Poor Substitutes? Counterfactual methods in IO and Trade compared*

Keith Head† Thierry Mayer‡

July 11, 2023

Abstract

Constant elasticity of substitution (CES) demand for monopolistically competitive firm-varieties is a standard tool for models in international trade and macroeconomics. Inter-variety substitution in this model follows a simple share proportionality rule. In contrast, the standard toolkit in industrial organization (IO) estimates a demand system in which cross-elasticities depend on similarity in observable attributes. The gain in realism from the IO approach comes at the expense of requiring richer data and greater computational challenges. This paper uses the data generating process of Berry et al. (1995), BLP, who established the modern IO method, to simulate counterfactual trade policy experiments. We use the CES model as an approximation of the more complex underlying demand system and market structure. Although the CES model omits key elements of the data generating process, the errors are offsetting, allowing it to fit BLP-based predictions closely. For aggregate outcomes, it turns out that incorporating non-unitary pass-through matters more than fixing over-simplified substitution patterns.

Keywords: Constant Elasticity of Substitution, Industrial Organization, Oligopoly, Trade, Tariffs, Counterfactual analysis.

*This research has received funding from the European Research Council (ERC) under the Grant Agreement No. 313522. We thank participants at the SMU-NUS-Paris Joint Trade Workshop (Merlion 2018 Workshop) and at the worksop on International trade and multilateralism (Villars 2023 Workshop). Very helpful comments were received during seminar presentations at Princeton University, Columbia University, Harvard University, University of Michigan, University of Oslo, Mannheim University, ULB, SED, Tilburg. The authors are particularly grateful to Laura Grigolon and Scott Orr for their detailed reading and advice. Comments from three anonymous referees and the editor Daniel Xu were also extremely helpful.

†University of British Columbia Sauder School of Business and CEPR, keith.head@sauder.ubc.ca

‡Sciences Po, CEPII, and CEPR, thierry.mayer@sciencespo.fr

1 Introduction

Tariffs, never completely absent, rose to the foreground of US economic policy again in 2018. The US imposed safeguard tariffs on washing machines and solar panels, followed by national security tariffs on steel and aluminum. The US president threatened Canada, Mexico, and Germany with national security tariffs on imported autos. Intensified tariff use led to renewed efforts by economists to quantify the impact of trade policies. The 2018 tariffs also reinforce the point that most trade policy is imposed at the industry level. This creates a dilemma for researchers. Trade economists have developed a toolkit for tariff counterfactuals that imposes minimal data and estimation requirements. Industrial organization (IO) economists have an even more established framework for conducting industry-level counterfactuals. It differs from the approach favored by trade economists in almost every important respect, but the most emphasized feature is *rich substitution*. Berry et al. (2004) state the main IO critique that applies to constant elasticity of substitution (CES) as well as other demand systems common in the models of monopolistic competition used in trade: "Models without individual differences in preferences for characteristics generate demand substitution patterns that are known to be *a priori* unreasonable (depending only on market shares and not on the characteristics of the vehicles)."

The IO structure pioneered by Berry et al. (1995) promises greater realism at the cost of more onerous data and estimation requirements. What can be said, systematically, about the suitability of the trade approach when the data are generated by the assumptions of the IO approach? This paper starts with the premise that IO economists have correctly specified the data generating process (DGP). The demand side of that DGP is mixed multinomial logit (MMNL). McFadden and Train (2000) proved that "any discrete choice model derived from random utility maximization has choice probabilities that can be approximated as closely as one pleases by a MMNL model." This theorem motivates the use of MMNL as the benchmark model of consumer behavior. This DGP presents several distinct challenges for the simpler representation of CES-monopolistic competition (CES-MC) offered by trade economists. After analyzing those problems, we investigate whether the CES-MC method can be viewed as an acceptable approximation. To do so, we use both the data and the parameter estimates from the seminal paper in the literature, Berry et al. (1995). Our counterfactual imposes a 10% tariff on foreign varieties and solves the model to obtain the new equilibrium. We then use the method of Exact Hat Algebra (EHA)—relying solely on initial market shares and on an estimate of the elasticity of substitution—to obtain a CES-MC prediction of *ex post* equilibrium market shares. This policy scenario is more general than it might appear, since it applies to any tax that discriminates against some across varieties (e.g. a tax based on fuel consumption as in Grigolon et al. (2018), or based on the sugar content of soft drinks as in Dubois et al. (2020)).

Throughout this paper, we evaluate the accuracy of the CES model at matching the changes in domestic market share caused by a tariff imposed on a data-generating process given by random coefficients discrete choice with multi-product firms (BLP for short).[1] In this first simulation, the CES prediction is astonishingly accurate, undershooting this target by only one quarter of a percentage point (8.00 vs 7.73). To understand this remarkable success, we proceed to a second set of simulations. Those are intended to investigate each of the methodological differences (errors) between the CES approximation and the BLP DGP. Trade economists typically estimate the CES as a constant price elasticity when BLP features own price elasticities that vary across firms. Within the single-market setup of the original data, we cannot easily estimate the CES parameter. A first change is therefore to augment the original model to consider multiple markets, each imposing import tariffs. Once we base the CES prediction on an estimated parameter rather than calibration, the central case performs a little less well, but alternative settings are more robust.

The simulations evaluating the CES approximation of the random coefficients logit DGP do not contain CES as a limiting case as heterogeneity goes to zero. This makes it hard to determine the role of rich substitution because that change is confounded by the logit versus CES differences in demand curvature. To remedy this, we include in our second set of simulations a mixed CES version of the data generating process. By nesting homogeneous CES as a polar case, we can see clearly the extent that random coefficients undermines the CES approximation.[2]

The main takeaway is that pass-through from tariffs into prices is a first-order consideration for the success of the CES approximation. When CES gets pass-through (close to) right, it tends to hit the aggregate target accurately. The CES success at approximating BLP in aggregate outcomes turns out to be a case of offsetting errors: assuming CES monopolistic competition, rather than logit oligopoly, overestimates the pass-through of tariffs into consumer prices. However, random coefficients on prices generates a selection effect that pushes in the opposite direction: As established in Nakamura and Zerom (2010), heterogeneous price sensitivity raises pass-through. In our context, when tariffs

[1]This paper is agnostic on the separate issue of whether BLP or CES models accurately predict the real-world impacts of tariffs. An important step towards evaluating this question is taken by Adao et al. (2022).

[2]Mixed CES finds support in work by Björnerstedt and Verboven (2016), who refer to it as the "constant expenditures specification." In a recent application to a pharmaceutical merger, mixed CES has a "more plausible range of elasticities, more reasonable markups, and yields more realistic average predicted price effects for the merging firms." Other applications include Adao et al. (2017), Redding and Weinstein (2019), and Piveteau and Smagghue (2021).

rise, the households who keep buying foreign varieties are the ones with low price responsiveness. This lowers the demand elasticity and increases the pass-through elasticity. Our simulations show that estimating the tariff elasticity (instead of calibrating demand parameters to fit literature estimates of the own-price elasticity) provides an additional degree of freedom to the CES approximation. This allows it to capture non-unitary pass-through and perform better in a broader range of DGP settings.

The large literature employing the Berry et al. (1995) framework motivates the use of the random coefficients demand by critiquing systems that fail to incorporate rich substitution. In a recent survey, Berry and Haile (2021) point to this crucial flaw:

> "[O]ne can go too far in the pursuit of parsimony. Some of the simplest demand specifications (e.g. the CES, multinomial logit, multinomial probit) impose strong *a priori* restrictions on demand elasticities—and therefore on markups, pass-through and other key quantities of interest—that are at odds with common sense and standard economic models."

Emphasis on the need to incorporate rich substitution, combined with multi-product oligopoly is particularly strong in the literature devoted to the car industry, an emblematic case studied from the beginning of the demand-centered IO literature (Berry et al. (1995, 1999), Goldberg (1995), Verboven (1996), Goldberg and Verboven (2001), Petrin (2002), Train and Winston (2007), Reynaert and Verboven (2014), and Coşar et al. (2018) for instance). Because of our use of the BLP structure, data, and parameters, we speak to this literature "on its playground", assessing when and why the approximation fails to predict aggregate outcomes. Using the same data and parameters as Berry et al. (1995, 1999) addresses the potential concern that an *ad hoc* DGP might not exhibit sufficiently rich substitution patterns or strong enough market power.

Notwithstanding the valid critiques made by IO economists, CES-MC has advantages that may not have been fully recognized. In addition to the parsimony point conceded in the quote above, CES allows us to exploit Exact Hat Algebra. This method, introduced by Dekle et al. (2007), allows the researcher to do without detailed data on product attributes and prices. It also does not require the inference of marginal costs from first order conditions. The tractability of models featuring CES was exploited by applied theorists and empiricists long before Exact Hat Algebra was developed.[3] EHA gives CES a new advantage in conducting quantitative research by providing a simple tool for running counterfactuals with minimal data requirements.

[3]For example, Krugman and Romer exploited CES modelling properties in their (respective and independent) work on trade and economic growth with increasing returns.

Relatedly, the IO literature has acknowledged that the random-coefficients models present serious challenges in computation (Knittel and Metaxoglou, 2014), identification (Gandhi and Houde, 2016), sensitivity to the choice of instruments (Reynaert and Verboven, 2014), data requirements, and transparency of estimation. Conlon and Gortmaker (2020) present a very complete coverage of the various practical challenges in BLP estimation, with different fixes to the original framework that have been proposed by the IO literature. Salanié and Wolak (2019) also note the estimation challenges of the BLP-based framework and propose an alternative estimation strategy, consisting in an approximation where consumer's tastes dispersion parameters can be estimated in a simple 2SLS procedure. Their Monte Carlo simulations show that their approximation result can be used to at least give very close starting values to a more elaborate but more challenging estimation technique. Our paper is also centered around Monte Carlo simulations, but we sidestep the estimation of issues related to the BLP model. Instead, our Monte Carlos assess the ability of the CES approximation to predict aggregate outcomes of BLP-generated data.

Our paper proceeds as follows. We first describe the BLP data and model structure in section 2. We then explain our two extensions to the Exact Hat Algebra method in section 3. After analyzing the three main causes of concern for the CES approximation in section 4, we assess in section 5 the relative importance of these issues using simulations that treat them one at a time.

2 The BLP data generating process

Berry et al. (1995, 1999) describe the data generating process in detail, but here we review the key equations and provide the necessary details on how we implement it in our simulations, together with some key statistics of the original data set used in both articles.

The key components of the BLP DGP are heterogeneous consumer choice probabilities and multi-product firms. Let each firm f own a set of varieties denoted $\mathcal{J}_f$. The union of these sets is $\mathcal{J}$ which we also partition into sets of domestic, $\mathcal{J}_H$ and foreign $\mathcal{J}_F$ varieties.[4] The total number of varieties, $|\mathcal{J}|$, is taken as fixed. The demand side consists of a large number, N, of households, with each h having its own indirect utility u_{mh} for variety m. The preferences of the households are unobserved in BLP, but we have data on the fraction, s_m, of the N consumers that select each model m within the set of new cars available for purchase, along with the fraction who purchase the outside good s_0 (used car or no

[4] In the BLP data, domestic models constitute 68% of new car sales.

purchase). With unit demand, the market share of variety m is[5]

$$s_m = \frac{\sum_h \mathbb{P}_{mh}}{N} \quad \text{where } \mathbb{P}_{mh} = \text{Prob}(u_{mh} > u_{m'h}, \forall m') \tag{1}$$

Assuming Gumbel-distributed additive shocks in u_{mh}, the choice probabilities are

$$\mathbb{P}_{mh} = \frac{\exp(\sum_{k=0}^{K} \beta_h^k x_m^k - \alpha_h p_m + \xi_m)}{1 + \sum_{j \in \mathcal{J}} \exp(\sum_{k=0}^{K} \beta_h^k x_j^k - \alpha_h p_j + \xi_j)}. \tag{2}$$

We will refer to β heterogeneity as the feature of the model that households value the physical characteristics (other than price) differently. There are $K = 4$ characteristics plus a constant ($x_h^0 = 1$). Since the indirect utility of the outside good is normalized to one, the coefficient β_h^0 tells how much the household prefers a new car relative to the outside good. The mean of these coefficients, $\bar{\beta}^0$ determines the share of the outside good. Reflecting the fact that only 9% of households buy new cars, Berry et al. (1995) estimate $\bar{\beta}^0$ to be -7.1. The standard deviation of β_h^0 is 3.6, suggesting considerable dispersion in appeal of new cars. The four other x_m^k are (1) acceleration(horsepower/weight), (2) fuel economy (miles per dollar), (3) space (width $\times$ length), and (4) air conditioning (as a standard feature). When we speak of β heterogeneity, we refer to the variance in the β_h^k. The means and standard deviations for each of these β_h^k are all obtained from Berry et al. (1995) and reported in the first column of Table 1.

Variance in the price responsiveness parameter α_h will be referred to as α heterogeneity. There are two important points. First, α heterogeneity is large because we follow Berry et al. (1999) in setting $\alpha_h = \alpha / y_h$ where $\ln y_h \sim \mathcal{N}(2.18, 1.72)$ in 1990.[6] While this specification imposes a negative relationship between income and price sensitivity (α_h), subsequent papers, such as Nevo (2001) and Nakamura and Zerom (2010), estimate the relationship using more flexible specifications. Second, as our simulations will illustrate, α heterogeneity changes the curvature of demand, leading to market outcomes that are qualitatively different from those generated by β heterogeneity.[7]

[5] Here we deviate slightly from the convention of expressing market shares as integrals over a continuum of consumers. Our summations over a finite number of consumers lead naturally to expressions of demand elasticities in terms of variances and covariances of household probabilities. The averages in equation (1) can also be thought of as a Monte Carlo integration, the method used in our simulations.

[6] We follow the recent literature that replicates the original BLP results by using the Berry et al. (1995) data and parameter values, combined with the Berry et al. (1999) approach to consumer-level heterogeneity in price sensitivity (α_h). Our approach follows Andrews et al. (2017) (with details contained in their replication package) and Conlon and Gortmaker (2020) (with code tutorial accessible at `https://pyblp.readthedocs.io/en/stable/_notebooks/tutorial/blp.html`).

[7] Heterogeneity in price sensitivity across firms can also be obtained using non-CES demand curves, often derived from non-homothetic preferences. See Mrázová and Neary (2017) and Matsuyama (2023).

Table 1: BLP data: estimated parameters and key statistics

Estimated parameters			Auto industry statistics in 1990	
Variable	Mean	Std. dev.	Statistic	Value
Constant	-7.061	3.612	Outside good share (%)	91
HP/WT	2.883	4.628	Domestic share (%)	68
Air con.	1.521	1.818	Concentration (CR5 in %) firms	86
Miles/USD	-0.122	1.050	Concentration (CR5 in %) models	18
Size	3.460	2.056	Number of firms	20
Price	43.501	91.906	Number of models	131

Note: Estimated parameters obtained from Table IV of Berry et al. (1995), with the exception of the standard deviation of the price parameter, calculated as $\alpha\sqrt{(\exp(\sigma^2)-1)\exp(-2\mu+\sigma^2)}$ with $\mu = 2.18$ and $\sigma = 1.72$, being the mean and standard deviation of log incomes in the United States in 1990 used by BLP. The Andrews et al. (2017) replication package provides these parameters as well as the data for our calculated statistics in the second column.

The right panel of table 1 summarizes some important industry statistics that guide our counterfactual experiments of sections 4.5 and 5. One important part of the BLP is the outside good, defined as the number of households minus the number of new cars purchased. The 91% share for the outside good means that actual market shares for new car models are very small. Including the outside good in the denominator of market shares, the mean s_m in 1990 is 0.07% and the maximum is 0.44%.

Such small market shares might seem to leave little room for oligopoly conduct. However there are two countervailing forces. First is the prevalence of multi-product firms. The Big 3 firms made half the 131 varieties sold in 1990. The five largest firms accounted for 86% of new car sales. Second, the large dispersion in taste for the outside good mentioned above also restores market power within the inside good (new cars). Intuitively, a sizable number of potential consumers are effectively outside the market for new cars because they have strong preference for the outside good. Car manufacturers therefore mostly compete for the remainder of the customer base, on which they have larger market share and therefore greater pricing power.

The firm's profit maximization problem chooses prices for each model accounting for the impact a rise in m's price would have on the profits earned for the other models $(j \neq m) \in \mathcal{J}_f$. The first order condition is

$$p_m = c_m - \frac{s_m + \sum_{(j\neq m)\in\mathcal{J}_f}(p_j - c_j)\frac{\partial s_j}{\partial p_m}}{\frac{\partial s_m}{\partial p_m}}. \tag{3}$$

The own- and cross-price derivatives of s_m are shown in Appendix B. Having data on

$\{s_m, p_m, \mathcal{J}_f\}$ and having estimates of the mean and standard deviation of $\{\alpha, \boldsymbol{\beta}\}$, we can infer ξ_m (via an inversion referred to by the literature as the contraction mapping). Then c_m can be obtained by moving c_m to the left hand side of equation (3). On the right hand side, s_m is known, $\partial s_m / \partial p_m$ is implied by the parameters and price data, leaving only the summation term as a function of the unknown c_m. Starting with $c_m = p_m$, we iterate until reaching a stable vector of marginal costs. At this stage, we have knowledge on all relevant characteristics of each car model m, and can use those combined with the parameters of consumer preferences to run counterfactual experiments.

In the next sections, we will consider several variants of the BLP data generating process. In order to make it very clear which variant we are referring to in different exercises, we adopt a systematic terminology. Starting with the DGP just described in that section, we use:

- *Mixed Logit* for the random coefficients logit DGP of Berry et al. (1995, 1999) including heterogeneity in β (characteristics) and α (price).
- *Mixed CES* for the random coefficients CES DGP described in appendix A, including heterogeneity in β (characteristics) and α (*log* price).
- *Logit* for homogeneous coefficient multinomial logit.
- β *heterogeneity* for random coefficients on characteristics. This can apply in either the logit or the CES functional forms as it is always clear by context.

3 Counterfactual calculations: true BLP vs CES hat algebra

The counterfactual policy we use to motivate this paper is a new tariff of 10% on imported varieties. Berry et al. (1999) consider quantitative restrictions on imports, but tariffs are much easier to model. Recent experience demonstrates that tariffs remain relevant, and they have a broader interpretation as discriminatory taxes. Because the tariff imposed on model m depends on the model's origin country, $i(m)$, and the market n where it is sold, we now move to a multi-market setup. Tariffs are modeled as shocks to the factory gate costs, c_m. Thus, tariff-inclusive costs are $c_{mn} = c_m \tau_{i(m)n}$. This makes sense since customs authorities apply tariffs to the customs value of the good at the border, not on the retail price. Rules on the value for duties are complex; we simplify matters by applying them to marginal costs. This allows our counterfactual to apply more broadly than the trade policy context as it can encompass any model-specific cost shock. Firms apply a discriminatory market-specific markup to set the retail price, p_{mn}. In the counterfactuals, marginal costs rise to $c_m \tau'_{i(m)n} = c_{mn}\hat{\tau}_{i(m)n}$, where $\hat{\tau}_{i(m)n} = \tau'_{i(m)n}/\tau_{i(m)n} = 1.1$ for all foreign models ($i(m) \neq n$) and $\hat{\tau}_{i(m)n} = 1$ for domestic models ($i(m) = n$).

The true new equilibrium is obtained by iterating equation (3) until a fixed point in new prices, $p^{\dagger}_{mn}$ is reached.[8] Then we substitute the prices into demand to obtain the new market shares, denoted $s^{\dagger}_{mn}$, which we aggregate to obtain the true change in the domestic share of new car production.

$$\Delta S_n^{\text{BLP}} = \sum_{m \in \mathcal{J}_H} \left(\frac{s^{\dagger}_{mn}}{1 - s^{\dagger}_{0n}} - \frac{s_{mn}}{1 - s_{0n}} \right). \tag{4}$$

Our goal is to investigate whether ΔS_n^{BLP} can be reasonably approximated with ΔS_n^{CES}, the prediction of market share changes implied by CES demand. In contrast to ΔS_n^{BLP}, the CES predictions are not obtained by solving the model in terms of its structural parameters. Rather, hat algebra methods predict new market shares using only the initial market shares s_{mn} and the single CES demand parameter, denoted η. The key idea is that the initial market shares are sufficient statistics for both observed and unobserved characteristics of the varieties and trade costs.

The CES market share for model m is given by[9]

$$s_{mn} = \frac{(p_{mn}/A_{mn})^{-\eta}}{1 + \sum_{j \in \mathcal{J}} (p_{jn}/A_{jn})^{-\eta}},$$

where A_{mn} is the demand shifter. Prices are given by $p_{mn} = \mu_{mn} c_m \tau_{i(m)n}$, where μ_{mn} is the markup (defined here as price divided by marginal costs). Defining $\varphi_{mn} \equiv A_{mn}/c_m$, we can re-express equilibrium market shares as

$$s_{mn} = \frac{\left(\mu_{mn}\tau_{i(m)n}/\varphi_{mn}\right)^{-\eta}}{1 + \sum_{j \in \mathcal{J}} (\mu_{jn}\tau_{i(j)n}/\varphi_{jn})^{-\eta}}. \tag{5}$$

In appendix D, we show that, with a constant markup of $\mu = \eta/(\eta - 1)$, Exact Hat Algebra calculates counterfactual proportional changes in market shares as

$$\hat{s}_{mn} = \frac{s'_{mn}}{s_{mn}} = \frac{\hat{\tau}_{i(m)n}^{-\breve{\eta}}}{s_{0n} + \sum_{j \in \mathcal{J}} s_{jn} \hat{\tau}_{i(j)n}^{-\breve{\eta}}}, \tag{6}$$

[8] Convergence of this process often requires use of a dampening factor $\nu < 1$ is to achieve convergence. Thus, if the new price implied by kth iteration of the first order condition is $p^{(k)}_{mn}$ we instead use $\nu p^{(k)}_{mn} + (1 - \nu)p^{(k-1)}_{mn}$. The choice of ν does not affect the equilibrium to which the iteration converges.

[9] As BLP work with quantity shares, we use the modification of the CES employed by Head and Mayer (2019), where η is the own-price elasticity holding constant the price index, P_n.

where $\breve{\eta}$ is an estimate of η.

The CES counterfactuals aggregate the new market shares obtained from hat algebra, $s'_{mn} = s_{mn}\hat{s}_{mn}$, to obtain the change in the domestic share of the new car market:

$$\Delta S_n^{\text{CES}} = \sum_{m \in \mathcal{J}_H} \left(\frac{s'_{mn}}{1 - s'_{0n}} - \frac{s_{mn}}{1 - s_{0n}} \right), \tag{7}$$

where $s'_{0n} = 1 - \sum_{m \in \mathcal{J}} s'_{mn}$.

We will consider two potential sources of $\breve{\eta}$. The first is the average own price elasticity implied by BLP data and estimated parameters (4.05).[10] The second estimate comes directly from a regression of log market shares on an *ad valorem* cost shifter such as the log of one plus the tariff rate.

We derive an equation for estimating η by taking logs of equation (5). Since μ_{mn} is constant in CES under monopolistic competition, it will be captured—along with the denominator in equation (5)—with n-specific fixed effects. This yields a firm-level version of the gravity equation:

$$\ln s_{mn} = -\eta \ln \tau_{i(m)n} + \text{FE}_m + \text{FE}_n + \upsilon_{mn}. \tag{8}$$

Here we have modeled $\eta \ln \varphi_{mn}$ as the sum of a model-specific fixed effect—capturing production cost (c_m) and the way the average consumer values the attributes of the car—and an idiosyncratic term, υ_{mn}. The latter is modeled as if it were a well-behaved error term capturing variation in A_{mn} across markets. In practice, it also contains the specification error from assuming CES when the underlying data comes from a BLP process. The last element of the specification is a market specific fixed effect capturing $-\eta \ln P_n$.

The estimation of (8) provides $\breve{\eta}$, which is the only parameter needed (besides observed market shares and changes in trade costs) to compute the counterfactual outcome in (6). Because of mis-specification, $\breve{\eta}$ does not estimate the underlying price elasticity of demand as it would have if the data were really generated by a CES-MC process. Instead, $\breve{\eta}$ recovers a rough estimate of the average elasticity of market shares with respect to *cost* shocks, building in non-unitary pass-through. Thus, if the underlying pass-through is less than one, $\breve{\eta}$ will be smaller than the price elasticity, which can compensate in part for the mis-specified functional form.

Equation (8) resembles the type of gravity regressions commonly estimated by trade economists using industry-level flows. It assumes the availability of variety-level (mi-

[10]This estimate, obtained from the 1990 data, hardly differs from the 3.928 pooled 1971–1990 estimate reported in the Conlon and Gortmaker (2020, Table 8) replication.

cro) market shares, the relevant assumption for our question which presumes that BLP-type estimation (involving model-level market shares and attributes) is a feasible option. However, many trade applications rely on country-to-country flows at the product-level to estimate the trade elasticity and run counterfactuals. In this case, the aggregate (macro) versions of (8) is

$$\ln s_{in} = -\eta \ln \tau_{in} + \mathrm{FE}_i + \mathrm{FE}_n + v_{in}, \tag{9}$$

where i is the origin country. The aggregate counterfactual version of equation (6) is

$$\hat{s}_{in} = \frac{s'_{in}}{s_{in}} = \frac{\hat{\tau}_{in}^{-\breve{\eta}}}{s_{0n} + \sum_i s_{in}\hat{\tau}_{in}^{-\breve{\eta}}}. \tag{10}$$

With the same estimate of η, and trade cost shocks that depend only on the country-pair, the macro counterfactual is just the sum of the micro counterfactuals ($\sum_{m\in i} \hat{s}_{mn} = \hat{s}_{in}$). In section 5, Table 3 provides results using both the micro (equations (8) and (6)) and macro (equations (9) and (10)) approaches of the CES approximation.

One strong restriction of EHA based on CES monopolistic competition is the constant markup, which implies unitary pass-through elasticities. In the appendix, we consider two ways to modify EHA to allow for incomplete pass-through. Within the homogeneous logit model, pass-through elasticities are much lower than one, even under monopolistic competition. In appendix G, we show how to conduct Exact Hat Algebra in a logit model with non-negligible market shares. Appendix A maintains CES demand but show how CES counterfactuals can be adjusted to allow large firms whose variable elasticities lead to non-unitary pass-through. The chief difference between the two oligopoly models is that the CES version has a pass-through elasticity just below one, while the Logit version pass-through elasticity is much smaller (0.67 on average when using BLP data, see second row in table 3).

In the next section we analyze three features of an equilibrium in the BLP model that the CES counterfactuals cannot match. Before continuing, we should acknowledge that Exact Hat Algebra's parsimony in terms of data requirements may come at a cost. Dingel and Tintelnot (2021) note that the method is equivalent to calibrating $|\mathcal{J}|$ unobserved parameters (here φ_{mn}) based on $|\mathcal{J}|$ market shares. When those market shares are based on small numbers of choosers (N in the model), granularity can lead to an overfitting problem. In the context of large consumer goods markets, like the US new car market, we do not see this as a major concern, given that millions of American households buy new cars each year.

4 Implications of the BLP data generating process

What behavioral predictions of the DGP used by Berry et al. (1995) present difficulties for the CES model? We have identified three main concerns.[11] The first, rich substitution, is well known but we offer a new way of quantifying its importance in the data. The second, niche market power, is probably familiar as well but we have a new analytic result and quantification. We believe the third result—on pass-through—has not received the attention it deserves, especially as we find it is the best indicator of when the CES approximation may be expected to miss the mark widely.

4.1 Rich substitution (cross-elasticities)

Implication 1. *Positive covariance in household choice probabilities raises cross-price demand elasticities.*

With heterogeneous α and β, the cross-price elasticity of demand is[12]

$$\frac{\partial \ln s_j}{\partial \ln p_m} = \frac{\sum_h \frac{\partial \mathbb{P}_{jh}}{\partial p_m}}{N} \frac{p_m}{s_j} = \left[\frac{\sum_h \alpha_h \mathbb{P}_{mh} \mathbb{P}_{jh}}{N} \right] \frac{p_m}{s_j}. \tag{11}$$

Similarity in the attributes of models m and j will make $\mathbb{P}_{mh}$ and $\mathbb{P}_{jh}$ covary positively, a feature that cannot be captured if all consumers value attributes identically. This implication of BLP arises from both α and β heterogeneity but as it does not require the former, it is easier to explain by focusing on β heterogeneity alone. Removing income variation by setting $y_h = 1$, the price coefficient is α and the factor in brackets is linear in the covariance of h probabilities, yielding a cross-price elasticity of

$$\epsilon_{jm}^{\beta\text{het}} \equiv \left. \frac{\partial \ln s_j}{\partial \ln p_m} \right|_{\alpha_h = \alpha} = \frac{\alpha p_m}{s_j} \sum_h \frac{\mathbb{P}_{jh} \mathbb{P}_{mh}}{N} = \alpha p_m s_m \left[1 + \frac{\text{cov}(\mathbb{P}_{jh}, \mathbb{P}_{mh})}{s_j s_m} \right].$$

Dividing $\epsilon_{jm}^{\beta\text{het}}$ by $\epsilon_{jm}^{\text{logit}} \equiv \alpha s_m p_m$, the cross-price elasticity with homogeneous consumers, the ratio of cross-price elasticities is

$$\frac{\epsilon_{jm}^{\beta\text{het}}}{\epsilon_{jm}^{\text{logit}}} = 1 + \frac{\text{cov}(\mathbb{P}_{jh}, \mathbb{P}_{mh})}{s_j s_m}. \tag{12}$$

[11] Here we consider only how parameter heterogeneity affects the true demand elasticities; the estimation of those parameters raises a different set of issues that lie outside the scope of this paper.

[12] Computation details to be found in Appendix B.

The cross-price elasticity of BLP (with only β heterogeneity) therefore depends on whether probabilities of buying varieties m and j covary positively or negatively across consumers. Two products with similar characteristics have similar attractiveness for each of the consumers, leading to positive covariance and therefore higher cross-price elasticity than under homogeneous logit.

What about the comparison with the CES cross elasticity that operates in the CES counterfactuals? Since the cross-price elasticity under CES is ηs_m, the β heterogeneity vs CES cross elasticity ratio is given by

$$\frac{\epsilon_{jm}^{\beta\text{het}}}{\epsilon_{jm}^{\text{CES}}} = \frac{\alpha p_m}{\eta} \times \left[1 + \frac{\text{cov}(\mathbb{P}_{jh}, \mathbb{P}_{mh})}{s_j s_m}\right]. \tag{13}$$

When quantifying the above expression, we run into the problem that the parameters α and η come from two different models. We resolve this by calibrating them both to match the average own-price elasticity implied by the BLP parameter estimates, that is $\bar{\epsilon}^{\text{BLP}} = 4.05$. Inverting the formula for the homogeneous logit own price elasticity ($\epsilon_m^{\text{logit}} \equiv \alpha p_m(1 - s_m)$), we isolate $\alpha = \frac{\epsilon_m^{\text{logit}}}{p_m(1-s_m)}$. Our calibration sets both $\bar{\epsilon}^{\text{logit}}$ and η equal to $\bar{\epsilon}^{\text{BLP}}$, implying that the ratio $\frac{\alpha p_m}{\eta}$ equals $\frac{1}{(1-s_m)}$. With the very small s_m associated with a 91% outside good share, the average value of $\alpha p_m/\eta$ is close to one. As a consequence, the β heterogeneity cross elasticity compared with both types of homogeneous tastes assumptions has the same sign and is roughly proportional to the covariance of probabilities. A further implication of the calibration equating average *own* price elasticities is that the average of the ratio of *cross*-price elasticities in the two homogeneous consumer models, $\epsilon_{jm}^{\text{logit}}/\epsilon_{jm}^{\text{CES}}$, will also be close to one.

4.2 Niche market power (own-elasticities)

The first implication relates to cross-price elasticities, and how models with homogeneous consumers will fail to account for the fact that the response in the demand for "proximate" varieties will be stronger for a given variety's increase in price. Introducing consumer heterogeneity in their preference for characteristics however presents a further challenge: it also changes the own price elasticity for each model m.

Implication 2. *Variance in household probabilities lowers own-price elasticities.*

Heterogeneity in the coefficient on product attributes and prices leads to consumers differing in their probabilities of choosing a model. This divergence in turn lower the own price elasticity for each car model compared to simple logit (or CES). We refer to this effect as niche market power. Own-price elasticities in mixed logit are

$$\frac{\partial \ln s_m}{\partial \ln p_m} = -\frac{p_m}{s_m} \times \frac{\sum_h \alpha_h \mathbb{P}_{mh}(1-\mathbb{P}_{mh})}{N}. \tag{14}$$

As with implication 1, niche market power is a consequence of both dimensions of consumers' heterogeneity, but exposition is simpler when restricting to the β heterogeneity case. Setting $\alpha_h = \alpha$, we obtain

$$\left.\frac{\partial \ln s_m}{\partial \ln p_m}\right|_{\alpha_h=\alpha} = -\alpha p_m \left(1 - \frac{\sum_h (\mathbb{P}_{mh})^2/N}{s_m}\right)$$

Let $V_m \equiv \sum_h (\mathbb{P}_{mh} - s_m)^2/N = \sum_h (\mathbb{P}_{mh})^2/N - s_m^2$ be the variance, for a given m of the household choice probabilities ($V_m = 0$ if $\beta_h = \beta$). Now the own price elasticity simplifies to

$$\left.\frac{\partial \ln s_m}{\partial \ln p_m}\right|_{\alpha_h=\alpha} = -\alpha p_m (1 - s_m - V_m/s_m).$$

This result is not specific to logit and the equation above holds for mixed CES as well (with the V_m redefined as the income-share weighted variance of $\mathbb{P}_{mh}$).

Dividing by $-\alpha p_m(1-s_m)$, the homogeneous counterpart of own price elasticity, the shrinkage of the own price elasticity due to β heterogeneity is given by

$$\frac{\epsilon_m^{\beta\text{het}}}{\epsilon_m^{\text{logit}}} = 1 - \frac{V_m}{s_m(1-s_m)} \leq 1, \tag{15}$$

with $\epsilon_m^{\beta\text{het}}$ and $\epsilon_m^{\text{logit}}$ being defined as $-\partial \ln s_m / \partial \ln p_m$, under β heterogeneity and logit cases respectively.

4.3 Non-unitary pass-through

The last, and quantitatively most important, implication relates to pass-through of cost changes into prices. Indeed, even assuming that the researcher can overcome Implication 2 and estimate the correct own price elasticity, the final effect on sales also depends on how the policy experiment (a cost shock) translates into prices. In this specific context, we want to know the elasticity of market shares with respect to tariffs:

$$\frac{\partial \ln s_m}{\partial \ln \tau_m} = \frac{\partial \ln s_m}{\partial \ln p_m} \times \frac{\partial \ln p_m}{\partial \ln c_m} \times \frac{\partial \ln c_m}{\partial \ln \tau_m}. \tag{16}$$

The first factor is the demand own-price elasticity. Under CES-MC, $\frac{\partial \ln s_m}{\partial \ln p_m} = -\eta$. The second factor is the pass-through elasticity (PTE). Under the constant markups assumed by CES-MC, $\frac{\partial \ln p_m}{\partial \ln c_m} = 1$. The last term is exactly one under the iceberg cost assumption (ad valorem tariffs) that we maintain throughout. Trade economists working with CES-MC therefore impose $\frac{\partial \ln s_m}{\partial \ln \tau_m} = -\eta$. However, if the DGP is mixed logit, true pass-through deviates from one, and the trade cost elasticity can be larger or smaller than the demand own-price elasticity.

Implication 3. *Logit demand without random coefficients has a pass-through elasticity strictly less than one, but random coefficients on prices can raise the pass-through elasticity over one. CES with monopolistic competition constrains the pass-through elasticity to be one.*

With multi-product firms, the calculation for the pass-through elasticity is too messy to be informative. Fortunately, in the single-product firms case, there is a very compact result, similar to one shown by Bulow and Pfleiderer (1983), that provides intuition on how demand curvature matters. Let ϵ and E be the own price elasticity ($\epsilon_m \equiv -\ln s_m/\partial \ln p_m > 0$) and super-elasticity ($E_m \equiv \partial \ln \epsilon_m / \partial \ln p_m$). Then the PTE is given by

$$\underbrace{\frac{\partial \ln p_m}{\partial \ln c_m}}_{\text{PTE}} = \underbrace{\frac{\partial p_m}{\partial c_m}}_{\text{PTR}} \times \frac{c_m}{p_m} = \frac{\epsilon_m}{\epsilon_m - 1 + E_m} \times \frac{\epsilon_m - 1}{\epsilon_m} = \frac{\epsilon_m - 1}{\epsilon_m - 1 + E_m}, \tag{17}$$

where PTR denotes the pass-through rate ($\frac{\partial p_m}{\partial c_m}$). Since $\epsilon_m > 0$, pass-through elasticities exceed one if and only if $E_m < 0$. Homogeneous logit has $E_m = 1 + \alpha p_m s_m = 1 - \epsilon_m s_m/(1 - s_m) > 0$ and hence $\partial \ln p_m / \partial \ln c_m < 1$. As the s_m become small (for example when the outside good has a high share), $E_m \to 1$, PTR $\to 1$, and PTE $\to (\epsilon_m - 1)/(\epsilon_m) < 1$, that is the inverse of the markup formula. On the other hand in CES monopolistic competition $E_m = 0$, so PTR $\to \epsilon/(\epsilon - 1) > 1$ and the PTE is one.

In mixed logit, the super-elasticity is given by

$$E_m = 1 + \epsilon_m - p_m \frac{\sum_h \alpha_h^2 \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})(1 - 2\mathbb{P}_{mh})}{\sum_h \alpha_h \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})}, \tag{18}$$

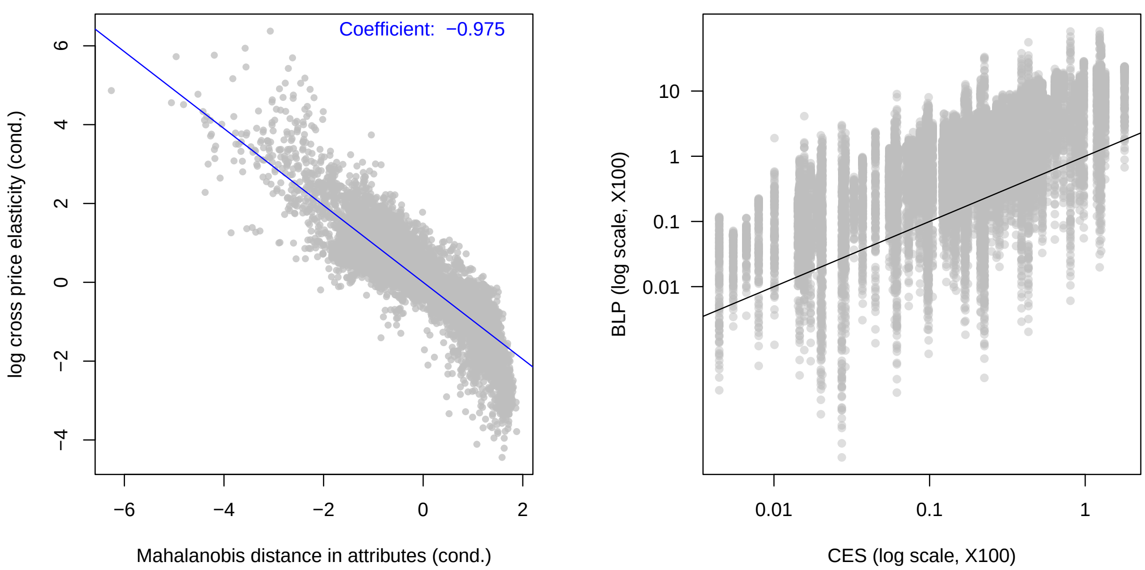
which is ambiguous in sign. With levels of α-heterogeneity across households implied by BLP estimates, we will see that the super-elasticity is negative and pass-through elasticities are greater than one (in the single product case).

4.4 Three implications illustrated

Figures 1 and 2 and Table 2 illustrate the quantitative relevance of the three implications in the context of the data set and parameter estimates of Berry et al. (1995, 1999). The figure and table contents are generated from one run of the BLP Data Generating Process drawing 100,000 consumers and using the parameters and data for 1990 described in Table 1.

Panels (a) and (b) of figure 1 display the rich substitution patterns involved in Implication 1. Equation (11) shows that the cross-price elasticity, $\frac{\partial \ln s_j}{\partial \ln p_m}$, is proportional to price m, and inversely proportional to market share s_j. We remove those effects by first computing the cross elasticities using the original data and estimated parameters and then regressing the log cross elasticity on fixed effects to capture the j and m terms. The residual from this regression is graphed against the dissimilarity in the characteristics vector, measured with the Mahalanobis distance in terms of the four x^k characteristics and price. As with the log cross-price elasticity, we purge the Mahalanobis distances of m and j effects by taking residuals from a fixed effects regression. The scatter plot reveals a striking fit: the within R^2 is 77%.

Figure 1: Cross-price elasticities in BLP data



(a) Cross elasticity and dissimilarity (b) Cross elasticities compared

Table 2 reports the coefficients of this regression in the third column, each of the rows corresponding to different degrees of consumer heterogeneity. With all sources of heterogeneity active, the coefficient is -0.98, while the relationship between cross elasticities and the distance in varieties' characteristics less steep at -0.53, but with an even larger fit at 90%. Since homogeneous logit predicts a zero slope and therefore a zero R^2, we see this regression as a useful way to quantify the amount of rich substitution conditional on a set of attributes and parameter estimates.

Table 2: BLP Data Generating Process: Key moments

	Elasticities		Rich substitution		
	ϵ_m	E_m	Maha.	$\frac{\text{cov}(\mathbb{P}_{jh},\mathbb{P}_{mh})}{s_j s_m}$	$\frac{V_m}{s_m(1-s_m)}$
Setting	Avg.	Avg.	Coef.	Avg.	Avg.
Mixed Logit	4.05	-0.49	-0.98	10.36	0.017
Logit	4.05	1.00	0.00	0.00	0.000
β het.	4.02	1.03	-0.53	7.09	0.010

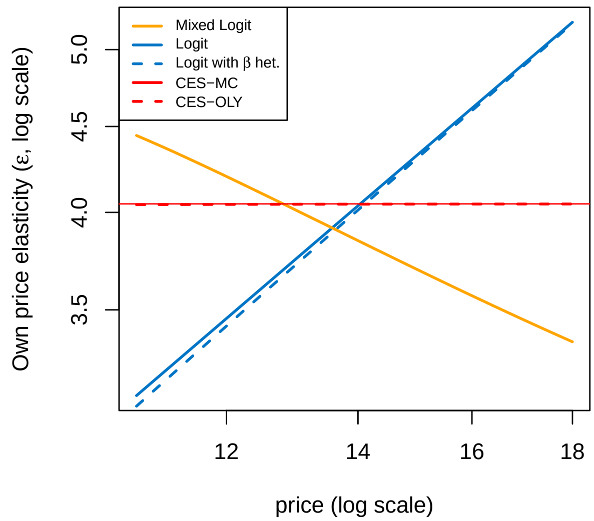
Note: ϵ_m is the (opposite of) the own price elasticity and E_m is the super-elasticity (the elasticity of ϵ_m wrt p_m, with formula given by (18). Maha. Coef. is the slope in a regression of the log of $\frac{\partial \ln s_m}{\partial \ln p_j}$ on D_{mj}, the Mahalanobis distance between characteristics of car models m and j. The "Mixed Logit" row is the original version of the model where both types of consumer heterogeneity are active. "Logit" sets $\alpha_h = \alpha$ (holding avg own price elas constant) and $\beta_h = \beta$. "β het." only sets $\alpha_h = \alpha$.

Depending on the sign of the covariance between household choice probabilities, equation (13) shows that the β heterogeneity cross-price elasticity can be higher or lower than the CES corresponding elasticity. Figure 1(b) illustrates the cross-elasticity comparison using the full BLP model including α heterogeneity. The intuition from equation (13) carries through, with BLP elasticities distributed on both sides of the 45-degree line representing equality with the CES approximation. An example of model pairs with an order of magnitude higher cross-elasticity than the CES is the Geo Metro and the Ford Escort. In the reverse direction, an increase in the price of the Yugo GV Plus has a tiny fraction of the cross elasticity with the Mercedes 560 under BLP as it does in CES (though both elasticities are very small due to the small share of the Yugo). The fourth column of Table 2 reports average value of the scaled covariance term $\frac{\text{cov}(\mathbb{P}_{jh},\mathbb{P}_{mh})}{s_j s_m}$ as 10, with all types of heterogeneity, and 7, when considering physical attributes only. The high scaled covariances imply, via (12), that cross elasticities average 8 to 11 times larger with heterogeneous consumers than for logit.[13]

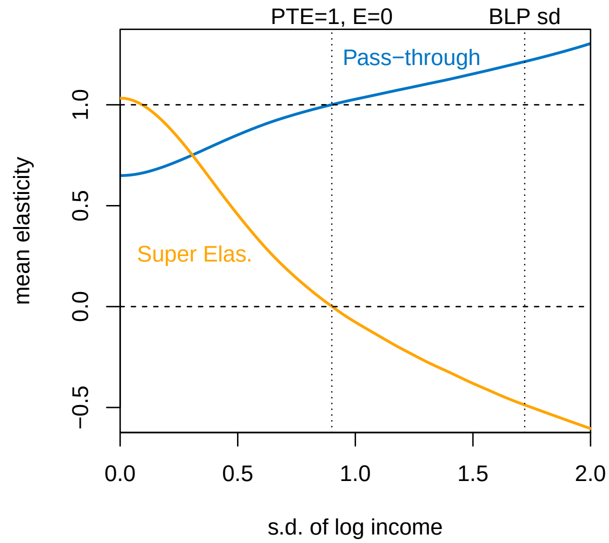
Figure 1 shows the "cross-section" of cross-price elasticities, i.e. evaluated for each pair of car models at the values of market shares and prices featuring in the original BLP data. In order to graphically represent Implication 3, we need to trace how the equilib-

[13]The same is true when compared to CES: As explained in section 4.1, when parameters α and η are calibrated to yield the same average own-price elasticity, the two versions of homogeneous consumer cross-price elasticities have approximately the same average values.

Figure 2: Own price elasticities



(a) Volvo 240 super-elasticity (E)



(b) Mean E and PTE

rium own price elasticity varies when price changes. Varying the assumptions regarding consumer heterogeneity makes the exercise also useful for evaluating the quantitative importance of Implication 2. In practice, we run a sequence of simulations in which we change the price of one car model, and then recompute the whole set of market shares, holding other prices fixed. This allows to calculate the implied new own price elasticity for the car model studied. While this is doable for all 131 car models in the data, the graphical display (in Figure 2(a)) is much clearer when focusing on one. The car model chosen (the 1989 Volvo 240) has a benchmark own price elasticity quite close to the average variety in the original BLP settings (4.05 as stated in the first column of Table 2). Starting with those settings, we then evaluate the own-price elasticity (14), varying the price of this car model by a range from -25% to $+25\%$ of the actual 1990 price. This evaluation involves recomputing household probabilities to buy each variety and therefore all car models' market shares. The results are displayed in Panel (a) of Figure 2. The downward-sloping orange line shows the case of full BLP. In that case, the log of own price elasticity falls with the log of price, implying a negative super-elasticity.

How general is the negative super-elasticity shown for the Volvo model in orange in Figure 2(a)? Figure 2(b) investigates this question by tracing out the average (across the 131 models) superelasticity as the standard deviation of log income (the source of heterogeneity in α_h) rises from zero to two.[14] Equation (17) tells us that in a single-product world, the $E_m < 0$ will lead to super-pass-through (PTE > 1). The blue line in panel (b),

[14]The two curves in Figure 2(b) hold the mean ϵ_m constant at 4.05 by adjusting α.

drawn using equation (17), shows that the average pass-through elasticity exceeds one at a standard deviation of 0.9, where the mean E becomes negative. This points to the generality of super-pass-through in random coefficients logit since the threshold amount of heterogeneity is well below the standard deviation of 1.72 used in Berry et al. (1999). A second aspect of generality is that the single-product predictions in the figure carry over to the multi-product firm data of BLP as shown in Table 3, where the average PTE of the tariff counterfactual is 1.13 (with an average super-elasticity of -0.49).

Returning to panel (a) of figure 2, the solid blue line illustrates the own-elasticity versus price relationship for the case of simple logit demand, i.e. canceling all sources of consumer heterogeneity (and adjusting α such that the average own price elasticity is the same as in the full BLP setup). The slope is *positive* as predicted by theory (where $E_m = 1 + \alpha p_m s_m$) and close to one, the limiting value as market shares go to zero. The dashed blue line adds β-heterogeneity to consumer behavior, which illustrates Implication 2: own price elasticities are systematically smaller in that case due to the niche market power. However, because market shares are so small in the BLP data (since the outside good share is 91%), the difference is quantitatively negligible.[15] Table 2 confirms that super-elasticities that are positive and very close to one are a feature of the logit model, even when allowing for heterogeneous tastes for attributes *other than price*.

Lastly, figure 2(a) also illustrates the two CES approximations used in the paper. First, the CES-MC case, with its continuum of negligible firms, gives a constant elasticity (η) chosen here to be at the average level of the BLP data (4.05). This is represented in solid red. In dashed red, we account for the fact that, with non-negligibly sized firms, the CES elasticity is $\eta(1-s_m)$, i.e. declining with the market share of model m. With low prices, this share increases and the own price elasticity falls. This is true in panel (a)'s representation, although the effect is very small quantitatively, again because of the very small market shares of all varieties in the data. Even without rich substitution, the positive super-elasticity implies that logit will have a very different pass-through from CES, a feature which will prove important in our simulations' results.

4.5 Benchmark counterfactual, known CES parameter

The first experiment we conduct asks a simple question: Would a CES monopolistic competition approximation of the US car industry be able to predict the response to a change in trade policy for data generated by mixed logit multi-product oligopoly? To pinpoint

[15] This generalizes the result of equation (15), where the ratio of β heterogeneity over logit own price elasticities is driven by $\frac{V_m}{s_m(1-s_m)}$, which has an average value of 0.01 (Table 2).

the role of functional forms, we first sidestep the issue of how to estimate the CES and simply assume we already know it to be 4.05 (the average own price elasticity coming from the BLP parameter and data). Keeping all parameters and data as in the original Berry et al. (1995) study, we then impose a 10% tariff on the foreign models, solve for the new BLP equilibrium and compare changes in outcomes to changes predicted by the CES approximation.

Table 3: Counterfactual 10% tariff using the BLP data

	Quantity shares		Pass-through		
	Agg. ΔS		rate	elasticity	
Setting	True	EHA	Avg.	Avg.	# 1
Mixed Logit	8.00	7.73	1.57	1.13	1.12
Logit	3.85	7.73	1.00	0.67	0.62
β het.	3.33	7.73	0.98	0.65	0.57

Note: CES EHA uses $\eta = 4.05$. ΔS is the change in aggregate share of domestic models in the new car market. The pass-through rate is the avg. $\Delta p_m/\Delta c_m$ evaluated for the foreign firms in the DGP. The elasticity averages the rate divided by the markup p_m/c_m. #1 gives the elasticity for the best-selling import variety. The "Mixed Logit" row is the original version of the model where both types of consumer heterogeneity are active. "Logit" sets $\alpha_h = \alpha$ (holding avg. own price elas. constant) and $\beta_h = \beta$, whereas "β het." only sets $\alpha_h = \alpha$.

The first line of Table 3 implements this counterfactual increase in tariffs imposed on all foreign cars. The "true" change in the domestic firms' collective share of the market for new cars is reported in the first column. The 10% tariff increases the domestic share by 8.00 percentage points (to 76%). As the CES approximation predicts a change of 7.73, the error is about one quarter of one percent. This extremely close fit is surprising in several respects. The CES approximation makes three deviations from the true DGP: 1) monopolistic competition rather than oligopoly, 2) a wrong functional form of demand (CES versus logit), 3) homogeneous consumers. We investigate the two first deviations (market structure and functional form) in the next section and focus here on the role of heterogeneity.

The second line of Table 3 (Logit) imposes homogeneity in consumer tastes. We calibrate α such that all consumers have the same price elasticity as the average one in the first line (BLP). The CES prediction remains the same (7.73 pp) since it still works with a price elasticity of 4.05. However, the true counterfactual falls drastically to 3.85pp. This comes from the fact that the logit demand system implies a pass-through *elasticity* substantially below one: here every one percent increase in costs by foreign firms triggers a

0.67 percent price increase.[16] Domestic firms therefore gain much less market share than in the first line (the BLP case), where the pass-through elasticity is close to one (the value predicted by the CES-monopolistic competition model).

We further investigate the role of heterogeneity in the third line (β heterogeneity). This is a hybrid case, as it imposes a single own-price effect α, but lets the β_h^k coefficients on the four physical car characteristics (as well as preference for the outside good) vary across households. The presence of β heterogeneity leads to a slight deterioration of the accuracy of the CES approximation as compared to logit. The pass-through elasticity is slightly lower (0.65 vs 0.67) than in the logit case and therefore exacerbates the deviation from CES. The rise in the bias from 3.88 to 4.40 also highlights Implication 1: the imposition of symmetric substitution patterns damages the quality of the CES approximation. This occurs because under β heterogeneity, the rising price of foreign cars leads to less substitution towards the outside good.[17] Thus, the new car market shrinks less under β heterogeneity because the foreign varieties do not fare as badly. The net result is a smaller increase of the share of domestic varieties as a share of new cars (the inside good).

The accuracy of the fit in the benchmark (BLP) case comes from a countervailing effect of α heterogeneity. When the price sensitivity of consumers is heterogeneous enough, a rise in prices triggers selection of consumers, such that only the less price sensitive ones continue to buy the most expensive varieties. In the BLP DGP this effect is so strong that it reverses the way own-price elasticity varies with price.[18] This raises the incentive to pass more of the tariff increase into final prices. This effect is so strong in the BLP data and estimates that the average pass-through elasticity, 1.13, is slightly larger than one, bringing it closer to the CES-MC assumption.

A potential concern is that our counterfactual imposes a tax based upon on a characteristic of the product, i.e. its foreign status, while this characteristic was not included in the original analysis of Berry et al. (1995). The omission of the mean effect for domestic/foreign cars has no effect since the unobserved demand shock ξ_m would capture this effect. However, a random coefficient on domestic status would alter patterns of substitution in a way that might be expected to undermine the performance of the CES approximation.

Some aspects of foreignness were already being incorporated through differences in

[16] Appendix C shows that logit models with small market shares have unitary pass-through *rates*. The elasticity divides the rate by the markup.

[17] The fact that random coefficients make consumers much less likely to switch to the outside good than in a homogeneous logit model is quantified in (Berry et al., 1995, Table VII).

[18] Birchall and Verboven (2022), examining the breakfast cereal industry, find that the random coefficients model yields a non-monotonic relationship between elasticity and price.

the observed attributes of foreign vs domestic varieties. Table 4 shows that the chief difference would come through higher-priced foreign cars, since the means of the other characteristics are very similar. Consumer tastes for domestic brands could be heterogeneous, even taking into account the fact that foreign cars are more expensive. As with other random coefficients (see our Implication 1), the greater the variance of tastes, the greater will be substitution within the import status relative to substitution between domestic and imported varieties. The CES approximation may therefore overestimate the effect of tariffs, even when it is calibrated to the correct average average own-price elasticity.

Table 4: Characteristics of car models in BLP data

Origin	count	price	HP/weight	air	miles/USD	size
Domestic	67	12.03	0.42	0.43	2.69	1.31
		(6.44)	(0.08)	(0.5)	(0.68)	(0.17)
Foreign	64	16.14	0.48	0.48	2.79	1.2
		(11.48)	(0.1)	(0.5)	(0.58)	(0.14)

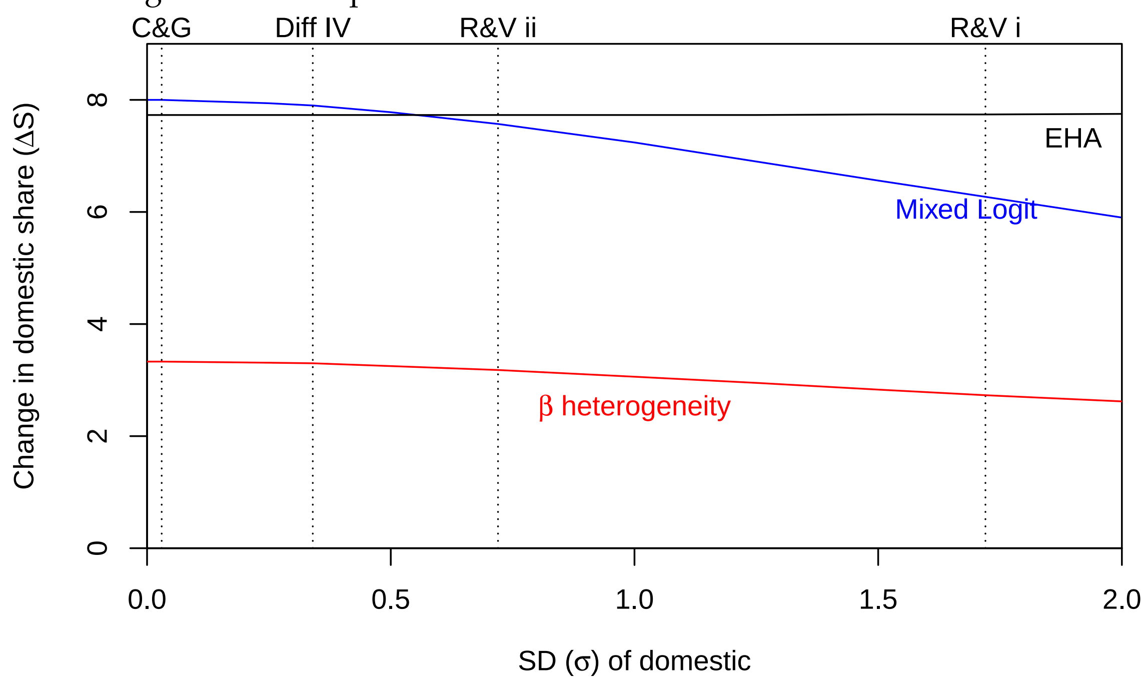
Note: 1990 average levels of prices (in $1000) and the four model characteristics used in Berry et al. (1999), with standard deviations in parentheses.

To investigate that potential concern, we rerun our simulations with a random coefficient on domestic status added to the other car characteristics. We then vary the standard deviation of the random coefficient on domestic status to encompass the plausible range of estimates. The key results are shown in figure 3, with details relegated to appendix H. Figure 2 plots ΔS^{BLP} and ΔS^{CES} against the standard deviation of the random coefficient on domestic status. Using vertical dotted lines, we call attention to four estimates, two that we conducted using the `PyBLP` software made available by Conlon and Gortmaker (2020), and two that are taken from Reynaert and Verboven (2014).

On the vertical axis (when evaluated at $\sigma = 0$), the heights correspond to the same change in aggregate domestic market share as shown in Table 3: 8 p.p. for the full BLP model (in blue), 7.73 for the CES approximation (in black), and 3.33 for the β-heterogeneity setting (in red). Both blue and red lines are monotonically decreasing as expected, when σ increases. For very large degrees of consumer heterogeneity regarding domestic status, the approximation starts to deteriorate substantially. However, CES continues to fit very closely the true DGP of BLP for the estimated values of σ using BLP data.

In the next section we proceed to a more complete investigation of the surprisingly good fit of CES to full BLP, where we vary all the relevant dimensions in sequence. Another important difference is that the counterfactuals we report in Table 3 assume the researcher knows the average own-price elasticity. In contrast, counterfactuals in the next section take the standard approach of trade economists, and use tariff variation to *estimate*

Figure 3: The impact of a random coefficient on Domestic status



Note: The vertical dotted lines correspond to estimates of the standard deviation of the random coefficient on domestic status of the car model. "C&G" refers to our estimate using the best practices recommended by Conlon and Gortmaker (2020). "Diff. IV" uses the differentiation instruments. Those two estimates are in columns 3 and 4 of Appendix Table H3. "RV i" and "RV ii" come from Reynaert and Verboven (2014) Table 6 (Optimal instruments panel, columns i and ii).

the elasticity of market shares to cost shocks.

5 Dissection via simulation: what makes CES work?

Our dissection exercises set up a simulated version of the BLP data generating process that is sufficiently close to be a valid representation of the original version, while having the flexibility needed to dig into the causes of the failures or successes of the CES approximation. Another important component of our approach is to bring it closer to the actual questions and methods of trade economists in that we estimate the key cost elasticity parameter ($\breve{\eta}$) required to run counterfactuals on trade costs variation.

5.1 Benchmark settings using estimated tariff elasticities

Our benchmark simulations involve the following steps:

1. Sample 90 varieties m from the original BLP data with their four observed attributes x_m, together with their unobserved quality, ξ_m, and marginal cost, c_m, that we backed out using the inversion methods described in section 2.

2. Assign ten varieties to nine firms, with three firms in each of three countries.

3. Trade costs consist of an initial 10% tariff and an *ad-valorem* equivalent of distance between countries, d_{AVE}.

4. We calibrate three parameters to comply with three moments of the BLP data.

    (a) α is chosen to set the average own price elasticity to 4,

    (b) d_{AVE} sets the domestic share equal to 68% (the domestic variety share of the new car market in 1990 in the BLP data),

    (c) $\bar{\beta}^0$ is adjusted so that the outside good share is 90%.

5. Compute the initial BLP equilibrium. This starts with using the first order condition (3) to solve for prices, followed by (2) and (1) to obtain equilibrium market shares s_{mn} in each country n.

6. Estimate the tariff elasticity, $\breve{\eta}$, using equation (8).

7. We then raise the tariff on foreign cars by 10 percentage points and compute new prices and ensuing $s^{\dagger}_{mn}$, i.e. the new market shares for all firm-destination combinations in the new equilibrium.

8. Compute s'_{mn}, the EHA counterfactual prediction, using equation (6). Then aggregate over the domestic varieties in one country to compute ΔS^{CES}, which we compare to the true changes ΔS^{BLP}.

We repeat the steps above 1000 times, reporting averages and standard deviations in the next subsection.

To investigate which features of the BLP initial setup make the CES approximation succeed or fail, we consider four modifications of the benchmark simulation described above:

Aggregation: Sum the market shares of varieties from a common origin country to create s_{in}, which we use to estimate (9). The aggregate approximation uses equation (10). This reflects the common limitation faced by researchers lacking firm/variety data.

Reduced outside good share: Increase the mean β_h^0 to generate smaller shares (50%, 10%) of the outside good. This leads to higher market shares for the inside firms and thus more scope for oligopoly forces.

Mixed CES: Each household spends y_h on a preferred vehicle, with household choice probabilities being the same as equation (2) except $-\alpha_h p_m$ is replaced by $-\alpha_h \ln p_m$. In this specification s_m is measured in values instead of quantities. In the enumerated list describing the DGP, the same steps are involved except the two computations of equilibrium (steps 5 and 7) use the mixed CES equations to solve for the equilibrium. Appendix section A gives a complete description of this setup.

Oligopoly estimation and EHA: Even without random coefficients, standard EHA is incorrect because it does not capture the variable markups of oligopolists. This can be fixed by modifying estimation to equation (A6) and adjusting EHA to the oligopoly case shown in equation (A7)—both equations being displayed in Appendix A.

5.2 Benchmark case and decreasing outside good share

The benchmark results, depicted in the third line of Table 5 and in Figure 4(a), show that, under BLP heterogeneity settings, EHA continues to predict Mixed Logit tariff counterfactuals accurately. The CES approximation overpredicts the change in domestic market share by only one third of a percentage point.[19] A fundamental difference from the simulation reported in Table 3 is that we now estimate η rather than assuming the average

[19]Even though this simulation samples from the underlying car models and allocates them to nine firms in three countries, it still retains the market structure of the original data: the average concentration ratio (an untargeted moment) is 84% on average in our simulations, just below the 86% in the original data.

own-price elasticity is known. The cross-country tariff variation in equation (8) estimates $\breve{\eta} = 4.29$ on average. This is larger than the calibrated own-price elasticity of 4 because α heterogeneity causes firms to pass on to consumers more than 100% of their costs increases. The average pass-through rate and elasticity reported in the last two columns of Table 5 are 1.65 and 1.14, respectively. This is because α heterogeneity creates a force that selects consumers according to their individual elasticity, raising the pass-through elasticity from around two thirds to a level just over unity.

As a consequence, the CES approximation works *better* with the BLP full dimensions of consumer heterogeneity than in cases of no heterogeneity or only β heterogeneity (first two lines of Table 5). This is the same pattern as depicted in Table 3. However, the bias of the CES approximation now falls to less than one percentage point as opposed to about four in the BLP data counterfactuals of Table 3. The primary reason is that the estimated $\breve{\eta}$ (shown in the fourth results column) falls to 2.64 and 2.26 in those cases respectively. Recall from equation (16) that $\breve{\eta}$ combines the effects of both own-price ($\frac{\partial \ln s_m}{\partial \ln p_m}$) and pass-through ($\frac{\partial \ln p_m}{\partial \ln c_m}$) elasticities. By capturing the much lower pass-through implied by logit demand, the estimation step gives the CES approximation greater flexibility to fit the underlying true data generating process. Rich substitution under the form of β heterogeneity lowers the fit of the approximation, but the first order issue is the functional form of demand.

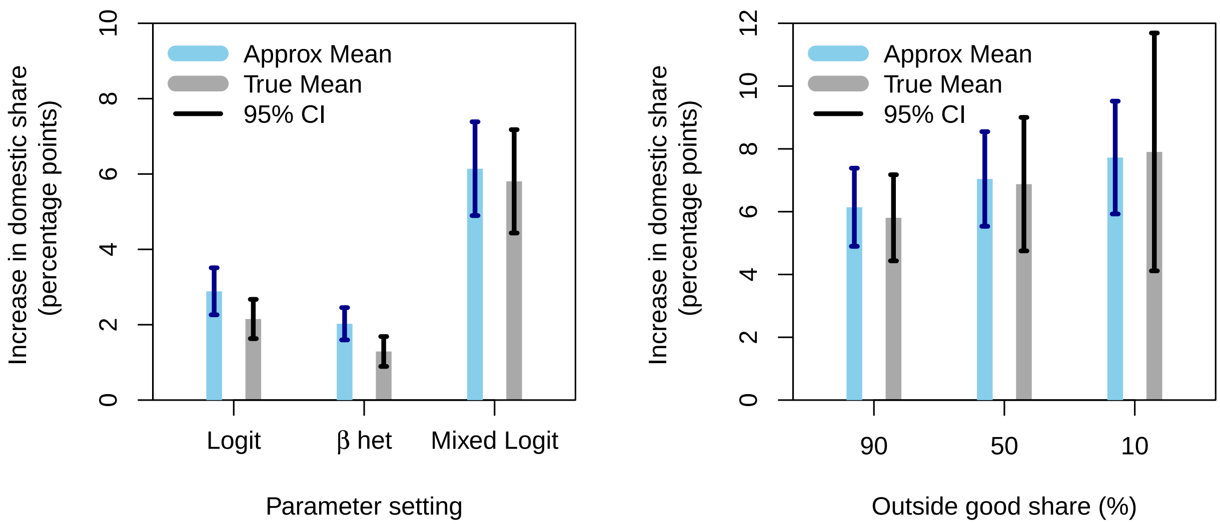
Table 5: The role of consumer heterogeneity assumptions

Setting	ΔS domestic varieties			Trade elas ($\breve{\eta}$)		Passthrough	
	True	CES Approx					
		micro	macro	micro	macro	rate	elas
Logit	2.15	2.89	2.40	2.64	2.19	0.99	0.68
β heterogeneity	1.29	2.02	1.48	2.26	1.65	0.98	0.66
Mixed Logit	5.81	6.14	6.35	4.29	4.45	1.65	1.14

Note: Each row applies different heterogeneity settings to Logit demand. As in the BLP data, the share of outside goods is calibrated to 90%, the share of domestic cars is 68%, and we set average ϵ_m to 4 for 1000 repetitions. The simulation has 9 firms with 10 models each; the combined market share of the top 5 firms is 84% (as compared to 86% in BLP).

Table 5 shows results using the aggregated gravity estimation described in section 3. The third column reports the aggregate change in domestic market share, with the corresponding tariff elasticity—obtained from aggregated gravity equation (9)—reported in the fifth column. In the benchmark Mixed Logit case, the macro estimate of η is large, therefore amplifying the overprediction of the micro-based CES approximation (6.35 vs 6.14, when the true change is 5.81). In the other cases, the macro gravity estimate is smaller than when using micro flows. Since the micro ΔS^{CES} overshoots, the macro approach has the surprising effect of improving the fit of the approximation.

Figure 4: Benchmark case and outside good share



(a) Benchmark case (b) Increasing oligopoly

Note: As in the BLP data, the share of domestic cars is 68%, and we set average ϵ_m to 4. The simulation has 9 firms with 10 models each; the combined market share of the top 5 firms is 84% (as compared to 86% in BLP). The error bars are 1.96 standard *deviations* of the simulation outcomes for 1000 repetitions. Panel (a) calibrates the OG share at 90%. OG shares of 50 and 10 in panel (b) yield larger market shares for the inside firms.

Table 6: Decreasing the share of the outside good (OG)

OG	Agg. ΔS			Passthrough	
(%)	True	Approx	$\breve{\eta}$	rate	elas
90	5.81	6.14	4.29	1.65	1.14
50	6.88	7.04	4.78	1.78	1.19
10	7.90	7.72	5.14	1.90	1.21

Note: 1000 repetitions. The demand system is mixed logit with both dimensions of heterogeneity in all 3 lines. OG shares of 50 and 10 yield larger market shares for the inside firms.

Intuitively, the 90% outside good share in the BLP data should contribute to the good performance of the monopolistic competition assumption used in the CES-MC counterfactuals.[20] Would CES-MC work as well for industries dominated by "inside" goods?

[20] The outside good share is so large in the baseline that it essentially wipes out oligopoly forces in the DGP. One way to verify this is to modify the DGP to allocate each of the 90 car models to individual firms. This moves the "true" prediction close to monopolistic competition. The results are so similar to the benchmark case shown in figure 4(a) that we relegate them to the appendix figure F1(b) and table F2(b).

Figure 4(b) shows that as we decrease the outside good share (from 90% on the left to 10% on the right), there is a greater increase in domestic market share, both for the true (gray) and CES-approximated (blue) outcomes. In the case of BLP, this is because more oligopoly power induces firms to adjust markups more, thus passing through a higher multiple of the tariff increase.[21] The pass-through elasticity can be seen to rise in Table 6 from 1.14 to 1.21. In the Exact Hat Algebra, the higher change comes from a larger estimated η; the tariff elasticity rises from 4.29 to 5.14. The true outcome rises faster than the approximation, with CES first over-predicting and then under-predicting, but never more than a fifth of a percentage point.

5.3 Mixed CES and oligopoly estimation

So far we have seen that CES-MC can approximate the aggregate predictions of the BLP DGP quite precisely. Essentially, high variance in the α_h price coefficients "solves" the problem that the logit form leads to low pass-through. It is therefore illuminating to move to random coefficients (and hence rich substitution) while holding the functional form of household choice probabilities constant. We present the mixed CES results in Table 7 and figure 5. The mixed CES has the advantage of containing homogeneous CES as the limiting case when the variance of the β_h and α_h go to zero. In this case the only difference between the DGP and the CES approximation is the latter assumes monopolistic competition.

In the first line of Table 7, the CES-MC approximation is almost perfect (up to rounding). This is because a market share of 90% for the outside good leaves little room for oligopoly to make a noticeable difference. The micro and macro estimates of η are almost identical, allowing the aggregate counterfactual to fit extremely well. As before, adding β-heterogeneity worsens the prediction, but now instead of improving the fit, α-heterogeneity exacerbates the problem. However, the main takeaways from panel (a) of Figure 5 are the stability of the BLP outcomes and the accuracy of the CES approximation across all three heterogeneity settings.

Our last dissection investigates whether CES can predict the BLP outcome better if the estimations and Exact Hat Algebra calculations are modified to account for oligopoly (CES-OLY).[22] Recall that in Table 7 (last row) the CES approximation has an upward bias of about one percentage point. That is for the benchmark case where the inside firms

[21] The higher amount of markup adjustment as the inside good market shares increase is a general feature. However, without α heterogeneity, the adjustments would be *downward*, leading to less complete pass-through and lower aggregate changes.

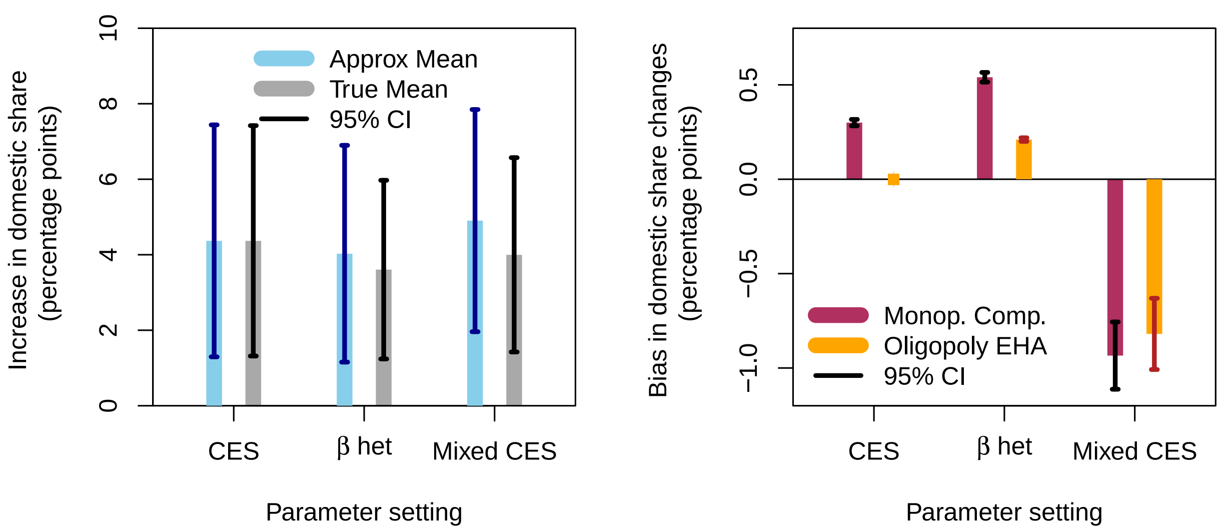
[22] Specifically, we estimate equation (A6) and use equations (A7) and (A8) for Exact Hat Algebra.

Table 7: Mixed CES results

Setting	ΔS domestic varieties			Trade elas ($\breve{\eta}$)		Passthrough	
	True	CES Approx					
		micro	macro	micro	macro	rate	elas
CES	4.37	4.37	4.34	3.96	3.94	1.25	1.00
β heterogeneity	3.61	4.03	3.71	3.78	3.46	1.23	0.98
Mixed CES	4.00	4.90	4.75	4.13	4.02	1.44	1.04

Note: Each row applies different heterogeneity settings to CES demand. As in the BLP data, the share of outside goods is calibrated to 90%, the share of domestic cars is 68%, and we set average ϵ_m to 4 for 1000 repetitions. The simulation has 9 firms with 10 models each; the combined market share of the top 5 firms is 84% (as compared to 86% in BLP).

Figure 5: Mixed CES and the oligopoly correction



(a) Mixed CES (b) Oligopoly est/EHA

Note: 1000 repetitions. As in the BLP data, the share of outside goods is calibrated to 90%, the share of domestic cars is 68%, and we set average ϵ_m to 4. The error bars in panel (a) are 1.96 standard *deviations* of the simulation outcomes, but those in panel (b) are standard *errors* of the bias.

Table 8: Adapting estimation & EHA to oligopoly helps

Setting	Agg. ΔS True	Agg. ΔS Approx	Trade elas ($\breve{\eta}$)	Passthrough rate	Passthrough elas
Monopolistic competition approximation					
CES	3.74	4.04	3.64	1.23	0.97
β heterogeneity	3.48	4.02	3.62	1.23	0.97
Mixed CES	8.41	7.47	4.95	1.98	1.28
Oligopoly approximation					
CES	3.75	3.75	4.00	1.23	0.97
β heterogeneity	3.52	3.73	3.96	1.23	0.97
Mixed CES	8.41	7.59	5.75	1.98	1.28

Note: 1000 repetitions. The outside good share is calibrated to 10% (instead of the 90% in the BLP data). All settings are calibrated to hold the average brand-level own price elasticity at 4.

have only a 10% share of the market. Here we maximize the role of oligopoly forces by using the lowest setting for the outside good, 10%. Figure 5(b) shows the average bias (the difference between the blue and gray lines in the preceding figures) in each heterogeneity setting.[23] The CES-OLY counterfactual predicts perfectly with homogeneous consumers, correcting the upward bias in monopolistic competition. With β heterogeneity, the oligopoly adjustments on the estimation and counterfactual calculation reduce bias without fully eliminating it.

The oligopoly adjustment offers the lowest improvement in the setting with α heterogeneity. As seen in Table 8, the reason for this is that CES-OLY estimates a larger $\breve{\eta}$ (5.75 instead of 4.95), which is going in the right direction because the pass-through elasticity exceeds one. The EHA partially undoes this by imposing a change in markups that entails incomplete pass-through (since it assumes CES under oligopoly). Thus, the "mistake" that the monopolistic competition version of CES makes (omitting oligopoly markup adjustment) is actually helpful in the presence of large amounts of α heterogeneity. The bottom line here is that the oligopoly markup adjustment helps—but only when Marshall's second law of demand holds. A DGP with super-passthrough, such as the mixed CES case with high α heterogeneity, is hard for the CES oligopoly approximation to fit.

[23] Another change is that the error bars in this figure correspond to standard errors for the mean rather than standard deviations of outcomes as in the previous figures.

6 Conclusion

A recurrent need in policy analysis is to predict the outcome of a cost shock (such as a tax or subsidy) applied to a subset of competing products. Examples span the fields of trade (e.g., tariffs on imported cars), environment (e.g., subsidies for electric vehicles), and health (e.g., taxes on sugary drinks). Two major challenges are how to model market structure and patterns of substitution between alternatives. Our most striking finding is that a simple, and admittedly unrealistic, model—the constant elasticity of substitution under monopolistic competition—is able to closely approximate aggregate outcomes of the much richer go-to model of IO economists.

The CES approximation fits aggregate market share changes well when the pass-through elasticity for cost shocks has an elasticity near one. Fajgelbaum and Khandelwal (2022) survey the growing literature estimating this elasticity for tariff-generated costs shocks (the focus of our paper). While early work on sugar and trucks found low rates of tariff pass-through, five recent investigations of the price effects of the US-China trade war estimate pass-through elasticities near one in both directions. Given these mixed results in the literature, we recommend estimation of the pass-through elasticity as a diagnostic. A near unit elasticity suggests the CES approximation will provide a better fit in a given application.

When pass-through elasticities differ from one, the CES approximation can still work well with an estimated elasticity of how market shares respond to tariffs. There is something of a paradox here. When the DGP departs from CES monopolistic competition, the coefficient on tariffs no longer estimates the structural demand elasticity. Instead, the estimand becomes the mean of a product of two elasticities: demand and pass-through. This combination of parameters can be thought of as a correction that mitigates the pass-through problem in counterfactual predictions.

The results showing the good ability of the CES model to approximate models with rich substitution do not necessarily apply in all contexts. For example, a prominent use of the Berry et al. (1995) methodology is to compute the consequences of mergers. Flexible substitution seems likely to play a larger role in counterfactual merger analysis than in the cost shock analysis contained in this paper.

References

Adao, R., A. Costinot, and D. Donaldson (2017). Nonparametric counterfactual predictions in neoclassical models of international trade. *American Economic Review 107*(3), 633–89.

Adao, R., A. Costinot, and D. Donaldson (2022). Putting quantitative trade models to the test: Evidence from the Trump tariffs. mimeo.

Anderson, S., A. De Palma, and J. Thisse (1992). *Discrete choice theory of product differentiation*. MIT Press.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the Sensitivity of Parameter Estimates to Estimation Moments. *The Quarterly Journal of Economics 132*(4), 1553–1592.

Atkeson, A. and A. Burstein (2008). Pricing-to-market, trade costs, and international relative prices. *The American Economic Review 98*(5), 1998–2031.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2018). Global firms. *Journal of Economic Literature 56*(2), 565–619.

Berry, S. and P. Haile (2021). *Foundations of Demand Estimation*, Volume 4, pp. ?–? Elsevier.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (1999). Voluntary export restraints on automobiles: evaluating a trade policy. *American Economic Review 89*(3), 400–430.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy 112*(1), 68–105.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Birchall, C. and F. Verboven (2022). Estimating substitution patterns and demand curvature in discrete-choice models of product differentiation. Discussion Paper 16981, Centre for Economic Policy Research.

Björnerstedt, J. and F. Verboven (2016). Does merger simulation work? Evidence from the Swedish analgesics market. *American Economic Journal: Applied Economics 8*(3), 125–64.

Breinlich, H., H. Fadinger, V. Nocke, and N. Schutz (2021). Gravity with granularity. LSE Research Online Documents on Economics 114419, London School of Economics and Political Science, LSE Library.

Bulow, J. I. and P. Pfleiderer (1983). A note on the effect of cost changes on prices. *Journal of political Economy 91*(1), 182–185.

Coşar, K., P. Grieco, S. Li, and F. Tintelnot (2018). What drives home market advantage? *Journal of International Economics 110*, 135–150.

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with PyBLP. *The RAND Journal of Economics 51*(4), 1108–1161.

Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced trade. *American Economic Review 97*(2), 351–355.

Dingel, J. I. and F. Tintelnot (2021). Spatial economics for granular settings.

Dubois, P., R. Griffith, and M. O'Connell (2020). How well targeted are soda taxes? *American Economic Review 110*(11), 3661–3704.

Fajgelbaum, P. D. and A. K. Khandelwal (2022). The economic impacts of the US–China trade war. *Annual Review of Economics 14*(1), 205–228.

Gandhi, A. and J.-F. Houde (2016). Measuring substitution patterns in differentiated products industries.

Gandhi, A. and J.-F. Houde (2019). Measuring substitution patterns in differentiated-products industries. Working Paper 26375, National Bureau of Economic Research.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the US automobile industry. *Econometrica 63*(4), 891–951.

Goldberg, P. K. and F. Verboven (2001). The evolution of price dispersion in the European car market. *The Review of Economic Studies 68*(4), 811–848.

Grigolon, L., M. Reynaert, and F. Verboven (2018). Consumer valuation of fuel costs and tax policy: Evidence from the European car market. *American Economic Journal: Economic Policy 10*(3), 193–225.

Head, K. and T. Mayer (2019). Brands in motion: How frictions shape multinational production. *American Economic Review 109*(9), 3073–3124.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics 131*(3), 1291–1364.

Knittel, C. R. and K. Metaxoglou (2014). Estimation of random-coefficient demand models: Two empiricists' perspective. *Review of Economics and Statistics 96*(1), 34–59.

Matsuyama, K. (2023). Non-CES aggregators: A guided tour. *Annual Review of Economics 15*(1), ??–??

McFadden, D. and K. Train (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics 15*(5), 447–470.

Mrázová, M. and J. P. Neary (2017, December). Not so demanding: Demand structure and firm behavior. *American Economic Review 107*(12), 3835–74.

Nakamura, E. and D. Zerom (2010). Accounting for incomplete pass-through. *The Review of Economic Studies 77*(3), 1192–1230.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Nocke, V. and N. Schutz (2018). Multiproduct-firm oligopoly: An aggregative games approach. *Econometrica 86*(2), 523–557.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy 110*(4), 705–729.

Piveteau, P. and G. Smagghue (2021). Foreign competition along the quality ladder.

Redding, S. J. and D. E. Weinstein (2019). Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences. *The Quarterly Journal of Economics 135*(1), 503–560.

Reynaert, M. and F. Verboven (2014). Improving the performance of random coefficients demand models: the role of optimal instruments. *Journal of Econometrics 179*(1), 83–98.

Salanié, B. and F. A. Wolak (2019). Fast, "robust", and approximately correct: Estimating mixed demand systems. Working Paper 25726, National Bureau of Economic Research.

Train, K. E. and C. Winston (2007). Vehicle choice behavior and the declining market share of US automakers. *International economic review 48*(4), 1469–1496.

Verboven, F. (1996). International price discrimination in the European car market. *The RAND Journal of Economics 27*(2), 240–268.

Appendix

A Mixed CES

The most common setup for random coefficients models is the unit demand mixed logit introduced in Berry (1994). More recently Björnerstedt and Verboven (2016), Piveteau and Smagghue (2021), Adao et al. (2017), Redding and Weinstein (2019) have worked with what the latter two papers refer to as mixed CES (MCES). The model assumes individual consumers have CES utility but that their price elasticity is heterogeneous. It is micro-founded by starting with the variable consumption discrete choice model of Anderson et al. (1992) (section 3.7), before extending it to include heterogeneity in the price responsiveness parameter. As in the mixed logit, the MCES also allows for random coefficients on the consumers' indirect utility derived from product attributes. The key difference is that households spend constant income shares rather than buying a single unit. Several well-known models can be thought of as special cases of mixed CES. As the variance across households of price elasticity and preference for characteristics goes to zero, mixed CES can reach three different limiting cases. First, with many single-variety firms it becomes the Dixit-Stiglitz model which we have also referred to as CES-MC. Second, with a small number of single-variety firms, the limiting case is a version of Atkeson and Burstein (2008) with the upper level CES set to one. Finally with several large multiproduct firms, MCES converges on models used by Hottman et al. (2016) and Bernard et al. (2018).

Birchall and Verboven (2022) estimate a random coefficients model that nests mixed logit and mixed CES and find the best-fit parameter is midway between those two polar cases. They also find that the simple logit predictions on how own price elasticities should vary with price are flatly inconsistent with the evidence from the breakfast cereal industry.

A.1 The Mixed CES data generating process

Denoting household income with y_h, the (indirect) utility of household h is given by

$$U_{mh} = \ln y_h - \tilde{\alpha}_h \ln p_m + \sum_{k=0}^{K} \tilde{\beta}_h^k x_m^k + \tilde{\xi}_m + \varepsilon_{mh}. \tag{A1}$$

With an outside good whose indirect utility is normalized to zero and ε_{mh} distributed Gumbel with scale parameter $1/\eta$, the choice probability of household h for model m

takes the form:

$$\mathbb{P}_{mh} = \frac{\exp(\sum_{k=0}^{K} \beta_h^k x_m^k - \alpha_h \ln p_m + \xi_m)}{1 + \sum_i \exp(\sum_k \beta_i^k x_i^k - \alpha_h \ln p_i + \xi_i)}. \tag{A2}$$

where $\alpha_h = \eta\tilde{\alpha}_h$, $\beta_h = \eta\tilde{\beta}_h$, and $\xi_m = \eta\tilde{\xi}_m$. Note that the specification of the random coefficients does not impose a relationship between α_h and β_h but it does imply that all buyers view the unobserved quality ξ_m in the same way. We adopt this approach to parallel the one taken by IO economists in the mixed logit models. An alternative, considered by Redding and Weinstein (2019), places the household heterogeneity in the η parameter. This has the consequence of making consumers who are more price sensitive also more sensitive to differences in quality, both observed and unobserved. This approach is attractive in many respects, but we have not pursued it in this version to limit the number of permutations to consider.

Each individual spends y_h on their preferred variety. Total expenditures on m are therefore $s_m Y$, where $Y \equiv \sum_h y_h$ and s_m is the variety's market share—defined in value. This market share is given by the expenditure-weighed average of the individual probabilities from equation (2):

$$s_m = \frac{\sum_h \mathbb{P}_{mh} y_h}{Y}, \tag{A3}$$

In the CES and β-heterogeneity cases, $\tilde{\alpha}_h = 1\ \forall h$, and therefore $\alpha_h = \eta$. With both types of heterogeneity active, $\tilde{\alpha}_h = 1/y_h$ where, as in BLP y_h is log-normally distributed using the distributional parameters from the BLP replication file. As before, we calibrate η to match the average own-price elasticity of 4.

The multi-product firm's profit maximization problem is very similar to that used in the mixed logit case, but it is important to note that the market shares, s_m are all measured in values, rather than in units. Let $\epsilon_m \equiv -\frac{\partial \ln s_m}{\partial \ln p_m}$ denote the elasticity of value market share with respect to own price and recalling that the Lerner index is $L_m = (p_m - c_m)/p_m$, the first order condition implies a price rule of

$$p_m = c_m \times \frac{(\epsilon_m + 1)}{\left[\epsilon_m - \frac{1}{s_m} \sum_{(j \neq m) \in \mathcal{J}_F} \frac{\partial \ln s_j}{\partial \ln p_m} L_j s_j\right]}. \tag{A4}$$

The formulas for own and cross price elasticities needed to compute prices are in section B of this appendix. This computation is done with the same fixed point iteration as for the mixed logit case.

A.2 Homogeneous CES approximation with variable markups

This subsection generalizes EHA to allow for markups determined inside a CES multi-product oligopoly such as that studied in Hottman et al. (2016) and Nocke and Schutz (2018). The optimal markup in this CES-OLY approximation varies over both models and markets. Assuming, as in BLP, that firms compete in prices (Bertrand), the markup equation at the model level depends on market shares at the firm level:

$$\mu_{mn} = \mu_{fn} = \frac{\eta(1 - s_{fn}) + 1}{\eta(1 - s_{fn})}, \quad \forall m \in \mathcal{J}_f, \quad \text{with} \quad s_{fn} = \sum_{m \in \mathcal{J}_f,} s_{mn}. \tag{A5}$$

The markup converges to $(\eta+1)/\eta$ as firm-level market shares go to zero.[24] Except in that limit case, there is no closed-form solution to the market share equation and estimation requires an iterative procedure to estimate η. Start with a guess of η^0. Since we observe firm market share s_{fn}, we can compute the equilibrium markup μ^0_{fn} using equation (A5). This markup is passed to the left-hand-side, and combined with the log of market shares to yield the following regression for the kth iteration

$$\ln s_{mn} + \eta^k \ln \mu^k_{fn} = -\eta^{k+1} \ln \tau_{i(m)n} + \text{FE}_m + \text{FE}_n + \upsilon_{mn}, \tag{A6}$$

The coefficient on trade costs provides a new estimate η^{k+1}, with which we can recalculate markups. The process iterates from $k = 0$ until $\eta^{k+1} = \eta^k$ (within tolerance) at which point we have an estimate $\check{\eta}$, consistent with Bertrand oligopoly pricing.[25]

Once the estimate $\check{\eta}$ is obtained, one can also work with Exact Hat Algebra to compute counterfactual market shares that account for changes in markups. The changes in market shares for the inside goods ($m > 1$) are

$$\hat{s}_{mn} = \frac{\left(\hat{\mu}_{mn}\hat{\tau}_{i(m)n}\right)^{-\check{\eta}}}{s_{0n} + \sum_{j \in \mathcal{J}} s_{jn}\left(\hat{\mu}_{jn}\hat{\tau}_{i(j)n}\right)^{-\check{\eta}}}. \tag{A7}$$

The change in markup is computed as

$$\hat{\mu}_{mn} = \hat{\mu}_{fn} = \frac{1}{\mu_{fn}} \frac{\check{\eta}[1 - \hat{s}_{fn}s_{fn}] + 1}{\check{\eta}[1 - \hat{s}_{fn}s_{fn}]}, \quad \forall m \in \mathcal{J}_f, \quad \text{with} \quad \hat{s}_{fn} = \frac{\sum_{m \in \mathcal{J}_f,} \hat{s}_{mn}s_{mn}}{s_{fn}}, \tag{A8}$$

[24] The reason the limiting monopolistic competition markup is no longer $\eta/(\eta - 1)$ is that we have now switched to using value market shares. Thus η is no longer a quantity-price elasticity but a value-price elasticity. The redefinition of η changes the markup formula.

[25] Breinlich et al. (2021) estimate an equation similar to (A6), using the Cournot equivalent of equation (A5) to construct the markup adjustment term. They however use calibrated (or separately estimated from an auxiliary regression) values of the needed parameters rather than internally estimate like we do in (A6).

We have all the elements to iterate over the EHA predictions. Start with $\hat{s}_{mn} = 1$, aggregate to obtain the firm-level market shares $\hat{s}_{fn}$. Using initial markup (A5), one can retrieve its change from (A8). The new vector of market share changes is finally obtained with (A7). The process stops when the vector of $\hat{s}_{mn}$ stops changing.

B Derivatives and elasticities with random coefficients

B.1 Mixed logit

Since the individual partial effect of a change in p_m is

$$\frac{\partial \mathbb{P}_{mh}}{\partial p_m} = -\alpha_h \mathbb{P}_{mh}(1 - \mathbb{P}_{mh}),$$

we obtain the partial derivative of market share with respect to price:

$$\frac{\partial s_m}{\partial p_m} = \frac{\sum_h \frac{\partial \mathbb{P}_{mh}}{\partial p_m}}{N} = -\frac{\sum_h \alpha_h \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})}{N}.$$

The own price elasticity is:

$$\frac{\partial \ln s_m}{\partial \ln p_m} = -\frac{p_m}{s_m} \times \frac{\sum_h \alpha_h \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})}{N} = -p_m \sum_h \omega_{mh} \alpha_h (1 - \mathbb{P}_{mh}), \quad \text{with} \quad \omega_{mh} \equiv \frac{\mathbb{P}_{mh}}{\sum_h \mathbb{P}_{mh}}.$$

Model m's own elasticity therefore is a weighted average of the individual household elasticities, which write

$$\frac{\partial \ln \mathbb{P}_{mh}}{\partial \ln p_m} = -\alpha_h (1 - \mathbb{P}_{mh}) p_m.$$

The weight ω_{mh} applied to each of those elasticities is the share of each household in total sales of the model. Note that in the individual elasticity, a low p_m will be associated with a high purchasing probability $\mathbb{P}_{mh}$, both contributing to a lowering of $\frac{\partial \ln \mathbb{P}_{mh}}{\partial \ln p_m}$. The individual response to price increases is therefore unambiguously concave, getting more and more pronounced as the price goes up. At the model level, however, a composition effect enters the picture. Low price models are preferred by low income individuals which are assumed to have a larger sensitivity for prices (a high α_h). Those low price models therefore face high α_h households with larger weight ω_{mh}, which raises the overall price elasticity. This introduces an element of convexity, which can dominate the individual-level concavity.

Let us turn to cross-price effects: the impact of an increase in the price of model m on demand for j. The partial effect of m's price on $\mathbb{P}_{jh}$ is

$$\frac{\partial \mathbb{P}_{jh}}{\partial p_m} = \alpha_h \mathbb{P}_{jh} \mathbb{P}_{mh},$$

which yields

$$\frac{\partial s_j}{\partial p_m} = \frac{\sum_h \frac{\partial \mathbb{P}_{jh}}{\partial p_m}}{N} = \frac{\sum_h \alpha_h \mathbb{P}_{jh} \mathbb{P}_{mh}}{N}.$$

The cross-price elasticity is then

$$\frac{\partial \ln s_j}{\partial \ln p_m} = p_m \sum_h \omega_{jh} \alpha_h \mathbb{P}_{mh}, \quad \text{with} \quad \omega_{jh} \equiv \frac{\mathbb{P}_{jh}}{\sum_h \mathbb{P}_{jh}}.$$

Again, this is a weighted average of the individual choice probability cross elasticities,

$$\frac{\partial \ln \mathbb{P}_{jh}}{\partial \ln p_m} = \alpha_h \mathbb{P}_{mh} p_m.$$

B.2 Mixed CES

The individual partial effect of a change in p_m is

$$\frac{\partial \mathbb{P}_{mh}}{\partial p_m} = -\frac{\alpha_h}{p_m} \mathbb{P}_{mh}(1 - \mathbb{P}_{mh}),$$

The partial derivative of market share with respect to price is

$$\frac{\partial s_m}{\partial p_m} = \frac{\sum_h \frac{\partial \mathbb{P}_{mh}}{\partial p_m} y_h}{Y}$$

The own price elasticity is:

$$\frac{\partial \ln s_m}{\partial \ln p_m} = -\sum_h \omega_{mh} \alpha_h (1 - \mathbb{P}_{mh}), \quad \text{with} \quad \omega_{mh} \equiv \frac{\mathbb{P}_{mh} y_h}{\sum_h \mathbb{P}_{mh} y_h}. \tag{B9}$$

Model m's own elasticity therefore is a weighted average of the individual elasticities,

$$\frac{\partial \ln \mathbb{P}_{mh}}{\partial \ln p_m} = -\alpha_h (1 - \mathbb{P}_{mh}),$$

where the weight ω_{mh} is the share of each household in total sales of the model.

Turning to cross-price effects: the impact of an increase in the price of model m on demand for j. The partial effect of m's price on $\mathbb{P}_{jh}$ is

$$\frac{\partial \mathbb{P}_{jh}}{\partial p_m} = \frac{\alpha_h}{p_m} \mathbb{P}_{jh} \mathbb{P}_{mh},$$

which yields a partial derivative of market share as

$$\frac{\partial s_j}{\partial p_m} = \frac{\sum_h \frac{\partial \mathbb{P}_{jh}}{\partial p_m} y_h}{Y} = \frac{\sum_h \alpha_h \mathbb{P}_{jh} \mathbb{P}_{mh} y_h}{p_m Y}.$$

Lastly, multiplying by p_m/s_j, where $s_j = (\sum_h \mathbb{P}_{jh} y_h)/Y$, the cross price elasticity is

$$\frac{\partial \ln s_j}{\partial \ln p_m} = \frac{\sum_h \alpha_h \mathbb{P}_{jh} \mathbb{P}_{mh} y_h}{s_j Y} = \sum_h \omega_{jh} \alpha_h \mathbb{P}_{mh} \quad \text{with} \quad \omega_{jh} \equiv \frac{\mathbb{P}_{jh} y_h}{\sum_h \mathbb{P}_{jh} y_h}. \tag{B10}$$

Again, this is a weighted average of the individual choice probability cross-elasticities,

$$\frac{\partial \ln \mathbb{P}_{jh}}{\partial \ln p_m} = \alpha_h \mathbb{P}_{mh}.$$

C Pass-through rates and elasticities

The derivation of theoretical pass-through starts from FOC for model m:

$$s_m + (p_m - c_m) \frac{\partial \ln s_m}{\partial \ln p_m} = p_m - (p_m - c_m)\epsilon_m = 0,$$

with $\epsilon_m \equiv -\frac{\partial \ln s_m}{\partial \ln p_m} > 0$ being the own price elasticity. Implicit differentiation gives

$$\frac{\partial p_m}{\partial c_m} = \frac{-\epsilon_m}{-\epsilon_m + 1 - (p_m - c_m)\frac{\partial \epsilon_m}{\partial p_m}}.$$

Using the first order condition to replace $(p_m - c_m) = p_m/\epsilon_m$, the pass-through rate simplifies to

$$\frac{\partial p_m}{\partial c_m} = \frac{\epsilon_m}{\epsilon_m - 1 + E_m}, \quad \text{where} \quad E_m \equiv \frac{\partial \ln \epsilon_m}{\partial \ln p_m}. \tag{C11}$$

E_m is the super-elasticity of demand, i.e. the elasticity of own price elasticity with respect to a change in own price.[26] Under CES demand and monopolistic competition, ϵ_m is a constant. Hence, $E_m = 0$, and the pass-through rate is a constant equal to $\epsilon/(\epsilon - 1)$. The

[26] Bulow and Pfleiderer (1983) appear to have been the first to show, in their equation (3′), the relationship between the pass-through rate and this measure of the curvature of the demand curve; Mrázová and Neary (2017) consider the role of curvature in many different families of demand curves.

pass-through elasticity is

$$\frac{\partial \ln p_m}{\partial \ln c_m} = \frac{\epsilon_m}{\epsilon_m - 1 + E_m} \times \frac{c_m}{p_m} = \frac{\epsilon_m - 1}{\epsilon_m - 1 + E_m}. \tag{C12}$$

The sign of E_m is therefore the determinant of whether the pass-through elasticity is greater or smaller than one. In the Dixit-Stiglitz case, $E_m = 0$ implies a unitary pass-through elasticity.

Under homogeneous logit, $\epsilon_m = \alpha p_m(1 - s_m)$, and $E_m = [1 + \alpha p_m s_m]$. Since $\alpha > 0$, the super-elasticity is positive (greater than one, its value when the market share of m approaches 0) and the pass-through elasticity (PTE) is less than one. Substitution into equation (C11) implies a pass-through rate (PTR) of $1 - s_m$. With large outside good shares $s_m \approx 0$, so the PTR ≈ 1 the PTE is $\approx c_m/p_m < 1$.

The mixed logit case is more complex. Recall that BLP demand at the household-model level implies $\frac{\partial \mathbb{P}_{mh}}{\partial p_m} = -\alpha_h \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})$, and therefore the following own elasticity:

$$\epsilon_m = \frac{p_m}{s_m} X_m, \quad \text{with} \quad X_m \equiv \frac{\sum_h \alpha_h \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})}{N} = -\frac{\partial s_m}{\partial p_m}.$$

Taking the derivative of ϵ_m with respect to price,

$$\frac{\partial \epsilon_m}{\partial p_m} = \frac{X_m}{s_m} + \frac{\partial X_m}{\partial p_m}\frac{p_m}{s_m} - \frac{p_m X_m}{s_m^2}\frac{\partial s_m}{\partial p_m}.$$

Using $\frac{\partial s_m}{\partial p_m} = -X_m$, one can re-write

$$\frac{\partial \epsilon_m}{\partial p_m} = \frac{X_m}{s_m}\left[1 + \frac{\partial X_m}{\partial p_m}\frac{p_m}{X_m} + \frac{p_m X_m}{s_m}\right] = \frac{X_m}{s_m}\left[1 + \frac{\partial \ln X_m}{\partial \ln p_m} + \epsilon_m\right].$$

Hence the super-elasticity is

$$E_m = \frac{\partial \epsilon_m}{\partial p_m}\frac{p_m}{\epsilon_m} = \left[1 + \epsilon_m + \frac{\partial \ln X_m}{\partial \ln p_m}\right],$$

where $\frac{\partial \ln X_m}{\partial \ln p_m}$ is the elasticity of the slope of demand to a change in price. One therefore needs to study how X_m varies with p_m

$$\frac{\partial \ln X_m}{\partial \ln p_m} = \frac{\sum_h \frac{\partial \mathbb{P}_m}{\partial p_m}\alpha_h(1 - 2\mathbb{P}_{mh})}{N}\frac{p_m}{X_m} = -\frac{p_m}{X_m}\frac{\sum_h \alpha_h^2 \mathbb{P}_{mh}(1 - \mathbb{P}_{mh})(1 - 2\mathbb{P}_{mh})}{N}$$

hence,

$$\frac{\partial \ln X_m}{\partial \ln p_m} = -p_m \frac{\sum_h \alpha_h^2 \mathbb{P}_{mh}(1-\mathbb{P}_{mh})(1-2\mathbb{P}_{mh})}{\sum_h \alpha_h \mathbb{P}_{mh}(1-\mathbb{P}_{mh})},$$

and the super-elasticity is

$$E_m = \left[1 + \epsilon_m - p_m \frac{\sum_h \alpha_h^2 \mathbb{P}_{mh}(1-\mathbb{P}_{mh})(1-2\mathbb{P}_{mh})}{\sum_h \alpha_h \mathbb{P}_{mh}(1-\mathbb{P}_{mh})}\right],$$

D Exact Hat Algebra for CES

Here we derive the Exact Hat Algebra formula using an inversion method inspired by Berry (1994). The core idea is to retrieve the unobserved mean utility parameter of each variety based on observed market shares.

The EHA approach starts with a specification of CES equilibrium market shares (5), combined with constant markups $\mu_{mn} = \mu = \eta/(\eta-1)$, yielding

$$s_{mn} = \left(\frac{\tau_{i(m)n}/\varphi_{mn}}{P_n}\right)^{-\eta}, \text{ where } \quad P_n \equiv \left[\sum_{j=0}^{|\mathcal{J}|}(\tau_{i(j)n}/\varphi_{jn})^{-\eta}\right]^{-1/\eta}. \tag{D13}$$

The above equation is a slight generalization of equation (5) in the text, where we follow the IO tradition of normalizing the outside good ($j=0$) to have an indirect utility of one.

In the counterfactual equilibrium, we maintain the unobservable φ_{mn} unchanged, and obtain new market shares denoted with s'_{mn}:

$$s'_{mn} = \left(\frac{\tau'_{i(m)n}/\varphi_{mn}}{P'_n}\right)^{-\eta}, \text{ where } \quad P'_n \equiv \left[\sum_{j=0}^{|\mathcal{J}|}(\tau'_{i(j)n}/\varphi_{jn})^{-\eta}\right]^{-1/\eta}. \tag{D14}$$

Recalling that $\hat{x} \equiv x'/x$,

$$s'_{mn} = \left(\frac{\hat{\tau}_{i(m)n}\tau_{i(m)n}/\varphi_{mn}}{\hat{P}_n P_n}\right)^{-\eta}, \text{ where } \quad \hat{P}_n P_n = \left[\sum_{j=0}^{|\mathcal{J}|}(\hat{\tau}_{i(j)n}\tau_{i(j)n}/\varphi_{jn})^{-\eta}\right]^{-1/\eta}. \tag{D15}$$

Let us denote geometric means over all choices ($|\mathcal{J}|$ car models and the outside good) by $\tilde{x} \equiv \exp((|\mathcal{J}|+1)^{-1}\sum_{j=0}^{|\mathcal{J}|}\ln x_j)$. We can invert the market share equation to obtain :

$$\tau_{i(m)n}/\varphi_{mn} = s_{mn}^{-1/\eta} \times K_n, \text{ where } \quad K_n \equiv \frac{\tilde{\tau}_n \tilde{s}_n^{1/\eta}}{\tilde{\varphi}_n}. \tag{D16}$$

With CES demand, the unobservable determinants of model-destination competitiveness, $\tau_{i(m)n}/\varphi_{mn}$, can be expressed as a power function of observable market share s_{mn} and of a market specific constant K_n (that will cancel when computing new market shares). The definition of K_n follows the approach of Hottman et al. (2016) using geometric means, although there is an equivalent approach using the outside good as a normalization.

We can then replace this expression in the counterfactual market share to obtain

$$s'_{mn} = \left(\frac{\hat{\tau}_{i(m)n} s_{mn}^{-1/\eta} K_n}{\hat{P}_n P_n}\right)^{-\eta}, \text{ where } \quad \hat{P}_n P_n = \left[\sum_{j=0}^{|\mathcal{J}|} (\hat{\tau}_{i(j)n} s_{jn}^{-1/\eta} K_n)^{-\eta}\right]^{-1/\eta}. \tag{D17}$$

K_n factors out of $\hat{P}_n P_n$ and cancels with the K_n in the numerator. Since s_{mn} has a power of 1 it can be passed to the denominator of the left-hand-side (yielding $\hat{s} = s'/s$), so that we finally obtain the CES EHA expression for the proportional change in market shares:

$$\hat{s}_{mn} = s'_{mn}/s_{mn} = \frac{\hat{\tau}_{i(m)n}^{-\eta}}{\sum_{j=0}^{|\mathcal{J}|} s_{jn} \hat{\tau}_{i(j)n}^{-\eta}} = \frac{\hat{\tau}_{i(m)n}^{-\breve{\eta}}}{s_{0n} + \sum_{j \in \mathcal{J}} s_{jn} \hat{\tau}_{i(j)n}^{-\breve{\eta}}}, \tag{D18}$$

where the last equality imposes that, in our counterfactuals, there is no change in the trade costs applied to the outside good.

For completeness, we briefly sketch the typical method used by trade economists to derive EHA. Starting from equation (D15), and using the market share equation (D13)

$$s'_{mn} = s_{mn} \left(\frac{\hat{\tau}_{i(m)n}}{\hat{P}_n}\right)^{-\eta}, \text{ where } \quad \hat{P}_n = \left[\sum_{j=0}^{|\mathcal{J}|} s_{jn} \hat{\tau}_{i(j)n}^{-\eta}\right]^{-1/\eta}. \tag{D19}$$

Dividing both sides by s_{mn}, yields (D18).

E Approximate Hat Algebra

The CES-OLY approach just described computes pass-through of cost changes into prices based on strong assumptions about conduct. We also consider a third approach to counterfactuals in CES that is agnostic on market structure and instead relies on empirical estimates of the pass-through elasticity. Let $\breve{\rho} = |\mathcal{J}_F|^{-1} \sum_{m \in \mathcal{J}_F} \partial \ln p_{mn} / \partial \ln c_{mn}$, be an estimate of the average rate at which foreign varieties pass through increases in their marginal costs. What we refer to as "approximate" hat algebra (AHA) computes the counterfactual as

$$\hat{s}_{mn} = \frac{s'_{mn}}{s_{mn}} = \frac{[1 + (\hat{\tau}_{i(m)n} - 1)\breve{\rho}]^{-\breve{\eta}}}{s_{0n} + \sum_{j \in \mathcal{J}} s_{jn}[1 + (\hat{\tau}_{i(j)n} - 1)\breve{\rho}]^{-\breve{\eta}}}. \tag{E20}$$

This is not exact since almost any model of imperfect pass-through will have differential pass-through across models and markets, rather than the scalar $\breve{\rho}$ used here.

Table E1: Counterfactual 10% tariff using the BLP data

	Quantity shares			Pass-through		
	Agg. ΔS			rate	elasticity	
Setting	True	EHA	AHA	Avg.	Avg.	# 1
Mixed Logit	8.00	7.73	8.57	1.57	1.13	1.12
Logit	3.85	7.73	5.39	1.00	0.67	0.62
β het.	3.33	7.73	5.27	0.98	0.65	0.57

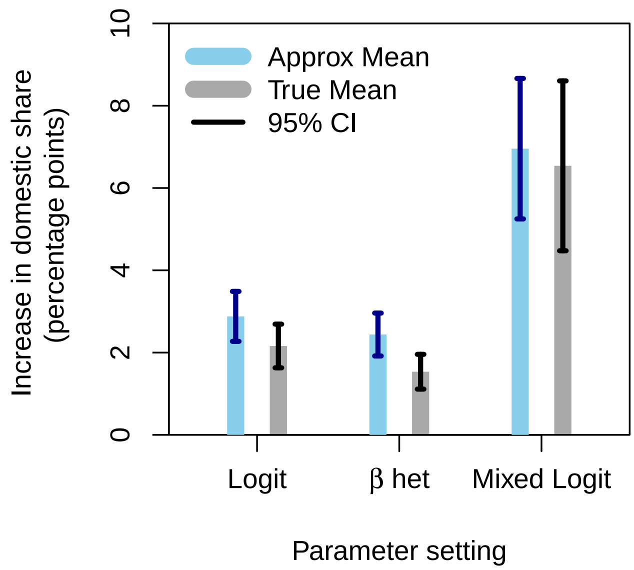
Note: CES EHA uses $\eta = 4.05$. AHA (approximate hat algebra) uses $\eta = 4.05$ and the average pass-through elasticity as in equation (E20). ΔS is the change in aggregate share of domestic models in new car market. PT rate $= \partial p_m / \partial c_m$. The "BLP" row is the original version of the model where both types of consumer heterogeneity are active. "Logit" sets $\alpha_h = \alpha$ (holding avg own price elas constant) and $\beta_h = \beta$. "β het." only sets $\alpha_h = \alpha$.

The pass-through issue suggests a relatively easy way to improve the counterfactuals assuming the CES model is true. Supposing one has a good estimate of the average pass-through elasticity, equation (E20) shows how to incorporate this moment to give an approximation to a more complex model of variable markups. These counterfactuals appear in the AHA column, showing the mean change in domestic market share and the average bias. As expected, AHA reduces the bias for the logit and for β heterogeneity. The halving of bias we see in those cases is not replicated in the BLP setting. Since EHA was already very accurate, AHA's increase in pass-through leads to overshooting the target.

F Random coefficient logit with monopolistic competition

Of the two key features of the BLP setup, rich substitution and multiproduct oligopoly, we have so far emphasized the former. How detrimental to the CES approximation is it to assume Dixit-Stiglitz market structure? In the first line of Table F2, we assign each of the 90 models to an individual firm. Hence, the market structure moves close to monopolistic competition for the "true" prediction. Note that the pass-through rate is 1, as predicted by monopolistic competition under logit demand. The pass-through elasticity equals the pass-through rate divided by markup μ, hence smaller than 1 (it averages at 0.69 over our 1000 replications). The CES-MC prediction of unitary elasticity implies an overprediction of the reaction of foreign firms and of domestic market share increase. In the second line of Table F2, the niche market power created by β-heterogeneity reinforces that overestimation of the change in market share.

Figure F1: Random coefficient logit with monopolistic competition



Note: As in the BLP data, the share of outside goods is calibrated to 90%, the share of domestic cars is 68%, and we set average ϵ_m to 4. The simulation has 90 firms with 1 model each; the combined market share of the top 5 firms is 34% (as compared to 86% in BLP). The error bars are 1.96 standard *deviations* of the simulation outcomes for 1000 repetitions.

Table F2: Random coefficient logit with monopolistic competition

Setting	Agg. ΔS True	Agg. ΔS Approx	Tariff elas ($\breve{\eta}$)	Passthrough rate	Passthrough elas
Logit	2.16	2.88	2.65	1.00	0.69
β heterogeneity	1.53	2.44	2.55	1.00	0.68
Mixed Logit	6.54	6.96	4.87	1.72	1.21

Note: As in the BLP data, the share of outside goods is calibrated to 90%, the share of domestic cars is 68%, and we set average ϵ_m to 4 for 1000 repetitions. The simulation has 90 firms with 1 model each; the combined market share of the top 5 firms is 34% (as compared to 86% in BLP).

G Exact Hat Algebra for logit

The derivation starts from an adapted version of the equation in Anderson et al. (1992, p. 45). Let us first state the market share equation for m in n under logit (no consumer heterogeneity):

$$s_{mn} = \frac{\exp(\sum_{k=0}^{K} \beta^k x_m^k - \alpha p_{mn} + \xi_{mn})}{1 + \sum_j \exp(\sum_k \beta^k x_j^k - \alpha p_{jn} + \xi_{jn})}. \tag{G21}$$

Denote the change in m's price in n as $\Delta p_{mn} = p'_{mn} - p_{mn}$, the counterfactual market share of m is

$$s'_{mn} = \frac{s_{mn}(\exp(-\alpha \Delta p_{mn}))}{s_{0n} + \sum_j s_{jn}(\exp(-\alpha \Delta p_{jn}))}.$$

Denoting the proportional change $\hat{x} = x'/x$ and with the additive markups $p_{mn} = c_{mn} + \mu_{mn}$ implied by logit demand, we obtain

$$\hat{s}_{mn} = \frac{\exp(-\alpha[\Delta c_{mn} + \Delta \mu_{mn}])}{s_{0n} + \sum_j s_{jn}[\exp(-\alpha[\Delta c_{jn} + \Delta \mu_{jn}])]}. \tag{G22}$$

The most natural counterfactual tariff change under logit demand is a specific duty of d dollars per unit, in which case $\Delta c_{mn} = d_{i(m)n}$, i being the country where firm m is located. In the monopolistic competition case, the markup is constant, and equation (G22) is enough to compute the new equilibrium based on three requirements: initial market shares, the structural parameter driving price response (α), and the policy change d. With ad valorem tariff rate of t per dollar, the change in unit costs becomes $\Delta c_{mn} = t_{i(m)n} c_m$. This makes the cost change variety-specific. With price, characteristics, and market share data, c_m can be obtained by inversion of the first order condition (and assuming there is an estimate of α). This increases the informational requirements relative to the CES case or the logit case with specific duties.

With non-atomistic varieties, we have to account for endogeneous markup adjustment. Under Bertrand oligopoly, the additive markup of m only depends on the market share of firm f to which m belongs (Nocke and Schutz, 2018, study more generally the properties under which the market share of a multi-product firm is sufficient to compute its markup and ensuing market power):

$$\mu_{mn} = \mu_{fn} = \frac{1}{\alpha(1 - s_{fn})}, \quad \forall m \in \mathcal{J}_f, \quad \text{with} \quad s_{fn} = \sum_{m \in \mathcal{J}_f,} s_{mn}. \tag{G23}$$

The change in markup is computed as

$$\Delta\mu_{mn} = \frac{1}{\breve{\alpha}}\left[\frac{1}{1-\hat{s}_{fn}s_{fn}} - \frac{1}{1-s_{fn}}\right], \quad \forall m \in \mathcal{J}_f, \quad \text{with} \quad \hat{s}_{fn} = \frac{\sum_{m\in\mathcal{J}_f,} \hat{s}_{mn}s_{mn}}{s_{fn}}. \quad \text{(G24)}$$

Combining (G22) with (G24), the elements needed to compute $\hat{s}_m$ are the initial observed initial market shares s, the policy change d, and α. With these formulae for $\hat{s}_{mn}$ and $\hat{\mu}_{mn}$ in hand, the rest of the Exact Hat Algebra algorithm proceeds as with the CES case, iterating until a fixed point is reached.

We can estimate α with an iterative procedure following the logic of the mixed CES case seen in appendix A. We start by taking logs of (G21) in the case of specific tariffs where $p_{mn} = \mu_{fn} + c_m + d_{i(m)n}$:

$$\ln s_{mn} = -\alpha d_{i(m)n} - \alpha\mu_{fn} + \text{FE}_m + \text{FE}_n + \xi_{mn},$$

where the structural interpretation of fixed effects are $\text{FE}_m = \sum_{k=0}^{K}\beta^k x_m^k - \alpha c_m$, and $\text{FE}_n = -\log\left[1 + \sum_j \exp(\sum_k \beta^k x_j^k - \alpha p_{jn} + \xi_{jn})\right]$. Start with a guess called α^0. With firm market share s_{fn}, we can compute the equilibrium markup μ_{fn}^0 using equation (G23). This markup is passed on the left-hand-side, and combined with the log of market shares to yield the following regression for the lth iteration

$$\ln s_{mn} + \alpha^l\mu_{fn}^l = -\alpha^{l+1}d_{i(m)n} + \text{FE}_m + \text{FE}_n + \xi_{mn}. \quad \text{(G25)}$$

The coefficient on per-unit trade costs d provides a new estimate α^{l+1}, with which we can recalculate markups. The process iterates from $l = 0$ until $\alpha^{l+1} = \alpha^l$ (within tolerance) at which point we have an estimate $\breve{\alpha}$, consistent with Bertrand oligopoly pricing.

H Random coefficient on domestic status

We augment the BLP estimation to include a random coefficients using the `PyBLP` package described in Conlon and Gortmaker (2020). This package includes the original BLP data and the authors provided us the code to estimate the best practices specification (optimal instruments and 10,000 scrambled Halton draws). Our Table H3 first shows the published parameters from Berry et al. (1995) and then reproduces the best practices results in Conlon and Gortmaker (2020) Table 8, column 3. In the third column we show the new estimation results. Several parameters change when adding domestic status. The key estimate for our purposes is that the standard deviation (σ) for the domestic status random coefficient is just 0.028. We also conducted the estimation using the *differentiation instruments*, introduced by Gandhi and Houde (2019). As seen in the fourth column, this procedure gives larger heterogeneity, 0.34, in the demand for domestic cars. These results, using the original BLP car data, are lower than those obtained by Reynaert and Verboven (2014) using data for nine European countries from 1998 to 2010. They estimate a standard deviation parameter of 0.72 in one specification and 1.72 in another.[27]

Table H3: BLP estimated parameters (CG2020)

Variable	Base		Domestic RC	
IV:	BLP1999	CG2020	CG2020	Diff. IV
β constant	-7.061	-6.679	-6.079	-3.861
β HP/weight	2.883	2.774	2.798	5.133
β air	1.521	0.572	-0.454	1.683
β miles per USD	-0.122	0.340	0.172	0.078
β size	3.460	3.920	3.114	4.289
β domestic	NA	NA	0.164	-0.589
σ constant	3.612	2.962	1.731	2.441
σ HP/weight	4.628	1.388	0.311	2.030
σ air	1.818	1.424	2.696	0.079
σ miles per USD	1.050	0.072	0.015	0.258
σ size	2.056	0.231	0.180	3.413
σ domestic	NA	NA	0.028	0.336
term on price (α)	-43.501	-45.898	-39.374	-76.581
γ constant	0.952	2.785	2.406	2.127
γ log(HP/weight)	0.477	0.731	0.583	0.231
γ air	0.619	0.528	0.549	0.431
γ log(miles per gallon)	-0.415	-0.651	-0.479	-0.323
γ log(size)	-0.046	-0.472	0.249	0.776
γ domestic	NA	NA	-0.304	-0.348
γ trend	0.019	0.018	0.021	0.021
mean own elas.	-3.93	-3.45	-2.92	-5.17

[27]Reynaert and Verboven (2014) obtain the mean domestic bias is 0.85, much larger than our 0.164 with best practices and -0.589 with differentiation instruments.

Table H4: Counterfactual 10% tariff using the BLP data

Setting	Elasticities ϵ_m	E_m	Pass-through rate	elas.	Agg. ΔS True	EHA
			BLP 1999 published parameters			
BLP	4.05	-0.49	1.57	1.13	8	7.73
Logit	4.05	1.00	1.00	0.67	3.85	7.73
β het.	4.02	1.03	0.98	0.65	3.33	7.73
			CG2020 Baseline			
BLP	3.43	-0.24	1.56	1.07	6.41	6.63
Logit	3.43	1.00	1.00	0.61	3.00	6.63
β het.	3.41	1.01	0.98	0.60	2.78	6.63
			CG2020 Baseline + Domestic RC			
BLP	2.95	-0.15	1.64	1.05	5.06	5.77
Logit	2.95	1.00	0.99	0.54	2.33	5.77
β het.	2.94	1.01	0.97	0.54	2.25	5.77
			Diff. IV + Domestic RC			
BLP	5.07	-0.42	1.41	1.08	9.48	9.47
Logit	5.07	1.00	1.00	0.73	5.20	9.47
β het.	5.05	1.02	0.99	0.72	4.65	9.47

Note: CES EHA uses the average own price elasticity (column 1) for η. ΔS is the change in aggregate share of domestic models in new car market. PT rate $= \partial p_m / \partial c_m$. The "BLP" row is the original version of the model where both types of consumer heterogeneity are active. "Logit" sets $\alpha_h = \alpha$ (holding avg own price elas constant) and $\beta_h = \beta$. "β het." sets $\alpha_h = \alpha$ but allows for heterogeneity on the β for all attributes (including constant).

Table H5: Effects of increasing SD on domestic, for general and targeted tariffs

	Elasticities		Pass-through		Agg. ΔS	
Setting	ϵ_m	E_m	rate	elas.	True	EHA
			BLP 1999 published parameters			
BLP	4.05	-0.49	1.57	1.13	8	7.73
Logit	4.05	1.00	1.00	0.67	3.85	7.73
β het.	4.02	1.03	0.98	0.65	3.33	7.73
			BLP 1999 + RV2014 (ii) Domestic RC			
BLP	4.05	-0.48	1.56	1.12	7.57	7.73
Logit	4.05	1.00	1.00	0.67	3.85	7.73
β het.	4.01	1.04	0.98	0.65	3.18	7.73
			BLP 1999 + RV2014 (i) Domestic RC			
BLP	4.05	-0.47	1.51	1.07	6.27	7.74
Logit	4.05	1.00	1.00	0.67	3.86	7.74
β het.	4.01	1.04	0.99	0.65	2.73	7.74

Note: CES EHA uses the average own price elasticity (column 1) for η. ΔS is the change in aggregate share of domestic models in new car market. PT rate $= \partial p_m / \partial c_m$. The "BLP" row is the original version of the model where both types of consumer heterogeneity are active. "Logit" sets $\alpha_h = \alpha$ (holding avg own price elas constant) and $\beta_h = \beta$. "β het." sets $\alpha_h = \alpha$ but allows for heterogeneity on the β for all attributes (including constant). RV2014 (i)/(ii) Domestic RC refers to the standard deviation of the random coefficient on domestic status taken from column (i) or (ii) in Table 6 of Reynaert and Verboven (2014).