From Discrete Choice to Distributions: Equivalence Results in Selection Models with Correlated Shocks

Sergey Nigai*

August, 2025

Abstract

This paper develops a unified framework for modeling discrete allocation across alternatives together with continuous heterogeneity within each allocation, when outcomes are governed by two widely used heavy-tailed families: the extreme-value (Fréchet) and power-law (Dagum) distributions. I establish *equivalence results* that deliver deterministic, one-to-one mappings between choice probabilities and the conditional distributions of post-allocation outcomes. These mappings operate in both directions, hold in general equilibrium, and accommodate correlated shocks, while preserving the tractability of canonical discrete-choice and heterogeneous-agent models within a single framework. The approach nests and extends a wide range of existing settings. Applications to international trade, economic geography, and assignment models demonstrate how the equivalence results offer new analytical tools for studying selection in general equilibrium under heavy-tailed heterogeneity.

Keywords: *International trade, Selection models, Heterogeneous firms, Discrete choice, Fréchet distribution, Dagum distribution, Economic geography, Correlated shocks*

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

*University of Colorado Boulder and CESIfo.

Acknowledgements and Disclosures: I am grateful to Carlos Martins-Filho for insightful discussions on multivariate distributions and copula dependence. I received no external funding for this research and have no relevant financial relationships or potential conflicts of interest to disclose.

1 Introduction

Closed-form discrete-choice models are central to quantitative work across economics. Separately, models of heterogeneity based on heavy-tailed distributions are equally influential for studying individual and aggregate outcomes. This paper develops a tractable framework that integrates these two approaches, establishing a closed-form link between the discrete-choice margin and the distributional margin.

I establish equivalence results for two widely used heavy-tailed families, Fréchet and Dagum, that produce deterministic, one-to-one mappings between choice probabilities and the conditional distributions of post-selection outcomes within each alternative. These mappings work in both directions, hold in general equilibrium, and allow dependence across alternatives while preserving the analytical transparency that makes canonical discrete-choice and distribution-based heterogeneous-agent models valuable.

The equivalence renders the discrete-choice and distributional margins jointly tractable: changes in fundamentals generate predictable co-movements of shares and within-choice distributions without new numerical machinery. By accommodating correlated shocks, the general structure applies broadly. Applications to international trade, spatial models, and occupational sorting show how standard environments can be specified so that allocation and within-choice heterogeneity are jointly governed by a single, closed-form relationship.

A useful way to see the contribution is to contrast two familiar approaches. In one, agents draw idiosyncratic shocks and select the option with the highest payoff, yielding smooth choice probabilities but leaving the composition within each choice unspecified. In the other, each alternative is endowed with a heterogeneous population, specifying within-choice distributions while treating allocation across alternatives as exogenous.The framework introduced here unifies these perspectives within a single tractable structure, exactly reproducing the choice probabilities of the first while simultaneously delivering the within-choice distributions of the second.

Section 3 introduces the multivariate environments and payoff structures. Section 4 states and proves the equivalence results in full detail. For Fréchet shocks with logistic-type dependence, choice shares take a logit-like form, and the conditional distribution within each chosen alternative is also Fréchet, with its scale determined by

that share. For Dagum shocks with Clayton-type dependence, an analogous equivalence holds with power-law tails. In both cases, the same closed-form link between shares and conditional distributions applies regardless of the number of alternatives and in the presence of correlation across them.

Section 5 applies the results in three settings. In trade, the framework delivers endogenous location choice by firms in 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 keeping congestion and agglomeration effects explicit. In a Roy model of occupational choice, as in Hsieh et al. (2019), it produces closed-form sorting patterns with correlated abilities across occupations. Together, these applications show how the equivalence results open new analytical margins in established general-equilibrium models while retaining tractability.

2 Related Literature

The equivalence results developed here rely on two well-established strands of the literature: (i) discrete-choice models, which deliver closed-form allocation shares and are widely used for modeling selection across discrete alternatives, and (ii) distribution-based approaches to heterogeneity, which characterize the distribution of outcomes within alternatives, often using extreme-value or power-law families. Discrete-choice models provide tractable representations of selection behavior, while heavy-tailed distributions capture empirically relevant patterns of heterogeneity in a parsimonious way. The contribution here is to unify these frameworks in a single structure that preserves their respective advantages. Accordingly, the literature review focuses on key contributions that make this integration possible.

First, this paper builds on the modern discrete-choice literature rooted in the random-utility framework of McFadden (1974) for conditional logit models with independent alternatives, and its extension to correlated choices via the generalized extreme value and nested logit models in McFadden (1978), where agents select from a finite set of alternatives based on latent utilities perturbed by idiosyncratic shocks. These frameworks accommodate rich patterns of correlation across alternatives but typically

do not yield explicit post-selection outcome distributions – a feature central to the equivalence results developed here. 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). This literature also draws on Anderson et al. (1992), who demonstrate how discrete-choice models integrate naturally into spatial and differentiated-product contexts. The equivalence results in this paper extend these approaches by embedding post-selection heterogeneity in wages, productivity, or other outcomes – dimensions typically left implicit in standard formulations.

The second relevant strand is the literature that employs extreme-value and power-law distributions to model heterogeneity in economics. In particular, the multivariate 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; Lind and Ramondo, 2023), 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). The equivalence results in this paper provide a structural foundation for interpreting Fréchet-distributed outcomes as the result of discrete selection over 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.

The equivalence results also encompass models based on power-law distributions, building on a longstanding 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;

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

Arkolakis et al., 2018), and the distributional effects of globalization (Helpman et al., 2010; Nigai and Yang, 2024). 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). As in the Fréchet case, the equivalence results here provide a structural link between discrete selection and continuous, within-group heterogeneity when the latter follows a power law.

Taken together, these literatures furnish the tractable allocation rules and heterogeneity patterns that the equivalence results combine in a single unified framework.

3 Selection Margins in Models with Heterogeneity

Many economic environments feature either a discrete-choice margin, which determines allocation across occupations, sectors, or locations, or a distributional margin, which captures heterogeneity in productivity, income, or skills within a given allocation. It is relatively uncommon to model both margins jointly, because linking them in a tractable way is analytically challenging. Yet both are essential for understanding how selection patterns interact with within-group heterogeneity and for mapping changes in fundamentals into aggregate and distributional outcomes. Conceptually, this process takes place in two distinct but potentially interdependent stages: a *discrete selection stage* and a *post-selection stage*.

Stage 1: Discrete-Choice Margin. Let the finite set of alternatives be $\mathcal{K} = \{1, \ldots, K\}$. Each alternative $k \in \mathcal{K}$ yields a payoff

$$V_k(\varepsilon_k), \quad \text{with } V_k : \mathbb{R}_+ \to \mathbb{R},$$

a deterministic, strictly increasing function mapping a k-specific shock ε_k into a realized payoff. The shock vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ is drawn from a joint distribution $F(\boldsymbol{\varepsilon})$ capturing latent heterogeneity across alternatives. Agents select the alternative with the highest payoff:

$$k^* = \arg\max_{k \in \mathcal{K}} \{V_k(\varepsilon_k)\}. \tag{1}$$

This selection rule determines how agents sort across $\mathcal{K}$, with patterns shaped jointly by the distribution $F(\boldsymbol{\varepsilon})$ and the payoff functions V_k. Much of the literature assumes

independence, $F(\boldsymbol{\varepsilon}) = \prod_k F_k(\varepsilon_k)$. In many economic settings, however, shocks are likely to be correlated, with nonzero tail dependence, making it important to allow for richer joint distributions. This possibility is precisely accommodated in the equivalence framework developed below.

Stage 2: Distributional Margin. Conditional on selecting k, agents' shocks follow the *conditional marginal CDF*

$$F_{\varepsilon_k|k^*=k}(x) \equiv \mathbb{P}(\varepsilon_k \leq x \mid k^* = k).$$

This object is central for analyzing the within-alternative distribution of outcomes, such as productivity, income, or skills, among agents who sort into alternative k.

Core Challenge. A fundamental recurring challenge in this class of models is jointly modeling both discrete selection and continuous within-choice heterogeneity in a tractable manner. The selection operator $\arg\max_k\{V_k(\varepsilon_k)\}$ mechanically induces dependence across alternatives, either directly through the maximization or indirectly through general equilibrium interactions. This dependence makes it analytically difficult to characterize $F_{\varepsilon_k|k^*=k}(x)$ while also retaining closed-form expressions for choice probabilities. As a result, most existing models achieve tractability at either the discrete-choice margin or the distributional margin, but rarely at both simultaneously.

This paper develops a unified distributional framework that resolves this tension. I construct two classes of multivariate heavy-tailed distributions – Fréchet with logistic dependence and Dagum with Clayton dependence – that simultaneously deliver:

(i) closed-form choice probabilities,

(ii) closed-form conditional distributions $F_{\varepsilon_k|k^*=k}(x)$, and

(iii) shocks across alternatives that may be positively correlated.

An *equivalence result* precisely characterizes the conditions under which probabilistic discrete choice is analytically equivalent to observing continuous conditional outcome distributions, allowing both margins to be modeled jointly in a tractable way.

4 Equivalence Results

Before presenting the formal results for the two heavy-tailed families, I first state the general version and provide the intuition. The equivalence result establishes a closed-form link between the discrete-choice margin and the distributional margin. Let $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ be drawn from a joint CDF $F : \mathbb{R}_+^K \to [0,1]$. Agents choose the alternative

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

with choice shares

$$\mu_k \equiv \Pr(k^* = k).$$

When the equivalence result holds, the joint distribution F implies a shape parameter $\varphi > 0$ and positive constants $\{m_k\}_{k=1}^K$. In this case, the following three properties are satisfied simultaneously:

(i) Choice shares: $\mu_k = \dfrac{(m_k W_k^{1/\omega})^{\varphi}}{\sum_{j=1}^K (m_j W_j^{1/\omega})^{\varphi}}$,

(ii) Post-selection shocks: $F_{\varepsilon_k | k^*=k}(x) = G(x; T_k(\mu_k), \varphi)$, where T_k depends on μ_k,

(iii) Equalized payoffs: $\mathbb{E}[V_k \mid k^* = k] = \Xi(W_1, \ldots, W_K) \quad \forall k.$

These properties imply that the set of choice shares $\{\mu_k\}$ and the set of conditional distributions $\{F_{\varepsilon_k | k^*=k}\}$, are informationally equivalent: knowing one determines the other exactly. This equivalence is valuable because it creates a direct mapping between discrete-choice models and within-group outcome distributions, allowing the two margins to be analyzed jointly without additional numerical machinery. It enables researchers to recover one margin from data on the other, to conduct counterfactuals that track both selection and composition effects in closed form, and to embed richer heterogeneity structures into established general-equilibrium models while preserving tractability.

Before turning to the formal propositions, it is useful to develop intuition with a simple example. Consider two alternatives, A and B, with with shocks ε_A and ε_B

drawn independently from identical Fréchet distributions,

$$F(\varepsilon) = \exp(-\varepsilon^{-\theta}), \quad \theta > 0.$$

While this example assumes independence for simplicity, the full framework accommodates correlation. Payoffs are given by:

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

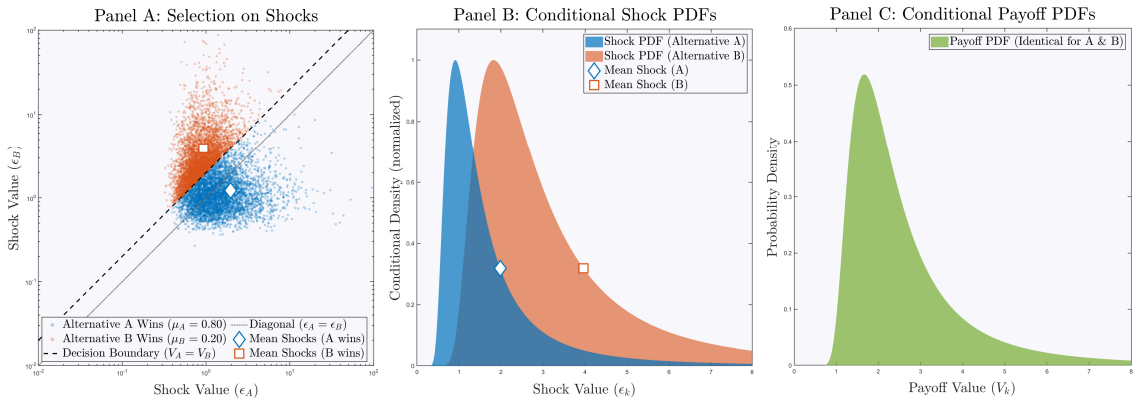
Let $W_A > W_B$, so that A is the "favorite" and B the "underdog." In this specific example, the parameters are chosen to generate choice probabilities of 80% for the favorite and 20% for the underdog.

Figure 1 provides a visual walkthrough of the equivalence result in this setting. Its three panels trace the journey from the latent shocks and the selection process (Panel A), to the resulting distribution of the winning shocks (Panel B), and finally to the distribution of the realized payoffs (Panel C).

Panel A visualizes the selection process in the two-dimensional space of shocks. Each dot represents a single agent's draw $(\varepsilon_A, \varepsilon_B)$, colored by their chosen alternative. The dashed black line represents the decision boundary where payoffs are equal ($V_A = V_B$). Because the favorite (Alternative A) has a higher multiplier W_A, this boundary is steeper than the main diagonal (the dotted gray line, where $\varepsilon_A = \varepsilon_B$). Because the selection operates on payoffs and not raw shocks, an agent choosing the underdog must draw a relatively large shock ε_B to overcome both the competing shock ε_A and the intrinsic advantage of the favorite as $W_A > W_B$. Hence, the conditional mean is higher for alternative B.

Panel B reveals the consequence of this selection process for the distribution of the winning shocks. The conditional distribution for the favorite (blue) is relatively close to the origin. In contrast, the conditional distribution for the underdog (red) is shifted to the right. Because the underdog is intrinsically disadvantaged, it only wins in those rare instances when it receives an exceptionally large shock. The favorite, on the other hand, can win with more moderate draws. The equivalence result formalizes this by showing that a lower choice share (μ_k) implies a stronger selection effect, which endogenously generates a conditional shock distribution with a higher mean.

The final panel illustrates that despite the differences in conditional distributions (as shown in Panel B), the resulting conditional distributions of the winning *payoffs* are identical. The higher average shock required for the underdog to win is exactly offset by its lower deterministic payoff shifter (W_B). This selection-on-shocks effect perfectly equalizes the selection-on-payoffs. This is the core of the equivalence result: the model remains tractable precisely because the complex, choice-specific selection on shocks collapses into a simple, choice-invariant distribution of equilibrium payoffs.



Notes: The figure illustrates the equivalence result with two alternatives, A ($\mu_A = 0.8$) and B ($\mu_B = 0.2$). 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 the conditional shock PDFs, normalized to a peak of 1. Panel C plots the conditional payoff PDFs.

Figure 1: Simulated Draws from a Bivariate Fréchet Model

This bivariate case illustrates the three properties of the unified framework – closed-form choice shares, post-selection shock distributions linked to those shares, and equalized conditional payoffs – in their simplest, independent-shock setting. I now turn to the general case with *positively dependent* shocks. The next two propositions give exact conditions under which the equivalence result holds for two widely used heavy-tailed families: multivariate Fréchet with logistic dependence and multivariate Dagum with Clayton dependence. In both cases, the dependence structure can be represented via a copula, as detailed in Appendices A and B.

Proposition 1 (Equivalence Result under Multivariate Fréchet with Logistic Dependence)**.** Suppose the vector of shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ has the joint cumulative

distribution function:

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

where $\theta, T, h_j > 0$ and $\rho \geq 1$. Let agents choose the alternative k that maximizes the payoff $V_k = W_k\varepsilon_k^{\omega}$, with $W_k > 0$ and $\omega > 0$. Then:

(i) The choice probabilities (shares) are given by

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

(ii) The conditional distribution function of ε_k given selection is

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

(iii) If $\theta > 1$, the expected value of shock ε conditional on choosing k is

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

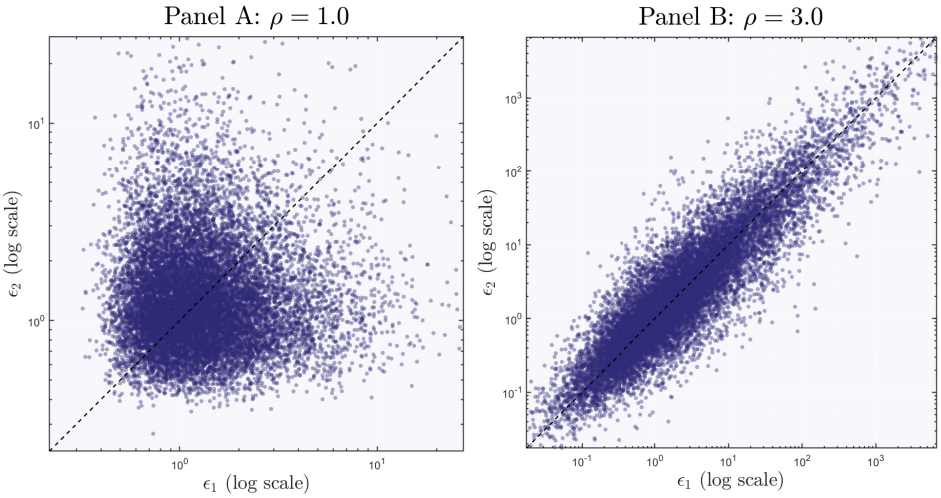
and if $\theta > \omega$ the expected payoff conditional on selection satisfies

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

where $W = \left[\sum_{j=1}^{K}\left(W_j^{1/\omega}h_j\right)^{\theta\rho}\right]^{1/\rho}$.

Proof. The proof is available in the Appendix A. □

Proposition 1 establishes a tractable and economically meaningful link between micro-level heterogeneity and aggregate outcomes. The multivariate Fréchet distribution with logistic-type dependence yields closed-form expressions for choice shares, conditional distributions, and expected payoffs in terms of interpretable parameters. This structure captures key features of many economic settings – heavy-tailed heterogeneity and positive dependence across alternatives – making it useful in multiple applications. The analytical transparency of the equivalence result provides a robust



Notes: The figure illustrates 15,000 simulated draws from a bivariate Fréchet distribution with common parameters $T = 1$, $h_k = 1$, 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

foundation for empirical implementation and counterfactual analysis in environments where extreme-value heterogeneity and selection are central. As shown in Figure 2, the framework accommodates both independent shocks ($\rho = 1$, Panel A) and positively correlated shocks ($\rho > 1$, Panel B). In both cases, the equivalence result holds: choice shares, post-selection distributions, and conditional payoffs all admit closed-form expressions governed by the same combination of parameters $\{\theta, \rho\}$.

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 model in international trade due to its sharp predictions regarding selection, sorting, and aggregate outcomes among heterogeneous firms. Despite capturing empirically observed fat-tailed behavior effectively, the Pareto distribution typically does not yield an equivalence result with closed-form choice shares. To address this limitation, I consider the Dagum distribution, which preserves power-law tail decay yet offers greater flexibility in modeling dependence across alternatives. Crucially, this distribution admits a closed-form equivalence structure, making it a valuable benchmark for studying selection and sorting under power-law heterogeneity while remaining analytically transparent. The corresponding equivalence result is formally stated in Proposition 2.

Proposition 2 (Equivalence Result under Multivariate Dagum with Clayton Dependence)**.** Let the vector of shocks $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_K)$ have the joint cumulative distribution function:

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

where $\xi, \delta, d_j > 0$. Let agents choose the alternative k that maximizes the payoff $V_k = W_k \varepsilon_k^{\omega}$, with $W_k > 0$ and $\omega > 0$. Then:

(i) The choice probabilities (shares) are given by

$$\mu_k = \frac{\left(W_k^{1/\omega} d_k\right)^{\xi}}{\sum_{j=1}^{K} \left(W_j^{1/\omega} d_j\right)^{\xi}}.$$

(ii) The conditional distribution function of ε_k given selection is

$$F_{\varepsilon_k | k^*=k}(x) = \left[1 + \left(\frac{x}{d_k \, \mu_k^{-1/\xi}}\right)^{-\xi}\right]^{-1/\delta}.$$

(iii) If $\xi > 1$, the expected value of shock ε_k conditional on choosing k is

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

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

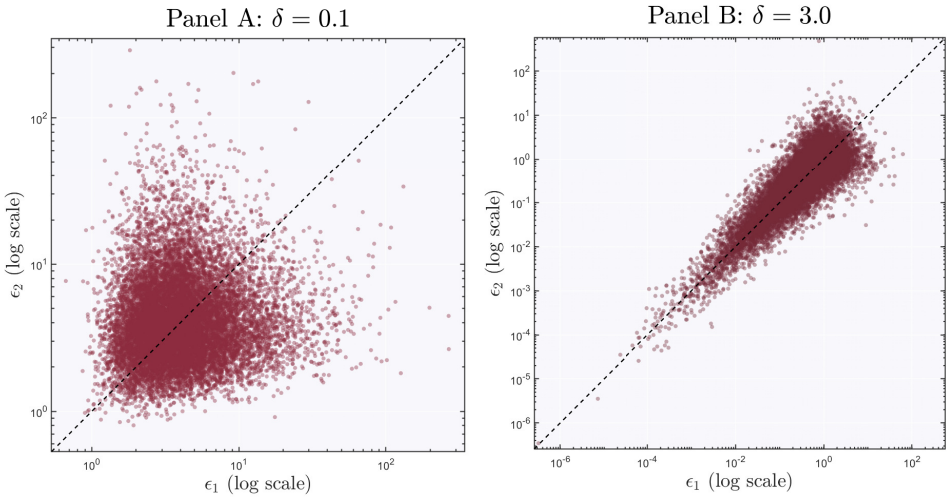
$$\mathbb{E}[V_k \mid k^* = k] = W^{\omega/\xi} \cdot \frac{\Gamma(1 - \omega/\xi)\, \Gamma(1/\delta + \omega/\xi)}{\Gamma(1/\delta)},$$

where $W = \sum_{j=1}^{K} \left(W_j^{1/\omega} d_j\right)^{\xi}$.

Proof. The proof is available in the Appendix B. □

Proposition 2 establishes that under multivariate Dagum shocks with Clayton dependence, the equivalence result holds: choice shares, post-selection distributions, and conditional payoffs all have closed-form expressions governed by the same shape parameter ξ, capturing power-law tails and flexible dependence across alternatives.

It applies whether shocks are independent or positively dependent, as illustrated in Figure 3, and provides a tractable basis for modeling selection in settings such as trade, spatial allocation, and occupational sorting where heavy-tailed heterogeneity is empirically relevant.



Notes: The figure illustrates 15,000 simulated draws from a bivariate Dagum distribution with common parameters $d_k = 1$ 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

4.1 Why Fréchet and Dagum? The Unifying Principle

Both the Fréchet and Dagum families, when paired with logistic (Gumbel–Hougaard) and Clayton copulas respectively, preserve tractability in selection models with correlated shocks. They share a distinctive structural property that underlies the equivalence results: the shocks can be decomposed multiplicatively into a *common, choice-irrelevant factor* that generates dependence and an *independent, choice-relevant component* that governs selection across alternatives. This decomposition is particularly valuable among widely used heavy-tailed families because it preserves closed-form expressions for both choice shares and post-selection distributions, even when shocks are positively dependent.

Formally, in both cases the shocks take the form

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

where Z_k is i.i.d. across k, C is a common factor capturing dependence, and $\varphi > 0$ is the shape parameter. Naturally, the distributions of Z_k and C differ across the two families.

Fréchet with logistic dependence. For 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 i.i.d. standard Fréchet random variables with shape parameter $\theta > 0$, and $S \sim \text{Stable}_\rho^+$ is a positive ρ-stable random variable ($\rho \geq 1$, independence when $\rho = 1$) capturing upper-tail dependence (see Resnick, 1987).

Under the selection rule

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

the *realized* value of S cancels from all alternatives for a given draw. However, the dependence parameter ρ remains economically relevant: it enters the closed-form choice probabilities through the effective shape parameter $\theta\rho$, and it rescales the post-selection distributions. Thus, while S does not drive selection at the level of a single draw, its distribution shapes both margins in expectation. This decomposition is mathematically similar to the "power-sum" form in the Generalized Extreme Value (GEV) framework of McFadden (1978), though here the shocks are multivariate Fréchet, so the framework retains both the discrete choice and the continuous distribution of the selected outcome in closed form.

Dagum with Clayton dependence. For 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 i.i.d. standard Fréchet random variables with shape parameter $\xi > 0$, and $Y \sim \text{Gamma}(1/\delta, 1)$ is gamma-distributed with shape $1/\delta > 0$ and unit scale, capturing lower-tail dependence ($\delta > 0$, independence as $\delta \to 0$) (see Nelsen, 2006).

Under the selection rule, the *realized* value of Y cancels from all alternatives for a given draw. Nonetheless, the dependence parameter δ remains relevant: it appears

in the closed-form choice probabilities through the shape of the marginal distribution of Z_k and rescales the post-selection distributions. The gamma factor generates a Clayton copula in the joint distribution, which together with the Fréchet marginals imply the resulting Dagum distribution.

4.2 The Equivalence Result and Selection Margins

The equivalence result delivers fully closed-form comparative statics in heterogeneous-agent models for both the *discrete-choice margin* and the *distributional margin*. These responses are (i) globally consistent with the underlying joint distribution and (ii) analytically tractable. Because both margins shape aggregate outcomes, the framework allows many distributional statistics (means, variances, quantiles) to be computed in closed form. By jointly characterizing allocation and composition, the framework provides a transparent, structural tool for tracing how economic shocks – whether from policy or exogenous factors – affect both aggregate and distributional outcomes.

Aggregate Outcomes. Let the aggregate outcome associated with alternative k (e.g., total output or income) be

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

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. From the equivalence result,

$$Y_k = \underbrace{\mu_k}_{\text{Discrete-choice Margin}} \cdot \underbrace{\mu_k^{-1/\varphi} M_k}_{\text{Distributional Margin}},$$

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 from Propositions 1–2. This decomposition separates the *extensive* margin (share of entrants) from the *intensive* margin (average quality of entrants).

Elasticity Decomposition. The elasticity of Y_k with respect to W_k, holding other parameters fixed, is

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

Closed-form expressions for both terms follow directly from the equivalence result:

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

Thus

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

The first term (*choice elasticity*) captures how the mass of agents selecting k responds to changes in W_k; the second (*distributional elasticity*) captures the composition effect, as marginal entrants are lower-drawn types. The dependence parameters ρ and δ influence φ and hence both margins simultaneously. When $\varphi > 1$, the discrete-choice margin dominates and increases in W_k raise Y_k despite the negative composition effect.

Distributional Outcomes. The equivalence result also yields the full conditional distribution $F_{\varepsilon_k|k^*=k}(x)$ in closed form. In both models, the conditional distribution remains in the same parametric family as the primitive shocks, with shape parameter preserved and scale linked to the choice share:

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

This parametric structure makes it straightforward to compute moments, quantiles, Lorenz curves, and Gini coefficients, and to trace the impact of changes in W_k or μ_k on the entire distribution. The same closed-form mapping governs both selection and composition, enabling unified analysis of aggregate and distributional effects of shocks—whether from trade, policy, or technology.

5 Quantitative Applications

Having established the main theoretical results and their economic interpretation, I now illustrate how the equivalence framework can be effectively applied in quantitative contexts. The applications presented here are selected to address three long-standing challenges in economics: (i) incorporating endogenous location choices by firms within general equilibrium models of international trade; (ii) explicitly modeling heterogeneity in discrete choice frameworks of worker location; and (iii) capturing labor assignment decisions when ability shocks are correlated across occupational alternatives. The equivalence results allow these extensions to be introduced in an analytically transparent manner, requiring minimal structural modifications to standard economic models.

5.1 Application A: Trade Model with Location Choice

In the canonical Eaton and Kortum (2002) framework, the country–specific scale parameter T_k of the Fréchet productivity distribution, as well as the number of potential producers in each country, are exogenously fixed. The equivalence result in Proposition 1 allows us to reinterpret T_k as the *conditional* scale parameter for firms that have chosen to locate in k, with the location choice governed by the same primitives that determine within–country productivity. This preserves the closed–form tractability of the Eaton–Kortum model while endogenizing the *mass of firms* μ_k across countries. Changes in fundamentals now affect both the *distributional margin* (the conditional productivity distribution within each country) and the *discrete–choice margin* (the share μ_k of global entrepreneurs locating in each country).

The dependence parameter ρ plays a pivotal economic role. When $\rho = 1$, location choices are independent across countries, and the discrete–choice and distributional margins exactly offset. In this knife–edge case, the model collapses to the standard Eaton–Kortum structure. When $\rho > 1$, positive dependence implies that locations are closer substitutes: small differences in fundamentals trigger large shifts in μ_k, amplifying all standard intensive–margin responses of trade flows, price indices, and welfare.

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:

$$c_i = w_i^{\gamma} P_i^{1-\gamma}, \tag{2}$$

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:

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

and the aggregate CES price index is:

$$P_j = C \left(\sum_{k=1}^{K} T_k \left(c_k \tau_{kj}\right)^{-\theta} \right)^{-\frac{1}{\theta}}, \tag{4}$$

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

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

and trade is balanced:

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

A *trade equilibrium* is a set of wages and prices such that (2)–(5) hold for all $i, j \in \mathcal{K}$.

Introducing Discrete Choice. In the canonical EK 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

$$T_k = T\, h_k^{\theta}\, \mu_k^{-1/\rho},$$

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 (3)–(4) yields:

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

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

where the factor $\mu_k^{1-1/\rho}$ is the *discrete–choice margin.*

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. 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:

$$\mu_k = \frac{(h_k/c_k)^{\theta\rho}}{\sum_{j=1}^{K} (h_j/c_j)^{\theta\rho}}.$$

Counterfactuals. The equivalence 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 $\rho = 1$, equation (v) is irrelevant because the discrete–choice margin has no effect on equilibrium. When $\rho = 1$, equation (v) is irrelevant because the discrete–choice margin has no effect on equilibrium. When $\rho > 1$, closer substitution between locations implies that small differences in fundamentals induce large reallocations of μ_k, amplifying all intensive–margin responses. Introducing location choice via the equivalence result thus endogenizes the mass of firms in each country and adds an extensive margin that can magnify the impact of trade-cost shocks on trade flows, prices, and welfare – a channel absent in the canonical Eaton-Kortum framework. Appendix C presents a simple simulation that quantifies the importance of this additional adjustment margin as a function of ρ and the initial allocation of μ_k.

Discussion of Other Trade Models The equivalence result developed in this paper can also be applied to introduce endogenous location choice in other class of models based on the heterogeneous firm framework of Melitz (2003). In these models, it also is generally assumed that the mass of firms in each country is fixed. This assumption can be imposed explicitly, as in Chaney (2008), or emerge endogenously when firm productivity is Pareto-distributed, as shown in Arkolakis et al. (2008).

The most common parameterization of the productivity distribution in this class of models is the unbounded Pareto distribution due to its tractability properties. Proposition 2 establishes the equivalence result for Dagum's distribution which is

related to Pareto as its tail decays at the same power rate preserving the overall tail behavior. Moreover, the Dagum distribution also possesses several analytical properties that make it immediately applicable to the Melitz (2003) framework. Recall from Proposition 2 that the conditional distribution given selection is defined as:

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

where $b_i = d_k \, \mu_k^{-1/\xi}$.

In the Melitz (2003) setting, firms incur fixed cost of exporting and only exports whenever its draw ϕ exceeds a cutoff ϕ^* pinned down by the fixed cost versus potential profits. First, as in the Pareto case, under the Dagum distribution the share of exporters remains analytic:

$$\Pr(\phi > \phi^*) = 1 - \left[1 + \left(\tfrac{\phi_i^*}{b_k}\right)^{-\xi}\right]^{-1/\delta},$$

The second selection statistics typically involves evaluating the expected value of left-truncated function of productivity shocks ϕ^r. Under the Dagum distribution, this expression also remains tractable and is given by:

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

where $\mathbb{B}(\cdot)$ is an upper incomplete Beta function and provided that $\xi > r$. Because both the export probability and the conditional moments stay analytic, the Dagum distribution furnishes a tractable alternative to the unbounded Pareto while preserving a power-law tail. Coupled with the equivalence result, it therefore offers a convenient route to embed *endogenous location choice* into Melitz-type heterogeneous-firm models without sacrificing closed-form solutions for entry cut-offs, export shares, or aggregate productivity.

5.2 Application B: Economic Geography Model with Heterogeneity

Many influential economic geography models, such as Allen and Arkolakis (2014), embed discrete location choice within a general equilibrium setting, treating the spatial distribution of agents as the outcome of individual optimization. While these

frameworks rigorously analyze spatial sorting and aggregate welfare, they typically abstract from *heterogeneity in outcomes conditional on location*. As a result, they are silent on within-region dispersion in productivity, income, or welfare, and offer limited insight into distributional issues such as inequality or local concentration.

The framework developed here closes this gap by enriching the *post-selection* margin. The equivalence result delivers closed-form solutions for both equilibrium location shares and the conditional distributions of outcomes within each location. This preserves the tractability of canonical models while introducing a second, *distributional margin*, governed by the same parameter $\theta\rho$ that determines substitution patterns across locations. Under the special case $\varsigma = \theta\rho$, the discrete-choice rule here is isomorphic to the canonical logit form in Allen and Arkolakis (2014), but the framework now accommodates within-location heterogeneity in a fully structural way.

The framework proposed in this paper complements this canonical approach by structurally enriching the post-selection margin. Specifically, my equivalence result provides closed-form solutions not only for equilibrium location shares but also for the conditional distributions of outcomes within each location. This extension preserves analytical tractability while significantly broadening the range of questions economic geography models can address.

Structurally, my approach parallels that of Allen and Arkolakis (2014), but whereas their spatial allocations arise from ex-post utility equalization, I derive allocations from idiosyncratic shocks to location-specific payoffs that also lead to utility equalization. Further, under certain parameter restrictions, the two perspectives become isomorphic: the equilibrium shares obtained from my probabilistic discrete choice framework replicate exactly the spatial allocation implied by utility-equalizing models. Accordingly, this formulation also 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:

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

$\sigma > 1$ denote 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:

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

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 constraints hold.

Introducing Heterogeneity. I use the equivalence result 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:

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

where scale parameter $T > 0$ is chosen to ensure that the total labor endowment equals that in the representative-agent benchmark, denoted L. While I am using the CDF from Proposition 1, it is worth noting that one could equivalently use the Dagum distribution in Proposition 2.

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:

$$\mu_k = \frac{W_k^{\theta\rho}}{\sum_{j=1}^{K} W_j^{\theta\rho}}.$$

Under the restriction $\varsigma = \theta\rho$, this discrete choice rule matches the canonical form in equation (8). The key innovation here is to decouple the extensive margin (how many individuals select into each location) from the intensive margin (how much effective labor they supply), introducing a distributional layer of heterogeneity while maintaining the 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:

$$\begin{aligned} M_k &= \mu_k M, \\ L_k &= \mathbb{E}[\ell_k \mid k^* = k] M_k = M T^{1/\theta} \mu_k^{1-\frac{1}{\theta\rho}} \Gamma(1 - 1/\theta). \end{aligned}$$

The equivalence result also immediately pins down the welfare level in all locations as the expected value of the payoff:

$$\mathbb{E}[V_k \mid k^* = k] = \left[\sum_{j=1}^{K} W_j^{\theta\rho} \right]^{1/(\theta\rho)} \Gamma(1 - \omega/\theta).$$

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 $\varsigma = \xi\theta$, the system of equations is:

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

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:

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

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 $\theta\rho \to \infty$, equations $(v.a)$ and $(v.b)$ coincide, recovering the representative-agent benchmark. As $\theta\rho$ 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:

$$\theta\rho > \sigma - 1.$$

For instance, if $\sigma = 5$, the admissible parameter range becomes $\theta\rho \in (4, \infty)$. 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 via the equivalence result is that the model delivers tractable, closed-form post-selection distributions. Unlike canonical economic geography frameworks—which typically remain silent about the distribution of outcomes conditional on location choice—the equivalence approach permits closed-form computation of dispersion statistics in each location and comparative statics with respect to amenities, productivity, and trade costs. This links spatial general equilibrium forces with distributional outcomes in a fully structural and analytically tractable manner.

5.3 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 impact of external shocks or entry barriers on economic outcomes. Most quantitative applications assume that occupational abilities are drawn independently from Fréchet distributions, as in Hsieh et al. (2019) and Galle et al. (2023).[2] While the independence assumption ensures tractability, it rules out economically plausible *positive tail dependence*—the idea that exceptionally high productivity in one occupation is likely to be accompanied by above-average productivity in related occupations.

The equivalence result in Proposition 1 directly addresses this limitation. It accommodates positive dependence across occupations and delivers closed-form expressions for both equilibrium occupational shares and conditional ability distributions. This preserves the analytical transparency of the standard Roy model while allowing the study of within-occupation heterogeneity alongside sorting patterns.

Consider groups $g \in \mathcal{G}$ and occupations $k \in \mathcal{K}$. Individuals from group g draw abilities $\boldsymbol{\varepsilon}_g = (\varepsilon_g(1), \ldots, \varepsilon_g(K))$ from

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

[2] Lagakos and Waugh (2013) allow for correlated abilities but focus on binary choice. My framework preserves tractability even when $|\mathcal{K}|$ is large.

where $\rho \geq 1$ governs tail dependence.

Let the payoff in k be

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

with w_k the wage, $\tau_{gk} \in [0, 1)$ a group-occupation barrier, and h_{gk} a fundamental scale term. The equivalence result implies the share of group-g members choosing k is

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

The average ability of group-g workers in k is

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

As in Hsieh et al. (2019), consider the effect of reducing τ_{gk}. When $\rho > 1$, the removal of barriers produces larger μ_{gk} adjustments than in the independent-draw case. This expands group-g participation in k but draws in marginal entrants with lower ability. The labor supply from group g to occupation k is

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

where N_g is the size of group g.

Because $V_{gk} \mid k^* = k, g$ is Fréchet with shape θ and scale $w_k(1 - \tau_{gk})\, T_g^{1/\theta} h_{gk}\, \mu_{gk}^{-1/(\theta\rho)}$, its variance (for $\theta > 2$) is

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

Hence, within-group income variance can be decomposed via the law of total variance 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} affect occupational allocation, average abilities, and both within- and between-occupation inequality.

6 Conclusion

This paper develops a unified, analytically tractable framework for modeling both *discrete allocation* across alternatives and *continuous heterogeneity* within the selected group. At its core are equivalence results for extreme–value (Fréchet) and power–law (Dagum) specifications that deliver closed–form mappings between choice shares and conditional outcome distributions. These results connect canonical discrete–choice models to distribution–based heterogeneity frameworks while accommodating dependence across alternatives.

Applications to international trade, economic geography, and assignment models illustrate how the equivalence results help to address longstanding modeling challenges. In each case, they yield transparent expressions for both allocation shares and post–selection distributions, enabling comparative statics that link fundamental parameters to observable economic outcomes in closed form. The framework accommodates correlated shocks, a feature that is important in many applications.

Finally, the results point to several directions for future work. First, they provide a transparent basis for identifying structural primitives from allocation data, allowing for richer counterfactual analyses in trade policy, spatial economics, and labor markets. Second, by preserving tractability under correlated shocks, the approach offers a general foundation for embedding multi-dimensional selection mechanisms into heterogeneous-agent models. Taken together, these features make the framework a useful building block for linking aggregate allocations to the distributional outcomes they generate.

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.

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.

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.

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 (2023): "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.

Resnick, Sidney I. (1987): *Extreme Values, Regular Variation, and Point Processes*, Springer.

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

First, I construct the multivariate Fréchet distribution with logistic (Gumbel-Hougaard) dependence from its marginal distributions and dependence structure. Then, I prove Proposition 1.

A1: Multivariate Fréchet with Logistic Dependence

From Proposition 1, the marginal distribution of each shock ε_k is a Fréchet distribution with shape $\theta > 0$ and scale $(Th_k^\theta)^{1/\theta}$. Its cumulative distribution function (CDF) is:

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

Define $u_k \equiv F_k(e_k)$, which are uniformly distributed on $[0,1]$ by the probability integral transform. The dependence structure is the Gumbel-Hougaard copula (also known as logistic in extreme value theory (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 $\prod u_j$, corresponding to independence; as $\rho \to \infty$, it approaches perfect positive dependence.

By Sklar's Theorem, the joint CDF is $F(e_1, \ldots, e_K) = C(F_1(e_1), \ldots, F_K(e_K))$. Substitute the marginal CDFs from (A1) into (A2):

$$\begin{aligned}
-\log u_k &= -\log\left(F_k(e_k)\right) = -\log\left(\exp\left(-Th_k^\theta e_k^{-\theta}\right)\right) = Th_k^\theta e_k^{-\theta}, \\
(-\log u_k)^\rho &= \left(Th_k^\theta e_k^{-\theta}\right)^\rho = T^\rho h_k^{\theta\rho} e_k^{-\theta\rho}, \\
\sum_{j=1}^{K}(-\log u_j)^\rho &= T^\rho \sum_{j=1}^{K}\left(\frac{e_j}{h_j}\right)^{-\theta\rho}.
\end{aligned}$$

Substitute into (A2):

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

This matches the multivariate Fréchet CDF in Proposition 1, confirming the construction.

A2: Proof of Proposition 1

The shocks $\boldsymbol{\varepsilon}=(\varepsilon_1,\ldots,\varepsilon_K)$ follow the joint CDF:

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

where $T>0$, $h_j>0$, $\theta>1$, $\rho\geq 1$. Agents select k^* to maximize:

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

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, equivalent to maximizing the log-payoff $U_k=\log(V_k)$:

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

forming 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:

$$\begin{aligned} F_{\varepsilon_k}(x) &= \lim_{e_j \to \infty, \forall j \neq k} \exp\left(-T \left[\left(\frac{x}{h_k} \right)^{-\theta\rho} + \sum_{j \neq k} \left(\frac{e_j}{h_j} \right)^{-\theta\rho} \right]^{1/\rho} \right) \\ &= \exp\left(-T \left(\frac{x}{h_k} \right)^{-\theta} \right) = \exp\left(-T h_k^{\theta} x^{-\theta} \right), \end{aligned}$$

a Fréchet distribution with shape θ and scale $(T h_k^{\theta})^{1/\theta}$. For $\tilde{\epsilon}_k = \omega \log(\varepsilon_k)$:

$$\begin{aligned} F_{\tilde{\epsilon}_k}(y) &= \Pr\left(\varepsilon_k \leq \exp\left(\frac{y}{\omega} \right) \right) = \exp\left(-T h_k^{\theta} \exp\left(-\frac{\theta y}{\omega} \right) \right) \\ &= \exp\left(-\exp\left(\log(T h_k^{\theta}) - \frac{\theta y}{\omega} \right) \right), \end{aligned}$$

a Gumbel distribution with location $\psi_k = \frac{\omega}{\theta} \log(T h_k^{\theta})$ and scale $\beta = \frac{\omega}{\theta}$.

A2.2: Derivation of Choice Probabilities

Standardize the shocks: $\eta_k = (\tilde{\epsilon}_k - \psi_k)/\beta$, so $\tilde{\epsilon}_k = \psi_k + \beta\eta_k$. The utility is $U_k = v_k + \psi_k + \beta\eta_k$, and maximizing U_k is equivalent to:

$$k^* = \arg\max_k \left(u_k + \eta_k \right), \quad u_k = \frac{v_k + \psi_k}{\beta}.$$

The joint CDF of $\boldsymbol{\eta}$ is derived from $F_{\boldsymbol{\varepsilon}}$:

$$\varepsilon_j = \exp\left(\frac{\psi_j + \beta\eta_j}{\omega} \right), \quad \psi_j = \frac{\omega}{\theta} \log(T h_j^{\theta}), \quad \beta = \frac{\omega}{\theta}.$$

Substitute into $F_{\boldsymbol{\varepsilon}}$:

$$\begin{aligned} \left(\frac{\varepsilon_j}{h_j} \right)^{-\theta\rho} &= \left(\frac{\exp((\psi_j + \beta\eta_j)/\omega)}{h_j} \right)^{-\theta\rho} = h_j^{\theta\rho} \exp\left(-\frac{\theta\rho}{\omega} (\psi_j + \beta\eta_j) \right) \\ &= h_j^{\theta\rho} \exp\left(-\rho \log(T h_j^{\theta}) - \rho\eta_j \right) = \frac{1}{T^{\rho}} \exp(-\rho\eta_j). \end{aligned}$$

Thus:

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

This is a GEV distribution with generator $G(\mathbf{y}) = \left[\sum_j y_j^{\rho}\right]^{1/\rho}$, corresponding to correlation parameter $1/\rho$ (McFadden, 1978; Train, 2009). The choice probability is:

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

Thus:

$$\mu_k = \frac{(W_k^{\theta/\omega}Th_k^{\theta})^{\rho}}{\sum_j (W_j^{\theta/\omega}Th_j^{\theta})^{\rho}} = \frac{(W_k^{1/\omega}h_k)^{\theta\rho}}{\sum_j (W_j^{1/\omega}h_j)^{\theta\rho}}.$$

This concludes the proof of Proposition 1(i).

A2.3: Derivation of Conditional CDFs

For $k^* = k$, $V_k = \max_j V_j$. Thus:

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

Substitute into the joint CDF:

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

Simplify:

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

With $W = \left[\sum_{j=1}^{K}(W_j^{\theta/\omega}h_j^{\theta})^{\rho}\right]^{1/\rho}$:

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

For ε_k:

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

Define $b_k = W_k^{\theta/\omega}h_k^{\theta}$, so $\mu_k = b_k^{\rho}/W^{\rho}$, $\mu_k^{-1/\rho} = W/b_k$. Then:

$$WW_k^{-\theta/\omega} = W/(b_k/h_k^{\theta}) = h_k^{\theta}\mu_k^{-1/\rho},$$

yielding:

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

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 θ and scale $s = (Th_k^{\theta}\mu_k^{-1/\rho})^{1/\theta} = T^{1/\theta}h_k\mu_k^{-1/(\theta\rho)}$. For $\theta > 1$, the expectation is:

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

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

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

Simplify:

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

Using $\mu_k^{-\omega/(\theta\rho)} = W^{\omega/\theta}/b_k^{\omega/\theta}$, $b_k^{\omega/\theta} = W_k h_k^{\omega}$:

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

This concludes the proof of Proposition 1(iii).

Appendix B: Multivariate Dagum with Clayton Dependence

First, I construct the multivariate Dagum distribution with Clayton dependence from its marginal distributions and dependence structure. Then, I prove Proposition 2.

B1: Multivariate Dagum with Clayton Dependence

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

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

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 expressed as $F(e_1, \ldots, e_K) = C(F_1(e_1), \ldots, F_K(e_K))$, where $F_k(e_k) = \left[1 + \left(\frac{e_k}{d_k}\right)^{-\xi}\right]^{-1/\delta}$ 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 $u_k = F_k(e_k)$. Compute $u_k^{-\delta}$:

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

Thus:

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

Substitute into the copula:

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

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

B2: Proof of Proposition 2

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

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

where $\xi > 0$, $\delta > 0$, $d_j > 0$. 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 with CDF $F_\zeta(z) = \exp(-z^{-\xi})$ for $z > 0$. Let $V \sim \Gamma(1/\delta, 1)$, with PDF $f_V(v) = [\Gamma(1/\delta)]^{-1} v^{1/\delta - 1} e^{-v}$ for $v > 0$.

Define the shocks by

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

Then, conditioning on $V = v$,

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

Integrating over V yields

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

B2.2. Derivation of Choice Probabilities

Maximizing V_k is equivalent to maximizing $V_k^{1/\omega} = W_k^{1/\omega} \varepsilon_k$. With $\varepsilon_k = d_k V^{1/\xi} \zeta_k$, define $\phi_k = W_k^{1/\omega} d_k$. 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

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

This concludes the proof of Proposition 2(i).

B2.3. Derivation of the Conditional CDFs

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

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

Therefore, using $\varepsilon_k = d_k V^{1/\xi}\zeta_k$ and conditioning on V,

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

which is Dagum with scale $d_k\,\mu_k^{-1/\xi}$. As an intermediate step, the conditional CDF for $V_k = W_k\varepsilon_k^{\omega}$ (for $v>0$) is

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

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 $b=d_k\mu_k^{-1/\xi}$. 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 $\xi>1$,

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

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

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

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

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

Hence,

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

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:

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

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	(1,10)	1	(1,3)

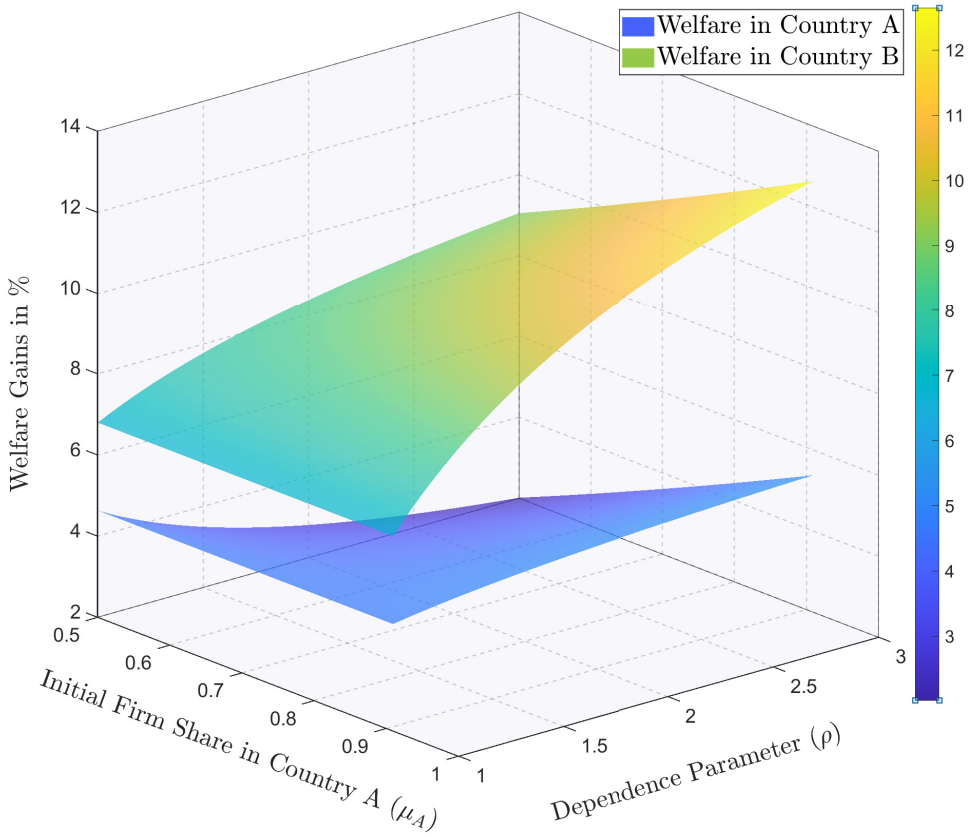
In the benchmark, the two countries are identical in all respects except for the initial firm shares μ_A and μ_B. I fix $\mu_B = 1$ and vary μ_A between 1 and 10. The dependence

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.



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 and the dependence parameter ρ.

Figure C1: Simulated Welfare Gains from Trade

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 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.