of within-alternative heterogeneity is endogenously determined by the equilibrium choice share.

(iii) If $\theta > 1$, the expected shock conditional on choosing k satisfies

$$\mathbb{E}[\varepsilon_k \mid k^* = k] = H_k^{1/\theta} \mu_k^{-1/(\theta\rho)} \Gamma(1 - 1/\theta).$$

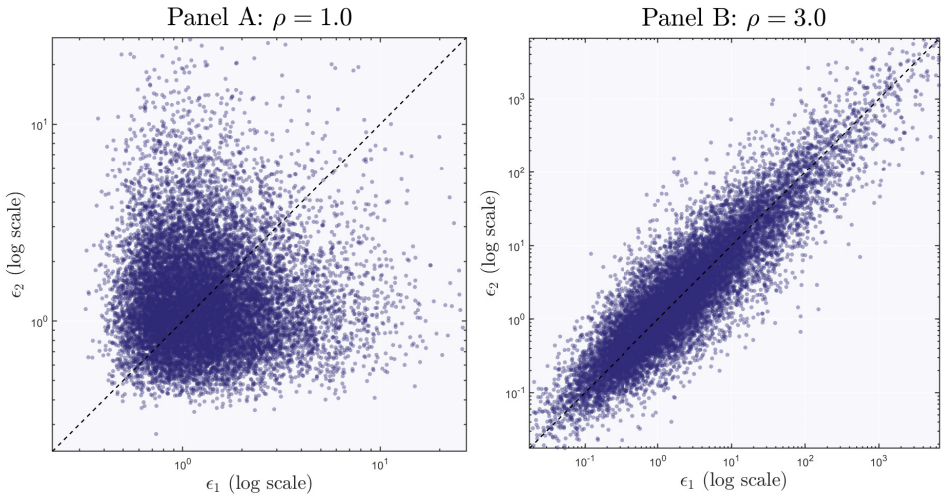
If $\theta > \omega$, the expected payoff conditional on selection is

$$\mathbb{E}[V_k \mid k^* = k] = \left(\sum_{n=1}^{K} \left(H_n W_n^{\theta/\omega} \right)^{\rho} \right)^{\omega/(\theta\rho)} \Gamma(1 - \omega/\theta),$$

which is identical across alternatives.

Proof. See Appendix A. □

Proposition 1 shows how Fréchet heterogeneity within each alternative can be endogenized as an equilibrium outcome of a payoff-based selection problem. Under a multivariate Fréchet distribution with logistic dependence, equilibrium choice shares determine the scale of the conditional distribution of heterogeneity within each alternative, while expected payoffs are equalized across choices. As a result, both the discrete-choice selection and the distribution of heterogeneity within alternatives are pinned down by the same underlying economic structure.



Notes: The figure illustrates 15,000 simulated draws from a bivariate Fréchet distribution with common parameters $H_k = 1$, and tail parameter $\theta = 2$. Panel A shows the independence case where the dependence parameter $\rho = 1$. Panel B shows the case with positive dependence where $\rho = 3$. Both panels are plotted on a log-log scale.

Figure 2: Simulated Draws from Bivariate Fréchet with Logistic Dependence

This structure captures two features that are empirically prominent in many economic environments: heavy-tailed heterogeneity and positive dependence across alternatives. The closed-form expressions delivered by the framework make explicit

how changes in economic conditions that affect selection shares necessarily reshape the distribution of heterogeneity within each alternative, while preserving analytical tractability. This endogeneity is particularly valuable for empirical and counterfactual analysis in settings where selection and heterogeneity co-move in response to shocks. As illustrated in Figure 2, the framework accommodates both independent shocks ($\rho = 1$, Panel A) and positively correlated shocks ($\rho > 1$, Panel B). In all cases, equilibrium choice shares, conditional distributions, and payoffs are jointly determined and admit closed-form characterizations.

A complementary strand of the literature emphasizes power-law heterogeneity (see, e.g., Gabaix, 2009). The Pareto distribution, routinely used to parameterize Melitz (2003) model and its extensions, has become a workhorse in international trade because it delivers sharp predictions for selection, sorting, and aggregate outcomes among heterogeneous firms. While Pareto distributions capture empirically relevant fat-tailed behavior, they are typically imposed exogenously and offer limited tractability for modeling discrete choice with rich dependence across alternatives. To address these issues, I consider the Dagum distribution, which preserves power-law tail decay while allowing for flexible dependence structures. Importantly, this distribution admits a closed-form characterization of equilibrium selection and conditional heterogeneity, making it a natural benchmark for endogenizing power-law heterogeneity in analytically tractable selection models. The formal result is stated in Proposition 2.

Proposition 2 (Endogenizing Power-Law Heterogeneity)**.** Consider a power-law distribution of heterogeneity of the Dagum form

$$F_k(\varepsilon) = \left[1 + \left(\frac{\varepsilon}{D_k}\right)^{-\xi}\right]^{-1/\delta}, \qquad D_k > 0,\ \delta > 0,\ \xi > 0,$$

interpreted as the distribution of heterogeneity among agents selecting alternative k.

There exists a payoff-based discrete-choice environment in which this distribution arises endogenously in equilibrium. In particular, suppose agents choose among alternatives $k \in \mathcal{K}$ by solving

$$k^* = \arg\max_{k \in \mathcal{K}} \{V_k\}, \qquad V_k = W_k \varepsilon_k^{\omega}, \quad W_k > 0,\ \omega > 0,$$

where the vector of shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ has joint cumulative distribution function

$$F(\boldsymbol{\varepsilon}) = \left(1 + \sum_{k=1}^{K} \left[\frac{\varepsilon_k}{H_k}\right]^{-\xi}\right)^{-1/\delta},$$

with parameters $H_k > 0$, $\xi > 0$, and $\delta > 0$. The parameter δ governs the dependence structure across alternatives: larger values of δ correspond to stronger positive dependence. Let $\mu_k \equiv \mathbb{P}(k^* = k)$ denote the equilibrium choice share of alternative k. Then the following results hold.

(i) The equilibrium choice shares are

$$\mu_k = \frac{\left(H_k W_k^{1/\omega}\right)^{\xi}}{\sum_{n=1}^{K} \left(H_n W_n^{1/\omega}\right)^{\xi}}.$$

(ii) The conditional distribution of ε_k given selection coincides with $F_k(\varepsilon)$:

$$F_k(\varepsilon) = \left[1 + \left(\frac{\varepsilon}{D_k}\right)^{-\xi}\right]^{-1/\delta} \qquad \text{such that} \qquad D_k = H_k\, \mu_k^{-1/\xi}.$$

Hence, the scale of power-law heterogeneity within each alternative is endogenously determined by the equilibrium choice share.

(iii) If $\xi > 1$, the expected shock conditional on choosing k satisfies

$$\mathbb{E}[\varepsilon_k \mid k^* = k] = H_k\, \mu_k^{-1/\xi} \cdot \frac{\Gamma(1 - 1/\xi)\, \Gamma(1/\delta + 1/\xi)}{\Gamma(1/\delta)}.$$

If $\xi > \omega$, the expected payoff conditional on selection is

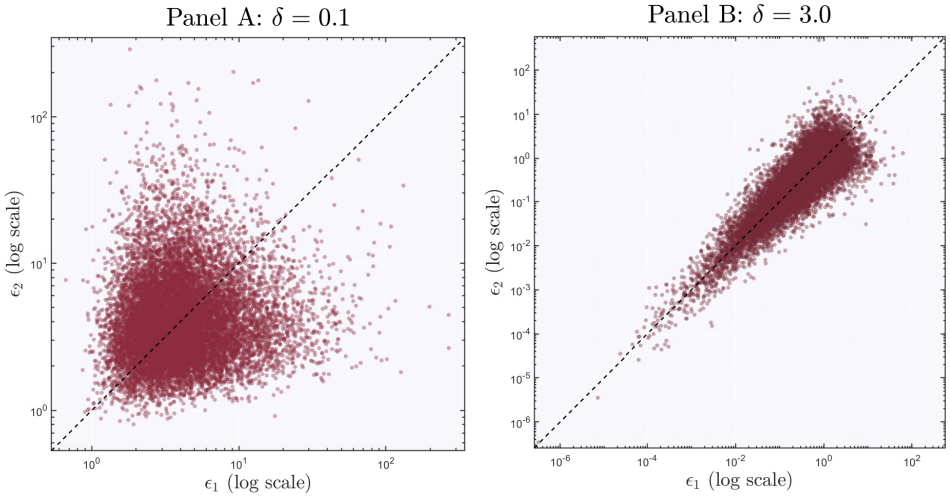
$$\mathbb{E}[V_k \mid k^* = k] = \left(\sum_{n=1}^{K} \left(H_n W_n^{1/\omega}\right)^{\xi}\right)^{\omega/\xi} \cdot \frac{\Gamma(1 - \omega/\xi)\, \Gamma(1/\delta + \omega/\xi)}{\Gamma(1/\delta)},$$

which is identical across alternatives.

Proof. See Appendix B. □

Proposition 2 shows that power-law heterogeneity of the Dagum form can be endogenized as an equilibrium outcome of payoff-based selection under Clayton depen-

dence. In this setting, equilibrium choice shares pin down the scale of heterogeneity within each alternative, while expected payoffs remain equalized across choices. The resulting structure delivers closed-form expressions for selection shares, conditional distributions, and payoffs, all governed by the same tail parameter ξ, which captures both power-law behavior and flexible dependence across alternatives. As illustrated in Figure 3, the framework applies under both independent and positively dependent shocks, providing a tractable foundation for modeling selection and sorting in environments where power-law heterogeneity is empirically salient.



Notes: The figure illustrates 15,000 simulated draws from a bivariate Dagum distribution with common parameters $H_k = 1$ and tail parameter $\xi = 2$. Panel A shows the case where the dependence parameter $\delta = 0.1$. Panel B shows the case where $\delta = 3.0$. Both panels are plotted on a log-log scale.

Figure 3: Simulated Draws from Bivariate Dagum with Clayton Dependence

3.1 Why Fréchet and Dagum? A Unifying Principle

The Fréchet and Dagum families, when combined with logistic and Clayton dependence respectively, play a central role in this paper because they preserve tractability in selection models with correlated shocks. Both families admit a structural representation that clarifies why heterogeneity can be endogenized while retaining closed-form equilibrium expressions. In particular, the shocks can be represented as the product of a common factor that generates dependence across alternatives and an idiosyncratic component that governs relative payoffs and selection. This representation is especially valuable among heavy-tailed families because it allows both equilibrium choice shares and the conditional distributions of heterogeneity within alternatives to

be characterized in closed form, even when shocks are positively dependent.

Formally, in both cases shocks admit a representation of the form

$$\varepsilon_k = \frac{Z_k}{C^{\varphi}},$$

where Z_k are i.i.d. across alternatives, C is a common random factor capturing dependence, and $\varphi > 0$ is a shape parameter. While the distributions of Z_k and C differ across the two families, the economic implication is the same: the realization of the common factor affects all alternatives symmetrically, while heterogeneity in relative payoffs is driven by the idiosyncratic components.

Fréchet with logistic dependence. In the Fréchet case,

$$\varepsilon_k \propto \frac{Z_k}{S^{1/\theta}},$$

where $Z_k \overset{\text{i.i.d.}}{\sim} \text{Fréchet}(\theta)$ are standard Fréchet random variables with shape parameter $\theta > 0$, and S is a positive ρ-stable random variable ($\rho \geq 1$) that captures upper-tail dependence (Hougaard, 1986). Independence obtains when $\rho = 1$.

Under the payoff-based selection rule

$$k^* = \arg\max_k \{W_k \varepsilon_k^{\omega}\},$$

the realized value of S cancels from the comparison across alternatives for a given draw. Nevertheless, the dependence parameter ρ remains economically relevant because it shapes equilibrium choice shares and governs how selection reshapes the distribution of heterogeneity within each alternative. In equilibrium, ρ enters the closed-form expressions for both margins through the effective tail parameter $\theta\rho$. This structure is closely related to the power-sum form underlying the Generalized Extreme Value framework of McFadden (1978), while retaining an explicit characterization of the distribution of outcomes within alternatives.

Dagum with Clayton dependence. In the Dagum case,

$$\varepsilon_k \propto \frac{Z_k}{Y^{1/\xi}},$$

where $Z_k \overset{\text{i.i.d.}}{\sim} \text{Fréchet}(\xi)$ are standard Fréchet random variables with shape parameter $\xi > 0$, and $1/Y$ follows a Gamma distribution with shape parameter $1/\delta > 0$ and unit scale, generating Clayton dependence across alternatives (Nelsen, 2006). Larger values of δ correspond to stronger positive dependence, while $\delta \to 0$ approaches independence.

As in the Fréchet case, the realized value of the common factor cancels from payoff comparisons at the level of individual draws. However, the dependence parameter δ governs how selection endogenously reshapes power-law heterogeneity within alternatives by rescaling the conditional distributions. Together with Fréchet marginals, this dependence structure yields the Dagum family as the equilibrium distribution of heterogeneity.

3.2 Selection and Endogenous Heterogeneity

The results in this section deliver fully closed-form comparative statics implications by jointly characterizing selection across alternatives and the resulting distribution of heterogeneity within each alternative. As a result, changes in economic conditions affect both the allocation of agents across alternatives and the distribution of outcomes among those who select each alternative.

Because heterogeneity is endogenized rather than imposed, the framework provides a transparent and disciplined way to trace how economic shocks, whether arising from policy interventions or exogenous changes in fundamentals, propagate through selection decisions and reshape heterogeneity within each alternative, with implications for both aggregate and distributional outcomes.

Aggregate Outcomes. Propositions 1 and 2 show that equilibrium selection generates systematic differences across alternatives in the conditional distribution of shocks, while expected payoffs are equalized in equilibrium. The conditional shocks can be interpreted as productivity or skill, whereas payoffs reflect profits or welfare that agents arbitrage across alternatives through selection. As a result, aggregate and distributional outcomes are jointly shaped by heterogeneity in fundamentals and by the no-arbitrage condition that equalizes expected returns across alternatives.

For example, suppose ε_k is interpreted as worker skill and labor income takes the form $y(\varepsilon_k) = w\varepsilon_k$. Total income generated in alternative k then depends on both the share of workers selecting that alternative and their average skill level:

$$Y_k = \mu_k \cdot w\,\mathbb{E}[\varepsilon_k \mid k^* = k],$$

where μ_k is the share of agents selecting k and $\mathbb{E}[\varepsilon_k \mid k^* = k]$ is the conditional mean of ε_k. Using Propositions 1 and 2, this expression can be written as:

$$Y_k = \underbrace{\mu_k}_{\text{Selection Margin}} \cdot \underbrace{\mu_k^{-1/\varphi} M_k}_{\text{Endogenous Heterogeneity}},$$

where φ is the effective shape parameter ($\varphi = \theta\rho$ in the Fréchet case, $\varphi = \xi$ in the Dagum case) and M_k is a constant.

This decomposition highlights that aggregate income reflects the interaction between the selection margin and the heterogeneity margin arising from endogenous sorting on skill within that alternative. Changes in economic conditions therefore affect aggregate outcomes not only by reallocating workers across alternatives, but also by reshaping the distribution of skills among those who select each alternative. The elasticity of aggregate income Y_k with respect to the payoff shifter W_k, holding other parameters fixed, can be written as:

$$\eta_k \equiv \frac{\partial \log Y_k}{\partial \log W_k} = \underbrace{\frac{\partial \log \mu_k}{\partial \log W_k}}_{\text{Selection Elasticity}} + \underbrace{\frac{\partial \log \mathbb{E}[\varepsilon_k \mid k^* = k]}{\partial \log W_k}}_{\text{Heterogeneity Elasticity}}.$$

Closed-form expressions for both terms follow directly from Propositions 1 and 2:

$$\frac{\partial \log \mu_k}{\partial \log W_k} = \frac{\varphi}{\omega}(1 - \mu_k), \qquad \frac{\partial \log \mathbb{E}[\varepsilon_k \mid k^* = k]}{\partial \log W_k} = -\frac{1}{\omega}(1 - \mu_k).$$

The first term captures how the mass of agents selecting alternative k responds to changes in relative payoffs, while the second captures the composition effect arising from marginal entrants having lower realizations of shocks. The dependence parameters ρ and δ affect aggregate elasticities through φ, thereby shaping both margins simultaneously. Combining these terms yields:

$$\eta_k = \frac{\varphi - 1}{\omega}(1 - \mu_k).$$

Hence, when $\varphi > 1$, the selection margin dominates, so that increases in W_k raise aggregate income Y_k despite the offsetting composition effect.

Distributional Outcomes. The framework also allows to examine distributional effects within each alternative in closed form. In both the Fréchet and Dagum cases, the conditional distribution belongs to the same parametric family as the underlying shocks: tail behavior is preserved, while the scale parameter adjusts endogenously with the equilibrium choice share. Specifically,

$$\text{Fréchet: scale } \propto \mu_k^{-1/(\theta\rho)}, \qquad \text{Dagum: scale } \propto \mu_k^{-1/\xi}.$$

As a result, the entire distribution of outcomes within each alternative responds systematically to changes in economic conditions through their effect on selection. This structure makes it straightforward to compute moments and inequality measures such as the Gini coefficient in closed form, and to trace how shocks to payoffs or fundamentals reshape the full distribution of heterogeneity and its responses to various shocks.

4 Quantitative Applications

Having established the main theoretical results and their economic interpretation, I now illustrate how the framework can be applied in quantitative settings. The applications are chosen to address three recurring challenges in economics: (i) allowing firm location choices to be endogenous within general-equilibrium models of international trade; (ii) modeling heterogeneity explicitly in discrete-choice frameworks of worker location; and (iii) modeling occupational choice and labor assignment under potentially correlated multidimensional heterogeneity. In each case, the framework delivers endogenous within-alternative heterogeneity as an equilibrium outcome of selection, while preserving analytical tractability and requiring only minimal departures from standard modeling approaches.

4.1 Application A: International Trade Model with Firm Location Choice

In the canonical Eaton and Kortum (2002) framework, both the country-specific scale parameter of the Fréchet productivity distribution and the mass of potential producers in each country are treated as exogenous primitives. Proposition 1 shows that this structure can be reinterpreted by viewing the Fréchet scale parameter T_k as an equilibrium outcome of firms' location choices: T_k reflects the endogenous productivity distribution of firms that select into country k, with location decisions governed by the same economic primitives that shape productivity heterogeneity. This reinterpreta-

tion preserves the closed-form tractability of the Eaton–Kortum model while endogenizing the cross-country allocation of firms μ_k across countries. Consequently, changes in fundamentals affect both the equilibrium distribution of productivity within countries and the allocation of firms across locations.

The dependence parameter ρ plays a central economic role in determining how these margins interact. When $\rho = 1$, shocks are independent across countries, and selection effects exactly offset the induced changes in within-country heterogeneity. In this knife-edge case, the model collapses to the standard Eaton–Kortum structure with exogenous productivity distributions. When $\rho > 1$, positive dependence across locations implies that countries are closer substitutes in firms' location choices. Small differences in fundamentals then generate larger shifts in the equilibrium distribution of firms across countries, amplifying the response of trade flows, price indices, and welfare relative to the baseline model.

Economic Environment. Consider a world economy composed of a finite set of countries $k \in \mathcal{K}$. There is a continuum of tradable goods, indexed by $n \in [0, 1]$. Consumers in all countries share identical CES preferences over this continuum of goods. Labor is immobile across countries with exogenous supply $L_k > 0$. Production in each country requires a bundle of labor and intermediate inputs with an aggregate cost $c_k > 0$. The cost of producing good n in country k is thus $c_k/z_k(n)$, where $z_k(n)$ is drawn from a Fréchet distribution with common shape parameter $\theta > 0$ and country–specific scale parameter $T_k > 0$. Trade between countries is subject to bilateral iceberg trade costs $\tau_{kj} \geq 1$ for goods shipped from k to j. Due to perfect competition, consumers in country j purchase each good from the lowest–cost supplier worldwide. The global economic environment is thus fully described by (L, T, c, τ, θ), where $L : \mathcal{K} \to \mathbb{R}_{++}$, $T : \mathcal{K} \to \mathbb{R}_{++}$, $c : \mathcal{K} \to \mathbb{R}_{++}$, $\tau : \mathcal{K} \times \mathcal{K} \to [1, \infty)$, and $\theta > 0$ is a scalar.

Following Eaton and Kortum (2002), I assume that production combines labor and intermediate inputs according to:

$$c_i = w_i^{\gamma} P_i^{1-\gamma}, \tag{1}$$

where $\gamma \in (0, 1)$ and P_i is the CES price index. Then the share of country j's total expenditure on goods from country i is:

$$\pi_{ij} = \frac{T_i \left(c_i \tau_{ij}\right)^{-\theta}}{\sum_{k=1}^{K} T_k \left(c_k \tau_{kj}\right)^{-\theta}}, \tag{2}$$

and the aggregate CES price index is:

$$P_j = C \left(\sum_{k=1}^{K} T_k \left(c_k \tau_{kj}\right)^{-\theta} \right)^{-\frac{1}{\theta}}, \tag{3}$$

where C is a constant. Total expenditure in country i is:

$$Y_i = \frac{1}{\gamma} L_i w_i,$$

and trade is balanced:

$$L_i w_i = \sum_k \pi_{ik} (L_k w_k). \tag{4}$$

A *trade equilibrium* is a set of wages and prices such that (1)–(4) hold for all $i, j \in \mathcal{K}$.

Introducing Selection Margin. In the canonical Eaton-Kortum environment, the productivity draw z_k in country k is sampled from the country–specific Fréchet CDF

$$F_{z_k}(x) = \exp\left(-T_k x^{-\theta}\right),$$

where T_k is fixed exogenously. Under Proposition 1, this CDF can be reinterpreted as the *conditional* distribution of firms that have chosen to locate in k, with the scale parameter

$$T_k = H_k \, \mu_k^{-1/\rho},$$

where μ_k is the equilibrium share of global entrepreneurs locating in k, and H_k reflects country fundamentals such as institutions, geography, or technology.

Substituting this endogenous T_k into (2)–(3) and accounting for the fact that now the measure of firms in k is given by μ_k yields:

$$\pi_{ij} = \frac{H_i \, \mu_i^{1-1/\rho} \left(c_i \tau_{ij}\right)^{-\theta}}{\sum_{k=1}^{K} H_k \, \mu_k^{1-1/\rho} \left(c_k \tau_{kj}\right)^{-\theta}}, \tag{5}$$

$$P_j = C \left(\sum_{k=1}^{K} H_k \, \mu_k^{1-1/\rho} \left(c_k \tau_{kj}\right)^{-\theta} \right)^{-1/\theta}, \tag{6}$$

where the factor $\mu_k^{1-1/\rho}$ captures selection and endogenous heterogeneity margins.

Payoff Function. In the Eaton-Kortum environment, perfect competition drives expected profits to zero. Location choice is therefore can be governed by *relative cost competitiveness*, not profit maximization. A natural specification is:

$$V_k = W_k \varepsilon_k^{\omega} = \frac{\varepsilon_k}{c_k},$$

so that H_k in $F_{\boldsymbol{\varepsilon}}(\mathbf{e})$ captures the fundamental appeal of country k such institutions and geography. This specification is one example of the payoff function and may be altered depending on the underlying structure. By Proposition 1, the equilibrium location share is:

$$\mu_k = \frac{\left(H_k c_k^{-\theta}\right)^{\rho}}{\sum_{n=1}^{K} \left(H_n c_n^{-\theta}\right)^{\rho}}.$$

Counterfactuals. This framework is fully compatible with the "hat algebra" approach of Dekle et al. (2007). Let a denote any variable, a' its counterfactual level, and $\hat{a} = a'/a$ its relative change. Given baseline $\{\mu_k, \pi_{kj}, Y_k\}$, a trade-cost shock $\hat{\tau}_{ij}$ leads to:

$$\begin{aligned}
&\text{(i)} \quad \hat{c}_k = \hat{w}_k^{\gamma} \hat{P}_k^{1-\gamma}, \\
&\text{(ii)} \quad \hat{\pi}_{ij} = \hat{\mu}_i^{1-1/\rho} \left(\hat{c}_i \hat{\tau}_{ij}\right)^{-\theta} \hat{P}_j^{\theta}, \\
&\text{(iii)} \quad \hat{P}_j = \left(\sum_k \hat{\mu}_k^{1-1/\rho} \left(\hat{c}_k \hat{\tau}_{kj}\right)^{-\theta} \pi_{kj} \right)^{-1/\theta}, \\
&\text{(iv)} \quad \hat{w}_j Y_j = \sum_k \hat{\pi}_{kj} \pi_{kj} (\hat{w}_j Y_j), \\
&\text{(v)} \quad \hat{\mu}_k = \frac{\hat{c}_k^{-\theta\rho}}{\sum_j \mu_j \hat{c}_j^{-\theta\rho}}.
\end{aligned}$$

When $\rho = 1$, the selection margin is exactly offset and equation (v) becomes redundant, so the equilibrium collapses to the standard case with fixed within-location heterogeneity. When $\rho > 1$, stronger substitution between locations implies that small differences in fundamentals induce large reallocations of μ_k, which in turn amplify equilibrium responses through endogenous changes in the distribution of productivity within each location.

One advantage of the framework developed here is that it introduces endogenous firm location choice without requiring an overhaul of the existing Eaton–Kortum structure. Apart from the new channel operating through changes in the equilibrium

allocation of firms $\hat{\mu}_i$, all other components of the model remain identical to those in the canonical environment. As a result, many familiar results continue to hold. For example, the structural relationship between changes in domestic trade shares $\hat{\pi}_{ii}$ and welfare gains from trade derived by Arkolakis et al. (2012) remains intact.

What changes is the mechanism through which trade shares respond to trade-cost shocks. By allowing firm location to adjust endogenously, trade-cost changes now affect not only relative competitiveness conditional on location, but also the equilibrium distribution of firms across countries. This additional selection channel amplifies the response of trade flows, price indices, and welfare to trade-cost shocks, relative to the standard Eaton–Kortum model in which the mass of firms in each country is fixed. Appendix C illustrates the quantitative importance of this channel through a simple simulation that traces how the magnitude of the response varies with the dependence parameter ρ and the initial allocation of μ_k.

Discussion of Melitz-type Trade Models. The framework developed in this paper can also be applied to introduce endogenous location choice in other classes of models based on the heterogeneous-firm framework of Melitz (2003). In these models, it is typically assumed that the mass of firms operating in each country is fixed. This restriction can be imposed explicitly, as in Chaney (2008), or emerge endogenously under specific assumptions on firm productivity, such as Pareto-distributed draws, as shown in Arkolakis et al. (2008).

The most common parameterization of productivity in this class of models is the unbounded Pareto distribution, due to its tractability properties. Proposition 2 shows that the Dagum distribution provides a natural alternative: it preserves the same power-law tail behavior as Pareto while allowing for additional flexibility in modeling dependence and selection. In particular, the conditional productivity distribution among firms selecting country i takes the form

$$F(\phi) = \left[1 + \left(\frac{\phi}{D_i}\right)^{-\xi}\right]^{-1/\delta},$$

where the scale parameter satisfies $D_i = H_i \, \mu_i^{-1/\xi}$ and μ_i denotes the equilibrium share of firms locating in country i.

In the Melitz (2003) environment, firms incur a fixed cost of exporting and export

only if their productivity draw ϕ exceeds a cutoff ϕ_i^* determined by the comparison of fixed costs and potential profits. As in the Pareto case, the share of exporting firms remains available in closed form under the Dagum distribution:

$$\Pr(\phi > \phi_i^*) = 1 - \left[1 + \left(\frac{\phi_i^*}{D_i}\right)^{-\xi}\right]^{-1/\delta}.$$

A second object of interest in Melitz-type models is the expected value of truncated productivity moments, which enter aggregate exports and revenues. Under the Dagum distribution, these moments also remain tractable and satisfy

$$\int_{\phi_i^*}^{\infty} \phi^r \, dF(\phi) = \frac{D_i^r}{\delta} \, \mathbb{B}\left(\frac{(\phi_i^*/D_i)^{\xi}}{1 + (\phi_i^*/D_i)^{\xi}}; \, \frac{1}{\delta} + \frac{r}{\xi}, \, 1 - \frac{r}{\xi}\right),$$

where $\mathbb{B}(\cdot)$ denotes the upper incomplete Beta function and the expression is defined for $\xi > r$.

Because both export probabilities and conditional moments remain analytic, the Dagum distribution furnishes a tractable alternative to the unbounded Pareto while preserving power-law tails. Combined with the selection mechanism developed in this paper, it therefore provides a convenient route to embedding *endogenous location choice* into Melitz-type heterogeneous-firm models without sacrificing closed-form solutions for entry cutoffs, export shares, or aggregate productivity.

4.2 Application B: Economic Geography Model with Heterogeneity

Many influential models in economic geography, such as Allen and Arkolakis (2014), embed discrete location choice within a general-equilibrium framework and determine the spatial distribution of agents as the outcome of individual optimization. These frameworks provide a rigorous analysis of spatial sorting, congestion, agglomeration, and aggregate welfare. However, heterogeneity conditional on location is typically abstracted from, with agents within a given location treated as homogeneous or summarized by representative types. As a result, these models offer limited guidance on within-location dispersion in productivity, income, or welfare, and are often uninformative about distributional outcomes such as local inequality or concentration.

The framework developed here extends these models by endogenizing heterogeneity in outcomes within each location. It delivers closed-form expressions for both equi-

librium location shares and the conditional distribution of outcomes among agents selecting each location, while preserving the tractability of canonical economic geography frameworks. As a result, location choice and within-location heterogeneity are jointly determined by the same underlying stochastic structure.

Structurally, the approach developed here extends the framework of Allen and Arkolakis (2014) by introducing power-law heterogeneity in worker skill, in the spirit of Gabaix (2009). While the model is formulated as a discrete-choice location problem in which idiosyncratic shocks interact with location-specific payoffs, Propositions 1 and 2 ensure that equilibrium spatial allocations remain consistent with utility equalization across locations. Under additional parameter restrictions, the two perspectives coincide exactly: the equilibrium location shares implied by the probabilistic discrete-choice formulation replicate those generated by utility-equalizing models. This correspondence provides a transparent bridge between discrete-choice and utility-equalization approaches to spatial sorting and extends interpretations consistent with frameworks such as Redding (2016).

Economic Environment. Consider an economy composed of a finite set of locations $k \in \mathcal{K}$, where each location produces a distinct variety. Consumers share identical CES preferences over these varieties, and trade between locations incurs iceberg trade costs. Each location is characterized by three exogenous primitives: fundamental productivity $\bar{A}_k > 0$, fundamental amenity level $\bar{u}_k > 0$, and bilateral iceberg trade costs $\tau_{kj} \geq 1$ for all $j \in \mathcal{K}$. The geography of the economy is thus fully described by the tuple $(\bar{A}, \bar{u}, \tau)$, where $\bar{A} : \mathcal{K} \to \mathbb{R}_{++}$, $\bar{u} : \mathcal{K} \to \mathbb{R}_{++}$, and $\tau : \mathcal{K} \times \mathcal{K} \to [1, \infty)$.

There is a measure M of individuals. Each individual is endowed with one unit of labor and total labor force is $L = M$. While the distinction is often trivial in economic geography models, as I show below it is necessary in models with heterogeneity. Following Allen and Arkolakis (2014), the model incorporates both congestion and agglomeration effects. Utility from local amenities decreases in local population M_k:

$$u_k = \bar{u}_k M_k^{\beta},$$

where $\beta < 0$. Individuals who choose to reside in k each supply one unit of labor. Total labor in k increases overall productivity in the location as:

$$A_k = \bar{A}_k L_k^{\alpha},$$

where $\alpha > 0$ and $L_k = M_k$ is the total effective labor supply in location k.

The CES function that defines utility from goods consumption takes the familiar form:

$$U_k = \left(\sum_i q_{ik}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where q_{ik} denotes the quantity from location i consumed in k and $\sigma > 1$ denotes the elasticity of substitution across varieties. I assume that varieties are produced in perfectly competitive setting and that labor is the only factor of production. Then, the price of the variety produced in location k and consumed in j is:

$$p_{kj} = \frac{w_k \tau_{kj}}{A_k},$$

where w_k is the wage per unit of labor in location k. Given total demand for the variety produced in location k, the value of bilateral trade shares from k to j is:

$$\lambda_{kj} = A_k^{\sigma-1} w_k^{1-\sigma} \tau_{kj}^{1-\sigma} P_j^{\sigma-1},$$

and the CES price index is:

$$P_j = \left(\sum_{k \in \mathcal{K}} A_k^{\sigma-1} w_k^{1-\sigma} \tau_{kj}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

Assuming log-additivity of amenities in total utility, individual deterministic welfare in location k is given by:

$$W_k = u_k \frac{w_k}{P_k}, \tag{7}$$

subject to the idiosyncratic preference shock ε_k. In standard models, $\log \varepsilon_k$ is assumed to be i.i.d. across k and distributed Gumbel with location zero and scale $1/\varsigma > 0$, leading to the canonical logit form of the share of agents that select location k:

$$\mu_k = \frac{W_k^{\varsigma}}{\sum_{j \in \mathcal{K}} W_j^{\varsigma}}. \tag{8}$$

The *spatial equilibrium* consists of the set of wages and populations such that the goods market clears, trade is balanced, labor market clears, and labor resource con-

straints hold.

Introducing Heterogeneity. I use the result in Proposition 2 to extend the benchmark model and introduce heterogeneity across individuals. For that, suppose a unit mass M of individuals, each drawing a vector of location-specific labor productivities in the form of labor endowments $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_K)$ from the following distribution:

$$F(\boldsymbol{\varepsilon}) = \left(1 + \sum_{k=1}^{K} \left[\frac{\ell_k}{H_k}\right]^{-\xi}\right)^{-1/\delta}.$$

While I am using the CDF from Proposition 2, it is worth noting that one could equivalently use the Fréchet distribution in Proposition 1.

Each individual is paid wage w_k per unit of effective labor ℓ_k, faces local price index P_k, and receives amenity flow u_k. The resulting payoff in location k is:

$$V_k = W_k \ell_k,$$

with W_k as defined in equation (7). Applying the equivalence result, the share of individuals selecting location k is given by:

$$\mu_k = \frac{(H_k W_k)^{\xi}}{\sum_{n=1}^{K} (H_n W_n)^{\xi}}.$$

Under the restrictions $\varsigma = \xi$ and $H_k = H$ for all k, this discrete choice rule matches the canonical form in equation (8). The key innovation of the present formulation is to separate the allocation of individuals across locations from the distribution of effective labor supplied by those who select each location. This separation introduces endogenous heterogeneity within locations as an equilibrium outcome of selection, while preserving the analytical tractability of the baseline framework.

Next, I examine the additional margins of adjustment introduced by worker heterogeneity. Because each worker now possesses a distinct effective labor endowment, the symmetry between population and labor supply no longer holds. In particular, the total population and total effective labor in location k are given by:

$$M_k = \mu_k M \;\text{ and }\; L_k = \mathbb{E}[\ell_k \mid k^* = k] M_k = \tilde{\Gamma} M \mu_k^{1-1/\xi},$$

where $\tilde{\Gamma} = \Gamma(1 - 1/\xi)\,\Gamma(1/\delta + 1/\xi)/\Gamma(1/\delta)$. Proposition 2 also immediately pins

down the welfare level in all locations as the expected value of the payoff. While agglomeration forces depend on total labor supply L_k, congestion effects depend on the number of individuals $M_k = \mu_k M$, regardless of their effective productivity:

$$u_k = \bar{u}_k (\mu_k M)^{\beta}.$$

To formally characterize additional margins of adjustment, let me again employ the hat algebra approach in Dekle et al. (2007). Conditional on benchmark values $\{\mu_k, \lambda_{kj}, Y_k\}$, and assuming $\varsigma = \xi$, the system of equations is:

$$\begin{aligned}
&\text{(i)} \quad \hat{\lambda}_{kj} = \hat{A}_k^{\sigma-1} \hat{w}_k^{1-\sigma} \hat{P}_j^{\sigma-1} \\
&\text{(ii)} \quad \hat{P}_j = \left(\textstyle\sum_k \hat{A}_k^{\sigma-1} \hat{w}_k^{1-\sigma} \lambda_{kj} \right)^{\frac{1}{1-\sigma}} \\
&\text{(iii)} \quad \hat{\mu}_j \hat{w}_j Y_j = \textstyle\sum_k \hat{\lambda}_{kj} \lambda_{kj} \left(\hat{\mu}_j \hat{w}_j Y_j \right) \\
&\text{(iv)} \quad \hat{\mu}_k = \frac{\left(\hat{\mu}_k^{\beta} \hat{w}_k / \hat{P}_k \right)^{\varsigma}}{\sum_j \mu_j \left(\hat{\mu}_j^{\beta} \hat{w}_j / \hat{P}_j \right)^{\varsigma}}
\end{aligned}$$

Equations (i)–(iv) are shared by both representative-agent and heterogeneous-agent models. The key difference lies in how agglomeration effects respond to changes in population. Specifically, the effective productivity term $\hat{A}_k$ evolves as:

$$\begin{aligned}
&\text{(v.a)} \quad \text{Representative-Agent:} && \hat{A}_k = \hat{\mu}_k^{\alpha}, \\
&\text{(v.b)} \quad \text{Heterogeneous-Agent:} && \hat{A}_k = \hat{\mu}_k^{\alpha\left(1-\frac{1}{\xi}\right)}.
\end{aligned}$$

The heterogeneous-agent model introduces an additional distributional margin: as more individuals select into location k, the lower tail of the labor endowment distribution is drawn in, reducing *average labor supply per capita*. This composition effect dampens the effective scale externality from agglomeration.

Note that in the limit as $\xi \to \infty$ and the distribution of ℓ's becomes degenerate, heterogeneity disappears and equations $(v.a)$ and $(v.b)$ coincide, recovering the representative-agent benchmark. As ξ decreases, the distribution effect becomes quantitatively more important: additional entrants into a location lower the average labor endowment per capita of the local labor force, dampening effective productivity gains. However, this adjustment is bounded from below by a structural requirement for equilibrium existence and finite aggregates:

$$\xi > \sigma - 1.$$

For instance, if $\sigma = 5$, the admissible parameter range becomes $\xi \in (4, \infty)$. Within this range, the elasticity of effective productivity with respect to location population can be reduced by up to 25% relative to the representative-agent case. This implies that omitting the post-selection margin may lead standard models to systematically overstate the strength of agglomeration externalities.

An additional advantage of incorporating within-location heterogeneity through the framework developed here is that it delivers tractable, closed-form characterizations of the distribution of outcomes among agents selecting each location. Unlike canonical economic geography models, which typically remain silent about outcome dispersion conditional on location choice, the present approach allows dispersion measures within each location to be computed in closed form and their comparative statics with respect to amenities, productivity, and trade costs to be analyzed transparently. In this way, spatial general-equilibrium forces are directly linked to distributional outcomes within locations, while preserving analytical tractability.

4.3 Application C: Roy Model with Correlated Abilities

Canonical occupational choice models in the tradition of Roy (1951) provide a powerful framework for analyzing how individuals sort across occupations or sectors and for quantifying the effects of shocks, policies, or entry barriers on economic outcomes. In most quantitative implementations, occupational abilities are drawn from Fréchet distributions and treated as exogenous primitives, often under the assumption of independence across occupations, as in Hsieh et al. (2019) and Galle et al. (2023).[3]

More recently, Fan (2025) employs multivariate Fréchet distributions to study labor market shocks and worker mobility across occupations, highlighting the empirical relevance of correlated ability draws in multi-occupation settings. The framework developed here allows the distribution of abilities within each occupation to emerge endogenously as an equilibrium outcome of selection, even when ability shocks are correlated across occupations. This structure preserves tractability while providing a systematic link between occupational sorting and the distribution of outcomes within occupations, yielding closed-form expressions for how shocks or policies reshape both

[3] Lagakos and Waugh (2013) allow for correlated abilities but focus on binary choice. The framework developed here remains tractable with many occupations.

allocations across occupations and distributional outcomes among workers within them.

I illustrate the value of the results using a standard Roy model, showing how occupational sorting and within-occupation heterogeneity can be analyzed jointly. Consider groups $g \in \mathcal{G}$ and occupations $k \in \mathcal{K}$. Individuals from group g draw occupation-specific abilities $\boldsymbol{\varepsilon}_g = (\varepsilon_g(1), \ldots, \varepsilon_g(K))$ from a multivariate Fréchet distribution with logistic dependence:

$$F_{\boldsymbol{\varepsilon}_g}(\mathbf{e}) = \exp\left(- \left(\sum_{k=1}^{K} \left[H_{gk}\, e_k^{-\theta} \right]^{\rho} \right)^{1/\rho} \right),$$

where $H_{gk} > 0$ scales average ability in occupation k for group g, $\theta > 0$ governs the tail thickness of abilities, and $\rho \geq 1$ controls the strength of positive dependence across occupations.

Let the payoff of an individual from group g in occupation k be:

$$V_{gk} = w_k(1 - \tau_{gk})\, \varepsilon_g(k),$$

where w_k is the occupation-specific wage and $\tau_{gk} \in [0, 1)$ captures group–occupation barriers.

Applying Proposition 1, the equilibrium share of group-g individuals selecting occupation k is:

$$\mu_{gk} = \frac{\left[H_{gk}\, w_k^{\theta} (1 - \tau_{gk})^{\theta} \right]^{\rho}}{\sum_{\ell=1}^{K} \left[H_{g\ell}\, w_\ell^{\theta} (1 - \tau_{g\ell})^{\theta} \right]^{\rho}}.$$

Conditional on selecting occupation k, ability remains Fréchet with shape parameter θ and endogenous scale:

$$T_{gk} = H_{gk}\, \mu_{gk}^{-1/\rho}.$$

Accordingly, the average ability of group-g workers in occupation k is:

$$\mathbb{E}[\varepsilon_g(k) \mid k^* = k,\, g] = H_{gk}^{1/\theta}\, \mu_{gk}^{-1/(\theta\rho)} \Gamma\left(1 - \frac{1}{\theta}\right), \qquad \theta > 1.$$

As in Hsieh et al. (2019), consider the effect of reducing τ_{gk}. When $\rho > 1$, barrier reductions induce larger reallocations in μ_{gk} than under independent draws. This expands group-g participation in occupation k but brings in marginal entrants with

lower ability, thereby reshaping within-occupation heterogeneity.

The effective labor supplied by group g to occupation k is:

$$L_{gk} = N_g \, \mu_{gk} \, \mathbb{E}[\varepsilon_g(k) \mid k^* = k, g] = N_g \, H_{gk}^{1/\theta} \, \mu_{gk}^{1-1/(\theta\rho)} \, \Gamma\left(1 - \frac{1}{\theta}\right),$$

where N_g is the size of group g.

Because $(\varepsilon_g(k) \mid k^* = k, g)$ is Fréchet with shape θ and scale parameter $T_{gk} = H_{gk}\mu_{gk}^{-1/\rho}$, the conditional payoff $(V_{gk} \mid k^* = k, g)$ is Fréchet with the same shape θ and scale:

$$w_k(1 - \tau_{gk}) \, T_{gk}^{1/\theta} = w_k(1 - \tau_{gk}) \, H_{gk}^{1/\theta} \, \mu_{gk}^{-1/(\theta\rho)}.$$

Its variance (for $\theta > 2$) is therefore:

$$\mathrm{Var}(V_{gk} \mid k^* = k, g) = \left[w_k(1 - \tau_{gk})\right]^2 H_{gk}^{2/\theta} \mu_{gk}^{-2/(\theta\rho)} \left[\Gamma\left(1 - \frac{2}{\theta}\right) - \Gamma\left(1 - \frac{1}{\theta}\right)^2\right].$$

Hence, within-group income variance admits a law-of-total-variance decomposition into within-occupation and between-occupation components, both available in closed form. This makes it straightforward to quantify how policy-induced changes in τ_{gk} reshape occupational allocations, conditional ability distributions, and both within- and between-occupation inequality.

5 Conclusion

This paper develops a tractable framework in which heterogeneity within economic environments is not imposed as a primitive, but arises endogenously as an equilibrium outcome of payoff-based selection. The central insight is that the same optimization and stochastic structures that generate closed-form allocation shares across discrete alternatives also determine the distribution of outcomes among agents selecting each alternative. For two widely used heavy-tailed families – Fréchet and Dagum – the framework delivers explicit mappings between equilibrium choice shares and the resulting distributions of productivity, skill, or earnings within alternatives, while allowing for correlated shocks.

Applications to international trade, economic geography, and occupational choice illustrate how endogenizing heterogeneity resolves longstanding modeling challenges. In each setting, standard general-equilibrium structures are preserved, but the distri-

bution of active agents within countries, locations, or occupations becomes responsive to economic fundamentals. As a result, counterfactual changes in policy, technology, or trade costs reshape not only allocations across alternatives but also the distribution of outcomes within them, yielding richer comparative statics.

More broadly, the framework provides a disciplined way to connect allocation decisions to distributional outcomes in heterogeneous-agent models. By maintaining analytical tractability even in the presence of correlated shocks, it offers a flexible foundation for quantitative work in environments where selection and heterogeneity jointly shape aggregate and distributional outcomes. These features make the approach a useful building block for future research in trade, spatial economics, labor markets, and related fields where understanding the endogenous formation of heterogeneity is central.

References

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf (2015): "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 83 (6), 2127–2189.

Allen, Treb and Costas Arkolakis (2014): "Trade and the Topography of the Spatial Economy *," *The Quarterly Journal of Economics*, 129 (3), 1085–1140.

Anderson, Simon P., André de Palma, and Jacques-François Thisse (1992): *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press, 1st ed.

Arkolakis, Costas (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118 (6), 1151–1199.

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

Arkolakis, Costas, Svetlana Demidova, Peter J. Klenow, and Andres Rodriguez-Clare (2008): "Endogenous Variety and the Gains from Trade," *American Economic Review*, 98 (2), 444–50.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple (2018): "Innovation and Production in the Global Economy," *American Economic Review*, 108 (8), 2128–73.

Arkolakis, Costas, Andres Rodríguez-Clare, and Zixuan Su (2025): "A Multivariate Distribution with Pareto Tails and Pareto Maxima," Working paper.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100 (3), 1008–45.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93 (4), 1268–1290.

Bryan, Gharad and Melanie Morten (2019): "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 127 (5), 2229–2268.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel (2019): "Changes in Between-Group Inequality: Computers, Occupations, and International Trade," *American Economic Journal: Macroeconomics*, 11 (2), 348–400.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87 (3), 741–835.

Caliendo, Lorenzo and Fernando Parro (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82 (1 (290)), 1–44.

Chaney, Thomas (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *The American Economic Review*, 98 (4), 1707–1721.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer (2011): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *The Review of*

Economic Studies, 79 (2), 581–608.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum (2007): "Unbalanced Trade," *American Economic Review*, 97 (2), 351–355.

di Giovanni, Julian and Andrei A. Levchenko (2012): "Country Size, International Trade, and Aggregate Fluctuations in Granular Economies," *Journal of Political Economy*, 120 (6), 1083–1132.

Dix-Carneiro, Rafael (2014): "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 82 (3), 825–885.

Eaton, Jonathan and Samuel Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70 (5), 1741–1779.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz (2011): "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 79 (5), 1453–1498.

Fajgelbaum, Pablo D and Cecile Gaubert (2020): "Optimal Spatial Policies, Geography, and Sorting*," *The Quarterly Journal of Economics*, 135 (2), 959–1036.

Fajgelbaum, Pablo D, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar (2018): "State Taxes and Spatial Misallocation," *The Review of Economic Studies*, 86 (1), 333–376.

Fan, Tianyu (2025): "The Labor Market Incidence of New Technologies," arXiv:2504.04047.

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

——— (2009): "Power Laws in Economics and Finance," *Annual Review of Economics*, 1 (Volume 1, 2009), 255–294.

——— (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79 (3), 733–772.

Galle, Simon, Andres Rodríguez-Clare, and Michael Yi (2023): "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade," *Econometrica*, 91 (1), 1–47.

Grogger, Jeffrey and Gordon H. Hanson (2011): "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, 95 (1), 42–57, symposium on Globalization and Brain Drain.

Helpman, Elhanan, Oleg Itskhoki, and Stephen J Redding (2010): "Inequality and unemployment in a global economy," *Econometrica*, 78 (4), 1239–1283.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple (2004): "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 94 (1), 300–316.

Holmes, Thomas J. and Holger Sieg (2015): "Chapter 2 - Structural Estimation in Urban Economics," in *Handbook of Regional and Urban Economics*, ed. by Gilles Duranton, J. Vernon Henderson, and William C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 69–114.

Hougaard, Philip (1986): "A class of multivariate failure time distributions," *Biometrika*, 73 (3), 671–678.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow (2019):

"The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 87 (5), 1439–1474.

Lagakos, David and Michael E. Waugh (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103 (2), 948–80.

Lind, Nelson and Natalia Ramondo (2023a): "Global Innovation and Knowledge Diffusion," *American Economic Review: Insights*, 5 (4), 494–510.

——— (2023b): "Trade with Correlation," *American Economic Review*, 113 (2), 317–53.

Luttmer, Erzo G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122 (3), 1103–1144.

McFadden, Daniel (1974): "Conditional logit analysis of qualitative choice behavior," in *Fontiers in Econometrics*, ed. by Paul Zarembka, New York: Academic press, 105–142.

——— (1978): "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, Amsterdam: North-Holland, 75–96.

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

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg (2018): "Commuting, migration, and local employment elasticities," *American Economic Review*, 108 (12), 3855–3890.

Nelsen, Roger B. (2006): *An Introduction to Copulas*, New York: Springer, second ed.

Nigai, Sergey (2017): "A tale of two tails: Productivity distribution and the gains from trade," *Journal of International Economics*, 104, 44–62.

Nigai, Sergey and Dongkyu Yang (2024): "Trade and Inequality: A Sufficient-Statistics Approach," *Journal of Political Economy Macroeconomics*, 2 (3), 508–552.

Ramondo, Natalia and Andrés Rodríguez-Clare (2013): "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 121 (2), 273–322.

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

Redding, Stephen J. and Esteban Rossi-Hansberg (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

Rossi-Hansberg, Esteban and Mark L. J. Wright (2007): "Urban Structure and Growth," *Review of Economic Studies*, 74 (2), 597–624.

Roy, Andrew D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (2), 135–146.

Simon, Herbert A. and Charles P. Bonini (1958): "The Size Distribution of Business Firms," *American Economic Review*, 48 (4), 607–617.

Tintelnot, Felix (2017): "Global Production with Export Platforms," *The Quarterly Journal of Economics*, 132 (1), 157–209.

Tombe, Trevor and Xiaodong Zhu (2019): "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 109 (5), 1843–72.

Train, Kenneth E. (2009): *Discrete Choice Methods with Simulation*, Cambridge University Press.

Appendix A: Multivariate Fréchet with Logistic Dependence

In this appendix, I illustrate the construction of the multivariate Fréchet distribution with logistic (Gumbel–Hougaard) dependence and provide proofs of the results in Proposition 1.

A1: Multivariate Fréchet with Logistic Dependence

To illustrate the construction of a multivariate Fréchet distribution with logistic dependence, start with the marginal distribution of each shock ε_k, which is Fréchet with shape $\theta > 0$ and scale H_k, with cumulative distribution function:

$$F_k(e_k) = \exp\left(-H_k e_k^{-\theta}\right). \tag{A1}$$

Define $u_k \equiv F_k(e_k)$, which are uniformly distributed on $[0, 1]$ by the probability integral transform.

The dependence structure is given by the Gumbel–Hougaard copula (also known as the logistic copula in extreme value theory; see Nelsen (2006)):

$$C(u_1, \ldots, u_K) = \exp\left(-\left[\sum_{j=1}^{K}(-\log u_j)^\rho\right]^{1/\rho}\right), \tag{A2}$$

where $\rho \geq 1$ is the dependence parameter. When $\rho = 1$, this simplifies to the product copula $\prod_j u_j$, corresponding to independence; as $\rho \to \infty$, it approaches perfect positive dependence.

By Sklar's Theorem, the joint CDF is:

$$F(e_1, \ldots, e_K) = C(F_1(e_1), \ldots, F_K(e_K)).$$

It is useful to derive the expression for the inner sum $\sum_{j=1}^{K}(-\log u_j)^\rho$. First, note that:

$$-\log u_k = -\log F_k(e_k) = -\log\left(\exp\left(-H_k e_k^{-\theta}\right)\right) = H_k e_k^{-\theta}.$$

Thus:

$$\sum_{j=1}^{K}(-\log u_j)^\rho = \sum_{j=1}^{K}\left(H_j e_j^{-\theta}\right)^\rho.$$

Substituting into (A2) yields:

$$F(e_1, \ldots, e_K) = \exp\left(- \left[\sum_{j=1}^{K} \left(H_j e_j^{-\theta} \right)^{\rho} \right]^{1/\rho} \right),$$

which matches the multivariate Fréchet distribution used in Proposition 1.

A2: Proof of Proposition 1

The shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ follow the joint cumulative distribution function

$$F_{\boldsymbol{\varepsilon}}(\mathbf{e}) = \exp\left(- \left[\sum_{j=1}^{K} \left(H_j e_j^{-\theta} \right)^{\rho} \right]^{1/\rho} \right),$$

where $H_j > 0$, $\theta > 1$, and $\rho \geq 1$. Agents select k^* to maximize

$$V_k = W_k \varepsilon_k^{\omega}, \qquad W_k > 0, \quad \omega > 0.$$

I derive the choice probabilities $\mu_k = \Pr(k^* = k)$, the conditional CDFs $F_{\varepsilon_k | k^*=k}(x)$ and $F_{V_k | k^*=k}(v)$, and the conditional expectations $\mathbb{E}[\varepsilon_k \mid k^* = k]$ $(\theta > 1)$ and $\mathbb{E}[V_k \mid k^* = k]$ $(\theta > \omega)$.

A2.1: Transformation and Shocks

The agent chooses k^* to maximize V_k, which is equivalent to maximizing the log-payoff $U_k = \log(V_k)$:

$$U_k = \log(W_k) + \omega \log(\varepsilon_k).$$

This defines a random utility model (RUM) with deterministic component $v_k = \log(W_k)$ and stochastic component $\tilde{\epsilon}_k = \omega \log(\varepsilon_k)$.

The marginal CDF of ε_k is obtained by letting $e_j \to \infty$ for all $j \neq k$:

$$F_{\varepsilon_k}(x) = \lim_{e_j \to \infty, \forall j \neq k} \exp\left(- \left[\left(H_k x^{-\theta} \right)^{\rho} + \sum_{j \neq k} \left(H_j e_j^{-\theta} \right)^{\rho} \right]^{1/\rho} \right) = \exp\left(-H_k x^{-\theta} \right),$$

which is a Fréchet distribution with shape parameter θ and scale parameter H_k.

For the transformed shock $\tilde{\epsilon}_k = \omega \log(\varepsilon_k)$, its distribution satisfies:

$$F_{\tilde{\epsilon}_k}(y) = \Pr\left(\varepsilon_k \leq \exp\left(\frac{y}{\omega} \right) \right) = \exp\left(-H_k \exp\left(-\frac{\theta y}{\omega} \right) \right) = \exp\left(- \exp\left(\log H_k - \frac{\theta y}{\omega} \right) \right),$$

which is a Gumbel distribution with location parameter $\psi_k = \frac{\omega}{\theta} \log H_k$ and scale

parameter $\beta = \frac{\omega}{\theta}$.

A2.2: Derivation of Choice Probabilities

Once the distribution of the transformed shock is known, it is useful to rewrite U_k as:

$$U_k = v_k + \tilde{\epsilon}_k = \beta \left(\frac{v_k + \psi_k}{\beta} + \frac{\tilde{\epsilon}_k - \psi_k}{\beta} \right) = \beta(u_k + \eta_k).$$

Since β is a positive constant, maximizing U_k is equivalent to:

$$k^* = \arg\max_k \left(u_k + \eta_k \right).$$

This representation is useful because η_k follows standard type-I extreme value marginal distribution with CDF $F_{\eta_k}(z) = \exp\left(\exp(-z)\right)$. To derive the joint distribution of $\boldsymbol{\eta}$, note that for any $\mathbf{z} \in \mathbb{R}^K$:

$$F_{\boldsymbol{\eta}}(\mathbf{z}) = \Pr(\eta_1 \leq z_1, \ldots, \eta_K \leq z_K).$$

By the definition $\eta_j = (\tilde{\epsilon}_j - \psi_j)/\beta$ and the monotonicity of the exponential transformation, the event $\{\eta_j \leq z_j\}$ is equivalent to:

$$\varepsilon_j \leq \exp\left(\frac{\psi_j + \beta z_j}{\omega} \right).$$

Hence:

$$F_{\boldsymbol{\eta}}(\mathbf{z}) = \exp\left(- \left[\sum_{j=1}^{K} \left(H_j \exp\left(-\theta \frac{\psi_j + \beta z_j}{\omega} \right) \right)^{\rho} \right]^{1/\rho} \right).$$

Using identities for ψ_j and β yields:

$$\left(H_j \exp\left(-\theta \frac{\psi_j + \beta z_j}{\omega} \right) \right)^{\rho} = H_j^{\rho} \exp(-\rho \log H_j - \rho z_j) = \exp(-\rho z_j).$$

Therefore,

$$F_{\boldsymbol{\eta}}(\mathbf{z}) = \exp\left(- \left[\sum_{j=1}^{K} \exp(-\rho z_j) \right]^{1/\rho} \right).$$

For $\rho \geq 1$, this is a multivariate type-I extreme value (GEV) distribution with generator:

$$G(\mathbf{y}) = \left(\sum_{j=1}^{K} y_j^{\rho} \right)^{1/\rho}, \qquad y_j = \exp(-z_j),$$

corresponding to the logistic GEV family with dependence parameter $1/\rho$ (McFadden, 1978; Train, 2009).

The implied choice probabilities are therefore:

$$\mu_k = \frac{\exp(u_k)^\rho}{\sum_{j=1}^K \exp(u_j)^\rho}, \qquad u_k = \frac{\log W_k + \frac{\omega}{\theta}\log H_k}{\omega/\theta} = \log\left(W_k^{\theta/\omega} H_k\right).$$

Thus,

$$\mu_k = \frac{\left(W_k^{\theta/\omega} H_k\right)^\rho}{\sum_{j=1}^K \left(W_j^{\theta/\omega} H_j\right)^\rho}.$$

This completes the proof of Proposition 1(i).

A2.3: Derivation of Conditional CDFs

For $k^* = k$, $V_k = \max_j\{V_j\}$ where $V_j = W_j \varepsilon_j^\omega$, the conditional CDF is defined as

$$F_{V_k|k^*=k}(v) = \Pr\left(\varepsilon_j \leq \left(\frac{v}{W_j}\right)^{1/\omega} \ \forall j\right).$$

Substitute $e_j = (v/W_j)^{1/\omega}$ into the joint CDF:

$$F_{V_k|k^*=k}(v) = \exp\left(-\left[\sum_{j=1}^K \left(H_j \left(\frac{v}{W_j}\right)^{-\theta/\omega}\right)^\rho\right]^{1/\rho}\right).$$

Simplify the inner summation term:

$$\left(H_j \left(\frac{v}{W_j}\right)^{-\theta/\omega}\right)^\rho = H_j^\rho W_j^{\theta\rho/\omega} v^{-\theta\rho/\omega},$$

so that

$$\sum_{j=1}^K H_j^\rho W_j^{\theta\rho/\omega} v^{-\theta\rho/\omega} = v^{-\theta\rho/\omega} \sum_{j=1}^K \left(H_j W_j^{\theta/\omega}\right)^\rho.$$

For convenience, define the aggregate term:

$$W \equiv \left[\sum_{j=1}^K \left(H_j W_j^{\theta/\omega}\right)^\rho\right]^{1/\rho},$$

and substitute into the CDF to obtain:

$$F_{V_k|k^*=k}(v) = \exp\left(-W\, v^{-\theta/\omega}\right).$$

To derive the conditional CDF of ε_k, use the relation $V_k \leq W_k \varepsilon_k^{\omega}$:

$$F_{\varepsilon_k|k^*=k}(x) = \Pr(V_k \leq W_k x^{\omega} \mid k^* = k) = \exp\left(-W(W_k x^{\omega})^{-\theta/\omega}\right).$$

Using the definition of choice shares from Proposition 1(i):

$$\mu_k^{-1/\rho} = \frac{W}{H_k W_k^{\theta/\omega}}, \qquad \text{or equivalently} \qquad W W_k^{-\theta/\omega} = H_k \mu_k^{-1/\rho},$$

yields:

$$F_{\varepsilon_k|k^*=k}(x) = \exp\left(-H_k \mu_k^{-1/\rho} x^{-\theta}\right).$$

This concludes the proof of Proposition 1(ii).

A2.4: Derivation of Conditional Expectations

The conditional distribution $\varepsilon_k \mid k^* = k$ is Fréchet with shape parameter θ and scale:

$$s = \left(H_k \mu_k^{-1/\rho}\right)^{1/\theta} = H_k^{1/\theta}\, \mu_k^{-1/(\theta\rho)}.$$

For $\theta > 1$, the expectation is:

$$\mathbb{E}[\varepsilon_k \mid k^* = k] = H_k^{1/\theta}\, \mu_k^{-1/(\theta\rho)}\, \Gamma\left(1 - \frac{1}{\theta}\right).$$

For expected payoffs:

$$\mathbb{E}[V_k \mid k^* = k] = W_k\, \mathbb{E}[\varepsilon_k^{\omega} \mid k^* = k].$$

When $\theta > \omega$:

$$\mathbb{E}[\varepsilon_k^{\omega} \mid k^* = k] = \left(H_k^{1/\theta}\, \mu_k^{-1/(\theta\rho)}\right)^{\omega} \Gamma\left(1 - \frac{\omega}{\theta}\right).$$

Substituting yields:

$$\mathbb{E}[V_k \mid k^* = k] = W_k\, H_k^{\omega/\theta}\, \mu_k^{-\omega/(\theta\rho)}\, \Gamma\left(1 - \frac{\omega}{\theta}\right).$$

Using the identity implied by Proposition 1(i):

$$\mu_k^{-\omega/(\theta\rho)} = \frac{W^{\omega/\theta}}{b_k^{\omega/\theta}}, \qquad b_k^{\omega/\theta} = W_k\, H_k^{\omega/\theta},$$

it follows that:

$$\mathbb{E}[V_k \mid k^* = k] = W^{\omega/\theta}\Gamma\Big(1 - \frac{\omega}{\theta}\Big),$$

which is identical across alternatives. This concludes the proof of Proposition 1(iii).

Appendix B: Multivariate Dagum with Clayton Dependence

In this appendix, I illustrate the construction of the multivariate Dagum distribution with Clayton dependence from its marginal distributions and dependence structure and provide proofs of the results in Proposition 2.

B1: Multivariate Dagum with Clayton Dependence

Assume the marginal distribution of each shock ε_k is a Dagum distribution with shape parameters $\xi > 0$, $1/\delta > 0$, and scale $H_k > 0$. Its cumulative distribution function (CDF), for $e_k > 0$, is:

$$F_k(e_k) = \left[1 + \left(\frac{e_k}{H_k}\right)^{-\xi}\right]^{-1/\delta}. \tag{B1}$$

Define $u_k \equiv F_k(e_k)$, which are uniformly distributed on $[0, 1]$ by the probability integral transform. The dependence structure is the Clayton copula, which is the multivariate Archimedean extension of the bivariate Clayton family (Nelsen, 2006), defined as:

$$C(u_1, \ldots, u_K) = \left(\sum_{j=1}^{K} u_j^{-\delta} - (K-1)\right)^{-1/\delta}, \tag{B2}$$

where $\delta > 0$ is the dependence parameter. As $\delta \to \infty$, this approaches perfect positive dependence (comonotonicity); as $\delta \to 0$, it approaches the product copula (independence).

By Sklar's Theorem, the joint CDF of the shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ is:

$$F(e_1, \ldots, e_K) = C(F_1(e_1), \ldots, F_K(e_K)),$$

where $F_k(e_k) = \left[1 + \left(\frac{e_k}{H_k}\right)^{-\xi}\right]^{-1/\delta}$ is the Dagum marginal CDF for $e_k > 0$, and $C(u_1, \ldots, u_K) = \left(\sum_{j=1}^{K} u_j^{-\delta} - (K-1)\right)^{-1/\delta}$ is the Clayton copula with $\delta > 0$.

Define $u_k = F_k(e_k)$. Then:

$$u_k^{-\delta} = \left(\left[1 + \left(\frac{e_k}{H_k} \right)^{-\xi} \right]^{-1/\delta} \right)^{-\delta} = 1 + \left(\frac{e_k}{H_k} \right)^{-\xi}.$$

Thus:
$$\sum_{j=1}^{K} u_j^{-\delta} = \sum_{j=1}^{K} \left[1 + \left(\frac{e_j}{H_j} \right)^{-\xi} \right] = K + \sum_{j=1}^{K} \left(\frac{e_j}{H_j} \right)^{-\xi}.$$

Substituting into the copula yields:

$$F(e_1, \ldots, e_K) = \left(\left[K + \sum_{j=1}^{K} \left(\frac{e_j}{H_j} \right)^{-\xi} \right] - (K-1) \right)^{-1/\delta} = \left(1 + \sum_{j=1}^{K} \left(\frac{e_j}{H_j} \right)^{-\xi} \right)^{-1/\delta}.$$

This confirms that the joint CDF matches the multivariate Dagum specification used in Proposition 2 for $e_j > 0$.

B2: Proof of Proposition 2

The shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ follow the joint cumulative distribution function (for $e_j > 0$):

$$F_{\boldsymbol{\varepsilon}}(\mathbf{e}) = \left[1 + \sum_{j=1}^{K} \left(\frac{e_j}{H_j} \right)^{-\xi} \right]^{-1/\delta},$$

where $\xi > 0$, $\delta > 0$, and $H_j > 0$. Agents select k^* to maximize:

$$V_k = W_k \varepsilon_k^{\omega}, \qquad W_k > 0, \quad \omega > 0.$$

I derive the choice probabilities $\mu_k = \Pr(k^* = k)$, the conditional CDF $F_{\varepsilon_k | k^* = k}(x)$ (for $x > 0$), and the conditional expectations $\mathbb{E}[\varepsilon_k \mid k^* = k]$ (assuming $\xi > 1$) and $\mathbb{E}[V_k \mid k^* = k]$ (assuming $\xi > \omega$). The conditional CDF $F_{V_k | k^* = k}(v)$ (for $v > 0$) is derived as an intermediate step.

B2.1. Transformation and Shocks

The multivariate Dagum distribution with Clayton dependence admits a gamma–mixture representation. Let $\zeta_1, \ldots, \zeta_K$ be i.i.d. standard Fréchet random variables with CDF $F_\zeta(z) = \exp(-z^{-\xi})$ for $z > 0$. Let $V \sim \Gamma(1/\delta, 1)$, with PDF:

$$f_V(v) = [\Gamma(1/\delta)]^{-1} v^{1/\delta - 1} e^{-v}, \qquad v > 0.$$

Define the shocks by:

$$\varepsilon_j = H_j \, V^{1/\xi} \, \zeta_j.$$

Then, conditioning on $V = v$:

$$\Pr(\varepsilon_j \le e_j \ \forall j \mid V = v) = \prod_{j=1}^{K} \exp\left(- \left(\frac{e_j}{H_j v^{1/\xi}} \right)^{-\xi} \right) = \exp\left(-v \sum_{j=1}^{K} \left(\frac{e_j}{H_j} \right)^{-\xi} \right).$$

Integrating over v yields:

$$F_{\boldsymbol{\varepsilon}}(\mathbf{e}) = \frac{1}{\Gamma(1/\delta)} \int_0^\infty v^{1/\delta - 1} e^{-v\,[\,1 + \sum_j (e_j/H_j)^{-\xi}\,]}\, dv = \left[1 + \sum_{j=1}^{K} \left(\frac{e_j}{H_j} \right)^{-\xi} \right]^{-1/\delta}.$$

B2.2. Derivation of Choice Probabilities

Maximizing V_k is equivalent to maximizing $V_k^{1/\omega} = W_k^{1/\omega} \varepsilon_k$. With $\varepsilon_k = H_k V^{1/\xi} \zeta_k$, define $\phi_k = W_k^{1/\omega} H_k$. Since $V^{1/\xi} > 0$ is common across k, it cancels in the arg max:

$$k^* = \arg\max_k \{ \phi_k \zeta_k \}.$$

As ζ_k are i.i.d. Fréchet(ξ), the choice probabilities are:

$$\mu_k = \Pr(k^* = k) = \frac{\phi_k^\xi}{\sum_{j=1}^K \phi_j^\xi} = \frac{(W_k^{1/\omega} H_k)^\xi}{\sum_{j=1}^K (W_j^{1/\omega} H_j)^\xi}.$$

This concludes the proof of Proposition 2(i).

B2.3. Derivation of the Conditional CDFs

For the winner's underlying shock, one has (for $z > 0$):

$$F_{\zeta_k \mid k^* = k}(z) = \frac{\Pr(\zeta_k \le z,\ k^* = k)}{\Pr(k^* = k)} = \exp\left(-\mu_k^{-1} z^{-\xi} \right).$$

Using $\varepsilon_k = H_k V^{1/\xi} \zeta_k$ and conditioning on V:

$$\begin{aligned} F_{\varepsilon_k \mid k^* = k}(x) &= \mathbb{E}_V \left[\Pr\left(\zeta_k \le \tfrac{x}{H_k V^{1/\xi}} \,\middle|\, k^* = k, V \right) \right] = \mathbb{E}_V \left[\exp\left(-\mu_k^{-1} \left(\tfrac{x}{H_k V^{1/\xi}} \right)^{-\xi} \right) \right] \\ &= \mathbb{E}_V \left[\exp\left(-\mu_k^{-1} (x/H_k)^{-\xi} V \right) \right] = \left[1 + \mu_k^{-1} \left(\frac{x}{H_k} \right)^{-\xi} \right]^{-1/\delta} = \left[1 + \left(\frac{x}{H_k \mu_k^{-1/\xi}} \right)^{-\xi} \right]^{-1/\delta}, \end{aligned}$$

which is Dagum with scale $H_k\, \mu_k^{-1/\xi}$.

As an intermediate step, the conditional CDF of $V_k = W_k \varepsilon_k^\omega$ (for $v > 0$) is:

$$F_{V_k|k^*=k}(v) = \left[1 + \mu_k^{-1} \left(\frac{v}{W_k}\right)^{-\xi/\omega} H_k^\xi\right]^{-1/\delta} = \left[1 + \left(\frac{v}{W_k H_k^\omega \mu_k^{-\omega/\xi}}\right)^{-\xi/\omega}\right]^{-1/\delta}.$$

This concludes the proof of Proposition 2(ii).

B2.4. Derivation of Conditional Expectations

The conditional distribution $\varepsilon_k \mid k^* = k$ is Dagum with shape $a = \xi$, second parameter $p = 1/\delta$, and scale $b = H_k \mu_k^{-1/\xi}$. Its m-th moment exists for $\xi > m$ and equals:

$$\mathbb{E}[X^m] = b^m \frac{\Gamma(p + m/a)\,\Gamma(1 - m/a)}{\Gamma(p)}.$$

Thus, for $\xi > 1$:

$$\mathbb{E}[\varepsilon_k \mid k^* = k] = H_k \mu_k^{-1/\xi} \cdot \frac{\Gamma(1/\delta + 1/\xi)\,\Gamma(1 - 1/\xi)}{\Gamma(1/\delta)}.$$

For $\mathbb{E}[V_k \mid k^* = k] = W_k \mathbb{E}[\varepsilon_k^\omega \mid k^* = k]$, with $\xi > \omega$:

$$\mathbb{E}[\varepsilon_k^\omega \mid k^* = k] = \left(H_k \mu_k^{-1/\xi}\right)^\omega \frac{\Gamma(1/\delta + \omega/\xi)\,\Gamma(1 - \omega/\xi)}{\Gamma(1/\delta)}.$$

Let $\phi_k = W_k^{1/\omega} H_k$ and $W = \sum_{j=1}^K \phi_j^\xi$. Since $\mu_k = \phi_k^\xi / W$:

$$W_k \left(H_k \mu_k^{-1/\xi}\right)^\omega = W_k H_k^\omega \frac{W^{\omega/\xi}}{\phi_k^\omega} = W^{\omega/\xi}.$$

Hence:

$$\mathbb{E}[V_k \mid k^* = k] = W^{\omega/\xi} \cdot \frac{\Gamma(1/\delta + \omega/\xi)\,\Gamma(1 - \omega/\xi)}{\Gamma(1/\delta)},$$

which is independent of k. This concludes the proof of Proposition 2(iii).

Appendix C: Simulated Eaton–Kortum Model

This appendix presents a simple quantitative simulation to assess the potential importance of the additional margin introduced by the equivalence result in an otherwise standard Eaton–Kortum model of international trade. The counterfactual system of equations is as outlined in the main text. To simulate the model, I take as given $\{\mu_k, \pi_{kj}, Y_k\}$ in the benchmark equilibrium. For tractability, I consider two symmetric

countries, A and B, with the following parameterization:

γ	θ	Y_A	Y_B	π_{AA}	π_{AB}	π_{BA}	π_{BB}	μ_A	μ_B	ρ
0.5	4	2	2	0.5	0.5	0.5	0.5	(0.5,1)	(0,0.5)	(1,3)

In the benchmark, the two countries are identical in all respects except for the initial firm shares μ_A and μ_B. I vary initial share μ_A between 0.5 and 1 such that $\mu_B = 1-\mu_A$ varies between 0 and 0.5. The dependence parameter ρ ranges between 1 and 3, with $\rho = 1$ making equation (v) irrelevant since the discrete–choice margin has no effect on equilibrium.

For each (μ_A, ρ) pair, I compute the counterfactual equilibrium following a 10% reduction in trade costs from A to B:

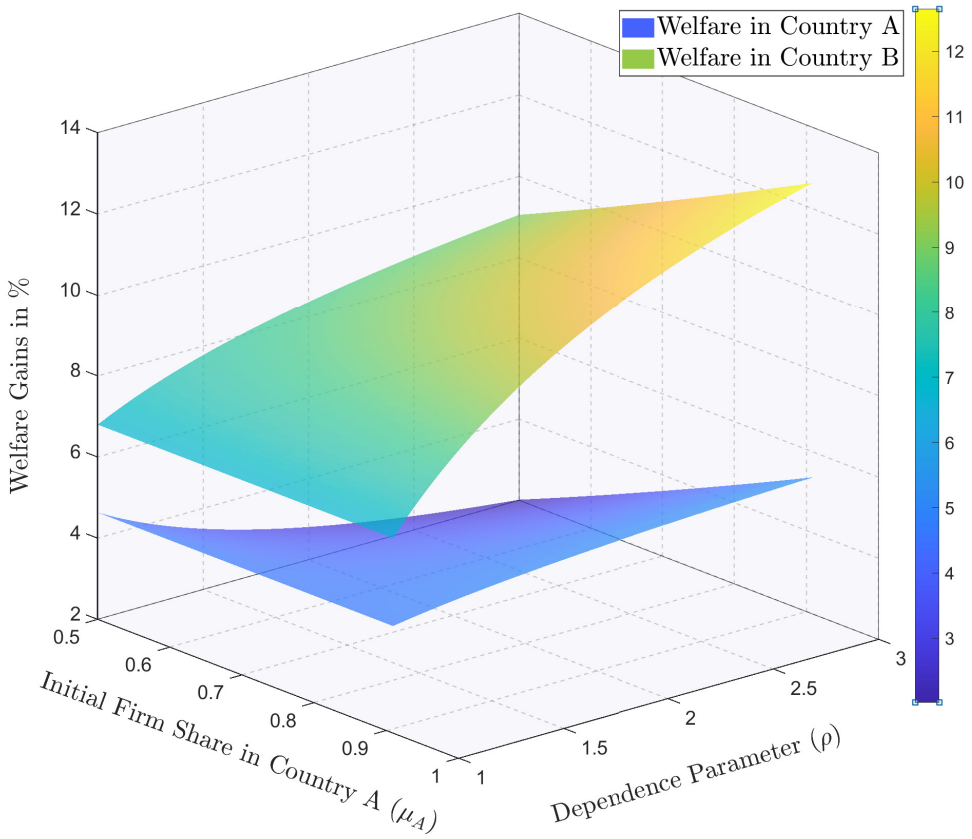
$$\widehat{\tau}_{AB} = 0.9.$$

The resulting welfare changes are plotted in Figure C1.

First, consider the benchmark Eaton–Kortum case in which the discrete–choice margin is absent ($\rho = 1$). In this case, the welfare gains from trade are independent of the initial allocation of μ's. The gains are slightly higher in Country B (6.8% vs. 4.6%) because imports into B become directly cheaper when $\widehat{\tau}_{AB} = 0.9$, whereas the effects for Country A operate only through general equilibrium feedbacks.

Next, fix a symmetric initial allocation of firms, $\mu_A = \mu_B = 0.5$, and vary ρ. As the correlation parameter increases from 1 to 3, the welfare gains from trade rise in Country B (from 6.8% to 9.1%) but fall in Country A (from 4.6% to 2.1%). This pattern is intuitive: higher values of ρ make firms more responsive to relative payoff differences, so the same counterfactual shock $\widehat{\tau}_{AB}$ induces a larger reallocation of firms toward B. The relocation is driven by lower production costs in B due to cheaper intermediate imports. As a result, the additional discrete–choice margin becomes quantitatively important when $\rho > 1$.

Finally, the quantitative impact of the additional discrete–choice margin is most pronounced when the initial allocation of firms is highly asymmetric. Suppose that over 90% of firms are initially located in Country A. In this case, as ρ increases from 1 to 3, the welfare gains from trade in Country B rise markedly from 6.8% to 12.7%, and they also increase for Country A from 4.6% to 5.4%. This occurs because a high



Notes: The figure plots counterfactual welfare changes for two symmetric countries in an Eaton–Kortum model with endogenous firm location, following a 10% reduction in trade costs from Country A to Country B. The surfaces show the welfare change for each country as a function of the initial firm share μ_A ($\mu_B = 1 - \mu_A$) and the dependence parameter ρ.

Figure C1: Simulated Welfare Gains from Trade

initial μ_A magnifies the effect of lowering τ_{AB} on the price of imports in Country B, which in turn generates positive general equilibrium spillovers back to Country A.

Although the exercise here is based on a stylized two–country setting, it highlights that even modest positive correlation ($\rho > 1$) can introduce sizable new adjustment margins through the equivalence result, without altering the rest of the Eaton–Kortum structure. While estimating ρ from data is beyond the scope of this paper, the example illustrates how the framework provides a clear path for such estimation in future work.