Economic Theory
https://doi.org/10.1007/s00199-021-01384-3

RESEARCH ARTICLE

Parallel inverse aggregate demand curves in discrete choice models

Kory Kroft[1,2] · René Leal-Vizcaíno[3] · Matthew J. Notowidigdo[2,4] · Ting Wang[5]

Received: 6 May 2021 / Accepted: 17 August 2021

Abstract
This paper highlights a previously unnoticed property of commonly-used discrete choice models, which is that they feature parallel demand curves. Specifically, we show that in additive random utility models, inverse aggregate demand curves shift in parallel with respect to variety if and only if the random utility shocks follow the Gumbel (Type 1 Extreme Value) distribution. Using results from Extreme Value Theory, we provide conditions for other distributions to generate parallel demands asymptotically, as the number of varieties increases. We establish these results in the benchmark case of symmetric products, illustrate them using numerical simulations and show that they hold in extended versions of the model with correlated tastes and asymmetric products. Lastly, we provide a "proof of concept" of parallel demands as an economic tool by showing how to use parallel demands to identify the change in consumer surplus from an exogenous change in product variety.

We thank Simon Anderson, Costas Arkolakis, Alex Frankel, Marti Mestieri, and Rob Porter for helpful comments. We gratefully acknowledge funding from the Social Sciences and Humanities Research Council (SSHRC). Any opinion, findings, and conclusions or recommendations in this material are those of the authors(s) and do not necessarily reflect the views of the SSHRC.

✉ Kory Kroft
kory.kroft@utoronto.ca

René Leal-Vizcaíno
rlealv@banxico.org.mx

Matthew J. Notowidigdo
noto@chicagobooth.edu

Ting Wang
tingwang@u.northwestern.edu

1 University of Toronto, Toronto, Canada

2 NBER, Cambridge, USA

3 Bank of Mexico, Mexico City, Mexico

4 University of Chicago Booth School of Business, Chicago, USA

5 Northwestern University, Evanston, USA

Published online: 31 August 2021

Keywords Logit model · Discrete choice · Parallel demands · Product variety

JEL Classification D00 · D01 · D11 · D6

1 Introduction

This paper shows that some commonly-used discrete choice models satisfy a parallel inverse aggregate demand property–hereafter referred to as “parallel demands”. Specifically, inverse aggregate demand curves shift vertically in parallel in response to an exogenous change in the number of varieties in a market. In this paper we show that this property holds for the Logit model and some of its Generalized Extreme Value (GEV) distribution variants. In additive random utility models (ARUM) featuring i.i.d. random utility shocks, this means that the random utility shocks are distributed according to the Gumbel (Type 1 Extreme Value) distribution. In fact, we show that the Gumbel distribution is both a necessary and sufficient condition for parallel demands in random utility models. As far as we know, this is a previously-unnoticed feature of this class of models, and as a result this paper focuses on characterizing this property theoretically and showing how it can be used in an economic application to identify the change in consumer surplus associated with an exogenous change in product variety.

In order to develop and build intuition, Sect. 2 considers an additive random utility model with symmetric products and prices and an outside option. Theorem 1 establishes that the Gumbel distribution is necessary and sufficient for parallel demands. Next, we show that for a broad set of distributions of the random utility shock, inverse aggregate demand curves are asymptotically parallel; that is, the aggregate demand curves approach parallel demands as the number of varieties increases (Theorem 2). This result comes directly from Extreme Value Theory (EVT): when the random utility shocks are independent and identically distributed, the distribution of the maximum order statistic converges to a Gumbel distribution for a wide range of distributions. This means that assuming parallel demands may be a useful approximation in many markets featuring a large number of varieties. We illustrate the accuracy of this approximation result using numerical simulations, and we find that convergence happens fairly quickly.

In Sect. 3, we extend the results in Theorems 1 and 2 in several ways. First, we extend the baseline model to allow for correlated tastes, which allows for differential substitutability within the market that has product variety, relative to the outside option. This extension allows us to accommodate the standard Nested Logit model as a special case (Cardell 1997; McFadden 1978). We show that in this extended model the Gumbel distribution is necessary and sufficient for parallel demands (Proposition 1). Second, we extend our results to allow for asymmetric products, since our baseline model assumes symmetric products and prices for simplicity. This extension allows us to accommodate a random utility model with unobserved product heterogeneity as in Berry (1994). The inverse aggregate demand curve is straightforward to define in the symmetric products model. When prices are asymmetric, however, we instead rely on the distribution of the maximal willingness-to-pay for any of the available varieties rather than the aggregate demand curve, and we provide necessary and sufficient conditions for when this distribution shifts in parallel, just as the inverse aggregate demand curve shifts in parallel in our baseline symmetric products model (Theorem 3). Theorem 4 extends the the asymptotic result of Theorem 2 to the asymmetric case.

Lastly, in Sect. 4, we show how to use the parallel demands property to identify the change in consumer surplus from an exogenous change in variety. In our baseline model with symmetric products, graphically the change in consumer surplus is the area between the inverse aggregate demand curves before and after a change in variety. Thus, the change in consumer surplus–what we call the "variety effect"–is the area between these curves. Intuitively, a key feature of the parallel demands property is that identifying the "vertical gap" between the two inverse aggregate demand curves (at two different variety levels) at any one location on the demand curve is sufficient to identify the full area between the two demand curves. Proposition 3 provides a graphical representation of the identification of this vertical gap under parallel demands. It shows that several parameters are sufficient to calculate the variety effect. First, one needs to identify the sensitivity of demand to price, holding variety fixed. Second, one needs to identify the change in price and output in response to an exogenous change in variety. Jointly, under parallel demands, these parameters are sufficient to identify the change in consumer surplus. Thus, the parallel demands property–which has a rigorous microfoundation based on the theoretical results in this paper–can be used to identify the change in consumer surplus stemming from a change in variety.[1] We next extend these results to cover the case of asymmetric products. When products are heterogeneous, we require an additional technical assumption that prices move uniformly after a change in variety. We show that under this assumption, a similar set of parameters identify the variety effect (Proposition 5). Since our approach to identifying changes in consumer surplus is based on aggregate demand, it is perhaps not surprising that we obtain identification by either assuming symmetric products or correlated prices–these are precisely the two scenarios highlighted in Nevo (2011) when discussing identification of aggregate demand and the problem of dimensionality.

This paper contributes to research that explores the theoretical properties of discrete choice models and the theoretical connections between these models and other economic properties. Perhaps most closely related to this paper is Anderson and Bedre-Defolie (2019) who consider a multi-product monopolist who chooses variety and price. They show that for asymmetric Multinominal Logit demand, the inverse demand shifts in parallel when the total variety increases and use this property to show that the monopolist chooses the socially optimal variety for a given total quantity. In terms of Spence's analysis of optimal quality provision (here phrased as product line length), the average and marginal consumer valuations coincide so that the monopolist chooses the right number of products under the Spence criterion, given total output. Another related paper is Anderson et al. (1987), which describes the formal connection between a Logit random utility model and an aggregate demand system featuring a representative agent with Constant Elasticity of Substitution (CES) preferences. This paper provides a formal connection between specific assumptions on the distribution

[1] One might speculate that since the assumed parallel shift in " aggregate demand" in an ARUM model amounts to assuming Logit demand, it is more direct to compute the effect on consumer surplus using the utility function directly. However, our results show that parallel demands are a good approximation for a larger set of distributions of the random utility shock beyond Logit.

of the shocks in additive random utility models and the resulting aggregate inverse demand curve that shifts in parallel with exogenous changes in product variety. Our theoretical approach makes use of Extreme Value Theory, which has been used in an additive random utility context in Gabaix et al. (2016) to show that there can exist high markups in large markets in equilibrium that are insensitive to the degree of competition. Our paper also relates to results in Kroft et al. (2021) who show that the parallel demands property is useful for identifying the love-of-variety from the passthrough of taxes under free entry, and our paper contributes to the literature that studies the foundations and properties of widely used logit model in a discrete choice setting (Breitmoser 2021; Echenique and Saito 2019; Matejka and McKay 2015). Lastly, our application of these theoretical results to identifying the benefits to consumers from greater variety relates to a large theoretical and empirical literature in international trade and industrial organization (see Arkolakis et al. 2008; Berry and Waldfogel 1999; Broda and Weinstein 2006; Dhingra and Morrow 2019; Dixit and Stiglitz 1977; Feenstra 1994; Mankiw and Whinston 1986; Romer 1994; Spence 1976a, b).

2 Parallel demands: symmetric products

In this section, we consider a discrete choice model with symmetric products and derive necessary and sufficient conditions under which inverse market demand curves, evaluated at different levels of product variety, are exactly parallel. Next, we characterize a class of models where parallel demands is likely to be a good approximation.

2.1 Necessary and sufficient conditions

Consider a unit mass population of ex ante identical and independent consumers indexed by i. Consumers either choose to purchase a single product in the market $j \in \{1, \ldots, J\}$, where J is defined as the number of product varieties available, or choose the outside option $j = 0$.

Preferences. The indirect utility of individual i who purchases product j is given by:

$$u_{ij}(y_i, p_j) = \alpha(y_i - p_j) + \delta_j + \varepsilon_{ij} \tag{1}$$

where the scalar α is the marginal utility of income, y_i is consumer $i's$ income, p_j is the price of good j, δ_j is the quality of product j which captures vertical differentation and ε_{ij} is an idiosyncratic match value between consumer i and product j which captures heterogeneity in tastes across consumers and products and the degree of horizontal differentiation. The utility of individual i who chooses the outside option is given by $u_{i0} = \alpha y_i + \varepsilon_{i0}$.

Product-Level Demand. The indirect utility function in equation (1) generates demand for product j, $q_j(p_1, \ldots, p_J) : \mathbb{R}^J_+ \to \mathbb{R}_+$, which we express as

$$q_j(p_1, \ldots, p_J) = \mathbb{P}\left(u_{ij}(y_i, p_j) = \max_{j' \in \{0, \ldots, J\}} u_{ij'}(y_i, p_{j'})\right) \tag{2}$$

Aggregate Demand. We express aggregate demand for all products excluding the outside good, when J varieties are available, as $Q(p_1, \dots, p_J) : \mathbb{R}_+^J \to \mathbb{R}_+$, which takes the form

$$Q(p_1, \dots, p_J) = \sum_{j=1}^{J} q_j(p_1, \dots, p_J) \tag{3}$$

The share of the outside good is $q_0 = 1 - Q$. We now impose the following symmetry assumption.

Assumption 1 We assume that (1) the random utility shocks $(\varepsilon_{ij})_{j=1}^{\infty}$ are continuously, independently, and identically distributed (i.i.d.), and are independent of the distribution of ϵ_{i0}, y_i, and $(\delta_j)_{j=1}^{\infty}$; (2) product qualities are symmetric, $\delta_j = \delta$.

Assumption 1 implies that product prices will be identical in equilibrium ($p_j = p_k$, $\forall j, k \in \{1, \dots, J\}$) under the additional assumption of identical production costs.[2] With symmetric prices, we can express the demand function as $q(p, J) : \mathbb{R}_+^2 \to \mathbb{R}_+$ and the aggregate demand function $Q(p, J) : \mathbb{R}_+^2 \to \mathbb{R}_+$ respectively as

$$\begin{aligned} q(p, J) &= \mathbb{P}\left(u_{ij}(y_i, p) = \max_{j' \in \{0, \dots, J\}} u_{ij'}(y_i, p)\right) \\ Q(p, J) &= J q(p, J) \end{aligned}$$

Next, noting that $Q(p, J)$ is a strictly decreasing function with respect to p, we can invert it to obtain the inverse aggregate demand function $P(Q, J) : \mathbb{R}_+^2 \to \mathbb{R}_+$. We now introduce our definition of parallel demands with symmetric products.

Definition 1 The discrete choice model with symmetric products is said to give rise to parallel demands if for all J_0, $J_1 \neq J_0$, and Q

$$\frac{\partial P}{\partial Q}(Q, J_0) = \frac{\partial P}{\partial Q}(Q, J_1) \tag{4}$$

where $P(Q, J_t), t \in \{0, 1\}$ is the inverse aggregate demand function, and J_0 and J_1 are any numbers of product varieties. An equivalent definition of parallel demands that we will make use of below is $Q(p, J_0) = Q(p + d(J_0, J_1), J_1)$; in other words, there exists some index $d(J_0, J_1)$, such that output is the same when the price is p with J_0 varieties or the price is $p + d(J_0, J_1)$ with J_1 varieties.

We now state our first theorem using Definition 1 and Assumption 1.

Theorem 1 *Suppose that Assumption 1 holds, prices are symmetric and ε_{i0} follows a continuous distribution. Then a necessary and sufficient condition for parallel demands (Definition 1) is that the random utility shocks $(\varepsilon_{ij})_{j=1}^{\infty}$ in equation (1) follow a Gumbel distribution $G(x) = e^{-e^{-\frac{x-\mu}{\beta}}}$ for some location and scale parameters $\mu \in \mathbb{R}$ and $\beta > 0$.*

[2] We do not explicitly model market equilibrium in this paper, but symmetric prices are achieved in equilibrium (Nash in prices) when firms have identical costs as shown in Anderson and Palma (1992).

Proof See Appendix. □

As an illustration, in equation (1), if ε_{i0} is also Gumbel, then this model corresponds to a multinomial Logit model in which there are $J_0 + 1$ products including the outside option. For any $j \in \{1, \ldots, J_0\}$

$$q(p, J_0) = \frac{e^{\delta - \alpha p}}{1 + J_0 e^{\delta - \alpha p}}$$

Aggregate demand is equal to

$$Q(p, J_0) = \frac{J_0 e^{\delta - \alpha p}}{1 + J_0 e^{\delta - \alpha p}}$$

Thus, the inverse aggregate demand curve of the multinomial Logit model is given by

$$P(Q, J_0) = \frac{\delta}{\alpha} + \frac{1}{\alpha} \log J_0 - \frac{1}{\alpha} \log \left(\frac{Q}{1 - Q} \right)$$

We verify that $\frac{\partial P}{\partial Q}(Q, J_0) = -\frac{1}{\alpha}\frac{1}{Q}\frac{1}{1-Q} = \frac{\partial P}{\partial Q}(Q, J_1)$ and so Definition 1 is satisfied. Equivalently, note that $Q(p, J_0) = \frac{J_0 e^{\delta-\alpha p}}{1+J_0 e^{\delta-\alpha p}} = \frac{J_1 e^{\delta-\alpha(p+d(J_0,J_1))}}{1+J_1 e^{\delta-\alpha(p+d(J_0,J_1))}} = Q(p + d(J_0, J_1), J_1)$ for $d(J_0, J_1) = \frac{1}{\alpha} log \left(\frac{J_1}{J_0}\right)$.

2.2 Asymptotic approximation as *J* grows large

The previous section showed that Gumbel random utility shocks are both necessary and sufficient for parallel demands in the case of symmetric products. Using Extreme Value Theory, we now show that there is a large class of random utility shocks beyond Gumbel that admit parallel demands asymptotically (as J grows large). The additive random utility models in this class have in common that the distribution of the maxima of the shocks is asymptotically Gumbel, which implies that the inverse aggregate demand curves are asymptotically parallel. We now define a class of models that admit this asymptotic approximation, and we provide a sufficient condition to show that a given additive random utility model is in this class.

Definition 2 Let (ε_{ij}) be i.i.d. distributed according to a continuous CDF F. Following Resnick (1987), F is in the domain of attraction of the Gumbel distribution if and only if there exist sequences (a_n, b_n) of real numbers such that $F^n(a_n x + b_n) \to G(x)$ for all x, where $G(x) = e^{-e^{-x}}$ is the standard Gumbel distribution.

Lemma 1 *Let x_0 be the supremum of the support of a CDF F that is twice continuously differentiable. If F satisfies $\lim_{x \to x_0} \frac{F''(x)(1-F(x))}{F'^2} = -1$ then F is in the domain of attraction of the Gumbel distribution.*

See Resnick (1987) for a proof of Lemma 1 and a full characterization of the domain of attraction of the Gumbel distribution. Although the characterization of the domain of attraction is outside the scope of the paper, it is worth mentioning the important result in statistics (the Fisher–Tippett–Gnedenko theorem) that plays a role akin to the Central Limit Theorem for Extreme Value theory. The result states that for a sequence of i.i.d. random variables X_i, letting $M_n = \max\left\{X_{1,} X_{2,} \ldots, X_n\right\}$ then if a sequence of real numbers $(a_{n,} b_n)$ exists such that $\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n}\right) = F(x)$, where F is a non-degenerate CDF, then F is either Gumbel, Fréchet or Weibull. A useful intuition is that if the tails of the random utility shocks (ε_{ij}) are very "thin" the resulting converging distribution is Weibull, while if they are "heavy" the distribution of the maxima converges to Fréchet. Gumbel is the intermediate case that gives rise to parallel demands. For our purposes Lemma 1 is enough to show that some common distributions fall into the Gumbel domain of attraction.

The domain of attraction of the Gumbel distribution includes the Normal $N(\mu, \eta^2)$, Exponential, Lognormal, Gamma, Chi-square, and Weibull distributions, but does not include heavy-tailed distributions like the Cauchy, Fréchet, Pareto or Student distributions nor does it include short-tailed distributions like the Beta and Uniform distributions. The next theorem shows that inverse aggregate demands become "asymptotically" parallel as variety increases, for any additive random utility model with shocks in the Gumbel domain of attraction.

Theorem 2 *Let the random utility shocks (ε_{ij}) be i.i.d. and distributed according to F in the domain of attraction of the Gumbel distribution. Then for any $\epsilon > 0$, there exists large enough J_0 such that for all $J_1 > J_0$, there exists $d(J_0, J_1)$ such that for all $p \in \mathbb{R}_+$ we have*

$$|Q(p, J_0) - Q(p + d(J_0, J_1), J_1)| < \epsilon.$$

Therefore the inverse demands are approximately parallel $P(Q, J_1) \approx P(Q, J_0) + d(J_0, J_1)$ for all Q, for large enough J_0 and J_1.

Proof See Appendix. □

Later in the paper we assess the approximation result in Theorem 2 by numerically simulating different additive random utility models and considering the effect of an exogenous change in the number of varieties on consumer surplus, using the exact formulas for consumer surplus in additive random utility models and using a reduced-form approach that assumes demands are parallel.

3 Generalizations and extensions: correlated tastes and asymmetric products

In this section, we generalize the model in 2.1 to a Logit model with correlated tastes, and we also consider a model with asymmetric products. While preserving the Extreme Value distribution of consumers' tastes within the inside market, the model with correlated tastes in many cases better captures the substitution patterns of products by allowing different substitutability within the variety market relative to the outside option and correlated tastes across products within the variety market. We show that in this model, we continue to obtain parallel demands when the distribution of random utility shocks satisfies the necessary and sufficient condition in Theorem 1. When we extend to asymmetric products, we are also able to obtain analogous results.

3.1 Logit model with correlated tastes

Similar to the multinomial Logit model, we consider a population of statistically identical and independent consumers indexed by i of mass unity who choose to purchase a single product $j \in \{1, \ldots, J\}$ or the outside option $j = 0$. We extend the baseline model to allow preferences across products to be correlated within individuals.

Preferences. The indirect utility of individual i who purchases product j is given by:

$$u_{ij}(y_i, p_j) = \alpha(y_i - p_j) + \delta_j + (1 - \sigma)\nu_i + \sigma\varepsilon_{ij} \tag{5}$$

where $(1 - \sigma)\nu_i + \sigma\varepsilon_{ij}$ is the idiosyncratic match value between consumer i and product j, which captures heterogeneity in tastes across consumers and products, and correlation in tastes across products. When $\sigma = 1$ and ε_{ij} follows the Gumbel distribution, we obtain the Logit model and when $\sigma = 0$, consumer tastes for all products in the inside market are perfectly correlated. Thus, the parameter σ captures the degree of correlation in consumer preferences across products of the inside market. The utility of individual i who chooses the outside option is given by $u_{i0} = \alpha y_i + \varepsilon_{i0}$. Similar to the Logit model, we make the following assumption.[3]

Assumption 2 We assume that (1) for $j \neq 0$, the random utility shocks (ε_{ij}), $j = 1 \ldots J$ are continuously, independently and identically distributed (i.i.d.) and independent of ε_{i0}, y_i, ν_i, and δ_j, $j = 1 \ldots J$, but we allow ε_{i0} to be correlated with ν_i; (2) product qualities and prices are symmetric $\delta_j = \delta$ and $p_j = p$. The next proposition extends the result in Theorem 1 to cover correlated tastes.

Proposition 1 *Suppose that Assumption 2 holds. Then, a necessary and sufficient condition for parallel demands (Definition 1) is that the random utility shocks* (ε_{ij}) *in equation (5) follow a Gumbel distribution.*

Proof See Appendix. □

The logic of the proof is the following: since the term $(1 - \sigma)\nu_i$ does not vary across products, we can use a location normalization for utility and move this term into the outside option. Then, we can apply the same arguments in the proof of Theorem 1. This explains why it is not necessary to impose a specific functional form assumption on the distribution for $(1 - \sigma)\nu_i$. While the proposition does not require a specific distribution, we can use the Nested Logit model as a special case of this model to illustrate this result.

[3] As in Sect. 2, we do not model the market equilibrium. Instead, we assume symmetric prices directly, which would be achieved as an equilibrium outcome in a Nested Logit demand model when firms have identical costs, following Anderson and Palma (1992).

In the Nested Logit model, the random utility shocks (ε_{ij}) in equation (5) are drawn from the Gumbel distribution, and $(1-\sigma)\nu_i$ has the distribution derived in Cardell (1997). In our case, there are two nests: one which includes $j=1,\ldots,J$, and another nest which includes only the outside option $j=0$.[4] In the Nested Logit model, product demand is:

$$q(p,J)=\frac{J^{\sigma-1}e^{\delta-\alpha p}}{1+J^{\sigma}e^{\delta-\alpha p}}.$$

In turn, aggregate demand is equal to:

$$Q(p,J)=\frac{J^{\sigma}e^{\delta-\alpha p}}{1+J^{\sigma}e^{\delta-\alpha p}}.$$

Inverting aggregate demand, the inverse aggregate demand curve is given by:

$$P(Q,J)=\frac{\delta}{\alpha}+\frac{\sigma}{\alpha}\log J-\frac{1}{\alpha}\log\left(\frac{Q}{1-Q}\right).$$

Thus, we see that the Nested Logit model (like the symmetric products Logit model above) satisfies Definition 1 since $\frac{\partial P}{\partial Q}(Q,J_0)=-\frac{1}{\alpha}\frac{1}{Q}\frac{1}{1-Q}=\frac{\partial P}{\partial Q}(Q,J_1)$ or equivalently $d(J_0,J_1)=\frac{\sigma}{\alpha}log\left(\frac{J_1}{J_0}\right)$.

3.2 Asymmetric products

Assumptions 1 and 2 impose symmetric products and prices, which leads to clean results but may limit the generality of the model. We now extend our results by considering asymmetric products so that $\delta_j\neq\delta_k$ and $p_j\neq p_k$ for $j\neq k$, and we continue to allow for an outside option as in the previous sections. In order to characterize parallel demands in this general model, we impose a technical assumption that we use in Theorem 3 below.

Assumption 3 We assume that (1) for $j\neq 0$, the random utility shocks $(\varepsilon_{ij})_{j=1}^{\infty}$ are continuously, independently and identically distributed (i.i.d.) and independent of ε_{i0} which has a continuous distribution; (2) $(\delta_j)_{j=1}^{\infty}$ is a deterministic sequence of real numbers, and there exists a real number $K>0$ such that all the quality parameters are bounded: $\delta_j\in[0,K]$ for all j.

In the case of symmetric products and prices considered above, we were able to invert the aggregate demand since there was a mapping from aggregate quantity to a single (uniform) price at a given level of product variety. This inverse aggregate demand curve corresponded to the distribution across consumers of their maximum willingness-to-pay (WTP) for any level of product variety. When prices and products

[4] See Cardell (1997) for the class of distributions, termed $C(\cdot)$ distributions, which makes the combined idiosyncratic shocks distributed Type I Extreme Value, and thus allows us to write the demand function in a closed form.

are asymmetric it is no longer straightforward to characterize the inverse aggregate demand curve. Thus, with asymmetric products we instead state our results in terms of the distribution of WTP rather than in terms of aggregate demand. In particular we study the distribution of the random variable $\max_{j\in\{0,\ldots,J\}} wtp_{ij}$, where $wtp_{ij}(\delta_j) \equiv \frac{\delta_j+\varepsilon_{ij}-\varepsilon_{i0}}{\alpha}$. We now introduce the definition of parallel shifts in WTP.

Definition 3 Let $WTP_i(J) \equiv \max_{j\in\{1,\ldots,J\}} wtp_{ij}(\delta_j)$. The discrete choice model in equation with asymmetric products is said to give rise to parallel shifts in willingness-to-pay (WTP) if for all $J_1 \neq J_0$, there exists $d(J_0, J_1) \in \mathbb{R}$, such that for all $x \in \mathbb{R}$:

$$\mathbb{P}\left(WTP_i(J_0) \leq x\right) = \mathbb{P}\left(WTP_i(J_1) \leq x + d(J_0, J_1)\right).$$

In particular, when $J_1 > J_0$, if consumers value variety, then we expect that $d(J_0, J_1) > 0$.

With this definition of parallel WTP shifts, we can now state the theorem that generalizes Theorem 1 to the case of asymmetric products.

Theorem 3 *A discrete choice model with asymmetric products satisfying Assumption 3 gives rise to parallel shifts in WTP (Definition 3) for all models satisfying Assumption 3 if and only if the random utility shocks $(\varepsilon_{ij})_{j=1}^{\infty}$ follow a Gumbel distribution (independently of the distribution of ε_{i0}).*

Proof See Appendix. □

Note that Theorem 3 lets us reinterpret Definition 3 in terms of aggregate demand. Assuming Gumbel shocks, Theorem 3 implies that we also get parallel shifts in consumer surplus $\max_{j\in\{1,\ldots,J\}} wtp_{ij}(\delta_j - \alpha p_j)$ (by substituting $\hat{\delta}_j = \delta_j - \alpha p_j$) and so the shift $d(J_0, J_1)$ can be seen as either a horizontal shift in the CDF of $WTP_i(J_0)$ or a vertical shift of the following function:

$$Q(s) \equiv Q(p_1 + s, \ldots, p_J + s, J_0) = \mathbb{P}\left(\max_{j\in\{1,\ldots,J\}} wtp_{ij}(\delta_j - \alpha\left(p_j + s\right)) \geq 0\right)$$

which maps aggregate demand as a function of the price index s. Lastly, as in the symmetric case, we can also use Extreme Value Theory to show that there is a larger class of models that admit parallel WTP asymptotically.

Theorem 4 *Suppose Assumption 3 holds. Let $(\epsilon_{ij})_{j=1}^{\infty}$ be i.i.d. and distributed with CDF F in the domain of attraction of the Gumbel distribution. Furthermore, assume there exists (α_n, β_n) and a nondegenerate CDF H such that $\Pi_{j=1}^{n} F(\alpha_n x + \beta_n - \delta_j) \to H(x)$ for all x.[5] Then for any $\epsilon > 0$, there exists large enough J_0 such that for all $J_1 > J_0$, there exists $d(J_0, J_1)$ such that for all $x \in \mathbb{R}$*

$$\left|\mathbb{P}\left(WTP_i(J_0) \leq x\right) - \mathbb{P}\left(WTP_i(J_1) \leq x + d(J_0, J_1)\right)\right| < \epsilon.$$

[5] This second assumption is satisfied automatically for sequences where $(\delta_j)_{j=1}^{\infty}$ is non-increasing or non-decreasing. The condition may also be violated for alternating sequences. A counterexample can be constructed by taking $\delta = 0$ or $\delta = K$ for alternate periods of increasing length. Therefore this assumption constrains the variation in the vertical differentiation parameter of the new varieties that can enter the market.

Proof See Appendix.

The technical result in Theorem 4, extends Theorem 2 to the maxima of non i.i.d. random variables. In the mathematics and statistics literature, it has proven difficult to extend the Fisher–Tippett–Gnedenko theorem to non i.i.d sequences of random variables. In particular, Kreinovich et al. (2015) show the impossibility of a simple generalization of the Fisher–Tippett–Gnedenko theorem when random variables are not identically distributed and contrast it to the Central Limit Theorem where this is possible. In our particular case, we are able to show that when the sequence of random variables is composed of mean shifts of the same CDF in the domain of attraction of the Gumbel distribution, the asymptotic theorem obtains.

The results thus far demonstrate a connection between discrete choice models featuring Gumbel-type preferences and parallel demands. The next section provides an example where parallel demands are valuable as an economic tool.

4 Parallel demands as an economic tool: identification of the variety effect

In this section, we show how to use parallel demands to identify the change in consumer surplus from an exogenous change in variety. Measuring the change in consumer surplus due to a change in variety has been studied in many branches of economics, including international trade and industrial organization (see Berry and Waldfogel 1999; Broda and Weinstein 2006; Dhingra and Morrow 2019; Feenstra 1994). We begin with the standard definition of consumer surplus and derive the variety effect in the general case. When there are new varieties introduced into the market, the variety effect depends on all of the demands for the new goods. When there are many differentiated products, as is typically the case in economic applications, this is a high dimension problem with a large number of parameters to be estimated and we need to impose some form of dimension reduction. We consider two complementary approaches: symmetry and aggregation. First, we consider a symmetric product environment, as is typically assumed in models in macro and trade, and show that we can characterize the variety effect as the area between two inverse aggregate demand curves. Second, we allow for heterogeneity in demands and prices but assume that prices are correlated which allows us to aggregate; specifically, we assume that prices shift by the same amount after the introduction of new varieties. This result relates to Hicks (1936) that in order to aggregate goods into commodities, prices of the goods must be highly correlated. The advantage of aggregation is that it permits one to be more flexible on functional forms without having to specify underlying preferences. The disadvantage is that prices may not be highly correlated.[6]

[6] See discussion in Nevo (2011) for the dimensionality problem and alternative approaches to identifying demand.

4.1 Variety effect

Consider the general discrete choice model in Sect. 3.2 with J asymmetric products and prices which are denoted by the vector $\mathbf{p_J}$. There are no income effects which means that consumer surplus is a valid measure of welfare and we can avoid the problem of path dependence of price changes.

Definition 4 Let $Q_J(\mathbf{p})$ be the aggregate demand when there are J differentiated products and prices are given by $\mathbf{p_J} = (p_1, p_2, \ldots, p_J)$. In this case, consumer surplus is defined:

$$CS(\mathbf{p_J}, J) = \int_0^\infty Q_J(\mathbf{p_J} + s\mathbf{1}_J)ds \tag{6}$$

When new varieties are introduced into the market, there are two effects on consumer surplus. First, there is a "price effect" that arises since market prices may change when firms enter or exit the market. Second, there is a "variety effect" which captures how much a new variety increases consumer surplus, holding prices constant. In this section, we focus on the "variety effect" which we define as follows.

Definition 5 Let $\mathbf{p_{J_0}} = (p_1, p_2, \ldots, p_{J_0})$ and $\mathbf{p_{J_1}} = (\mathbf{p_{J_0}}, p_{J_0+1}, \ldots, p_{J_1})$. The "variety effect" when the number of products goes from J_0 to J_1 (with $J_1 > J_0$) is defined as:

$$\Lambda = \int_0^\infty Q_{J_1}(\mathbf{p_{J_1}} + s\mathbf{1}_{J_1})ds - \int_0^\infty Q_{J_0}(\mathbf{p_{J_0}} + s\mathbf{1}_{J_0})ds \tag{7}$$

From Definition 5 we see identifying the variety effect requires identification of aggregate demand before and after the change in varieties. In order to make the problem more tractable, we focus on two special cases: symmetric products and asymmetric products with correlated prices (aggregation).

4.2 Symmetry

When all potentially existing products are symmetric, in the equilibium we have $p_j = p_k, \forall j, k$. Then we can use the definitions and foundations laid in Sect. 2 to simplify the expressions of consumer surplus and the variety effect as follows. First, consumer surplus is defined as the integral of aggregate demand:

$$CS(p, J) = \int_p^\infty Q(s, J)ds \tag{8}$$

Next, using the inverse demand $P(Q, J)$ we can adapt Definition 5 for the variety effect when variety in the market changes from J_0 to J_1 to:

$$\Lambda = \int_0^Q \big(P(s, J_1) - P(s, J_0)\big)ds \tag{9}$$

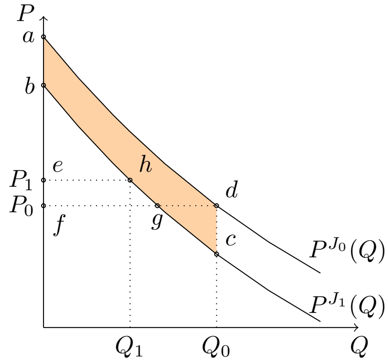

Fig. 1 Variety Effect. Notes: This figure shows the result of a decrease in variety (from J_0 to J_1). The shaded area *abcd* between the two demand curves represents the variety effect

where instead of holding fixed prices, we are holding fixed quantity, as this will prove more useful in this section. The next result shows that, the variety effect can be calculated exactly in a simple form when we assume parallel demands.

Proposition 2 *Starting from equilibrium quantity Q_0 and price p_0, under the assumption of parallel demands (Definition 1), when variety changes from J_0 to J_1, the variety effect can be equivalently expressed as:*

$$\Lambda = Q_0 * d(J_0, J_1) \tag{10}$$

where $d(J_0, J_1)$ is such that $P(Q, J_0) + d(J_0, J_1) = P(Q, J_1)$.

Proof See Appendix. □

The price effect and variety effect are illustrated in Fig. 1 which considers a reduction in product variety in the market from J_0 to J_1 . The price effect is represented by the area $efgh$ and the variety effect is given by the area $abcd$, so that $-\Delta CS = abcd - cdg + efgh$, where cdg is an adjustment that is second-order relative to $\Delta Q * \Delta J$. Intuitively, when the number of varieties is reduced, some consumers will no longer be able to purchase their most preferred option. Thus, the maximum willingness-to-pay for purchasing an inside good will be lower for these consumers. This is represented as a downward shift in the inverse aggregate demand curve. The area between the inverse aggregate demand curves $abcd$ before and after the change in variety up to initial quantity Q_0 corresponds exactly to the variety effect.

We can now state our next Proposition, which uses Definition 1.

Proposition 3 *Denote the equilibrium quantity Q_0 and market price p_0 when initial variety is J_0. Consider an exogenous increase in varieties from J_0 to J_1 and denote the new equilibrium quantity Q_1 and market price p_1. Under the assumption of parallel demands (Definition 1):*

$$d(J_0, J_1) = p_1 - P(Q_1, J_0) = \left(\frac{\frac{dp}{dJ}}{\frac{dQ}{dJ}} - \frac{\partial P(Q, J)}{\partial Q} \right) \frac{dQ}{dJ} \triangle J + O((\triangle J)^2) \tag{11}$$

where $\frac{\partial P(Q,J)}{\partial Q}$ *denotes the slope of inverse demand when variety* J *is held fixed and* $\frac{dp}{dJ}/\frac{dQ}{dJ}$ *denotes the slope of inverse demand when* J *is variable.*

Proof See Appendix. □

When variety changes from J_0 to J_1, prices change from p_0 to p_1. However, this is not sufficient to recover $d(J_0, J_1)$. This is because the counterfactual price $P(Q_1, J_0)$ is not directly observable since it depends on the market price that would prevail at the final level of output but on the original demand curve. To see how to recover an expression for $d(J_0, J_1)$, we note from Fig. 1 that it must satisfy the following relationship $Q(p_1, J_1) = Q(p_1 - d(J_0, J_1), J_0)$. Thus, we can identify $d(J_0, J_1)$ as follows:

$$\begin{aligned}
dQ &= Q(p_1, J_1) - Q(p_0, J_0)\\
dQ &= Q(p_1 - d(J_0, J_1), J_0) - Q(p_0, J_0)\\
dQ &\approx \frac{\partial Q(p, J)}{\partial p}|_{p=p_0}(-d(J_0, J_1) + p_1 - p_0)\\
dQ &= \left.\frac{\partial Q}{\partial p}\right|_{p=p_0}\left(-d(J_0, J_1) + \frac{dP}{dQ}dQ\right)
\end{aligned}$$

The first equality holds by definition. The second equality holds by Definition 1. The third approximation holds by doing a Taylor expansion of $Q(p, J)$ around p_0. The fourth equality holds by definition. Rearranging and solving for $d(J_0, J_1)$ yields:

$$d(J_0, J_1) \approx dp - \frac{\partial P}{\partial Q}dQ$$

In economic terms, $d(J_0, J_1)$ can be interpreted as the reduction in the willingness-to-pay for the marginal unit. Under Definition 1, it can further be interpreted as the change in willingness-to-pay for inframarginal units. In order to identify $d(J_0, J_1)$, two causal effects are required. First, one requires the effecs of an exogenous change in variety on prices $\left(\frac{dp}{dJ}\right)$ and output $\left(\frac{dQ}{dJ}\right)$. Second, one requires the effect of prices on demand, holding variety fixed, $\frac{\partial P}{\partial Q}$. Intuitively, when we multiply $\frac{\partial P}{\partial Q}\Delta Q$ we are implicitly calculating the counterfactual price that would hold when J_0 varieties are available and quantity is adjusted to Q_1.

4.3 Numerical simulations

From Theorem 1 we know that when we have preferences in the form of (1) and the random utility shocks follow the Gumbel distribution, we can apply the parallel demands and compute $\frac{\partial P}{\partial J}(Q', J)$ for any Q' on the support of the aggregate demand function. This saves us from integrating over the whole support. Moreover, from Theorem 2 we have that as long as the random utility shocks (ε_{ij}) are distributed according to F in the domain of attraction of the Gumbel distribution, for any large

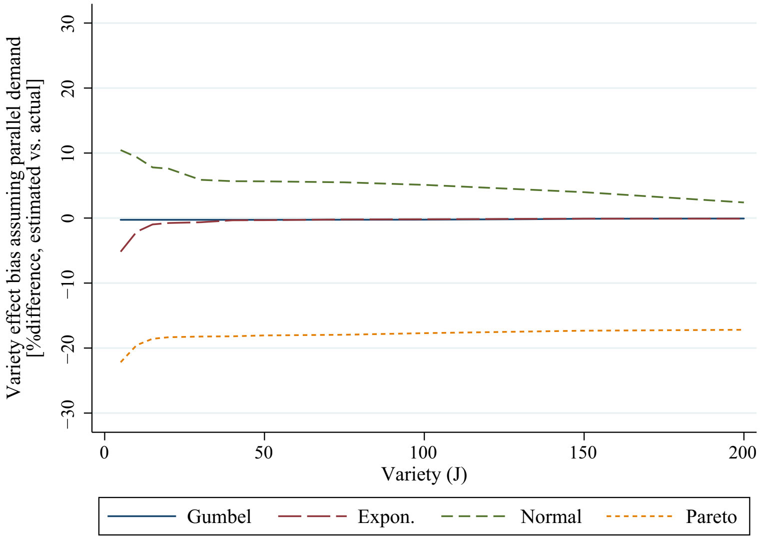

Fig. 2 Approximate Parallel Demand Curves. Notes: This figure reports results from numerical simulations that are designed to evaluate the quality of the key approximation theorem (Theorem 2) in the main text. By simulating simple discrete choice models under different assumptions about the distribution of the i.i.d. error terms and increasing the number of varieties in the market, we calculate the (exact) variety effect numerically and compare it to the variety effect that we would infer from assuming parallel demands. Consistent with the result of Theorem 2, for distributions that satisfy assumptions of theorem, as J increases, the bias in the variety effect from assuming parallel demands approaches zero

enough varieties, we have parallel demands as an approximation. In this subsection we show these results in Monte Carlo simulations. We assess the parallel demands assumption by simulating a model of a large number of consumers with utility over products given by equation (1). We choose $\alpha = 1$ and $y = 1$ in the simulation, and we consider four different shock distributions (Gumbel, Normal, Exponential, and Pareto). For each distribution, we consider a hypothetical 20 percent increase in the number of products (from an initial value of J), and we compute the exact change in consumer surplus resulting from this change in variety by numerically integrating the increase in consumer surplus across each consumer. Then, we compute the change in consumer surplus implied by assuming parallel demands following equations (10) and (11).

The results in Fig. 2 show that the bias that arises from assuming parallel demands is a function of the number of varieties in the market, measuring the bias as the difference between the estimated and the exact change in consumer surplus. The benchmark distribution is Gumbel, where we know from Theorem 1 that the demand curves are exactly parallel, and therefore the bias is 0 for all initial values of J. For both the Normal and Exponential distributions, we find that the bias is small in magnitude and converges to 0 fairly quickly as the number of varieties increase. On the other hand, with a Pareto distribution, there is a bias of roughly 20 percent, which does not vanish as varieties increase. In this case, the change in consumer surplus from assuming parallel demands is a lower bound on the true change in consumer surplus, and it does not converge to 0 because the Pareto distribution is not in the domain of attraction of the Gumbel distribution.

4.4 Aggregation

The previous results focus on the special case symmetric products, which allows for a clear graphical representation since the inverse aggregate demand curve can be defined for a uniform (symmetric) price. We now consider the case of asymmetric products. The main objective in what follows is to give plausible and parsimonious sufficient conditions to identify the variety effect using reduced-form methods based on local information.

We first note that under the assumption of parallel shifts in WTP (Definition 3), there exists some price index $d(J_0, J_1)$ such that $Q_{J_1}(\mathbf{p_{J_1}} + s\mathbf{1}_{\boldsymbol{J_1}}) = Q_{J_0}(\mathbf{p_{J_0}} + (s - d(J_0, J_1))\mathbf{1}_{\boldsymbol{J_0}})$ for all $s \in \mathbb{R}$.[7] In other words, increase prices starting from $\mathbf{p_{J_0}}$ by some constant amount $d = d(J_0, J_1)$ until total quantity demanded equals quantity demanded when there are J_1 products in the market. Under this assumption, it follows that the variety effect can be expressed as:

$$\Lambda = \int_0^d Q_{J_0}(\mathbf{p_{J_0}} + (s - d)\mathbf{1}_{\boldsymbol{J_0}})ds.$$

Next, by the mean value theorem for integrals, there exists $d' \in [0, d]$ such that

$$\Lambda = Q(\mathbf{p_{J_0}} - d'\mathbf{1}_{\boldsymbol{J_0}}) * d.$$

This leads to the following result.

Proposition 4 *Under the assumption of parallel shifts in WTP (Definition 3), when variety changes from J_0 to J_1 , there exists $d' \in [0, d(J_0, J_1)]$ such that*

$$\Lambda = Q_{J_0}(\mathbf{p_{J_0}} - d'\mathbf{1}_{\boldsymbol{J_0}}) * d(J_0, J_1). \tag{12}$$

All that remains is to develop a method to identify $d(J_0, J_1)$. To do this, we introduce an additional technical assumption.

Assumption 4 The prices of the existing products in the market ($j = 1, \ldots, J_0$) shift by the same amount after the introduction of new varieties, i.e. $p_j^1 - p_j^0 = p_k^1 - p_k^0$ for all products j, k available in both periods of time.

With this assumption in hand, we can now state our main result for asymmetric products.

Proposition 5 *Suppose that the assumption of parallel shifts in WTP (Definition 3) and Assumption 4 holds. Let the post-entry equilibrium prices be $\mathbf{p}_{J_1}$ and define $\Delta P \equiv$*

[7] This is related to the price index in Feenstra (1994). However, in Feenstra (1994), the price index is defined as the (common) price change that would have to occur when there are J_0 goods in the market in order to give the same *utility* as when there are J_1 goods.

$\rho \in \mathbb{R}$ to be the change in price of any of the existing products before and after entry of new varieties. Letting $\Delta Q = Q_{J_1}(\mathbf{p_{J_1}}) - Q_{J_0}(\mathbf{p_{J_0}})$, we have:

$$d(J_0, J_1) = \left(\frac{\Delta P}{\Delta Q} - \left.\frac{dP}{dQ_{J_0}}\right|_{J_0} \right) \Delta Q + O\left((\rho - d)^2\right) \tag{13}$$

where $\left.\frac{dP}{dQ_{J_0}}\right|_{J_0} = \left.\left(\frac{dQ_{J_0}(\mathbf{p_{J_0}}+t\mathbf{1_{J_0}})}{dt}\right)^{-1}\right|_{t=0}$.

Proof See Appendix. □

Several features of Proposition 5 are worth highlighting. First, observe that the key step for the Proposition to hold is to be able to find a ρ and d such that $Q_{J_1}(\mathbf{p_{J_1}}) = Q_{J_0}(\mathbf{p_{J_0}} + (\rho - d)\mathbf{1_{J_0}})$. This requires both that all prices adjust uniformly after the introduction of the new varieties (Assumption 4) and that aggregate demands shift in parallel (Definition 3). Restricting prices to adjust in the same direction $\mathbf{1_{J_0}}$ as the vertical shift d allows us to identify d by a simple application of the Taylor approximation theorem.

Second, we interpret the directional derivative $\left.\frac{dQ_{J_0}}{dP}\right|_{J_0} = \frac{dQ(\mathbf{p_{J_0}}+t\mathbf{1_{J_0}})}{dt} = \sum_{j=1}^{J_0} \frac{\partial Q_{J_0}}{\partial p_j}$ as the short-run slope of aggregate demand in the direction of uniform price changes, that connects the interpretation of (13) with equation (11) in the symmetric model. Furthermore, if we observe the change in aggregate demand Q_{J_0} when all prices are increased simultaneously, one does not need to identify each partial derivative separately; instead it is sufficient to identify $\left.\frac{dQ_{J_0}}{dP}\right|_{J_0}$.

5 Conclusion

This paper highlights a previously-unnoticed feature of a class of discrete choice models, which is that they feature parallel demand curves. Specifically, we show that in additive random utility models, inverse aggregate demand curves shift in parallel with respect to variety if and only if the random utility shocks follow the Gumbel distribution. While it may seem that the parallel demands property is a special case, our theoretical results suggest instead that parallel demands are a general property of many discrete choice models. Specifically, using results from Extreme Value Theory, we provide conditions for other distributions to generate parallel demand asymptotically, as the number of varieties increases. We illustrate these results using numerical simulations and extend them to cover correlated tastes and asymmetric products.

Given the generality of our theoretical results, we provide an application and show that parallel demands are useful to identifiy the change in consumer surplus from a change in variety. In this application, parallel demands provide a straightforward identification approach–intuitively, identifiying the "vertical gap" at one point in the aggregate demand curve is sufficient for identifying the entire area between the inverse aggregate demand curves before and after the change in variety. Because of this, we view the parallel demands property as a tool that can potentially be used for both producing theoretical results on the value of variety (which can be an input into theoretical analysis of whether the equilibrium level of variety is socially optimal) as well as a tool for empirical work, where the parallel demands assumption may be used as an alternative "reduced-form" identification approach (instead of relying on specific structural models of consumer demand for identification).

We conclude by speculating that parallel demands may also be a useful property when studying other economic questions. Discrete choice models are widespread in economics, and our theoretical results may therefore be useful in other economic settings, such as the choice of neighborhood (Bayer et al. 2007; McFadden 1978), occupation (Hsieh et al. 2013), firm (Card et al. 2018; Chan et al. 2019; Lamadon et al. 2020), and school (Dinerstein and Smith 2014). In all of these settings, as long as the parallel demands assumption holds, the welfare effects corresponding to changes in the number of available choices (or "varieties") may be calculated using the approach described in this paper.

Appendix

Proofs of Claims, Propositions, and Theorems

Proof of Theorem 1

Proof Assuming symmetric prices the inverse demands when there are J_0 and J_1 varieties are parallel if and only if there exists a $d(J_0, J_1)$ such that for all p then $Q(p, J) = Q(p + d(J_0, J_1), J_1)$; that is

$$\mathbb{P}(\varepsilon_{0m} < \delta - \alpha p + \max_{1 \leq j \leq J_0} \varepsilon_j) = \mathbb{P}(\varepsilon_{0m} < \delta - \alpha(p + d(J_0, J_1)) + \max_{1 \leq j \leq J_1} \varepsilon_j).$$

Since ε_{0m} is independent of $\max_{1 \leq j \leq J_0} \varepsilon_j$ this can only be true if the distribution of the maxima for J_0 and J_1 of ε_j for $j \geq 1$ is the same, that is

$$\max_{1 \leq j \leq J_0} \varepsilon_j \stackrel{d}{=} -\alpha d(J_0, J_1) + \max_{1 \leq j \leq J_1} \varepsilon_j$$

Let F be the CDF of ε, then the equation above implies that for all natural number n there exists $t(n)$ such that for all x:

$$F(x) = F^n(x + t(n)).$$

Iterating on both sides implies

$$F^{nm}(x + t(nm)) = F^{nm}(x + t(n) + t(m))$$

we recognize an instance of the functional equation $t(nm) = t(n) + t(m)$ which has the unique solution $t(n) = c\log(n)$.[8] Therefore:

$$F(x) = F^{y}(x + c\log y),$$

letting $x = 0$, $s = c\log y$, we get $F(0) = F^{e^{s/c}}(s)$, and so:

$$F(s) = e^{\log F(0) e^{-s/c}},$$

which is a Gumbel distribution with location parameter $c\log(-\log F(0))$ and dispersion parameter c. This derivation proves that the parallel demands condition implies that the random utility shocks $(\varepsilon_{ij})_{j=1}^{\infty}$ follow the Gumbel distribution. Moreover, if the random utility shocks $(\varepsilon_{ij})_{j=1}^{\infty}$ follow the Gumbel distribution then $F(x) = e^{\log F(0) e^{-x/c}}$ and $F^{n}(x) = e^{\log F(0) e^{log(n)-x/c}} = F\,(x - clog(n))$ and so parallel demands hold:

$$\begin{aligned}
&\mathbb{P}\left(\varepsilon_{0m} < \delta - \alpha p + \max_{1\leq j\leq J_0} \varepsilon_j\right) \\
&\quad = \mathbb{P}\left(\varepsilon_{0m} < \delta - \alpha(p + clog(J_1) - clog(J_0)) + \max_{1\leq j\leq J_1} \varepsilon_j\right).
\end{aligned}$$

□

Proof of Theorem 2

Proof Let the random utility shocks (ε_j) be i.i.d. and distributed according to F in the domain of attraction of the Gumbel distribution. Let $G(x) = exp[-exp(-x)]$ be the Gumbel distribution. Then there exist sequences (a_n, b_n) such that

$$F^{n}(a_n x + b_n) \to G(x),$$

Furthermore, $\lim_{n\to\infty} \frac{a_n}{a_{[nt]}} = 1$ and $\lim_{n\to\infty} \frac{b_n - b_{[nt]}}{a_{[nt]}} = -c\log(t)$ for any $t > 0$ and some $c \in \mathbb{R}$ where $[nt]$ is the integer part of nt (see Resnick (1987) Chapter 1). Since the convergence $F^{n}(a_n x + b_n) \to G(x)$ is uniform (see Resnick (1987) Chapter 0) and F^{n} is uniformly continuous, then for any $\epsilon > 0$ there exists η and $N(\eta, \epsilon)$ such

[8] It is easy to extend the formula for real numbers through rationals, note

$$F(x) = F^{n}(x + t(n)) = F^{m}(x + t(m))$$

implies

$$F(x) = F^{n/m}(x + t(n) - t(m)),$$

so we can consistently define $t(n/m) = t(n) - t(m)$.

that for all $x \in \mathbb{R}$ and all $J_0, J_1 > N(\eta, \epsilon)$ we have $\left|\frac{a_{J_1}}{a_{J_0}} - 1\right| \leq \eta$ and

$$\begin{aligned}\left|F^{J_0}(a_{J_0}x + b_{J_0}) - F^{J_1}(a_{J_0}x + b_{J_1})\right| &\leq \left|F^{J_0}(a_{J_0}x + b_{J_0}) - F^{J_1}(a_{J_1}x + b_{J_1})\right| \\ &\quad + \left|F^{J_1}(a_{J_1}x + b_{J_1}) - F^{J_1}(a_{J_0}x + b_{J_1})\right| \\ &< \epsilon\end{aligned}$$

Therefore, for any $p \in \mathbb{R}$

$$\begin{aligned}&\left|Q(p, J_0) - Q\left(p + b_{J_1} - b_{J_0}, J_1\right)\right| \\ &\quad = \left|\mathbb{P}\left(\max_{j\in\{1,\dots,J_0\}} u_{ij}(p) > u_{i0}\right) - \mathbb{P}\left(\max_{j\in\{1,\dots,J_1\}} u_{ij}(p + b_{J_1} - b_{J_0}) > u_{i0}\right)\right| \\ &\quad = \left|\int_{\mathbb{R}} \left(F^{J_1}\left(\varepsilon_{i0} - \alpha(y-p) - \delta + \alpha(b_{J_1} - b_{J_0})\right) - F^{J_0}\left(\varepsilon_{i0} - \alpha(y-p) - \delta\right)\right) f_0(\varepsilon_{i0}) d\varepsilon_{i0}\right| \\ &\quad < \epsilon\end{aligned}$$

where f_0 is the probability density of ε_{i0}. We conclude that the inverse aggregate demands are asymptotically parallel. □

Proof of Proposition 1

Proof Redefine $\widetilde{\varepsilon}_{i0} = \varepsilon_{i0} - (1-\sigma)\nu_i$. Then the proof follows from Theorem 1. □

Proof of Theorem 3

Proof Let F be the CDF of the random utility shocks. Define Condition A as: for all $(\delta_n)_{n=1}^{J_0}$ bounded vector of real non-negative numbers there exists $f((\delta_n)_{n=1}^{J_0})$ such that $F(x) = \Pi_{n=1} F\left(x - \delta_n + f((\delta_n)_{n=1}^{J_0})\right)$. Theorem 1 applies for vectors of constants $(\delta, \dots, \delta)$ of any size, and shows that the only possible candidate CDF F that satisfies condition A must be Gumbel. Therefore if Condition A is going to hold for any $(\delta_n)_{n=1}^{J_0}$ bounded vector of real non-negative numbers, then F must be Gumbel. Condition A is a rephrasing of parallel WTP CDFs and so, Gumbel is necessary for parallel WTP CDFs.

Moreover, if $\left(\epsilon_{ij}\right)_{j=1}^{\infty}$ are i.i.d. Gumbel then $\delta_j + \epsilon_{ij} \sim F_j(\mu_j, \beta)$ are also Gumbel, where μ_j is the position parameter of the Gumbel distribution and β is the scale parameter ($\{\mu_j\}$ is well defined, because $\{\delta_j\}$ is bounded.) Then

$$\begin{aligned}\mathbb{P}\left(\delta_j + \epsilon_{ij} < x\right) &= F_j(x) \\ &= exp\left(-exp\left(\frac{\mu_j - x}{\beta}\right)\right).\end{aligned}$$

Let $j^* = argmax_{j\in J_0}\{\delta_j + \epsilon_{ij}\}$, we have

$$F_{j^*}(x) = \Pi_{j\in J_0} F_j(x)$$

$$= exp(-\Sigma_{j \in J_0} exp(\frac{\mu_j - x}{\beta}))$$
$$= exp(-exp(\frac{\mu - x}{\beta})),$$

where $\mu = \beta log \sum_{j \in J_0} exp(\frac{\mu_j}{\beta})$.

Similarly, let $j^{**} = argmax_{j \in J_1}\{\delta_j + \epsilon_{ij}\}$ for $J_1 \neq J$. We have

$$F_{j^{**}}(x) = exp\left(-exp\left(\frac{\mu' - x}{\beta}\right)\right)$$

where $\mu' = \beta log \sum_{j \in J_1} exp(\frac{\mu_j}{\beta})$. The above derivation shows that we have parallel WTP distributions by letting

$$t_{J_1} = \mu' - \mu$$
$$= \beta log \frac{\sum_{j \in J_1} exp(\frac{\mu_j}{\beta})}{\Pi \sum_{j \in J_0} exp(\frac{\mu_j}{\beta})}.$$

□

Proof of Theorem 4

Proof Take (α_n, β_n) and the nondegenarate CDF H such that $\Pi_{j=1}^{n} F(\alpha_n x + \beta_n - \delta_j) \to H(x)$ for all x. Because

$$F^n(\alpha_n x + \beta_n) \leq \Pi_{j=1}^{n} F(\alpha_n x + \beta_n - \delta_j) \leq F^n(\alpha_n x + \beta_n)$$

and by continuity, there exists $\gamma_n \in [0, K]$ such that $\Pi_{j=1}^{n} F(\alpha_n x + \beta_n - \delta_j) = F^n(\alpha_n x + \gamma_n) \to H(x)$. But because F is in the domain of attratcion of the Gumbel, by Proposition 0.2 of Resnick (1987) there exists a and b such that $H(x) = G(ax+b)$ is a rescaling of the Gumbel distribution.

The rest of the proof follows exactly the same steps as the proof of Theorem 1, starting from $\Pi_{j=1}^{n} F(\alpha_n x + \beta_n - \delta_j) \to G(x)$. We have $\lim_{n\to\infty} \frac{a_n}{a_{[nt]}} = 1$ and $\lim_{n\to\infty} \frac{\gamma_n - \gamma_{[nt]}}{a_{[nt]}} = -c \log(t)$ for any $t > 0$ and some $c \in \mathbb{R}$ where $[nt]$ is the integer part of nt (see Resnick (1987) Chapter 1).

Since the convergence $\Pi_{j=1}^{n} F(\alpha_n x + \beta_n - \delta_j) = F^n(\alpha_n x + \gamma_n) \to G(x)$ is uniform (see Resnick (1987) Chapter 0) and F^n is uniformly continuous, then for any $\epsilon > 0$ there exists η and $N(\eta, \epsilon)$ such that for all $x \in \mathbb{R}$ and all $J_0, J_1 > N(\eta, \epsilon)$ we have $\left|\frac{a_{J_1}}{a_{J_0}} - 1\right| \leq \eta$ and

$$\left|F^{J_0}(a_{J_0} x + \gamma_{J_0}) - F^{J_1}(a_{J_0} x + \gamma_{J_1})\right| \leq \left|F^{J_0}(a_{J_0} x + \gamma_{J_0}) - F^{J_1}(a_{J_1} x + \gamma_{J_1})\right|$$
$$+ \left|F^{J_1}(a_{J_1} x + \gamma_{J_1}) - F^{J_1}(a_{J_0} x + \gamma_{J_1})\right|$$
$$< \epsilon$$

Therefore, for any $p \in \mathbb{R}$

$$
\begin{aligned}
&\left|\mathbb{P}\left(WTP_i(J_0) \le x\right) - \mathbb{P}\left(WTP_i(J_1) \le x + \frac{\gamma_{J_1} - \gamma_{J_0}}{\alpha}\right)\right| \\
&= \left|\mathbb{P}\left(\max_{j\in\{1,\dots,J_0\}}\{\frac{\delta_j + \varepsilon_{ij} - \varepsilon_{i0}}{\alpha}\} \le x\right) - \mathbb{P}\left(\max_{j\in\{1,\dots,J_1\}}\{\frac{\delta_j + \varepsilon_{ij} - \varepsilon_{i0}}{\alpha}\} \le x + \frac{\gamma_{J_1} - \gamma_{J_0}}{\alpha}\right)\right| \\
&= \left|\int_{\mathbb{R}} \left(F^{J_1}\left(\alpha x + \varepsilon_{i0} - \delta_j + \gamma_{J_1} - \gamma_{J_0}\right) - F^{J_0}\left(\alpha x + \varepsilon_{i0} - \delta_j\right)\right) f_0(\varepsilon_{i0}) d\varepsilon_{i0}\right| \\
&< \epsilon
\end{aligned}
$$

where f_0 is the probability density of ε_{i0}. We conclude that the willingness-to-pay densities are asymptotically parallel. □

Proof of Proposition 2

Proof Assume parallel demands (Definition 1) and let $d(J_0, J_1)$ be such that $P(Q, J_0) + d(J_0, J_1) = P(Q, J_1)$. Then $\Lambda = \int_0^Q \left(P(s, J_1) - P(s, J_0)\right) ds = d(J_0, J_1) * Q$. □

Proof of Proposition 3

Proof Observe:

$$
\begin{aligned}
d(J_0, J_1) &= p_1 - P(Q_1, J_0) \\
&= \left(\frac{p_1 - p_0}{Q_1 - Q_0} - \frac{P(Q_1, J_0) - p_0}{Q_1 - Q_0}\right)(Q_1 - Q_0)
\end{aligned}
$$

Now assume $(p(J), Q(J))_{J\in\mathbb{R}}$ is a continuously differentiable interpolation of $(p(J), Q(J))_{J\in\mathbb{N}}$ which exists by the Stone-Weierstrass theorem. Then by the Taylor approximation theorem:

$$
\begin{aligned}
d(J_0, J_1) &= \left(\frac{p_1 - p_0}{Q_1 - Q_0} - \frac{P(Q_1, J_0) - p_0}{Q_1 - Q_0}\right)(Q_1 - Q_0) \\
&= \left(\frac{\frac{dp}{dJ}}{\frac{dQ}{dJ}} - \frac{\partial P(Q, J)}{\partial Q}\right)\frac{dQ}{dJ}\triangle J + O((\triangle J)^2)
\end{aligned}
$$

□

Proof of Proposition 5

Proof Let $d = d(J_0, J_1)$. Observe by assumption $Q_{J_1}(\mathbf{p_{J_1}}) = Q_{J_0}(\mathbf{p_{J_0}} + (\rho - d)\mathbf{1_{J_0}})$, then the second part of the theorem follows directly from the first-order Taylor approximation:

$$
Q_{J_1}(\mathbf{p_{J_1}}) = Q_{J_0}(\mathbf{p_{J_0}}) + (\rho - d)\frac{dQ_{J_0}(\mathbf{p_{J_0}} + t\mathbf{1_{J_0}})}{dt} + O\left((\rho - d)^2\right)
$$

where $\frac{dQ_{J_0}(\mathbf{p_{J_0}}+t\mathbf{1_{J_0}})}{dt}$ is the directional derivative in the direction $\mathbf{1_{J_0}}$. And so

$$d = \left(\frac{\rho}{\Delta Q} - \left(\frac{dQ_{J_0}(\mathbf{p_{J_0}} + t\mathbf{1_{J_0}})}{dt} \right)^{-1} \right) \Delta Q + O\left((\rho - d)^2 \right)$$

□

References

Anderson, S.P., Bedre-Defolie, O.: Variety provision of a multiproduct monopolist. Working Paper. (2019)

Anderson, S.P., Palma, A.D.: The logit as a model of product differentiation. Oxf. Econ. Pap. **44**, 51–67 (1992)

Anderson, S.P., Palma, A.D., Thisse, J.F.: The CES is a discrete choice model? Econ. Lett. **24**(2), 139–140 (1987)

Arkolakis, C., Demidova, S., Klenow, P.J., Rodríguez-Clare, A.: Endogenous variety and the gains from trade. Am. Econ. Rev. Pap. Proc. **98**(2), 444–450 (2008)

Bayer, P., Ferreira, F., McMillan, R.: A unified framework for measuring preferences for schools and neighborhoods. J. Polit. Econ. **115**(4), 588–638 (2007)

Berry, S.: Estimating discrete-choice models of product differentiation. RAND J. Econ. **25**(2), 242–262 (1994)

Berry, S., Waldfogel, J.: Free entry and social inefficiency in radio broadcasting. Rand J. Econ. **30**(3), 397–420 (1999)

Breitmoser, Y.: An axiomatic foundation of conditional logit. Econ. Theor. **72**, 245–261 (2021)

Broda, C., Weinstein, D.E.: Globalization and the gains from variety. Q. J. Econ. **121**(2), 541–585 (2006)

Card, D., Cardoso, A.R., Heining, J., Kline, P.: Firms and labor market inequality: evidence and some theory. J. Law Econ. **36**(1), 13–70 (2018)

Cardell, S.: Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. Economet. Theor. **13**(2), 185–213 (1997)

Chan, M., Kroft, K., Mourifie, I.: An empirical framework for matching with imperfect competition. Working Paper. (2019)

Dhingra, S., Morrow, J.: Monopolistic competition and optimum product diversity under firm heterogeneity. J. Polit. Econ. **127**(1), 196–232 (2019)

Dinerstein, M., Smith, T.: Quantifying the supply response of private schools to public policies. SIEPR Working Paper. (2014)

Dixit, A.K., Stiglitz, J.E.: Monopolistic competition and optimum product diversity. Am. Econ. Rev. **67**(3), 297–308 (1977)

Echenique, F., Saito, K.: General luce model. Econ. Theor. **68**, 811–826 (2019)

Feenstra, R.C.: New product varieties and the measurement of international prices. Am. Econ. Rev. **84**(1), 157–177 (1994)

Gabaix, X., Laibson, D., Li, D., Li, H., Resnick, S., de Vries, C.G.: The impact of competition on prices with numerous firms. J. Econ. Theory **165**, 1–24 (2016)

Hicks, J.: Value and Capital. Oxford University Press, Oxford (1936)

Hsieh, C.T., Hurst, E., Jones, C.I., Klenow, P.J.: The allocation of talent and us economic growth. National Bureau Econ. Res. (2013)

Kreinovich, V., Nguyen, H.T., Sriboonchitta, S., Kosheleva, O.: Modeling extremal events is not easy: Why the extreme value theorem cannot be as general as the central limit theorem. Departmental Technical Reports. (2015)

Kroft, K., Leal-Vizcaíno, R., Laliberté, J.W.P., Notowidigdo, M.J.: A welfare analysis of tax structures with love-of-variety preferences. Working paper. (2021)

Lamadon, T., Mogstad, M., Setzler, B.: Imperfect competition, compensating differentials and rent sharing in the U.S. Working Paper. (2020)

Mankiw, N.G., Whinston, M.: Free entry and social inefficiency. RAND J. Econ. **17**, 48–58 (1986)

Matejka, F., McKay, A.: Rational inattention to discrete choices: a new foundation for the multinomial logit model. Am. Econ. Rev. **105**(1), 272–298 (2015)

McFadden, D.: Modeling the choice of residential location. Transpor. Res. Rec. **673**, 71 (1978)
Nevo, A.: Empirical models of consumer behavior. Annu. Rev. Econ. **3**, 51–75 (2011)
Resnick, S.I.: Extreme Values, Regular Variation and Point Processes. Springer, New York (1987)
Romer, P.: New goods, old theory, and the welfare costs of trade restrictions. J. Dev. Econ. **43**(1), 5–38 (1994)
Spence, M.: Product differentiation and welfare. Am. Econ. Rev. **66**(2), 407–414 (1976a)
Spence, M.: Product selection, fixed costs, and monopolistic competition. Am. Econ. Rev. **43**(2), 217–235 (1976b)