

Brands in Motion: How Frictions Shape Multinational Production[†]

By KEITH HEAD AND THIERRY MAYER*

Following the 2016 Leave vote in the referendum on UK membership in the EU and the election of Donald Trump, trade agreements have entered a period of great instability. To predict the impact of possible disruptions to existing arrangements requires counterfactual analysis that takes into account the complex set of factors influencing the production and marketing strategies of multinational corporations. We estimate a model of multinational decision-making in the car industry. This model predicts the production reallocation and consumer surplus consequences of changes in tariffs and non-tariff barriers induced by US-led protectionism, Brexit, transpacific, and transatlantic integration agreements. (JEL F13, F23, L21, L62, M31)

After decades in which free trade agreements proliferated and deepened in scope, 2016 appeared to mark a major turning point. The Leave vote in the UK referendum on EU membership and the election of Donald Trump brought long-standing integration arrangements to the brink of disruption. Pure trade models are ill-equipped for predicting the outcomes of regional dis-integration because they omit an increasingly important feature of the world economy: multinational production (MP). The foreign affiliate structures of multinational corporations (MNCs) complicate matters because they introduce new sets of bilateral relationships. In addition to the origin-destination goods flows of standard trade models, MP models feature interactions between headquarters and subsidiary locations. MNCs must decide which of their network of production facilities will serve each market. Furthermore, because MNCs are typically multiproduct firms, they face decisions over which subset of varieties to offer in each of the markets where they choose to operate distribution facilities. Each of these decisions is likely to be influenced by distinct bilateral frictions.

*Head: Sauder School of Business, University of British Columbia, 2053 Main Hall, Vancouver, BC V6T 1Z2, CEPR, and CEP (email: keith.head@sauder.ubc.ca); Mayer: Sciences Po, Département d'économie, 28 Rue des Saints-Pères, 75007 Paris, France, Banque de France, CEPR, and CEP (email: thierry.mayer@sciencespo.fr). Penny Goldberg was the coeditor for this article. This research has received funding from the European Research Council (ERC) under the grant agreement 313522. Participants at presentations at Banque de France, Berkeley, Bilkent, Bocconi, CREI (UPF), EITI, ERWIT, Harvard-MIT trade seminar, HKU, HKUST, Kiel, Princeton, Society of Economic Dynamics, UBC, University of Toronto, and Warwick provided insightful comments. Suggestions by Kerem Coşar, Swati Dhingra, Gene Grossman, Beata Javorcik, Jan de Loecker, Esteban Rossi-Hansberg, and David Weinstein led to significant modifications to earlier versions of the paper. Comments by the editor and three anonymous referees contributed to major revisions. Jules Hugot, David Rodrigues, Eve Sihra, Xiaonan Sun, and Chenying Yang performed valuable research assistance. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <https://doi.org/10.1257/aer.20161345> to visit the article page for additional materials and author disclosure statements.

Data limitations present a major challenge in estimating an economy-wide model of MP that encompasses these decisions and the corresponding frictions. We therefore study a single industry, cars, where multinational production is prevalent. Multinational brands, those produced in more than one country, account for 99.5 percent of cars sold in the OECD. It is also a sector where firm-level trade flows are available for all the main producing and consuming nations. This allows us to estimate the impacts of trade integration based on variation in tariffs on final cars and parts as well as the presence of integration agreements that go well beyond tariff cuts. The estimated model predicts the consequences for producers and consumers of the shocks to trade policy that are currently being debated: US threats to its NAFTA partners and other major producers to massively raise tariffs under Section 232 national security provisions, Brexit, transpacific, and transatlantic integration agreements. This paper offers the first quantitative assessment of such proposals that takes into account the micro-structure of multinational production.

To meet our goals of estimating structural parameters and solving for responses to counterfactual policies, we need a tractable MP framework. A salient feature of MP in the car industry is the prevalent use of export platforms: 50 percent of cars sold in OECD markets are assembled in locations that are neither the headquarter nor the consuming country. The recently developed quantitative framework, that we refer to as the double CES (constant elasticity of substitution) MP model, tractably incorporates export platforms. It combines a CES heterogeneous-firm product market structure as in Melitz (2003) with a constant-elasticity sourcing decision adapted from Eaton and Kortum (2002). Important contributors to the development of this framework include Ramondo (2014); Ramondo and Rodríguez-Clare (2013); Irarrazabal, Moxnes, and Oromolla (2013); Arkolakis et al. (2018); Antràs, Fort, and Tintelnot (2017); and (closest to our setup) Tintelnot (2017). Comparative statics in these papers generally hinge on two parameters: the first governs *substitutability between products* from the view of consumers, whereas the second describes the *interchangeability of potential production locations* from the firm's perspective. The double CES framework extends the gravity equation to a setting where it is costly for headquarters to coordinate foreign assembly and distribution affiliates. Gravity has proven to be a powerful tool for understanding international trade flows; its most attractive features being tractable estimation, good fit to the data, and the ability to conduct counterfactuals with minimal data requirements.¹ Our implementation of the double CES framework here maintains those three desirable features.

With the goal of making plausible predictions for the consequences of trade policy changes, we extend the double-CES MP model in three ways.

- (i) Multinational firms are also multiproduct firms who offer different sets of varieties in each market. In the car industry, the typical brand makes ten models, of which they only offer one-quarter in a typical market. They frequently opt out of serving markets at all (the typical brand serves only one-third of the market-year combinations). These facts motivate the inclusion of two

¹Dekle, Eaton, and Kortum (2007) were the first to use the CES structure of gravity to implement what Costinot and Rodríguez-Clare (2014) call the exact hat algebra approach to counterfactuals.

extensive margins of adjustment to policy shocks: which markets to serve and what subset of varieties to offer in each.²

- (ii) While frictions between production and consumption locations are standard in pure trade models, the possibility that the firm is headquartered in a third country calls for the consideration of two additional frictions. Costs of producing far from headquarters are a well-established feature of the MP framework (going back to Ramondo 2014). An innovation of our paper is to incorporate a third friction between headquarters and the market. We interpret this third friction as marketing costs, with both variable and fixed components. The new friction will play a crucial role in explaining our first extension of the MP framework, the market entry margins. A key motivation for incorporating the new frictions is that modern “deep” integration agreements contain whole chapters that do not operate on the origin-destination path traversed by goods. Rather, topics such as harmonization of standards, protection of investments, and facilitation of temporary movement of professionals, mainly affect the flows of *headquarters* services to production and distribution affiliates. Because our data track the three countries where a brand is headquartered, produces, and sells its products, we are able to econometrically identify the new frictions separately from traditional trade costs.
- (iii) The model incorporates *external* increasing returns to scale (IRS) by specifying the marginal costs of each plant in a given country as a function of total car production in that country. This contrasts with the internal scale economies modeled by Goldberg and Verboven (2001) and Antràs, Fort, and Tintelnot (2017) in related contexts. There are two benefits of introducing external IRS into the double-CES MP framework. First, there is ample evidence suggesting that external (Marshallian) scale economies are important in practice.³ One well-supported mechanism comes from forward and backward linkages between assemblers and parts suppliers.⁴ Second, building external IRS into the cost function allows for important interdependencies across markets without sacrificing tractability, one of the chief attractions of the double-CES MP model. As in Tintelnot (2017) and Antràs, Fort, and Tintelnot (2017), firm-level decisions of where to assemble cars remain independent from each other, facilitating estimation. Aggregate-level production decisions nevertheless become *interdependent*, with a change in frictions between two countries affecting outcomes in third countries. Tintelnot (2017) and Antràs, Fort, and Tintelnot (2017) build in interdependencies through the mechanism of firms paying a fixed cost to add countries to the choice set. This

²Prior work on multinational production omits the two market entry margins, as Arkolakis et al. (2018) assume single-product firms, Antràs, Fort, and Tintelnot (2017) consider a single market, and Tintelnot (2017) assumes that firms offer a unit mass of varieties in every market. The variety-entry extensive margin for multiple-product firms was incorporated in a pure trade model by Bernard, Redding, and Schott (2011).

³Spatial concentration of car production has been an important feature of the industry since its founding, as seen in the production clusters around Detroit and Paris. More recently, the examples of plant agglomerations in cities of Slovakia (4 plants), Central Mexico (11 plants), Northern France (6 plants), and the I-75 corridor in the United States (about 10 plants) point to the persistent importance of Marshallian economies.

⁴See Smith and Florida (1994), Klier and McMillen (2008), and Schmitt and van Biesebroeck (2013).

approach would not be computationally feasible in our context because of the above-mentioned extensive margins that are essential features of the industry.

The data we use come from an automotive industry consultancy that tracks production at the level of *brands* (Acura, BMW, Chevrolet) and *models* (RDX, X5, Corvette). Because our paper uses car data, it invites comparison to a series of papers that have considered trade and competition in this industry. Goldberg (1995); Verboven (1996); and Berry, Levinsohn, and Pakes (1999) investigate quantitative restrictions on imports of cars into the US and EU markets. More recently, an independent and contemporaneous paper by Coşar et al. (2018) combines a demand side from Berry, Levinsohn, and Pakes (1995) with the MP model of Tintelnot (2017). These papers feature multi-product oligopoly and use either nested or random coefficients differentiated products demand systems. The advantage of these approaches is that they allow for variable markups and yield more realistic substitution patterns than the monopolistic competition with symmetric varieties demand assumed in the double CES model. This method has two disadvantages in our context. First, it severs the connection to the gravity equation from trade. Second, to implement the rich substitution models, the researcher needs to know the prices and continuous characteristics of all the models. Such data are only available for a drastically reduced set of brands, models, and markets.⁵ This would make it impossible for us to consider the global production reallocations associated with the mega-regional agreements.

The chief concern about CES for the purposes of this paper is that it might exaggerate the degree of substitution between models in very different segments of the car market. This could lead to erroneously large responses to trade policy changes (e.g., a Brexit-induced tariff on Polish-made Fiat 500s would be unlikely to trigger much substitution toward UK-made Land Rovers). We mitigate this concern, while maintaining all the computational advantages of CES, by also estimating and simulating a version of the demand side that nests varieties within market segments.⁶ This follows the research line of Goldberg (1995) and Verboven (1996), with two important modifications. As in Björnerstedt and Verboven (2016), substitution within each nest takes the CES form (albeit with quantity shares). Second, to characterize the maximal extent of divergence from symmetric CES, our formulation restricts all substitution to occur *within segments*. The unified and segmented markets versions of the model therefore bracket the extent of substitution between car models in different segments. The interval between these extreme approaches turns out to be fairly narrow in terms of the outcomes of our counterfactual scenarios.

We recover the structural parameters of the MP model through the sequential estimation of four equations corresponding to four key decisions made by multinational firms: (i) from which of the firm's factories to source each variety, (ii) the quantity supplied to each market, (iii) which varieties to offer in each market the brand is distributed in, and (iv) where to distribute the brand. The two first

⁵The Coşar et al. (2018) dataset has 9 markets and 60 brands compared to the 76 markets and 138 brands in our estimating sample.

⁶Grigolon and Verboven (2014) show that nested logit can match fairly closely the cross-price elasticities of a random coefficients model.

equations deliver credible estimates of the two pivotal elasticities of the double CES framework: identifying from variation in car tariffs, we estimate a sourcing elasticity of 7.7. The sourcing equation delivers a brand's cost index for supplying models to a given destination. Variation in this index identifies the demand-side CES and estimates it as 3.87. We find that regional integration has substantial effects on all three dimensions of frictions. This is fully in line with the observation that export platforms are organized on a regional basis: 85 percent of export platform for OECD markets occurs within regional trade agreements. Combining all four equations, the double-CES framework performs well when applied to the global car industry data. The bilateral trade flows predicted by the model match the data with a correlation of 0.74. The new features that we incorporate into the MP framework, the market entry margins for models and brands and the marketing costs, prove to be quantitatively important. The median ad valorem equivalents of the combined variable and fixed components of marketing costs (68 percent) are larger than the trade costs (24 percent) and frictions between headquarters and assembly locations (31 percent) already standard in the MP literature.

The results from counterfactual trade policy changes improve our understanding of the impacts of the creation and dissolution of regional integration agreements. We predict substantial reallocations of production in response to a set of policies that have recently been proposed or implemented. Perhaps the most striking outcomes are the dramatic output reductions that the Canadian and Mexican car industries would have suffered if the United States had followed through on threats to abrogate NAFTA and impose 25 percent Section 232 (national security) tariffs on its neighbors. Our simulations show that even when the two countries impose retaliatory tariffs, there would be a 40 percent decline in Mexican production and an even larger 67 percent cut in Canada. Meanwhile, consumer surplus in all three countries would decline (by as much as 6 percent in Canada). Through the exemptions to Section 232 tariffs that Canada and Mexico obtained in NAFTA renegotiations, the two countries stand to increase production substantially if the United States imposes such tariffs on the rest of the world. Under this trade war scenario, US-based plants also increase production sharply, mainly at the expense of Korea, Japan, and Germany, which collectively lose between 2.4 and 2.8 million cars. The other major trade disruption we consider, Brexit, causes relatively minor production losses in the United Kingdom (about two-thirds of a typical size plant), but the country's consumers pay up to 8 percent more for cars.

A second set of counterfactuals predicts the consequences of new integration agreements that the United States opted not to join. We predict that joining the Trans-Pacific Partnership (TPP) or forming a deep trade agreement with the EU (TTIP) would have lowered production in the United States. On the other hand, plants in Canada stand to increase production by going ahead with both transoceanic agreements. Membership in the 11-member CPTPP boosts Canadian output by one-third, primarily because our estimates imply a more than 6 percent cost reduction for Japanese-owned plants in Canada. By far the greatest consumer benefits in all the policy scenarios go to the Vietnamese whose car price index falls by over 25 percent with transpacific integration. Most of this comes from eliminating 44 percent tariffs on imported Japanese cars, however. The reductions in marketing costs from transpacific integration are predicted to generate quite large reactions for the entry

margins introduced in our paper: for instance, the CPTPP (excluding the United States) expands model entry by Japanese brands in Canada by around 20 percent and reduces the probability that Chevrolet enters the Vietnamese market by more than 30 percentage points.

The paper continues in five main sections. We first discuss and display some of the important empirical features of the global car industry, using the nearly exhaustive firm-level information on where each variety is designed, assembled, and sold. Drawing on these facts, the next section generalizes the existing models to include marketing frictions and market-entry decisions at the model and brand level. We then show how the structural parameters of the MP model can be recovered from four estimating equations corresponding to four key decisions made by multinationals. Following estimation, we present the key methodological aspects of our counterfactual exercises. Finally, we use those methods to project the outcomes of (i) trade wars provoked by US imposition of national-security tariffs on cars, (ii) soft and hard versions of Brexit, (iii) transpacific, and (iv) transatlantic integration agreements.

I. Data and Model-Relevant Facts

Recent work on multinational production uses datasets that cover all manufacturing or even the universe of multinational activities (including services). The drawback of such datasets is the absence of complete micro-level flows. This forces the theory to do more of the work in the estimation process. We concentrate on a single activity within a single sector: the assembly of passenger cars. As this focus raises the issue of the external validity of our results, we think it worthwhile to emphasize compensating advantages of studying the car industry.

The first and foremost advantage of the car industry is the extraordinary richness of the data compiled by IHS Markit.⁷ IHS uses new car registration information (and probably other sources of information) to obtain annual flows at the level of individual models identifying the assembly plant and country of sale. From it we extract origin-destination flows for 4,791 car models sold by 138 brands over the 2000–2016 period.⁸

What we refer to as a “model” is a combination of three variables in the original dataset: (i) *sales nameplate*, which IHS defines as the “Name under which the vehicle is sold in the respective country”;⁹ (ii) the *bodytype* defined as “Vehicle silhouette without doors designation”;¹⁰ (iii) the *program*, which IHS defines as the “code used by OEMs to identify vehicle throughout design lifecycle.” Programs constitute redesigns, or new generations of a model.¹¹

⁷ Other attractive aspects of the car industry include its size (passenger cars alone constitute 4 percent of global trade and the broader industry accounts for 5 to 6 percent of employment in the United States and European Union) and prominence in public debate.

⁸ Due to entry and exit, there are fewer brands and models in each year. For instance, 2,128 models were offered by 120 brands in 2016. Online Appendix E explains the cuts we applied to the original IHS dataset.

⁹ Our sample includes 2,377 distinct nameplates such as the 500 (Fiat), Twingo (Renault), 3 (Mazda).

¹⁰ The bodytypes are sedan, SUV, hatchback, MPV, wagon, coupe, convertible, and roadster.

¹¹ The Renault Twingo, for instance, has had three generations to date: X06 (1993–2012), X44 (2007–2014), and X07 (2014–).

The empirical analysis in the main text maps the theoretical concept of varieties to models and the concept of firms to brands. Models appear to be the natural counterpart to the concept of varieties. As implied by our theory for individual varieties, we show that models sold in a particular market are almost always sourced from a single assembly location. There are several reasons we employ brands, rather than parent corporations, to correspond to the theoretical concept of the firm. First, the brand is the common identity across models that is promoted to buyers via advertising and dealership networks. This suggests that the brand's home is the one relevant for marketing frictions. Second, most of the brands under common ownership were originally independent firms (e.g., Chevrolet and Opel (GM), Ferrari and Chrysler (Fiat), Volvo (Geely), Mini (BMW)). Partly for historical reasons, brand headquarters often correspond to the location where models are designed. For example, while Jaguar is owned by Tata Motors, based in India, Jaguar's cars are designed at the brand's headquarters in Coventry in the United Kingdom. We think of the brand's headquarters as a principal source of tangible (e.g., engines) and intangible (e.g., designs, managerial oversight) inputs used by the assembly plants.

There are two potential sources of concern when using the brand/model concepts. The first is that headquarter inputs may originate mainly from a higher level than the brand headquarters. A second worry comes from the industry practice of re-badging: different brand/model combinations might cover what is essentially the same underlying car. The richness of the IHS data enables us to replicate all our analysis using an alternative approach that deals with those concerns. The alternative specifies varieties as particular car designs using the identifiers for the "platform" (the underbody of the car) to which we add the program and bodytype defined above. The concept of firm is the "Design Parent," the corporation that has managerial control over the design of the platform used by each variety. We discuss the results from implementing this approach, shown in full in online Appendix F.2, where relevant as we report stylized facts and regression results.

We identify the brand headquarters (i) as the country in which each brand was founded. In the case of spin-off brands like Acura, we use the headquarters of the firm that established the brand (Japan in this case). Unlike the few available government-provided datasets used in the literature, we are not restricted to parent firms or affiliates based in a single reporting country. Rather, our dataset is a nearly exhaustive account of global car headquarters, assembly, and sales locations. Our estimating sample comprises the shipments of cars assembled in 52 countries by brands headquartered in 21 countries and sold in 76 national markets.

Figure 1 illustrates some of the important aspects of our data using the case of the brand Fiat in 2013 for two of its main models, the Punto and the 500, and seven markets. Fiat makes the 199 program of the Punto in Italy, selling the cars to domestic and EU consumers (the dashed line reflects the 66 cars sold in Mexico). A budget version of the Punto (code 310) is made in Brazil for several markets in South America. The Fiat 500, mainly made in Poland (for EU markets) and Mexico (for the Americas), exemplifies the importance of regional export platforms in the car industry. A striking feature of the Fiat example is that no market is assigned to more than one assembly location for a given model. This pattern of single sourcing generalizes very broadly as we show in Fact 1. The absence of the Punto in the US market provides an example of selective model-market entry. Fiat does not distribute any

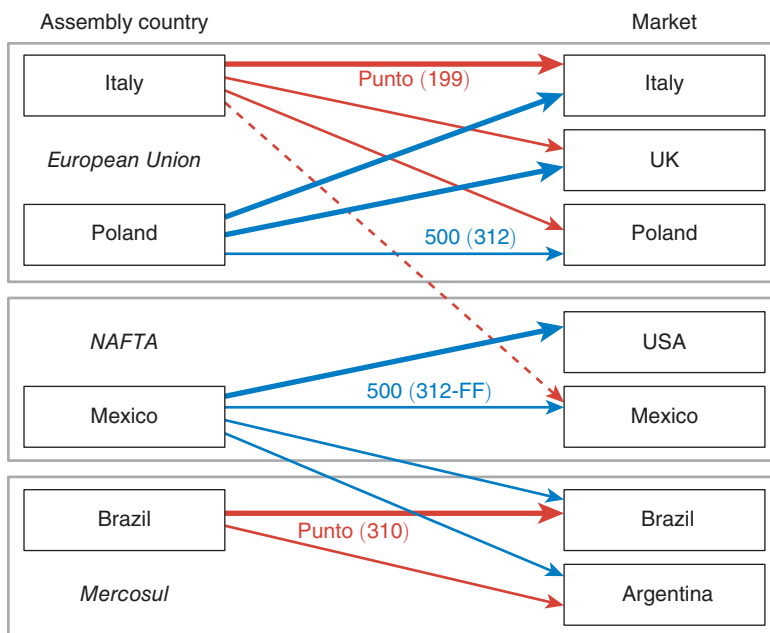


FIGURE 1. EXAMPLE: FIAT 500 AND PUNTO PRODUCTION ORGANIZATION

of its models in 11 of the 76 markets (mainly in Asia). We show in Fact 3 that most models and brands are offered in only a minority of the potential markets.

We now turn to describing three empirical facts that bear on the specific features of the model we estimate. The first two relate to key tractability assumptions of the existing model whereas the third represents a feature that we argue should be added to the standard model.

A. Fact 1: Almost All Models Are Single-Sourced

At the level of detail at which trade data are collected (six-digit HS), most large countries import from multiple source countries. This is part of the reason why the Armington assumption that products are differentiated by country of origin became so commonplace in quantitative models of trade.

In the car industry we have finer detail because specific models of a car are more disaggregated than tariff classifications. At the level of models, for a specific market, firms almost always source from a *single* origin country. This is not because all models are produced at single locations. In 2016, about one-fifth of all models are produced in more than one country and we observe four that are produced in ten or more countries. Rather, it is because firms match assembly sites to markets in a one-to-many mapping.

Table 1 shows that 98 percent of the model-market-year observations feature sourcing from a single assembly country. Sourcing from up to four countries happens occasionally but it is very rare. This is true for models produced by brands that have ten or more *potential* production countries, where potential sites are measured

TABLE 1—NUMBERS OF SOURCES FOR EACH MARKET-MODEL-YEAR

Number of sources	All model-markets			Brands with 10+ locations		
	Count	Col. %	Cum. %	Count	Col. %	Cum. %
1	320,069	97.7	97.7	194,901	97.1	97.1
2	7,386	2.3	99.9	5,651	2.8	99.9
3	243	0.1	100	208	0.1	100
4	5	0	100	5	0	100

by the number of countries where the brand conducts assembly (of any model). In 97 percent of the cases, these models are still single-sourced.

B. Fact 2: Most Markets Are Not Highly Concentrated

Firms in the car industry are not, of course, “massless” as assumed in the monopolistic competition model. The pertinent question is whether the monopolistic competition provides a useful approximation for answering the questions considered in this paper. The serious drawback of assuming oligopolistic price setting as in Atkeson and Burstein (2008) is that we would no longer be able to express flows as a closed-form multiplicative solution in terms of frictions. This would lose the connection to gravity and therefore also make it impossible to use the simple and direct estimation methods derived in the next sections.

One defense of the use of monopolistic competition is that, in some respects, the industry is not as concentrated as one might imagine. Table 2 shows that most markets feature many competitors. For three-quarters of market-years we consider, more than 191 models are available. Even at the highest level of aggregation, the sales parent,¹² three-quarters of the markets have at least 18 competing firms. Column 2 shows that median market shares are small (mainly under 5 percent), implying that oligopoly markups for the majority of firms would be close to those implied by monopolistic competition. Column 3 shows the concentration ratio for the top five actors at each level of aggregation. In three-quarters of the market-years, the top 5 brands account for less than 74 percent of the market. The last three columns show that at the highest levels of ownership (parent), EU merger guidelines would be “unlikely to identify horizontal competition concerns” for 72 percent of the market-year combinations.¹³ Even within segments, the majority of markets are moderately concentrated except in the case of MPVs and sport and luxury cars.

To be clear, we are not arguing that oligopoly is irrelevant in the industry. The largest firms are big enough to have endogenous markups that significantly exceed those implied by monopolistic competition. Nevertheless, even under a data generating processes that matches the level of concentration observed in the industry, an estimated CES monopolistic competition model can deliver surprisingly accurate predictions for trade policy counterfactuals. Head and Mayer (2018) simulate data

¹²“Sales parent” is defined by IHS as “The company who owns the brand at the current point in time.” For example, Volkswagen is the Sales Parent of Audi, Bentley, Bugatti, Lamborghini, SEAT, Skoda, and Volkswagen. There is a many-to-one mapping between brands and their sales parent.

¹³We use the EU threshold because it is intermediate between the corresponding Herfindahl thresholds used by the US Department of Justice (1,800) and the Federal Trade Commission (2,500).

TABLE 2—MARKET SHARE CONCENTRATION IN CAR SALES, 2000–2016

Level	Count	Interquartile-range			Concentration		
		Market shares			Percent market-years		
		Median	CR5	Top	Low	Moderate	High
Model	191–372	0.05–0.1	17–37	5–12	94	5	1
Brand	33–49	0.32–0.98	49–74	14–31	51	33	16
Parent	18–24	1–2.53	69–83	22–36	6	66	28
MPV	9–13	3.01–7.21	80–96	26–49	0	44	56
SUV	13–19	2.16–4.53	67–85	21–38	8	61	31
Big car	13–16	2.16–4.79	74–88	25–40	1	56	43
Small car	14–18	1.34–3.88	74–89	24–40	1	59	40
Sport lux	7–14	2.48–8.2	84–96	32–54	0	29	71

Notes: All figures are calculated over all market-year combinations (76 countries, 2000 to 2016). *CR5* is the combined share of the top five. Markets classified as low ($H < 1,000$), medium ($1,000 \leq H \leq 2,000$), and high ($H > 2,000$) concentration based on EU Commission thresholds. The first three rows calculate shares of the whole passenger car market; the last five rows use parent firm shares within market segments.

from a BLP framework featuring oligopoly, rich substitution in demand, and multi-product firms that internalize cannibalization effects. The CES-MC model is capable of closely approximating the aggregated counterfactuals for BLP-generated data under settings that replicate data moments for parent firms in Table 2 (5-firm concentration ratios of 70–80 percent and an average of 10 models per firm).¹⁴ There is no theorem guaranteeing the close fit we have found in these simulations generalizes to all situations. However, the simulations establish that the mere fact that CES-MC omits many theoretically desirable features does not systematically prevent it from being a useful tool for counterfactual policy exercises. The success of the CES-MC framework in these simulations reinforces its appeal for our purposes, given its tractability, low data requirements, and connection to the gravity equation.

C. Fact 3: Most Brands and Models Are Offered in a Minority of Markets

In the MP model presented in the next section, the firm decides where to establish distribution networks and which of its varieties to offer in each of those markets. Here we show that the model-level entry margin is very important for multi-model brands in the car industry. Panel A of Figure 2 depicts the histogram of $\bar{\mathcal{I}}_{mn}$, the model-level mean of the binary variable \mathcal{I}_{mnt} indicating model m is offered in market n in year t . The sample comprises model-market-years where the brand is available, the model is offered in more than one market, and the brand makes more than one model. We observe that brands almost never serve a market with *all* their models and only 15 percent of models are offered in the majority of the markets where the brand is available. With the average entry rate being just 23 percent, it seems clear that the standard MP framework should be augmented to include the extensive margin of model-level entry. A potential concern with these figures is that we may be underestimating entry due to the re-badging phenomenon. For example, Mazda sells the car design specified by platform “C1” and program “J68C” as the “Axela”

¹⁴On average, brands offer 6.8 models and parent firms offer 13.4 models.

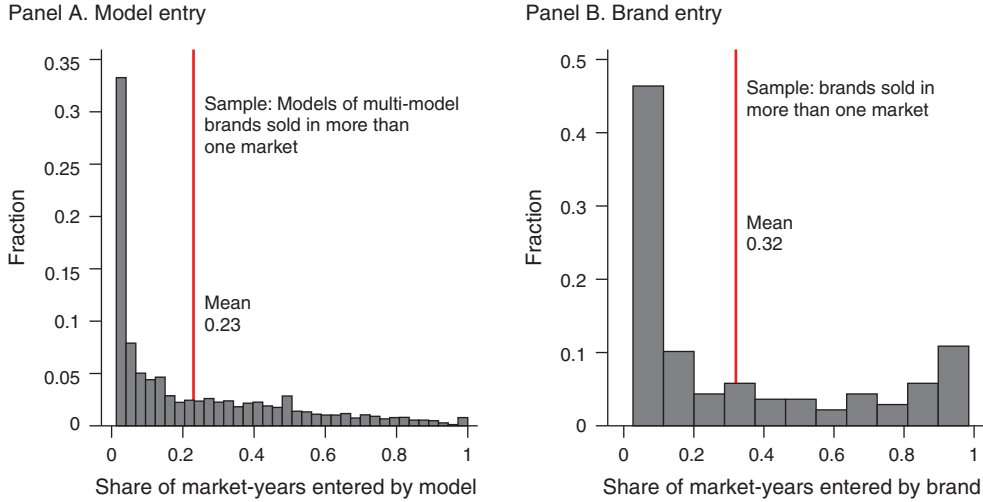


FIGURE 2. ENTRY RATES FOR MODELS AND BRANDS

in Japan but as the “3” everywhere else. We thus treat the Axela as being offered in just 1.5 percent of the market-years. Using the firm-variety methodology described in online Appendix F.2, we see that the hatchback version of C1-J68C has an 78 percent entry rate. However, looking across all varieties the average entry rate is just 24 percent, slightly larger than the average across all models. Online Appendix Figure F.1 shows that the whole distribution of entry rates is visually unchanged after removing the re-badging issue. The takeaway is that whether we define varieties as the consumer sees them or based on firm-level design distinctions, they tend to be offered in about one-quarter of the places where they might be offered.

Panel B of Figure 2 shows the distribution of entry rates by brands over markets. The key distinguishing feature with the model entry histogram is the existence of a local peak of high entry rates for a few brands. Ten brands enter more than 95 percent of the available market-years. At the other extreme, 43 brands enter 5 percent or fewer. On average, a brand enters less than one-third of countries it could sell in.

II. The Double CES Model of Multinational Production

There are B brands, M symmetrically differentiated car models, and N countries. Brand b is endowed with a headquarter country $i(b)$; a location-independent productivity φ_b ; a portfolio of $M_b = \sum_m \mathcal{M}_{bm}$ models, where $\mathcal{M}_{bm} = 1$ if b owns m ; and a set of production facilities, with $\mathcal{L}_{b\ell} = 1$ for countries where b can manufacture any of its models.

The sequence of decisions is depicted in Figure 3. At the beginning of period 1, brands are endowed with their defining attributes and learn F_{bn}^d , the fixed costs of creating a dealership network in each potential market. Brands then decide where to establish distribution facilities, corresponding to $\mathcal{D}_{bn} = 1$. For markets the brand has entered, it learns the model-entry fixed costs, F_{mn}^e , in period 2 and decides which models to offer in each market ($\mathcal{I}_{mn} = 1$). In the third period, brands learn the

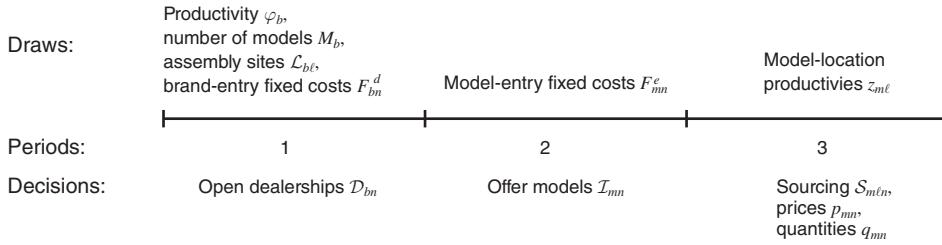


FIGURE 3. TIMELINE OF WHEN SHOCKS ARE REALIZED AND DECISIONS ARE MADE

model-location productivity shocks, $z_{m\ell}$, and select the source ℓ , that minimizes the delivered cost to market n for model m (denoted $c_{m\ell n}$), subject to $\mathcal{L}_{b\ell} = 1$. Selected sources have $\mathcal{S}_{m\ell n} = 1$. Based on realized costs, firms set prices, p_{mn} , and read quantities, q_{mn} , off their demand curves. In order to be able to estimate the model sequentially, we assume that, conditional on the observables (e.g., distance from market n to the headquarters of brand b), shocks to market entry costs (F_{bn}^d and F_{mn}^e) are uncorrelated with each other and with $z_{m\ell}$, the production cost shock.¹⁵

Brand-level profits are given by aggregating the model-level gross profits, π_{mn} , and netting out all fixed costs:

$$(1) \quad \Pi_b = \sum_{m=1}^M \mathcal{M}_{bm} \left[\sum_{n=1}^N \mathcal{I}_{mn} (\pi_{mn} - F_{mn}^e) \right] - \sum_{n=1}^N \mathcal{D}_{bn} F_{bn}^d$$

where

$$\pi_{mn} = \left(p_{mn} q_{mn} - \sum_{\ell=1}^N \mathcal{S}_{m\ell n} c_{m\ell n} q_{m\ell n} \right).$$

In summing over models above and throughout the paper, we follow our data and our programming in treating models as discrete entities. However, firm behavior is modeled as being monopolistically competitive: brands make pricing and entry decisions as if both brands and models were massless.

Profit maximization is constrained by $\mathcal{I}_{mn} \leq \mathcal{D}_{bn}$, and $\mathcal{S}_{m\ell n} \leq \mathcal{L}_{b\ell}$. That is, a brand distribution network in market n is necessary if any of the brands' models are to be offered there and the brand must have a plant in location ℓ if it is to be used as a source for any model.

This paper considers location choices as predetermined variables that constrain subsequent choices. This makes sense for analysis of the medium-run consequences of policy changes. As we have already mentioned, one reason we take this approach is to avoid the computational challenges of modeling plant location choice when the profit function is neither globally submodular nor supermodular. A second reason for taking locations as given is the strong persistence observed in the set of locations where each brand operates: three-quarters of all cars produced in 2016 (the year for which we run counterfactuals) were assembled in brand-country combinations that

¹⁵ Ciliberto, Murry, and Tamer (2018) consider more general patterns of correlations between cost and demand shocks in a framework that also involves oligopoly pricing and multiple equilibria in the entry game. These features lead to heavy computational burden and challenging parameter identification that go beyond the scope of our paper.

were already active in 2000. This fraction rises to 88 percent for OECD countries and to 94 percent if we draw the line in 2007, a decade before our counterfactuals.

A. Consumer Preferences and Demand

In our data we observe only quantities, not expenditures, and therefore wish to use a specification in which firm-level sales volumes are expressed as a share of total quantity demanded. As in the recent work of Fajgelbaum, Grossman, and Helpman (2011), we derive demand from the discrete choices across models by logistically distributed consumers. In contrast to that paper, however, our formulation retains the constant elasticity of substitution. Following Hanemann (1984), under conditions detailed in online Appendix A, households denoted h choose m to minimize $p_{mn(h)}/\psi_{mh}$, where $p_{mn(h)}$ is the price of model m in the market n where household h is located and ψ_{mh} is the quality that household perceives. With ψ_{mh} distributed Fréchet with shape parameter η (an inverse measure of customer heterogeneity), quantity demanded for model m in market n is given by

$$(2) \quad q_{mn} = \left(\frac{p_{mn}}{P_n} \right)^{-\eta} Q_n \quad \text{where} \quad P_n \equiv \left(\sum_j \mathcal{I}_{jn} p_{jn}^{-\eta} \right)^{-1/\eta},$$

where Q_n denotes aggregate new car purchases in the market. The Fréchet shape parameter η is the first of the two elasticities of substitution that drive outcomes in this framework. As with σ in the Dixit-Stiglitz framework, η is the own price elasticity of demand and also determines the sales to profit ratio. The key difference from the Dixit-Stiglitz setup is that here market shares q_{mn}/Q_n are expressed in terms of *quantities* rather than expenditures.

The delivered price of model m in n under monopolistic competition is a constant markup $\eta/(\eta - 1)$ of marginal cost. Substituting this price into the demand curve, sales are

$$(3) \quad q_{mn} = \left(\frac{\eta}{\eta - 1} c_{mn} \right)^{-\eta} Q_n P_n^\eta,$$

where c_{mn} is the marginal cost of model- m cars delivered to market n . Delivered costs depend on the sourcing decision, which in turn depends on a comparison of assembly and trade costs across candidate supply countries.

B. Costs (Including Frictions)

With marginal costs taken as given, each firm looks for the best site, conditional on its set of potential locations: $c_{mn} = \min\{c_{m\ell n}, \forall \ell \text{ such that } \mathcal{L}_{b\ell} = 1\}$. The marginal cost of assembling model m in country ℓ depends on the costs of inputs as well as three productivity determinants, which we elaborate on in successive paragraphs.

Plants combine inputs obtained locally with inputs imported from the headquarter country with Cobb-Douglas technology. Let w_ℓ and w_i denote the costs of a composite factor in the assembly and headquarter countries respectively. On top of worker wages and efficiency, w captures the price and variety of parts available in each country. Parameter $1 - \alpha$ denotes the cost share of headquarter-country

inputs.¹⁶ Frictions applicable to inputs sourced from the headquarter country are captured in $\tau_{i\ell}^H \geq 1$. The $\tau_{i\ell}^H$ term includes tariffs country ℓ imposes on key inputs (engines, transmissions) from country i . We use $\gamma_{i\ell} \equiv (\tau_{i\ell}^H)^{1-\alpha}$ as the composite measure of the costs of separating assembly from headquarters to emphasize the similarity to the corresponding friction in Arkolakis et al. (2018). The difference is that our γ reflects input transfers from headquarters whereas their γ is a penalty in terms of lost productivity associated with the transfer of operational methods from HQ to the assembly country. The combined cost of inputs can therefore be expressed as $w_\ell^\alpha w_i^{1-\alpha} \gamma_{i\ell}$.

Turning to productivity determinants, the first we consider is a brand-level shifter, φ_b , familiar from models of firm-level heterogeneity. The second productivity shifter, $z_{m\ell}$, is a model-location shock capturing how well-suited country ℓ is for assembly of model m . It is also a familiar component of the multinational production framework growing out of the product-country heterogeneity term in Eaton and Kortum (2002). Following the literature, $z_{m\ell}$ is distributed Fréchet with shape parameter θ . The final determinant of productivity is external economies of scale, a novel element in multinational production models, but recently incorporated in trade models by Kucheryavyy, Lyn, and Rodríguez-Clare (2016). As with that paper and the related empirical estimation carried out in Bartelme, Costinot, and Rodríguez-Clare (2018), productivity is specified as a power function of industry size in country ℓ . Parameter ς is the elasticity of costs with respect to the amount of production in the assembly country, q_ℓ . External economies of scale correspond to $\varsigma < 0$. While many micro-foundations could underlie the external IRS posited here, the Marshallian mechanism of downstream production attracting a denser network of parts suppliers is very plausible in this industry. Combining the input cost terms and the three productivity shifters, marginal costs are given by

$$(4) \quad c_{m\ell} = \frac{w_\ell^\alpha w_i^{1-\alpha} \gamma_{i\ell} q_\ell^\varsigma}{\varphi_b z_{m\ell}}.$$

Compared to other papers in the multinational production literature, the main innovation in equation (4) is that it allows for returns to scale that are external to the plant. Based on the large absolute size of car assembly plants and their tendency to cluster spatially, increasing returns that are both internal and external to the plant are likely to be important in the industry. Incorporating internal increasing returns would be computationally challenging. One critical problem in our context is that internal IRS would violate the conditions underlying the tractable equations for the multinational production model that we employ in the estimation and the solution of the model. This is especially true with respect to the case where marginal costs are decreasing in plant scale. Then the outcome of each sourcing decision affects the relative attractiveness of the sources in every other decision. This means there are vast sets of possible output configurations to evaluate. An alternative way to obtain internal IRS is to model it in the form of constant variable costs combined with a fixed cost per plant as in Tintelnot (2017) and Antràs, Fort, and Tintelnot (2017).

¹⁶The notion that plants rely heavily on inputs from their HQ country is consistent with de Gortari's (2017) observation that exports from German-owned plants in Mexico contain much higher German content than US-owned plants exports to the United States.

This retains the tractability of the sourcing equation but imposes a complex computation of optimal plant location. While those two papers have solved this problem within their frameworks, our extensive margins on the model and brand distribution side prevent us from applying their solutions.¹⁷

Fortunately, many of the key aggregate implications of internal IRS also carry over to the external IRS setup. For example, suppose the United States imposes a tariff on cars from Mexico. This will tend to lower production in Mexico, causing the plants there to contract, with each moving up its downward-sloping marginal cost curve. The rise in marginal cost will further reduce demand from the United States (an amplification effect) and also lower demand from Canada (an interdependency effect). Suppose instead that marginal costs are constant, but a fixed cost must be paid to keep each plant open. Then the least profitable plants will close and drop out of the brand's sourcing set for sales in both the United States and Canada. We would therefore see declines in realized sourcing from Mexico in both markets. These are the same qualitative industry-country-level predictions entailed by external increasing returns. As the assumption of external IRS allows the model to incorporate interdependencies and amplification effects in response to trade policy changes at the industry-country level, while still being able to exploit the attractive functional forms implied by the assumption of constant returns at the brand level, it seems like an attractive compromise. The outcomes under alternative formulations of IRS will vary across firms, of course, and the aggregate magnitudes need not be the same. Finding a way to handle IRS internal to the firm in MP models with extensive margins on the distribution side remains an interesting topic for future research.

The delivered marginal cost of model m from assembly country ℓ to market n is

$$(5) \quad c_{m\ell n} = c_{m\ell} \tau_{\ell n} \delta_{in},$$

where $\tau_{\ell n} \geq 1$ represents conventional trade costs such as tariffs and freight, and $\delta_{in} \geq 1$ captures variable distribution and marketing costs. The δ_{in} friction includes the added cost of operating dealership networks abroad, as they may be easier to manage over shorter distances, and with RTA visas (or free movement of labor in the case of economic unions) facilitating visits from head office managers. Increases in variable costs brought about by foreign regulatory requirements would also be reflected in δ_{in} .¹⁸

C. Sourcing Decision

Brands choose the optimal production locations for each model they intend to sell in a market from the set of countries where the brand has assembly facilities, i.e., $\mathcal{L}_{b\ell} = 1$. The firm's optimal strategy is to single-source for each model-market

¹⁷The algorithm employed by Antràs, Fort, and Tintelnot (2017) to solve for the optimal choice set requires a supermodular objective function. Although Arkolakis and Eckert (2017) generalizes the algorithm to handle submodular problems, our objective function has both super- and submodular regions.

¹⁸For example, foreign car makers complained about the additional costs of daytime running lamps when Canada mandated them for new cars in 1990. Another telling example comes from the 2018 renegotiation of the Korea-US RTA (<https://money.cnn.com/2018/03/27/news/economy/us-south-korea-trade-deal/index.html>). The revised deal allows US car makers to export up to 50,000 vehicles per year to South Korea that do not comply with South Korean safety rules (up from 25,000).

combination from the country offering the minimum $c_{m\ell n}$. The probability that ℓ is selected is the probability that $c_{m\ell n}$ is lower than the alternatives:

$$\begin{aligned} & \Pr(\mathcal{S}_{m\ell n} = 1 | \mathcal{L}_{b\ell} = 1) \\ &= \Pr(c_{m\ell n} \leq c_{mkn}, \forall k \text{ with } \mathcal{L}_{bk} = 1) \\ &= \Pr(\ln z_{m\ell n} - \alpha \ln w_\ell - \ln \gamma_{i\ell} - \ln \tau_{\ell n} - \varsigma \ln(q_\ell) \\ &\quad > \ln z_{mkn} - \alpha \ln w_k - \ln \gamma_{ik} - \ln \tau_{kn} - \varsigma \ln(q_k), \forall k \text{ with } \mathcal{L}_{bk} = 1). \end{aligned}$$

Firm-level productivity, φ_b , the friction δ_{in} , and the HQ cost factor w_i cancel out of this probability since they affect all ℓ locations the same way. The probability of selecting origin ℓ from the set of locations where the brand has a plant ($\mathcal{L}_{b\ell} = 1$) as the source of model m in market n is the same for all models of a given brand:

$$(6) \quad \Pr(\mathcal{S}_{m\ell n} = 1 | \mathcal{L}_{b\ell} = 1) = \left(\frac{w_\ell^\alpha \gamma_{i\ell} \tau_{\ell n} q_\ell^\varsigma}{C_{bn}} \right)^{-\theta},$$

$$\text{with } C_{bn} \equiv \left(\sum_k \mathcal{L}_{bk} (w_k^\alpha \gamma_{ik} \tau_{kn} q_k^\varsigma)^{-\theta} \right)^{-1/\theta}.$$

The parameter θ is the second CES in this framework, playing the same role as in Eaton and Kortum (2002); C_{bn} is the multinational production cost index, summarizing the firm's costs of serving market n . Versions of this equation appear in Arkolakis et al. (2018) as equation (6), Tintelnot (2017) as equation (9), and Antràs, Fort, and Tintelnot (2017) as equation (7).¹⁹ We are the first to estimate this equation directly with both $\tau_{\ell n}$ and γ_{in} frictions, because such estimation requires variety-level data on sourcing for multiple markets and for firms with many different headquarters.

D. Brand-Level Market Shares

All models are ex ante symmetric. Taking expectations over the z shocks implicit in c_{mn} we can use equation (3) to derive expected model-level sales in market n as

$$(7) \quad E[q_{mn}] = \mathcal{I}_{mn} \left(\frac{\eta}{\eta - 1} \right)^{-\eta} P_n^\eta Q_n E[c_{m\ell n}^{-\eta} | \mathcal{S}_{m\ell n} = 1].$$

Expected $c_{m\ell n}^{-\eta}$ is multiplicative in the expectation of $z_{m\ell n}^\eta$ conditional on ℓ being the lowest cost location for mn . Adapting a result from Hanemann (1984), this expectation is

$$E[z_{m\ell n}^\eta | \mathcal{S}_{m\ell n} = 1] = [\Pr(\mathcal{S}_{m\ell n} = 1 | \mathcal{L}_{b\ell} = 1)]^{-\frac{\eta}{\theta}} \Gamma\left(1 - \frac{\eta}{\theta}\right),$$

¹⁹Like Tintelnot (2017), we assume independent productivity shocks whereas the Arkolakis et al. (2018) formulation allows for them to be correlated.

where $\Gamma(\cdot)$ denotes the Gamma function. Combining this result with the cost function equations (4) and (5), the ℓn and $i\ell$ cost factors cancel with their counterparts in $\Pr(\mathcal{S}_{m\ell n} = 1 | \mathcal{L}_{b\ell} = 1)$. Substituting back into (7) leads to a simple multiplicative expression for expected market share:

$$(8) \quad E[q_{mn}/Q_n] = \mathcal{I}_{mn} \kappa_1 \left(\frac{\varphi_b P_n}{w_i^{1-\alpha} \delta_{in}} \right)^\eta C_{bn}^{-\eta},$$

where $\kappa_1 \equiv (\eta/(\eta-1))^{-\eta} \Gamma(1 - (\eta/\theta))$.

Summing over the models that b sells in n , the expected market share of brand b in market n (conditional on having a distribution network in n and offering M_{bn} models) is

$$(9) \quad E[q_{bn}/Q_n | \mathcal{D}_{bn} = 1, M_{bn}] = \sum_m \mathcal{M}_{bm} E[q_{mn}/Q_n] = \kappa_1 M_{bn} (\varphi_b / w_i^{1-\alpha})^\eta \delta_{in}^{-\eta} P_n^\eta C_{bn}^{-\eta},$$

where $M_{bn} = \sum_m \mathcal{M}_{bm} \mathcal{I}_{mn}$ and the price index is re-expressed as

$$(10) \quad P_n = \kappa_1^{-1/\eta} \left(\sum_b M_{bn} (\varphi_b / w_i^{1-\alpha})^\eta \delta_{in}^{-\eta} C_{bn}^{-\eta} \right)^{-1/\eta}.$$

The number of models that a brand offers in a market, M_{bn} , is endogenous but it can be moved to the left-hand side of (9) to obtain an expression for the brand's average market share in market n :

$$(11) \quad E\left[\frac{q_{bn}}{M_{bn} Q_n} | \mathcal{D}_{bn} = 1\right] = \exp(\ln \kappa_1 - \eta \ln \delta_{in} - \eta \ln C_{bn} + \eta \ln(\varphi_b / w_i^{1-\alpha}) + \eta \ln P_n).$$

This equation for expected average market shares is a generalized linear model of the determinants of $\ln \delta_{in}$. The γ or τ frictions enter via the multinational production cost index C_{bn} .

The coefficient on $\ln C_{bn}$ identifies the demand elasticity η . Our identification exploits the CES-monopolistic competition implication of complete pass-through. This allows us to identify the demand elasticity by substituting in a delivered cost index (C_{bn}) in place of the expected price. This index varies across markets and periods for a given brand mainly because of tariff and RTA variation. Therefore our key identifying assumption is that trade policy variation is orthogonal to variation in the quality of a brands' varieties. Because the cost index C_{bn} also depends on local productivity-adjusted wages in the set of countries where the brand assembles, we also need to assume that brands do not systematically source high-demand varieties from high-wage countries. We return to this issue when interpreting our elasticity estimates.

E. Model-Market Entry Decision

The incentive to enter a market depends on expected profitability. To explain why all models of a given brand do not always enter (or stay out of) a given market, we introduce mn heterogeneity in the form of fixed market-entry costs. Entry costs increase proportionately to a new set of frictions denoted δ_{in}^e , the fixed cost

counterpart of δ_{in} , representing systematic increases in fixed costs associated with separation between the headquarters country and the market. For example, regulations are often claimed to mandate product specifications that the home-based firms have already adopted. Redesigning a model to conform with foreign product regulations, and promoting the model to make consumers aware of it are two examples of costs that enter δ_{in}^e . With HQ and destination-market inputs used for marketing in proportions ζ and $1 - \zeta$, the fixed costs are given by $F_{mn}^e = w_i^\zeta w_n^{1-\zeta} \delta_{in}^e \epsilon_{mn}^e$. The fixed costs shock, ϵ_{mn}^e , is log-normal with parameters $\mu_n^e + \beta_b^e$ and σ_e . Country characteristics such as size and costs of registering a new product are captured in μ_n^e whereas brand-specific determinants of entry costs are in β_b^e .²⁰

The probability that entry occurs, $\mathcal{I}_{mn} = 1$, is the probability that model-level expected profits net of fixed costs are positive:²¹

$$(12) \quad \Pr(\mathcal{I}_{mn} = 1) = \Pr(E[\pi_{mn}] > F_{mn}^e).$$

With a constant demand elasticity, η , variable profits are given by

$$(13) \quad E[\pi_{mn}] = E[p_{mn} q_{mn}] / \eta = E[p_{mn}^{1-\eta}] P_n^\eta Q_n / \eta.$$

The brand foresees that it will choose the optimal assembly location after learning the realizations of the model-location productivity shocks, $z_{m\ell}$. Applying the moment generating equation from Hanemann (1984),

$$E[p_{mn}^{1-\eta}] = \kappa_2 \left(\frac{w_i^{1-\alpha} \delta_{in}}{\varphi_b} \right)^{1-\eta} C_{bn}^{1-\eta},$$

where $\kappa_2 \equiv (\eta/(\eta-1))^{1-\eta} \Gamma(1 + (1-\eta)/\theta)$. Substituting this expression into equation (13),

$$(14) \quad E[\pi_{mn}] = \frac{\kappa_2}{\eta} \left(\frac{w_i^{1-\alpha} \delta_{in}}{\varphi_b} \right)^{1-\eta} C_{bn}^{1-\eta} Q_n P_n^\eta.$$

Taking logs on both sides of the inequality in (12), substituting in expected profits, and incorporating the distributional assumptions for F_{mn}^e , the expected share of models offered is

$$(15) \quad E[M_{bn}/M_b] = \Pr(\mathcal{I}_{mn} = 1) \\ = \Phi \left[\left(\ln \kappa_2 - \ln \eta + (\eta-1) \ln(\varphi_b/w_i^{1-\alpha}) - (\eta-1) \ln C_{bn} - (\eta-1) \ln \delta_{in} \right. \right. \\ \left. \left. - \beta_b^e - \ln w_i^\zeta - \ln \delta_{in}^e + \ln Q_n + \eta \ln P_n - \ln w_n^{1-\zeta} - \mu_n^e \right) / \sigma_e \right].$$

²⁰ We implicitly assume β_b^e is linearly decreasing with $\ln \varphi_b$ so as to make the model invariant to the scale of φ_b . This allows to normalize one brand's productivity to be one when we extract the structural parameters.

²¹ A consequence of monopolistic competition is that concerns over cannibalization are absent from the model entry decision.

The terms of this equation indexed b or i will be captured collectively with a brand fixed effect. The last four terms on the second line go into a destination fixed effect. This entry equation produces the sensible prediction that the share of models offered in a market increases with its size, quality, and efficiency of the brand, and declines with frictions, fixed costs, and local competition (a low P_n). The entry decision also depends negatively on C_{bn} , the expected cost of serving n , which is lower for a brand if its plants are located in countries that have either low assembly costs or low transport costs to the market, both costs being part of C_{bn} . The probability of entry is invariant to an increase of wages everywhere by the same proportion.²²

F. Brand Entry (Distribution Networks)

Individual models differ in terms of comparative advantage and marketing fixed costs but each model has the same expected value of profits, $E[\pi_{mn}]$. The brand's total profit in market n gross of costs of a brand-level distribution network is

$$(16) \quad E[\pi_{bn}] = M_b \times (E[\pi_{mn}] - E[F_{mn}^e | \mathcal{I}_{mn}]) \times \Pr(\mathcal{I}_{mn} = 1),$$

where $E[\pi_{mn}]$ comes from equation (14) and $\Pr(\mathcal{I}_{mn} = 1)$ comes from equation (15). Expected fixed costs conditional on the model being profitable enough to offer is

$$(17) \quad E[F_{mn}^e | F_{mn}^e < E[\pi_{mn}]] \\ = \exp\left(\left[\mu_n^e + \beta_b^e + \ln(w_i^\zeta w_n^{1-\zeta}) + \ln \delta_{in}^e\right] + 0.5\sigma_e^2\right) \frac{\Phi(z_{bn} - \sigma_e)}{\Phi(z_{bn})},$$

where $z_{bn} \equiv (\ln E[\pi_{mn}] - [\mu_n^e + \beta_b^e + \ln(w_i^\zeta w_n^{1-\zeta}) + \ln \delta_{in}^e]) / \sigma_e$ is the “standardized” expected net profitability of an individual model, which is the same for all models of a given brand.²³ With this notation, $\Pr(\mathcal{I}_{mn} = 1) = \Phi(z_{bn})$. Using this equality and plugging expected fixed costs (17) into (16) yields expected brand profit gross of setting up distribution facilities in country n as

$$(18) \quad E[\pi_{bn}] = M_b [E[\pi_{mn}] \Phi(z_{bn}) - w_i^\zeta w_n^{1-\zeta} \delta_{in}^e \exp(\mu_n^e + \beta_b^e + 0.5\sigma_e^2) \Phi(z_{bn} - \sigma_e)].$$

The probability of brand entry is the probability that expected profits of b in n exceed fixed costs of establishing distribution facilities for the brand. As with model-level entry costs, we assume brand-level distribution costs are the product of headquarter wages, frictions, and a shock term: $F_{bn}^d = w_i^\zeta w_n^{1-\zeta} \delta_{in}^d \epsilon_{bn}^d$. The brand-destination shock to fixed costs of brand entry, ϵ_{bn}^d , is log-normal with parameters $\mu_n^d + \beta_b^d$ and σ_d . Country-level and brand-level determinants of the fixed costs associated with setting up a new business are captured by μ_n^d and β_b^d respectively.

²² Multiplying the composite factor cost w by λ lowers profits by $-[1 + (\eta - 1)(1 - \alpha)] \ln \lambda$ directly. There is also the $-(\eta - 1)\alpha \ln \lambda$ effect via C_{bn} and the $\eta \ln \lambda$ effect via the price index. These terms cancel each other.

²³ There is homogeneity of degree zero in z_{bn} with respect to wages for the same reason as equation (15).

Taking logs of the brand entry condition $E[\pi_{bn}] > F_{bn}^d$ leads to the following probability of brand entry:

$$(19) \quad \Pr(\mathcal{D}_{bn} = 1) = \Phi\left(\frac{\ln E[\pi_{bn}] - [\mu_n^d + \beta_b^d + \ln(w_i^\zeta w_n^{1-\zeta}) + \ln \delta_{in}^d]}{\sigma_d}\right).$$

III. Empirical Implementation

Equations (6), (11), (15), and (19) collectively describe firms' behavior in the model. We now consider the empirical implementation of those four equations.

A. Friction Determinants

We start by specifying the empirical content of frictions. The frictions governing trade costs (τ), HQ input transfer costs (γ), variable marketing costs (δ), and fixed model-entry (δ^e) and brand-entry (δ^d) costs, are exponential functions of the observable determinants (some of which vary over time, hence the new subscript t) denoted $\mathbf{X}_{\ell nt}$, $\mathbf{X}_{i\ell t}$, and \mathbf{X}_{int} :

$$(20) \quad \tau_{\ell nt} = \exp(\mathbf{X}'_{\ell nt} \boldsymbol{\rho}), \quad \gamma_{i\ell t} = \exp(\mathbf{X}'_{i\ell t} \mathbf{g}), \quad \delta_{int} = \exp(\mathbf{X}'_{int} \mathbf{d}),$$

$$\delta_{int}^e = \exp(\mathbf{X}'_{int} \mathbf{f}^e), \quad \delta_{int}^d = \exp(\mathbf{X}'_{int} \mathbf{f}^d),$$

where $\boldsymbol{\rho}$, \mathbf{g} , \mathbf{d} , \mathbf{f}^e , and \mathbf{f}^d are vectors of the primitive friction cost parameters.

The \mathbf{X} vectors include the standard explanatory variables used in gravity equations: home, distance, and common language. These variables have already been shown to matter for trade flows and affiliate sales. The differences in subscripts are of critical importance to the estimation. Thus $\text{home}_{\ell n}$ indicates that the assembly plant is in the same country as where the car is bought, whereas $\text{home}_{i\ell}$ equals 1 when the plant is located in the headquarters country, and finally home_{in} turns on when consumer and brand share the same home country. Distance is the average number of kilometers on great-circle route between the main cities in the corresponding countries. Language indicates that the countries share an official language.

In keeping with our focus on the role of trade policies in determining the pattern of multinational production, the \mathbf{X} vectors include additional determinants that are novel to our study. First, in $\mathbf{X}_{\ell nt}$ we have the log of 1 plus the tariff each country n imposes on ℓ -origin passenger cars in year t . We also include in $\mathbf{X}_{\ell nt}$ an indicator for a "deep" regional trading agreement between ℓ and n in year t , set equal to 1 if the agreement includes customs-related procedures or services.

In $\mathbf{X}_{i\ell t}$ we include tariffs on imported inputs (major components only) from the headquarters country. As with tariffs on assembled cars, the input tariffs enter with the functional form $\ln(1 + \text{tariff})$. As with the determinants of τ , we allow γ to depend on the existence of a deep integration agreement. In the $i\ell$ dimension, depth is obtained via an investment chapter, or if the RTA includes a services agreement or customs-related procedures. The last of these is likely to be important if the assembly factor relies on the headquarters country for car parts.

The frictions in the *in* dimension, δ_{in} , δ_{in}^e , and δ_{in}^d , differ from the previous \mathbf{X} vectors in two important respects. First, there is no analogue to tariffs in this dimension. To capture the idea that LDCs may be more protective in their regulations of domestic brands, we interact home_{in} with LDC_n , an indicator that the country in question is not a member of the OECD. Our distinctive indicator of depth for RTAs in the *in* dimension is the inclusion of a chapter on technical barriers to trade (TBTs), which often include provisions for mutual recognition of standards. As in the other dimensions, a sufficient condition to qualify as a deep agreement (in all dimensions) is the inclusion of services. The rationale here is that the operation of car dealerships is a service activity.

Online Appendix E provides more detail on measurement of the friction determinants, in particular the sources and procedures used for the tariffs and the deep RTA indicators.

B. Estimating Equations

We now express the four equations that identify the structural parameters in an estimable way in terms of observed variables with associated coefficients and fixed effects (denoted $FE^{(j)}$ where $j = 1, 2, 3, 4$).

1. Sourcing.—We transform the sourcing equation into its estimable version by substituting the τ and γ frictions from equation (20) into (6) and setting $\theta\alpha(\ln w_{\ell t} - \ln w_{\ell T}) = \mathbf{W}'_{\ell t}\mathbf{v}_1$, where $\mathbf{W}_{\ell t}$ comprises two proxies for changes in production cost: per capita income and the price level of GDP, and \mathbf{v}_1 is the set of associated coefficients.²⁴ Both proxies are expressed as logs of indices that take values of 1 in $T = 2016$.

In our setup, the probability that brand b sources model m from country ℓ to serve consumers in n in year t is the same across all b 's models. We can therefore aggregate the binary decisions into a count variable, summing over the number of models owned by b and sourced from ℓ : $S_{b\ell nt} \equiv \sum_m \mathcal{M}_{bm} S_{m\ell nt}$. Expected sourcing counts are

$$(21) \quad E[S_{b\ell nt} | \mathcal{L}_{b\ell t} = 1] \\ = \exp \left[FE_{\ell}^{(1)} - \mathbf{W}'_{\ell t}\mathbf{v}_1 - \theta\varsigma \ln q_{\ell t} - \theta \mathbf{X}'_{\ell nt}\mathbf{p} - \theta \mathbf{X}'_{i\ell t}\mathbf{g} + FE_{nt}^{(1)} \right].$$

The destination-time fixed effect is $FE_{nt}^{(1)} = -\ln \left(\sum_k \mathcal{L}_{bkt} \exp \left[FE_k^{(1)} - \mathbf{W}'_{kt}\mathbf{v}_1 - \theta\varsigma \ln q_{kt} - \theta \mathbf{X}'_{knt}\mathbf{p} - \theta \mathbf{X}'_{ikt}\mathbf{g} \right] \right)$, and the assembly-country fixed effects are interpreted as $FE_{\ell}^{(1)} = -\theta\alpha \ln w_{\ell T}$.

²⁴The sign of per capita income is ambiguous since it reflects productivity (cost-lowering) and wages (cost-raising). On the other hand, price level of GDP should have a negative influence on sourcing since it captures exchange rate over-valuation.

Brands select sources from the set of countries in which they currently have plants ($\mathcal{L}_{b\ell t} = 1$).²⁵ Equation (21) can be consistently estimated using Poisson PMLE.²⁶ Substituting the estimated coefficients and fixed effects into (6) yields C_{bnt} which we need in the next two estimation steps, market share and model entry.

2. Brand-Level Market Shares.—The second key equation to be estimated is the intensive margin of brand-level sales in each market n , year t . Including the measurable version of our δ_{int} frictions into (11), we obtain the following estimable equation of the brand's average market share over its models:

$$(22) \quad E\left[\frac{q_{bnt}}{M_{bnt}Q_{nt}} \mid \mathcal{D}_{bnt} = 1\right] = \exp\left[FE_b^{(2)} - \mathbf{W}'_{i(b)t}\mathbf{v}_2 + FE_{nt}^{(2)} - \eta\mathbf{X}'_{int}\mathbf{d} - \eta\ln C_{bnt}\right],$$

where $\eta(1 - \alpha)(\ln w_{i(b)t} - \ln w_{i(b)T}) = \mathbf{W}'_{i(b)t}\mathbf{v}_2$ captures the evolution of HQ-related costs through changes in income per capita and GDP price. Notation $i(b)$ designates the HQ country of brand b and $T = 2016$. The structural interpretation of the fixed effects becomes $FE_b^{(2)} = \eta\ln\varphi_b - \eta(1 - \alpha)\ln w_{i(b)T}$ and $FE_{nt}^{(2)} = \ln\kappa_1 + \eta\ln P_{nt}$. The C_{bnt} included as the last control comes from the sourcing probability results from equation (21) where $C_{bnt} = (\sum_k \mathcal{L}_{b\ell t}(w_{kt}^\alpha \tau_{knt} \gamma_{ikt} q_k^\varsigma)^{-\theta})^{-1/\theta}$. This regression allows us to estimate the δ_{int} determinants and provides our estimate of η . The natural way to estimate the moment condition shown in equation (22) is Poisson PML because it does not require an additional homoskedastic log-normality assumption for the error term.²⁷

3. Model Entry Decision.—As with the sourcing decision, we use the fact that our model predicts the entry probability of models to be constant for a given brand to specify the regression as a fractional probit with left-hand-side variable being the share of models offered by b in market n and year t .

Substituting $\delta_{int} = \exp(\mathbf{X}'_{int}\mathbf{d})$ and $\delta_{int}^e = \exp(\mathbf{X}'_{int}\mathbf{f}^e)$ into equation (15) and introducing fixed effects, we obtain the estimable version of the model-market entry equation,

$$(23) \quad E\left[\frac{M_{bnt}}{M_{bt}} \mid \mathcal{D}_{bnt} = 1\right] = \Phi\left[CST^{(3)} + \mathbf{X}'_{int}\mathbf{e} - (\eta - 1)\ln C_{bnt} + FE_b^{(3)} - \mathbf{W}'_{i(b)t}\mathbf{v}_3 + FE_{nt}^{(3)}\right],$$

²⁵ Our choice set assumption differs from Coşar et al. (2018) who estimate a cost function that assumes that only the countries currently producing a model enter the set of alternative sourcing locations. For example in the Coşar et al. (2018) approach, the choice set for the Renault Twingo would be France and Colombia in 2006, whereas in 2008 the choice set would switch to Colombia and Slovenia (because Renault relocated all its Twingo production for Europe from France to Slovenia in 2007). In our approach, all the countries where Renault is active in a given year are included in the choice. Thus, France, Slovenia, and Colombia (and Turkey, etc.) are sourcing options in every year. The distinction between these approaches could be seen as one of short and medium runs (in the long run, brands can expand the set of countries where they have factories).

²⁶ The fact that a multinomial discrete choice model can be estimated using Poisson with fixed effects on counts, and yielding identical results to conditional logit was discovered by Guimarães, Figueirdo, and Woodward (2003).

²⁷ Santos Silva and Tenreiro (2006) elaborate on this advantage in the context of gravity equations but it is equally applicable to the estimation of any constant elasticity relationship.

where the constant, $CST^{(3)}$ is given by $(\ln \kappa_2 - \ln \eta)/\sigma^e$. The coefficients on the gravity determinants in \mathbf{X}_{int} have structural interpretations given by $\mathbf{e} = -[(\eta - 1)\mathbf{d} + \mathbf{f}^e]/\sigma^e$. Thus, the coefficients on the *in* friction determinants combine the δ_{in} variable marketing cost effects with the δ_{in}^e fixed marketing costs. Changes in HQ-related costs also involve both determinants: $((1 - \alpha)(\eta - 1) + \zeta)/\sigma^e (\ln w_{i(b)t} - \ln w_{i(b)T}) = \mathbf{W}'_{i(b)t} \mathbf{v}_3$.

All the γ and τ geography effects are captured in the $\ln C_{bnt}$ term, the (inverse) index of how well-positioned brand b 's assembly plants are to serve market n in t . Structural interpretation of fixed effects are $FE_b^{(3)} = [(\eta - 1)\ln \varphi_b - (((1 - \alpha)(\eta - 1) + \zeta)/\sigma^e) \ln w_{i(b)T} - \beta_b^e]/\sigma^e$ and $FE_{nt}^{(3)} = [\ln Q_{nt} + \eta \ln P_{nt} - (1 - \zeta) \ln w_{nt} - \mu_n^e]/\sigma^e$.

4. Brand Entry Decision.—The empirical version of brand entry is obtained inserting $\delta_{int}^d = \exp(\mathbf{X}'_{int} \mathbf{f}^d)$ into (19), with the headquarter inputs needed for brand entry fixed costs specified as $(\zeta/\sigma_d)(\ln w_{i(b)t} - \ln w_{i(b)T}) = \mathbf{W}'_{i(b)t} \mathbf{v}_4$:

$$(24) \quad \Pr(\mathcal{D}_{bnt} = 1) = \Phi\left(\frac{1}{\sigma_d} \ln E[\pi_{bnt}] - \mathbf{X}'_{int} \mathbf{f}^d/\sigma_d + FE_b^{(4)} - \mathbf{W}'_{i(b)t} \mathbf{v}_4 + FE_{nt}^{(4)}\right).$$

Inverting the coefficient of our calculated profitability of brand b in market n gives a direct estimate of the standard deviation of log fixed costs, σ_d . The structural interpretations of the brand and destination-time fixed effects are $FE_b^{(4)} = -(\beta_b^d + \zeta \ln w_{i(b)T})/\sigma_d$ and $FE_{nt}^{(4)} = -(\mu_n^d + \ln w_{nt}^{1-\zeta})/\sigma_d$. We estimate equation (24) as a binary probit with the constructed $\ln E[\pi_{bn}]$ on the right-hand side, together with the friction determinants, brand, and destination fixed effects.

C. Identification of Structural Parameters

Equations (21), (22), (23), and (24) estimated sequentially, yield all the parameters needed to solve the model. Our model is specified such that there is only one estimate for each parameter of interest.

Sourcing.—Coefficients from the sourcing equation (21) have structural interpretations $-\theta \mathbf{p}$, $-\theta \mathbf{g}$, and $-\theta \zeta$. Thus we can calculate τ and γ friction parameters, as well as the scale elasticity, by dividing our estimates by $-\theta$, the coefficient on car tariffs. The fixed effects on origin countries combined with θ allows us to recover $w_\ell^\alpha = \exp(-FE_\ell^{(1)}/\theta)$. We recover the share of headquarters' country input in the total costs of production by using the ratio of our two direct price shifters in γ and τ . Recalling that $\ln \gamma_{ilt} = (1 - \alpha) \ln \tau_{ilt}^H$, we estimate $1 - \alpha$ by dividing the coefficient on $\ln(1 + \text{parts tariff}_{ilt})$ by the coefficient on $\ln(1 + \text{car tariff}_{lt})$.

Market Share.—The coefficients on the friction determinants correspond to $-\eta \mathbf{d}$. Dividing by $-\eta$, the coefficient on $\ln C_{bn}$ in equation (22), yields the vector of δ friction parameters \mathbf{d} . The price indices, P_n , are proportional to $\exp(FE_{nt}^{(2)}/\eta)$, the exponentiated destination fixed effects divided by the consumer elasticity. A combination

of brand-related parameters involving physical productivity and the factor costs in the headquarters of brand b is given by $\varphi_b/w_{i(b)}^{1-\alpha} = \exp(FE_b^{(2)}/\eta)$.

Model Entry.—We obtain σ_e as $1 - \eta$ divided by the coefficient on $\ln C_{bn}$ in equation (23). Model entry fixed costs depend on $\ln w_{nt}^{1-\zeta} + \mu_{nt}^e + \ln \delta_{in}^e$. Inverting the definition of the destination fixed effects, $\ln w_{nt}^{1-\zeta} + \mu_{nt}^e = \ln Q_{nt} + \eta \ln P_{nt} - FE_{nt}^{(3)}/\sigma_e$. The friction parameters for model entry (needed to compute δ_{in}^e) are obtained from the coefficients on friction determinants, \mathbf{e} , combined with variable version of δ costs obtained from the market share equation, using the formula $\mathbf{f}^e = -\mathbf{e}\sigma^e - \mathbf{d}(\eta - 1)$. The remaining components of model entry fixed costs F_{mnt}^e , namely $\beta_b^e + \ln w_{i(b)}^\zeta$, require more involved manipulations of $FE_b^{(2)}$ and $FE_b^{(3)}$, which we relegate to online Appendix B. There we also show how the fixed effects from market share and model entry regressions can be used to reconstruct $E[\pi_{mnt}]$, and then $E[\pi_{bnt}]$, which is needed in the brand entry equation.

Brand Entry.—Equation (24) estimates σ_d as the inverse of the coefficient on $\ln E[\pi_{bnt}]$. Destination and brand fixed effects yield estimates of $\mu_n^d + \ln w_{nt}^{1-\zeta}$ and $\beta_b^d + \ln w_{i(b)t}^\zeta$ respectively. Multiplying the coefficients on the friction determinants by σ_d yields the vector \mathbf{f}^d .

The market share and model entry equations both depend on $\ln C_{bnt}$, a variable generated using the estimated coefficients of the sourcing equation. The key determinant of brand entry, $\ln E[\pi_{bnt}]$, also depends on C_{bnt} as well as estimates of fixed costs and the price index extracted from the market share and model entry equations. Since the presence of these generated regressors has the potential to bias the standard errors, we report bootstrap standard errors for all equations. As brands make repeated decisions (sourcing, entry, etc.), it is important to make standard errors robust to possible correlation in the errors by clustering. In the bootstrap setting this is achieved by drawing *all* the observations for a given brand cluster in each of the four equations.²⁸

²⁸There are 45 brands that cannot enter the sourcing stage estimation since they only have a single country to source models from. We therefore conduct two separate draws in each repetition of the bootstrap. First, we draw (with replacement) among the 93 brands that can source from at least two countries in a given year. The estimated parameters allow us to construct $\ln C_{bm}$ for all brand-destination combinations relevant in the next three steps. We then draw (with replacement) from the full set of 138 car brands and use this bootstrap sample for brand market share, model entry, and brand entry decisions. This completes the procedure needed for one replication. In order to choose the number of replications, we follow procedures described in Andrews and Buchinsky (2000). This involves (i) running a number of initial bootstrap repetitions (500 in our case), (ii) calculating an excess kurtosis statistic for every parameter, (iii) running 500 additional bootstrap repetitions. The number of added repetitions is sufficient to set the percentage deviation bound (pdb) for all parameters of interest below 5 percent with a confidence level of 95 percent, taking into account the deviations from normality implied by the step (ii) calculation (see their equation 3.4).

IV. Results

A. Baseline Estimates

Table 3 reports the coefficients for each of the four estimating equations.

Sourcing Estimates.—Column 1 reports our sourcing results. The estimates reveal the importance of trade costs in selecting sources. Home effects are large: the implied increase in the odds of choosing a location is obtained by exponentiating the coefficient. Plants located in the market being served have odds of being chosen that are 2.6 times higher. Distance from the market also significantly reduces the share of models sourced from an assembly country.

The coefficient of -7.7 on the log of 1 plus the car tariff implies $\theta = 7.7$ as the critical elasticity of substitution between sources.²⁹ Deep regional trade agreements augment the odds of being chosen by 28 percent, even after accounting for the tariffs applied by the destination market to the different possible origins of the car. Both tariffs and deep RTA effects will be important for our counterfactuals where we experiment with scenarios involving different combinations of RTA and tariff changes.

The estimates of the γ frictions are much less precise, with standard errors several times those estimated for trade frictions. Two of the effects, distance and language, do not even enter with the expected sign, although neither is significantly different from zero. The significant effect is that assembly locations in the brand's home country are $\exp(2.248) \approx 9.5$ times more likely to be selected. The elasticity on the car parts tariff can be used to infer the share of assembly costs attributable to components from the headquarters country, $(1 - \alpha)$ in the cost equation, which is about 37 percent ($2.87/7.7$). While the precise value of this ratio should be taken with caution, we now have direct evidence of the importance of intermediate inputs from the headquarters country. This feature of the MP model has major qualitative and quantitative implications for the impact of trade liberalization, as we shall see in the counterfactuals. Deep RTAs between assembly and headquarter countries are estimated to have a larger effect on sourcing than deep RTAs between assembly and consumer countries, but the standard error is also larger.³⁰

Our method estimates external economies of scale based on the magnitude of the revealed preferences of brands for assembly locations with high aggregate output (q_{it}). We estimate $-\theta\varsigma = 0.27$ which, given our $\theta = 7.7$ estimate, implies $\varsigma = -0.035$. As pointed out by Goldberg and Verboven (2001), unobserved factors can both make a location attractive and increase its aggregate production. This would lead to upward bias in our estimate of $-\theta\varsigma$. Our estimation mitigates

²⁹The estimate of θ when using all the locations of the parent firm as options for sourcing also rounds to 7.7 as can be seen in online Appendix Table F.2. The firm-variety approach also shows similar coefficients for the other determinants of trade costs.

³⁰Online Appendix F.3 presents a set of estimates of γ frictions from an alternative moment conditions that are consistent with the double-CES MP model. The main takeaway from Table F.3 is that the coefficients on Deep RTA_{it} and on tariffs on car parts are stronger and more significant than in our baseline results. However, since the estimates of θ are also larger, the AVE of deep RTA remain very similar to the baseline. The ratio of coefficients between car and parts tariffs also provides comparable alternative estimates of $1 - \alpha$ ranging between 29 percent and 50 percent.

TABLE 3—BASELINE RESULTS

Decision:	Sourcing	Market share	Model entry	Brand entry
Dependent variable:	$S_{b\ell nt}$	$\frac{q_{bnt}}{M_{bnt}Q_{nt}}$	$\frac{M_{bnt}}{M_{bt}}$	\mathcal{D}_{bnt}
Method:	PPML	PPML	Frac. probit	Probit
	(1)	(2)	(3)	(4)
home _{ℓn}	0.973 (0.142)			
ln dist _{ℓn}	−0.323 (0.06)			
language _{ℓn}	−0.042 (0.068)			
ln(1 + car tariff _{ℓn})	−7.696 (0.398)			
Deep RTA _{ℓn}	0.246 (0.085)			
home _{$i\ell$}	2.248 (0.49)			
ln dist _{$i\ell$}	0.166 (0.136)			
language _{$i\ell$}	−0.218 (0.308)			
ln(1 + parts tariff _{$i\ell$})	−2.872 (3.197)			
Deep RTA _{$i\ell$}	0.495 (0.301)			
ln q_ℓ	0.27 (0.076)			
home _{in}		0.816 (0.249)	0.26 (0.068)	0.718 (0.422)
home _{in} × LDC _{n}		−0.028 (0.36)	0.829 (0.118)	1.328 (0.576)
ln dist _{in}		−0.339 (0.104)	−0.059 (0.017)	0.009 (0.069)
language _{in}		0.289 (0.14)	0.068 (0.036)	0.017 (0.122)
Deep RTA _{in}		−0.04 (0.14)	0.121 (0.042)	0.165 (0.117)
ln C_{bn}		−3.874 (0.97)	−0.512 (0.221)	
ln $E[\pi_{bn}]$				0.595 (0.067)
Observations	347,542	46,299	46,300	128,589
R^2	0.824	0.601	0.715	0.618
Fixed effects	ℓ, bnt	b, nt	b, nt	b, nt

Notes: Standard errors bootstrapped with brand b clusters over 1,000 replications (see footnote 28). R^2 is the squared correlation of fitted and true dependent variables except in specification (4) where the pseudo- R^2 is reported. Each regression controls for log per capita income and price level of the assembly country.

this through the inclusion of country-specific fixed effects (identifying the degree of external returns in the within dimension). However, our estimate of the external scale elasticity should still be seen as an upper bound since time-varying cost shocks could be correlated with changes in $q_{\ell t}$.

Two simple and common ways to mitigate the endogeneity bias that can be applied to our case are (i) lagging $q_{\ell t}$, (ii) constructing a Bartik-style prediction for $q_{\ell t}$ to be used in a control function approach. The Bartik prediction applies changes in total demand Q_{nt} to brand-origin-destination market shares fixed at their 2002 levels. We also report an even more demanding Bartik specification where the shares to be shifted are brand-origin market shares. In regressions reported in online Appendix F.1, we followed the two approaches. Lagging lowers the magnitude of increasing returns to $\varsigma = -0.032$, whereas the Bartik approach reduces it further to $\varsigma = -0.024$ (for the most demanding one, $\varsigma = -0.028$ for the other one). Neither approach is perfect, but both values support our use of $\varsigma = -0.035$ as the upper bound of the parameter governing the strength of interdependencies.

The sole estimate of the external returns elasticity in the motor vehicle industry that we know of is from a recent version of Bartelme, Costinot, and Rodríguez-Clare (2018). Using origin-destination trade flows, they estimate a compound parameter analogous to our ς as 0.15. This is four times larger than our non-instrumented “upper bound” estimate even though they employ a demand-based instrument. There are two quantitatively important reasons to expect their study to yield larger estimates. First, the trade elasticity entering their estimate is 5.7, which is 26 percent smaller than our $\theta = 7.7$. Second, our sourcing method neutralizes variation in the number, quality, and productivity of brands and models produced in each country. Since the Bartelme, Costinot, and Rodríguez-Clare (2018) method is identified by variation from more aggregated trade flows, their scale elasticity is more encompassing as it also captures any additional mechanisms linking the national scale of production to *firm*-level performance variables. As we hold such variables constant, $\varsigma = -0.035$ should be interpreted as the output scale elasticity on the costs of different countries in assembling a given variety.

The literature also provides some estimates of *internal* returns to scale, estimated from data on prices or costs. These papers omit the multinational production dimension that causes the combinatorial computational challenge. Our implied $\varsigma = -0.035$ lies within the interval of those estimates of internal returns to scale for the car industry. They are bigger than the range of values reported by Goldberg and Verboven (2001), -0.006 to -0.03 , but smaller than the -0.11 and -0.07 provided by Verboven (1996) and Fuss and Waverman (1990).

Brand-Level Market Share Estimates.—Determinants of a brand’s market share are estimated in column 2 of Table 3. The estimate of $\eta = 3.87$ (from the coefficient on $\ln C_{bnt}$) is substantially smaller than the θ obtained in the sourcing decision. It implies that there is considerably more heterogeneity in consumer evaluations of brands than in car maker evaluations of assembly locations. One concern with our estimate of η is that it could be biased toward zero if brands with high unobserved demand shocks systematically locate their assembly in high-wage countries. However, the markup implied by our estimate ($\eta/(\eta - 1) = 35$ percent) lies within the highly dispersed set of results found in the Industrial Organization literature on cars. The three pioneering papers, Goldberg (1995); Berry, Levinsohn, and Pakes (1995); and Feenstra and Levinsohn (1995) report average markups of 38 percent, 24 percent, and 18 percent, respectively. Verboven (1996) and Berry, Levinsohn, and Pakes (1999) show markups of specific models that range from 8 to 36 percent

in the former paper and 24 to 42 percent in the latter. Most recently, Coşar et al. (2018) report in their Table 11 average firm-market markups ranging from 6 percent (Peugeot in Brazil) to 12.4 percent (Peugeot in France).³¹

Among the determinants of marketing frictions, consumers are more than twice as likely to select a home brand, corroborating the large home bias found by Coşar et al. (2018). In addition, we estimate that increasing consumer distance from headquarters sharply lowers market shares, even controlling for distance from the consumer to the assembly location, which is captured by C_{bn} in the same regression. Sharing a common language reduces variable marketing costs, increasing the average market share of a brand in those markets by around one-third compared to destinations where consumers speak a different language.

The effect of deep regional agreements is perversely negative in this regression, but its magnitude is small, and is very imprecisely estimated. Deep RTA status operates very strongly on the extensive margins: sourcing in column 1, model entry in column 3, and brand entry in column 4.

Model Entry Estimates.—Column 3 of Table 3 shows that all the marketing cost determinants have the expected signs and are highly significant. More models are offered in the home country of the brand, especially when this country is a developing one. Spatial proximity promotes entry as well. Deep RTAs between the headquarter country (i) and the market (n) increase the fraction of models offered by 14 percent (calculated as the average semi-elasticity). As it seems unlikely that RTAs change preferences, we see the deep RTA_{int} effects as supporting the cost-shifter interpretation. Under this approach, our δ^e frictions include various types of marketing efforts, in particular managing dealership networks. This may be facilitated by the freer movement of skilled workers that is a commonly included provision of RTAs (e.g., NAFTA, EU). The RTA_{int} effect may also capture the greater ease of compliance with regulatory standards if the head office lies within the region and is therefore more able to exert influence on specific requirements in harmonized rules. Note also that the significance of this fixed cost dimension of RTA effects contrasts with the weak impact of the same variable on brand-level sales. This suggests that deep RTAs reduce the fixed costs of model entry between HQ and destination (δ_{int}^e), rather than the variable marketing costs (δ_{int}), that affect brand sales as well.

The overall cost of serving n for brand b (C_{bn} , constructed from sourcing estimates of the first column) strongly reduces the share of models offered in a market as expected. Dividing $(1 - \eta)$ by that coefficient provides our estimate of $\sigma^e = 5.61$. The fixed cost of introducing new models in a market therefore exhibits very large variation.

Brand Entry Estimates.—The last column of our baseline table shows that a number of determinants for model entry are also relevant for whether the brand is present altogether in a market. Domestic entry is naturally a dominant feature of the data (among the exceptions are Acura, Lexus, Infiniti, Isuzu, and Scion, which

³¹ The firm-variety estimate of $\eta = 1.9$ (standard error of 1) shown in online Appendix F.2 implies markups over 100 percent, well outside this range. We attribute this low, noisy estimate to measurement error in C_{bn} caused by the firm-variety approach.

TABLE 4—FRICTION PARAMETERS

Friction: Estimate:	Variable costs			Fixed costs	
	τ ρ	γ \mathbf{g}	δ \mathbf{d}	δ^e \mathbf{f}^e	δ^d \mathbf{f}^d
home	-0.126	-0.292	-0.211	-0.856	-1.206
home \times LDC			0.007	-4.674	-2.231
ln distance	0.042	-0.022	0.088	0.078	-0.016
Common language	0.005	0.028	-0.075	-0.165	-0.029
RTA (deep)	-0.032	-0.064	0.010	-0.708	-0.278

Notes: Elasticities used to obtain frictions: $\theta = 7.7$, $\eta = 3.87$, $\sigma^e = 5.61$, and $\sigma^d = 1.68$. Calculations, described in Section IIIC, use coefficients from Table 3.

are not sold in Japan, and Hummer which continued to be sold in Japan and Taiwan after ending sales in the United States). Deep RTAs also reduce the fixed costs of establishing distribution networks, resulting in a 27 percent larger probability of brand entry. The inverse of the coefficient obtained on $\ln E[\pi_{bn}]$ yields our estimate of $\sigma^d = 1.68$.

B. Interpreting the Structural Parameters

Using the procedures detailed in Section IIIC, we now proceed to report and interpret the structural parameters underlying our estimates. The three sets of parameters relevant for variable costs, trade costs τ , multinational production costs γ , and marketing costs δ , are reported in the first three columns of Table 4. For each, the coefficient on the k th element of \mathbf{X} maps to proportional increases in price (an ad valorem equivalent) of $\exp(\rho^{(k)} \Delta X^{(k)}) - 1$.

In the case of the τ frictions, we can relate our estimates to what is known from direct measurement of the frictions. The elasticity of τ with respect to distance is of particular interest to us since it has been estimated on its own using various types of data in the literature, including the effect of physical distance on freight costs. The \ln distance cost elasticity in column 1 of Table 4 is $\rho^{\text{distance}} = 0.042$. Coşar et al. (2018) report a somewhat smaller value of $\rho^{\text{distance}} = 0.016$ (Table 12, column IV). Both estimates of ρ^{distance} , fit in the “reasonable range” of 0.01 to 0.07 in the literature summarized by Head and Mayer (2013). Our results imply that the distance effects on trade flows can be fully explained without reference to the “dark matter” invoked by Head and Mayer (2013) to explain aggregate distance elasticities of -1 or higher. This is not surprising since the main candidate explanations for dark matter—poor information and differences in preferences—should be accounted for in the δ_{in} marketing cost parameters.

The estimates from the market share equation imply large variable costs in the headquarter-market dimension. The distance elasticity is 0.088, more than double the corresponding transport cost elasticity 0.042. Our elasticity is also larger than the Wang (2017) estimate of 0.044 based on export sales of foreign-owned manufacturers in China. By contrast, the home bias in marketing costs for cars is 0.21, quite comparable to Wang’s (2017) estimate of 0.24 for US-headquartered firms, but lower than his 0.95 for Japan and Korea affiliates.

How should we interpret the parameters shown in Table 4? The first thing to note is that the three chief variable cost frictions in the model, $\tau_{\ell n}$, $\gamma_{i\ell}$, and δ_{in} all require a normalization to be meaningful. Put another way, any of these frictions could be scaled up by a constant without changing any of the endogenous variables. The normalization we use is the internal friction within the United States.³² An estimated $\delta_{in} = 1.3$, for example means that firms headquartered in i inflate their delivered prices to consumers in n by 30 percent more than firms headquartered in the United States inflate costs for their home-country consumers.

The fixed cost parameters δ_{in}^e and δ_{in}^d are also defined relative to a reference dyad. Thus again if we estimate $\delta_{in}^e = 1.3$ it means that fixed costs of adding another model are 30 percent higher for firms from i offering models in n than US firms adding a model in the United States. This interpretation also holds for δ_{in}^d . To facilitate comparisons with the variable cost frictions, we want to convert δ_{in}^e and δ_{in}^d into their ad valorem equivalents (AVE). This can be accomplished using the following thought experiment: Let gross profits be a given ratio of fixed costs. Suppose we shock fixed costs by δ_{in}^e . Then, in order to keep the previous ratio (and thus model entry probability) unchanged, gross profits must rise by the same proportion. Using (15), we can find a δ'_{in} that would achieve this proportional increase, i.e., $(\delta'_{in})^{-(\eta-1)} = \delta_{in}^e$. Inverting we obtain $\delta'_{in} = (\delta_{in}^e)^{-1/(\eta-1)} < 1$. We define the AVE(δ_{in}^e) = $1 - (\delta_{in}^e)^{-1/(\eta-1)}$. Determining the AVE for δ_{in}^d is more complex because δ_{in}^d affects $E[\pi_{bn}]$ through multiple non-separable channels. We can still conduct an analogous thought experiment. Define δ'_{in} as the variable marketing cost that would magnify expected gross profits of the brand by the same factor as δ_{in}^d :

$$\frac{E[\pi_{bn}](\delta'_{in})}{E[\pi_{bn}](\delta_{in} = 1)} = \delta_{in}^d.$$

The δ'_{in} can be found by defining a function $g(\delta'_{in}) = E[\pi_{bn}](\delta'_{in}) - \delta_{in}^d E[\pi_{bn}](\delta_{in} = 1)$, and then solving for a root.

Results of those calculations are reported in Figure 4 adapted from Arkolakis et al. (2018) with the addition of an edge corresponding to variable and fixed marketing costs δ_{in} , δ_{in}^e , and δ_{in}^d . On each edge of the triangle we report the relevant frictions, which is the median value calculated in our sample for the year 2016. The variable frictions are $\tau_{\ell n} - 1 = 24$ percent, $\gamma_{i\ell} - 1 = 31$ percent, and $\delta_{in} - 1 = 33$ percent. The sales of a model produced in ℓ , headquartered in i , and sold in n would therefore face a total cost-increasing friction of 116 percent $(\tau_{\ell n} \gamma_{i\ell} \delta_{in} - 1)$. This is not out of line with the figures provided in Table 7, column 1 of Anderson and van Wincoop (2004), ranging from 91 to 174 percent. Being the largest of the three frictions, variable marketing costs are quantitatively important enough to warrant inclusion in the multinational production framework. Set on top of that, the extra burden of fixed marketing costs have AVEs of 9.7 percent and 26 percent for model and brand entry, reinforcing the finding that the new dimension of frictions we added to the MP model is quantitatively very important.³³ The only

³²This means countries with smaller internal distances than the United States can have $\tau_{nn} < 1$.

³³Wang (2017) estimates the first and third components of the marketing costs, using foreign affiliate trade data from China. Since Wang's sample includes all manufacturing industries, the large magnitudes he obtains suggest that these frictions are important beyond the car industry.

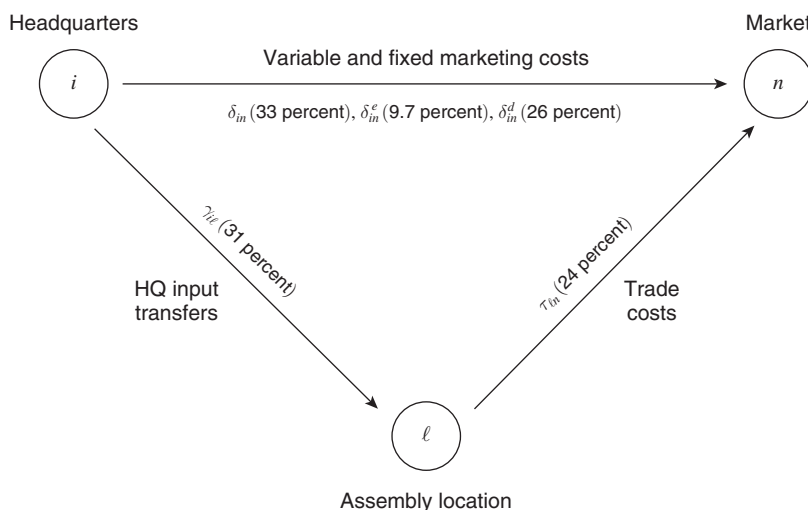


FIGURE 4. FRICTIONS IMPEDING MULTINATIONAL FLOWS

Notes: Friction ad valorem equivalents shown in parentheses. See text for explanation.

frictions that change in our counterfactual experiments are tariffs and RTAs. Tariffs are ad valorem already. The AVE of RTAs for the three dimensions of our friction triangle are 3.2 percent for τ , 6.4 percent for γ , and a total of 29 percent for the combined effects of δ , δ^e , and δ^d .

The friction estimates shown in Figure 4 do not distinguish cost-based interpretations of τ_{ln} , γ_{il} , and δ_{in} from preference-based interpretations. For example, a desire by consumers to “buy local” to support workers has the same effect on sourcing as an increase in τ_{ln} . Similarly, if Japanese workers had a reputation for quality control, then Toyota’s assembly facilities outside Japan would have their sourcing shares reduced in a way that would be isomorphic to an increase in γ_{il} . Finally, spatially correlated taste differences (e.g., for fuel economy, safety, or shape) could be equivalent in their effects on market shares to a rise in δ_{in} due to higher distribution costs in remote markets. Allowing for such preference effects in the utility function would just add more parameters that could not be identified separately from the ones existing in our specifications. To estimate separately the cost and demand-side effects would require a different estimation strategy that uses price information. Such a data requirement would severely limit the geographic scope of the study. For the purposes of our counterfactuals on how integration affects production and trade, we do not need to disentangle cost mechanisms from preference mechanisms.³⁴

The index of local assembly costs in each country, $c_\ell \equiv w_\ell^\alpha q_\ell^\zeta$, is a key parameter of the model because it tells us where production would gravitate in the absence of frictions. We obtain the 2016 levels as $c_\ell = \exp(-FE_\ell^{(1)}/\theta) q_\ell^\zeta$ from

³⁴ Goldberg and Verboven (2001) separate out the home bias into demand and cost components. Coşar et al. (2018) estimate cost-based (γ_{il}) frictions of distance from a brand’s home. They also have a home-brand effect in preferences that would operate as a δ_{in} effect in our model.

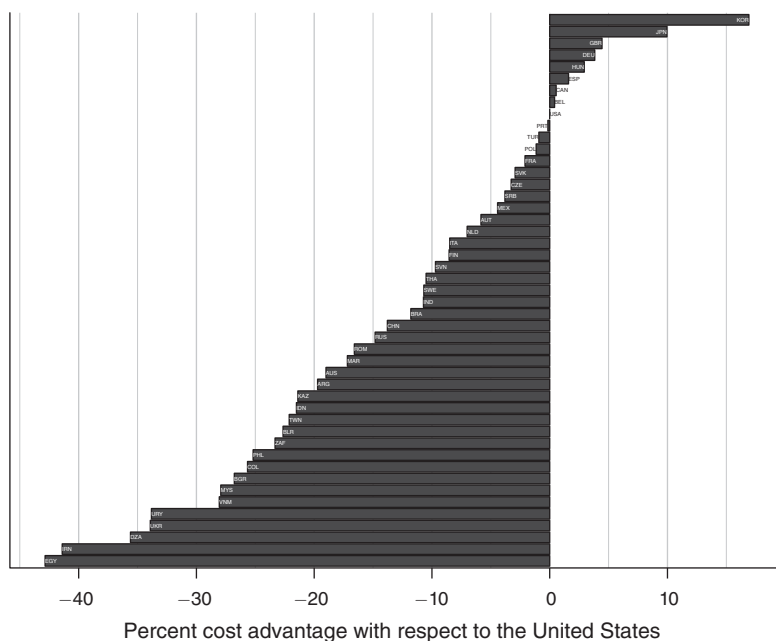


FIGURE 5. COST ADVANTAGE INFERRED FROM SOURCING DECISIONS

the estimates in the sourcing equation. The c_ℓ can only be identified up to a scalar so we express them all as cost advantages with respect to the United States, i.e., $100 \times (c_{USA} - c_\ell)/c_\ell$.

Figure 5 graphs the cost advantage of the 47 assembly countries we use in estimation. The clear “winner” for the car industry is South Korea with Japan as runner up. Egypt is the outlier in the other direction. The implied differences in unit assembly costs are quite small for the main European brand headquarters. France, the United Kingdom, and Germany are within a few percentage points from each other. Canada is also very similar to its southern neighbor. The similarity in costs between these countries suggests that friction changes have the potential to cause substantial reallocations in production.

C. Segment-Level Estimates

Up until this point we have treated the market for cars as one in which all car models substitute symmetrically for each other. In reality, we think the Camry and Accord mid-sized sedans are closer substitutes for each other than either would be for a Sienna minivan (all Honda models). To allow for more realistic substitution patterns, we follow the tradition initiated by Goldberg (1995) and Verboven (1996) which groups models according to their primary function. In this extension, households receive idiosyncratic utility shocks that determine the choice of the segment from which they then select a model. This approach leads to a modified

version of equation (22), in which market shares are measured *within segments* denoted s :

$$(25) \quad E\left[\frac{q_{bnst}}{M_{bnst}Q_{nst}} \mid \mathcal{D}_{bnt} = 1\right] \\ = \exp\left[FE_{bs}^{(2)} - \mathbf{W}'_{i(b)t}\mathbf{v}_{2s} + FE_{nst}^{(2)} - \eta_s \mathbf{X}'_{int}\mathbf{d}_s - \eta_s \ln C_{bnt}\right].$$

All of the demand parameters are segment-specific. While estimating an upper branch decision to determine Q_{nst} would be possible, the effect would simply be to bring back inter-segment substitution. Instead, we take Q_{nst} as exogenous in our counterfactual simulations, a modeling decision that restricts all substitution to take place *within* segments.³⁵ Bracketing the range of substitution is important for our counterfactuals since the unified market assumption has the potential to exaggerate the response to changes in frictions.

Allowing for segment-specific demand elasticities and nested substitution also affects the model-level expected profits: equation (14) features the size of the market and the price index, which are now both segment-specific. In addition, the size of η_s determines the response to the cost index C_{bn} . Model entry should therefore also be run for each segment. The dependent variable changes from M_{bnt}/M_{bt} to M_{bnst}/M_{bst} . Once parameters from market share and model entry regressions have been collected, $E(\pi_{bnt}) = \sum_s E(\pi_{bnst})$ is computed to yield an estimation of brand entry choices, that takes into account the underlying segmentation of the car market.³⁶

Table 5 provides segment-level estimates for the market share equation (upper panel) and model and brand entry (lower panel). We consider five segments: small cars, big cars, multi-purpose vehicles (MPVs, which includes minivans), sport utility vehicles (SUVs), and a combination of sport and luxury cars.³⁷ The most important set of coefficients obtained is the one on $\ln C_{bn}$, revealing η_s . The demand side elasticity for small cars, at -4.4 is reasonably close from the one obtained in the baseline regression (-3.87). The price response is larger for MPVs, big cars, and SUVs, all around -6 . The *Sport&lux* η of just 0.458 implies an infinite markup over costs, which makes it impossible to include this segment in aggregate profit calculations or counterfactuals. The average η_s , excluding *Sport&lux*, is -5.6 . This higher elasticity of market shares with respect to changes in C_{bn} will tend to increase responsiveness to tariff changes in the segment version of the model, offsetting to some extent the elimination of cross-segment substitution.

The segment specification exhibits a similar pattern of friction estimates to the ones obtained in Table 3. Market shares of home brands are higher in every segment; physical distance has a robust negative effect. Deep RTAs again lack significant effects on the intensive margin. However, the positive impact of deep RTAs

³⁵ Fixed Q_{nst} can be rationalized as the limiting value of a nested model in which the “segment shock” has very large variance.

³⁶ The precise steps to construct $E(\pi_{bnst})$ are the same as in online Appendix B, with the appropriate s subscript when the segment dimension is relevant.

³⁷ We based the categorization on a combination of information on the “Global sales sub-segment” (a functional categorization) and the “Global sales segment” (a size categorization) of the model specified in IHS original data. Our segments therefore represent roughly similar-sized sets of models, grouped by categories suggested by the industry consultancy from which we bought the data.

TABLE 5—SEGMENT-LEVEL MARKET SHARE AND MARKET ENTRY ESTIMATES

Dependent variable:	Average market share				
Method:	PPML				
Segment:	Small cars	Big cars	MPV	SUV	Sport&lux
	(1)	(2)	(3)	(4)	(5)
$home_{in}$	0.545 (0.379)	0.383 (0.338)	0.191 (0.362)	0.553 (0.297)	1.21 (0.174)
$home_{in} \times LDC_n$	0.198 (0.442)	-0.467 (0.749)	-0.366 (0.853)	-0.399 (0.31)	-1.143 (0.698)
$\ln dist_{in}$	-0.382 (0.117)	-0.269 (0.123)	-0.259 (0.102)	-0.121 (0.104)	-0.155 (0.081)
$language_{in}$	0.216 (0.179)	0.327 (0.147)	0.285 (0.216)	0.518 (0.19)	0.219 (0.105)
$Deep RTA_{in}$	0.042 (0.193)	-0.127 (0.147)	-0.083 (0.182)	-0.171 (0.132)	0.165 (0.169)
$\ln C_{bn}$	-4.442 (1.195)	-6.046 (1.045)	-6.257 (1.665)	-5.558 (0.795)	-0.458 (0.692)
Observations	29,564	26,692	16,415	29,379	17,225
R^2	0.476	0.494	0.388	0.434	0.542

Dependent variable:	Model entry (fraction)					Brand entry
Method:	Fractional probit					Binary probit
Segment:	Small cars	Big cars	MPV	SUV	Sport&lux	All
	(1)	(2)	(3)	(4)	(5)	(6)
$home_{in}$	0.029 (0.079)	0.218 (0.09)	0.222 (0.131)	0.152 (0.084)	0.224 (0.074)	0.845 (0.471)
$home_{in} \times LDC_n$	0.779 (0.132)	0.076 (0.226)	-0.094 (0.257)	0.541 (0.195)	-0.563 (0.162)	1.088 (0.614)
$\ln dist_{in}$	-0.064 (0.022)	-0.07 (0.03)	-0.154 (0.033)	-0.085 (0.03)	-0.045 (0.041)	-0.065 (0.073)
$language_{in}$	0.088 (0.052)	0.04 (0.038)	-0.022 (0.041)	0.049 (0.041)	0.08 (0.049)	-0.2 (0.149)
$Deep RTA_{in}$	0.047 (0.04)	0.155 (0.057)	0.045 (0.092)	0.089 (0.046)	0.052 (0.063)	0.365 (0.118)
$\ln C_{bn}$	-0.953 (0.311)	-0.529 (0.403)	-0.733 (0.497)	-0.148 (0.333)	0.178 (0.499)	
$\ln E[\pi_{bn}]$						0.554 (0.082)
Observations	27,871	24,864	15,838	25,794	15,916	113,207
R^2	0.747	0.681	0.757	0.711	0.464	0.650

Notes: Standard errors bootstrapped with brand b clusters over 1,000 replications (see footnote 28). R^2 is the squared correlation of fitted and true dependent variables. Each regression controls for log per capita income and price level of the assembly country.

on model and brand entry show up in the segment specification. The model entry effects are generally smaller in the segment specification whereas the brand entry effects of RTAs are estimated to be larger.

The model entry regressions have the expected negative signs on $\ln C_{bn}$ for each segment other than *Sport&lux*. These coefficients are used to identify the dispersion parameters on fixed costs distributions for model entry (σ_s^e). In brand entry, column 6, it is the coefficient on $\ln E(\pi_{bn})$ which plays this role for estimating σ_d^d . Both sets of dispersion parameters are reported in Table 6. The σ_s^e are larger for several segments but σ_d^d hardly changes when moving to segments. This table also

TABLE 6—FRICTION PARAMETERS (SEGMENT-LEVEL)

Friction: Estimate:	δ \mathbf{f}	δ^e \mathbf{f}^e	δ^d \mathbf{f}^d
<i>Small cars</i> ($\eta = 4.44, \sigma^e = 3.61$)			
home	−0.123	0.318	−1.524
home \times LDC	−0.045	−2.660	−1.964
ln distance	0.086	−0.066	0.118
Common language	−0.049	−0.149	0.361
RTA (deep)	−0.010	−0.138	−0.659
<i>Big cars</i> ($\eta = 6.04, \sigma^e = 9.54$)			
home	−0.063	−1.757	−1.524
home \times LDC	0.077	−1.119	−1.964
ln distance	0.044	0.443	0.118
Common language	−0.054	−0.112	0.361
RTA (deep)	0.021	−1.584	−0.659
<i>Multi-purpose vehicles</i> ($\eta = 6.26, \sigma^e = 7.17$)			
home	−0.031	−1.427	−1.524
home \times LDC	0.058	0.368	−1.964
ln distance	0.041	0.885	0.118
Common language	−0.046	0.397	0.361
RTA (deep)	0.013	−0.395	−0.659
<i>Sport utility vehicles</i> ($\eta = 5.56, \sigma^e = 30.73$)			
home	−0.100	−4.225	−1.524
home \times LDC	0.072	−16.950	−1.964
ln distance	0.022	2.517	0.118
Common language	−0.093	−1.074	0.361
RTA (deep)	0.031	−2.872	−0.659
<i>Sport&lux</i> ($\eta = 0.46, \sigma^e = 3.05$)			
home	−2.641	−2.115	−1.524
home \times LDC	2.496	3.071	−1.964
ln distance	0.339	0.321	0.118
Common language	−0.478	−0.502	0.361
RTA (deep)	−0.359	−0.355	−0.659

Notes: Elasticities used to obtain frictions: (i) the segment-specific ones are given in the table, (ii) for brand entry, $\sigma^d = 1.80$. Calculation of those frictions are described in the text and use coefficients from Tables 3 and 5.

calculates the structural parameters associated with all the frictions. A comparison with the corresponding frictions in the last three columns of Table 4 shows lower tariff-equivalents for home bias (ranging from 3–12 percent excluding the *Sport&lux* outlier). There are higher effects for deep RTAs on brand entry fixed costs, and heterogeneous deep RTA effects on model entry across the segments, with SUVs and big cars having very large fixed costs of adding models.

V. Counterfactual Methods

The counterfactuals investigate a set of trade policy scenarios involving shocks to tariffs and the deep RTA indicators. They treat as exogenous country-level new car purchases (Q_n), each brand's total number of models (M_b), and each brand's set of production locations, $\mathcal{L}_{b\ell}$. Since the data used for each variable comes from a single year (2016, the last available in our sample) we suppress the time subscripts in this section. We start by highlighting the features of our counterfactuals that require further elaboration.

A. Three Methodological Considerations

The first such feature is our treatment of country-level external returns to scale. Our counterfactuals solve for the equilibrium for a given q_ℓ and then update q_ℓ to deliver new unit costs and a new equilibrium. This iteration continues until a fixed point is reached. There is concern over existence and uniqueness of equilibria with increasing returns. However, our simulations suggest that for our parameter values, there is a unique fixed point. Kucheryavyy, Lyn, and Rodríguez-Clare (2016) find that a sufficient condition for uniqueness is that the trade elasticity multiplied by the scale elasticity ($-\varsigma\theta$ in our notation) should be less than 1. Our estimates $\theta = 7.7$ and $\varsigma = -0.035$ imply $-\varsigma\theta = 0.27 \ll 1$, suggesting a unique equilibrium. As we show in online Appendix D, $-\varsigma\theta < 1$ is a sufficient solution to avoid explosive outcomes, which is necessary for existence and uniqueness of an internal equilibrium. To isolate the impact of increasing returns and quantify the importance of interdependencies, we also conduct a non-IRS version of the counterfactuals. That setting treats q_ℓ as another proxy for local assembly costs (w_ℓ^α), along with GDP per capita and the exchange rate over-valuation index. Thus it is held constant at the observed level even as the policy variables are changed. We focus on the IRS results but comment on some interesting differences relative to the non-IRS case.

The second methodological aspects of our counterfactuals to be detailed regards the treatment of segments. We approach the issue of market segments with two boundary assumptions in order to ensure that how we handle demand substitution patterns does not exaggerate the response to policy changes. The unified car market assumption makes every car model a symmetric substitute for every other model. In contrast, the segmented market assumption shuts down between-segment substitution by fixing the Q_{ns} at the 2016 levels. To see how this could matter in counterfactuals, consider the response of Smart production in France to the ending of NAFTA preferences. The C_{bn} for Smart's rivals who produce in North America will rise whereas Smart's C_{bn} will be unchanged. Therefore Smart's sales are expected to rise. The difference is that, under segmented markets, Smart achieves a higher market share in small cars, a relatively unimportant segment in the United States, whereas in a unified market Smart also gains at the expense of North American SUV production. This will have aggregate effects on French car production since the brands that produce in France and also have a distribution presence in the United States all make small cars.

A third important aspect of the counterfactuals is the method of solving for changes relative to the factual policies. The method we report in the main text solves the full model under the current set of frictions, computing expected values of sales, sourcing shares, model-level and brand-level entry decisions. The same set of calculations is repeated under the counterfactual set of frictions to obtain a new set of expected values. We refer to the comparison of predictions under current and alternative sets of frictions as Difference in Expected Values (DEV). DEV requires estimates of 662 parameters.³⁸ An alternative approach, described in detail in online

³⁸There are 10 elasticities, 23 structural friction parameters, 47 production country FEs ($FE_\ell^{(1)}$), three times the 74 market FEs ($FE_n^{(2)}$, $FE_n^{(3)}$, and $FE_n^{(4)}$) and three times the 120 brand FEs ($FE_b^{(2)}$, $FE_b^{(3)}$, and $FE_b^{(4)}$). DEV is feasible here because all the parameters in our model are identified.

Appendix C, is called Exact Hat Algebra (EHA). The EHA approach requires just 10 parameters ($\eta, \theta, \sigma^e, \sigma^d, \alpha, \varsigma$, and the 4 friction parameters for deep RTAs). EHA allows observed levels of the endogenous variables to “stand in” for parameter estimates as well as unobservables. Thus, by definition, EHA replicates the actual data, whereas DEV sometimes errs by large amounts in predicting the factual levels of production in each country (q_ℓ).³⁹ EHA has two important disadvantages. The first is that brand entry cannot be handled by this method because it is a binary variable. The second concern in using EHA is that it does not allow a brand to start sourcing from an assembly country that was not used prior to the shock. Any zero remains zero, no matter how large the change in frictions. Online Appendix I provides a full set of results for this alternative approach to counterfactual simulations, together with a section discussing when and why EHA and DEV results differ.

B. The Solution Algorithm

Our DEV approach solves the model in levels twice: once at the factual level of frictions, and once for the same frictions evaluated under the counterfactual scenario. The equations summarizing the equilibrium are the sourcing decision (6), the market share (9), the price index (10), and the two entry equations (15) and (19). The identification of structural parameters needed for those equations is detailed in Section IIIC and in online Appendix B. Solving the model involves nested fixed point iterations with an inner, a middle, and an outer loop.

- (i) The inner loop solves a system of two non-linear equations obtained from the price index (10) and the model entry probability (15). It takes as given the vectors of brand entry (\mathcal{D}_{bn}) and the multinational cost index, C_{bn} , which is determined by the set of frictions and national production, q_ℓ . Fixed point iteration yields equilibrium values of P_n and M_{bn} .
- (ii) The middle loop takes these two variables and feeds them into expected market shares, equation (9). Combined with sourcing probabilities, $\Pr(\mathcal{S}_{b\ell n} = 1)$, from equation (6) the vector of equilibrium flows is given by

$$(26) \quad E[q_{b\ell n}] = E\left[\frac{q_{bn}}{Q_n} \mid \mathcal{D}_{bn} = 1, M_{bn}\right] \times \Pr(\mathcal{S}_{b\ell n} = 1) \times Q_n.$$

Next, we sum over $E[q_{b\ell n}]$ for all b and n to obtain the expected value of q_ℓ , then used to update C_{bn} and $\Pr(\mathcal{S}_{b\ell n} = 1)$. The inner loop (step (i)) is re-run with these new inputs to output a new vector of quantities. The process iterates until the vector of q_ℓ stops changing.

- (iii) The brand entry vector is then updated in an outer loop using the rule that entry occurs when expected profits exceed fixed entry costs, $E[\pi_{bn}] > F_{bn}^d$,

³⁹EHA was developed by Dekle, Eaton, and Kortum (2007) and Arkolakis, Costinot, and Rodríguez-Clare (2012). Eaton, Kortum, and Sotelo (2013) is an example of a paper using DEV even though two of the authors helped to originate the EHA approach. The reason was that the later paper abandoned the continuum assumption to consider granular data of the type that is important here as well.

with expected profits and fixed costs calculations being detailed in online Appendix B. Since the inner and middle loops depend upon this vector of entry, the algorithm iterates over the three loops until the set of profitable brand-market combinations stabilizes.

Handling the decisions of brands to enter or not in counterfactuals is far from straightforward. The theory specifies the distribution of brand entry fixed costs as log-normal. If we drew fixed costs from the complete distribution, there would be a large number of instances of brands that are in fact available in a country but would be absent even in the *factual* version of the solution of the model. Conversely, there would also be many false entrants in the simulation. If a major brand were falsely absent or present it would severely endanger the realism of the counterfactual results. We therefore follow König et al. (2017) in drawing from a distribution that is truncated such that fixed costs of factual entrants are *lower* than our predicted value of their gross profits. Brands that are absent from a particular market (Renault in the United States for example), take their fixed cost draws from the portion of the distribution where fixed costs *exceed* gross profits. The counterfactual policies maintain the exact draw of fixed costs for each brand-destination but recomputes expected gross profits. This allows for some brands to be drawn into or out of its factual entry decision. To obtain expectations we replicate the solution of the model with 1,000 draws of fixed costs.⁴⁰

The discrete nature of brand entry has to be taken into account when computing the equilibrium. In contrast to entry at the model level, where the only relevant object is the share of models the brand decides to offer, the identity of which brand enters matters for the outcomes of the counterfactual. This is because each brand has its own network of potential assembly locations ($\mathcal{L}_{b\ell}$) and its own mass of models (M_b) with associated productivity φ_b . This means we have to keep track of which particular brands have entered or exited as a result of a policy change. The algorithm iterates until a fixed point in the brand-entry vector is reached. At each iteration, the computation of expected profits takes the price index for each market (calculated in the inner loop where brand presences are held fixed) as a given. In practice, entrants do affect the price index. In the small car markets (e.g., Bulgaria, Ukraine) this decline in the price index can be large enough that it induces exit in the following iteration. This leads to a rise in the price index which can attract firms back into the market. This process repeats itself in an oscillatory pattern, at which point we terminate the iteration. Note that this oscillation only occurs because brand-level entry has to be considered as an integer problem. The expected fraction of models offered in a market ($E[M_{bn}/M_b]$) is a continuous variable, so there is no integer issue in the inner loop.

The segmented market version of DEV has a few important differences. In step (i) (the inner loop), price indices, the mass of models offered, and all three marketing costs (δ , δ^e , and δ^d) need to be defined at the segment level (the sourcing

⁴⁰ A subtle aspect of this approach is that it does not actually guarantee that entry choices in the simulation of factual policies match the real decisions 100 percent of the time. This is because the truncation is based on the gross profits calculated using price indexes extracted from the fixed effects in the market share regression. The P_n obtained in the solved model will differ, occasionally by enough to alter the entry decision of individual brands that were near the entry/exit threshold.

probabilities, as well as τ and γ frictions do not have a segment dimension). The inner loop therefore solves for equilibrium P_{ns} and M_{bns} . The middle loop (run as a second step) updates national output by summing up the segment-level sales of the brand that are expected to be sourced from different origin countries:

$$(27) \quad E[q_\ell] = \sum_b \sum_n E[q_{b\ell n}]$$

$$= \sum_b \sum_n \Pr(\mathcal{S}_{b\ell n} = 1) \sum_s E\left[\frac{q_{bns}}{Q_{ns}} \mid \mathcal{D}_{bn} = 1, M_{bns}\right] \times Q_{ns}.$$

Lastly, the decision of the brand to enter a market depends on the sum of the profits to be earned in each segment where the brand has models. Therefore, the outer loop updates the vector of brand entry decisions which are transformed as $E[\pi_{bn}] = \sum_s E[\pi_{bns}] > F_{bn}^d$.

Before turning to solutions of counterfactual policies, it is important to demonstrate that the endogenous variables solved for under factual trade policies do not depart too much from their data counterparts. Starting with the decision where brands are offered, the entry rate in fact is 36.2 percent, slightly lower than the 36.6 percent average in the simulation with a maximal difference of 0.82 percent. Online Appendix Figures G.1 and G.2 show the fit of one run of the DEV simulation (for unified and segmented markets respectively) under the factual set of policies. The correlations between simulated and actual are high for all four variables examined. In the unified markets with IRS case, we obtain the following correlations: 0.98 for the price index, 0.63 for brand-origin-destination sales, 0.74 for origin-destination flows, and 0.86 for aggregate origin-level output. The correlations are even higher for the non-IRS and segmented versions of the model.

VI. Counterfactual Results

The main motivation for estimating the model of this paper is the investigation of counterfactual trade policy changes. We report results on eight different scenarios: two different versions for each of four types of policy experiments.

(i) Trumpit/Section 232:

- The United States imposes a 25 percent tariff on cars and parts from Canada and Mexico and ends NAFTA (a deep RTA). Canada and Mexico retaliate with equal tariffs.
- The United States imposes the same 25 percent tariffs on all origins *except* Canada and Mexico. The targeted countries reciprocate.

(ii) United Kingdom exit from the European Union:

- Soft Brexit: a free trade agreement retains tariff-free trade between the United Kingdom and the EU27 but rescinds all deeper integration measures.
- Hard Brexit: in addition to rescinding the deep integration measures, the EU27 and United Kingdom impose the current EU MFN tariffs.

- (iii) Transpacific integration with and without the United States:
 - TPP: a deep integration agreement between Australia, Brunei, Canada, Chile, Japan, Malaysia, Mexico, New Zealand, Peru, Singapore, the United States, and Vietnam.
 - Comprehensive and Progressive Trans-Pacific Partnership (CPTPP): a revised TPP agreement omitting the US (in force since December 2018).
- (iv) Transatlantic integration:
 - Comprehensive Economic Trade Agreement (CETA): deep integration between the EU28 and Canada (provisionally applied since September 2017).
 - CETA + Transatlantic Trade and Investment Partnership (TTIP): EU28 deep agreements with Canada *and* the United States.

Figures 6, 7, 8, and 9 display predictions of the model for each of the four policy experiments. Each figure contains two panels showing changes in output on the left and percent changes in consumer surplus on the right. For each of the ten most affected countries, we plot outcomes for unified (circles) and segmented markets (squares) versions of the model. We contrast the two policy variants by showing one in red and the other in blue. The counterfactual underlying these figures solves the model to calculate the difference in expected values (DEV). We also present results using the EHA method for counterfactuals in online Appendix I. All numbers used for graphical displays of both DEV and EHA versions of counterfactual scenarios can be found from the detailed tables in online Appendix J.

A. Trumpit/Section 232

Renegotiation of the North American Free Trade Agreement began shortly after the inauguration of Donald Trump. In the final stages of the negotiation, the US president threatened to leave NAFTA and impose 25 percent tariffs in the auto sector under Section 232 national security provisions. We refer to this combined threat as “Trumpit.”

Figure 6 depicts in blue the imposition of 25 percent tariffs on both assembled cars and parts from Canada and Mexico. The latter are assumed to retaliate with identical tariffs on the United States (while maintaining full NAFTA preferences with each other). Our simulations point to disastrous outcomes for Canada and Mexico, whose combined losses sum to 1.2–1.4 million cars.⁴¹ These losses would be equivalent to shutting two median-sized plants in Canada (producing 243,000 cars per year) and five of the smaller (150,000) plants operating in Mexico. Production also declines in the United States but the loss of 42,000–223,000 cars represents less than 1 percent of the initial level. Canada stands to lose 67–68 percent of its car industry, while Mexico loses 38–41 percent of its production.⁴²

⁴¹ As the true amount of substitution between segments lies between the segmented and unified polar cases, we report ranges of outcomes except when they round to the same value.

⁴² The percentage changes in production are all based on the initial levels predicted by the model.

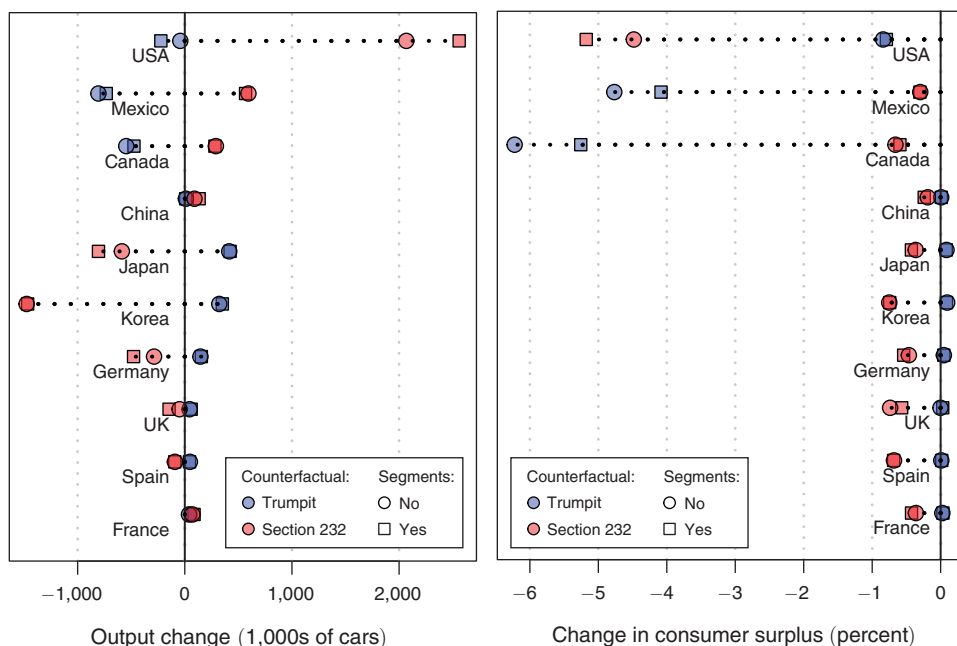


FIGURE 6. TRUMPIT/SECTION 232

Why do Canada and Mexico fare so badly under Trumpit? The first factor is that the shock is very large, including both the final and upstream tariffs and also the sizable tariff equivalents of lost deep integration. In total, τ frictions rise to 28.2 percent while γ frictions rise to 31.4 percent. The shock to τ applies to large shares of Canadian and Mexican production. In 2016 the United States purchased 83 percent of Canadian-made cars and 55 percent of Mexican cars. In contrast, those countries' combined purchases amounted to only 8.4 percent of the cars assembled in the United States. The shock to the γ friction also applies broadly in Canada and Mexico. Our estimates imply that parts from the United States account for 37 percent of the assembly costs of the US brands who account for 55 percent of Canadian and 36 percent of Mexican production in the 2016 data. The final key fact explaining our simulation predictions is that almost all the brands made in Canada (11 of 12) and Mexico (10 of 14) are also made in the United States. The large estimated sourcing elasticity ($\theta = 7.7$) implies these brands will heavily shift production for the US market to US assembly sites when Trumpit raises the production and delivery costs of using sites in Canada and Mexico.

The asymmetric damage to the factories of US brands in Canada and Mexico from higher γ frictions is absent in a pure trade model. The Canadian and Mexican governments would likely realize that putting tariffs on imported parts from the United States is a defective way to retaliate. To inflict the greatest damage on US production with the lowest losses to Canadian production, it would make more sense to limit retaliation to final goods. To investigate what part of the losses seen in the Trumpit scenario come from avoidable cost increases, we ran a version including this modification. Dropping the retaliatory tariffs on car parts does indeed shrink the

losses to Mexican and Canadian production by 104,000 and 76,000 cars, respectively in the segmented case.⁴³

The blue symbols in the right panel of Figure 6 show that car buyers in all three countries lose from Trumpit. Losses in the United States are minor (about 1 percent) but Canada loses 5–6 percent of consumer surplus while Mexico loses 4–5 percent. Consumer losses in both countries can be partially mitigated by “smart retaliation”: exempting parts lowers the losses by about 0.4 percent for both countries.

The Trumpit counterfactual hurts all three participants, both in terms of reduced production (cumulating to 1.4 million fewer cars made in North America) and lower consumer surplus. The disproportionate losses by Canada and Mexico help to explain why they were willing to agree to NAFTA revisions that were widely perceived as unfavorable. While the members of NAFTA lose from Trumpit, the seven other countries shown in Figure 6 gain both in terms of output and consumer surplus. This is specially true for Japan, Korea, and Germany, who collectively increase production by around 900,000 cars. Since our estimates rank those three countries among the four lowest cost assemblers, the reallocation of sourcing decisions for the US market heavily favors them.

The changes in production and consumer surplus displayed in Figure 6 take into account adjustments in sourcing, market shares, model entry, and brand entry. The last of these is of particular interest because it is a binary outcome that conditions the three subsequent decisions in any given replication. The simulations calculate the probability of opening a distribution network over 1,000 repetitions with and without the policy intervention. We find that Trumpit does not change entry probabilities for the major brands. However, US brands that serve few foreign markets substantially lower their likelihoods of entering Canada and Mexico. The most extreme case is Buick, which enters Mexico in every replication without Trumpit but only in 54 percent of the Trumpit replications. In contrast, a number of second-tier EU brands increase their entry propensity in both markets. For example, Citroën’s chance of entering Mexico rises from 11 percent to 27 percent due to Trumpit. Skoda, Dacia, and Opel more than triple their brand entry probabilities in Mexico. In practice, we find that brand entry issues do not have much quantitative impact because the brands that are on the margin of entering and staying account for very little output and also have limited sourcing alternatives.

The NAFTA revision (agreed upon in October 2018) exempted shipments originating in Canada and Mexico of up to 2.6 million cars each from any subsequent Section 232 tariffs. At the time of writing, this threat remains in effect for other countries, which motivates our second variant (labeled Section 232). The red symbols in Figure 6 apply to this drastic scenario, in which the United States imposes 25 percent tariffs on cars and parts imported from all non-North American partners, who retaliate with equivalent tariffs.⁴⁴

US imposition of Section 232 tariffs on non-North American car imports would cut production in Japan and Korea by large amounts, with German production also sharply reduced. Their combined losses sum to 2.4 million–2.8 million cars. The beneficiaries

⁴³ With unified markets, Mexican and Canadian production losses contract by 119,000 and 87,000 cars.

⁴⁴ A few countries (India, Iran, and Vietnam, for instance) have tariffs larger than 25 percent in 2016. The simulation leaves those tariffs unchanged.

would be North American plants, with those in the United States experiencing the largest increase in production (up to 2.6 million cars added in the segmented case). The large rises in predicted exports of Mexico and Canada to the United States, 416,000–431,000 and 228,000–230,000 respectively, may explain the US insistence on limiting the number of car imports to be exempted from Section 232. Consumers lose surplus in all countries, with the price index faced by US buyers rising by 4.5–5.2 percent. The losses for Mexican and Canadian consumers, who do not impose any tariffs in this scenario, stem mainly from rising production costs at the factories of Asian and EU brands in the United States, where car parts rise in cost by 25 percent.

Increasing returns to scale play an important role in the counterfactual predictions displayed in Figure 6. The estimated magnitude of scale economies is modest: a doubling of national output reduces costs by just 2 percent ($2^{-0.035} = 0.976$). However, the large tariff response elasticities lead to substantial *amplification* of production changes. To quantify the IRS-induced magnification effects, we re-solve the counterfactual holding domestic production (q_{lt}) inside the cost function constant at the factual levels. This is equivalent to treating country scale as an omitted cost term that does not vary in response to policy changes. IRS generates much larger predicted output losses for Mexico and Canada in the Trumpit scenario (8 and 11 percentage points larger, respectively, for unified markets). The amplification effects turn out to be strikingly systematic. Regressing the Trumpit country-level change in log output in the IRS case against the change in log output for the non-IRS case, we obtain a coefficient of 1.36 (for segments and unified markets). The fit (R^2) of this simple linear regression is 0.99. As explained in online Appendix D, a simplified version of our model predicts the amplification coefficient to be $1/(1 + \zeta\theta) = 1.37$. Similar values of the amplification effect show up in all our counterfactuals.

The presence of increasing returns exerts a second, more subtle, effect on the conduct of the counterfactual responses: it generates interdependencies across markets. China's production changes in response to Section 232 tariffs provide a useful illustration. China's increase in shipments to the domestic market under IRS (122,000–168,000 cars) is more than double its increase in the constant returns simulation. This occurs in part because nearby suppliers Japan and Korea experience rising production costs under IRS because of their diminished scale (shown in red in Figure 6). Why do China's domestic shipments rise (by 56,000–74,000 cars) even in the absence of IRS? Chinese car tariffs do not change because our experiment imposes 25 percent retaliatory duties across all countries and China's tariffs were already at 25 percent. The production increase in the home market stems from a second interdependency caused by γ effects. For example, BMW exports SUVs from the United States to China. These have substantial German content (37 percent according to our estimates) which implies that higher Section 232 tariffs on car parts raise the cost of German cars exported to China. Thus, even without higher Chinese tariffs on such cars, non-US brands assembled in the United States become less competitive in China compared to locally assembled cars. All in all, such brands export 45 percent fewer cars to China under the Section 232 tariffs (without IRS effects). In sum, the two interdependencies, one from IRS and the other from HQ-sourced parts, lead to the interesting prediction that US Section 232 tariffs imposed on non-North American producers would actually expand the number of cars assembled in China.

Figure 6 shows that, contrary to our initial expectations, the segment version of the model does not systematically dampen the impact of changes in trade policies. This is because switching from unified markets to segments has three effects. First, it limits substitution within segments, which acts as a dampener. Working in the opposite direction, the demand elasticities are larger for segments than for unified car markets. This implies a greater market share response to policy changes. Another factor that increases responses is that segmenting tends to reduce the number of competing brands. The net effect is larger absolute impacts in the segmented version for the United States and Japan under Trumpit and Section 232 but slightly smaller effects for Canada and Mexico. As the segmented and unified magnitudes tend to be similar, we comment only on the segmented results in the discussions of the rest of the policy experiments.

B. *Brexit*

Since the 2016 Leave vote, debate revolved around whether Brexit will be “soft” or “hard.” The soft Brexit case captures the scenario in which Britain retains tariff-free access to the European Union but loses the deep integration aspects of the RTA such as free mobility of professionals and the ability to influence EU regulations on car standards. We also simulate a hard Brexit scenario where UK exports face the European Union’s 10 percent MFN tariffs while the United Kingdom reciprocates at the same rates. In both scenarios the United Kingdom cannot “roll over” existing EU trade agreements (with South Korea, Mexico, and Turkey notably), so trade with all those countries reverts to MFN tariffs. The two post-Brexit scenarios set all the deep RTA dummies to 0 if they correspond to (i) dyads involving the United Kingdom and European Union, (ii) a dyad involving the United Kingdom and one of the countries having a preferential deep RTA with the European Union.

Figure 7 points to poor outcomes for both the UK car industry and the British car buyer. Workers in UK plants lose slightly more (83,000 cars) under Soft Brexit than with Hard Brexit (73,000 cars). To put those numbers in perspective, they amount to losses of between two-thirds and three-fourths of Honda’s Swindon factory, which in 2016 assembled 110,000 cars employing around 3,100 workers.⁴⁵ The relatively small net production losses in the United Kingdom mask large changes in the “gross flows.” Under Hard Brexit, UK factories increase production for the local market by 290,000 cars, while their sales on the continent fall by 259,000. Sales outside the European Union are predicted to fall by 103,000 cars because UK-made cars lose their tariff-free status in Turkey, Mexico, and other countries where the European Union signed FTAs.

The protective effect of 10 percent tariffs more than offsets the lost exports to the Continent under hard Brexit. This prediction contrasts with the Trumpit prediction that Canada loses sales in both the United States and at home. Two factors underlie this difference. First, US brands make the majority of cars in Canada and their plants experience production cost increases (γ effects) under Trumpit. Falling US-brand production even harms the Japanese makers, through Canada’s loss of external scale

⁴⁵ *Financial Times*, March 13, 2015. Swindon is an interesting plant because it is about average in size and accounts for the majority of Honda sales in the European Union.

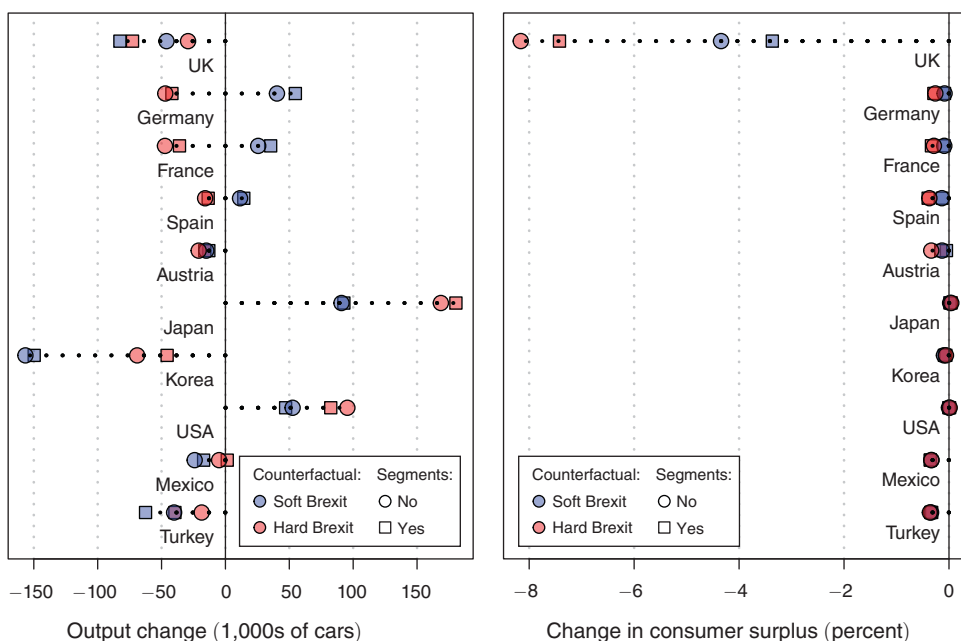


FIGURE 7. BREXIT

economies. Meanwhile in Britain, domestic brands accounted for 45 percent of production, whereas EU brands account for just 8 percent of production (2016 data). A second factor is that the EU27 runs a substantial trade surplus with the United Kingdom in cars. This means that as tariffs in both directions rise to 10 percent, the United Kingdom has less to lose.

While these simulation results do not predict disastrous consequences to auto workers from a “no deal” Brexit, the cost to consumers is much more severe. Consumers lose 3.4 percent of their surplus with soft Brexit but their losses more than double to 7.4 percent when 10 percent tariffs are imposed on EU-made cars. The tariffs imposed on UK cars exported to the European Union have very little effect on consumer surplus there. One problem the United Kingdom faces is the fact that major multinationals like Ford and Nissan have factories in Spain they can easily switch production to in order to continue to serve the EU27 tariff-free. On the other hand, none of the major EU brands operate factories in the United Kingdom, implying that British buyers will have to pay more for their cars.

C. Transatlantic Integration

While recent reversals on trade integration have dominated attention, some new deep integration agreements have recently been implemented. The Comprehensive Economic and Trade Agreement (CETA) between Canada and the EU28 still awaits ratification by different European national parliaments, but it has been applied provisionally since September 21, 2017. By contrast, negotiations on the Transatlantic Trade and Investment Partnership (TTIP), an integration agreement between the

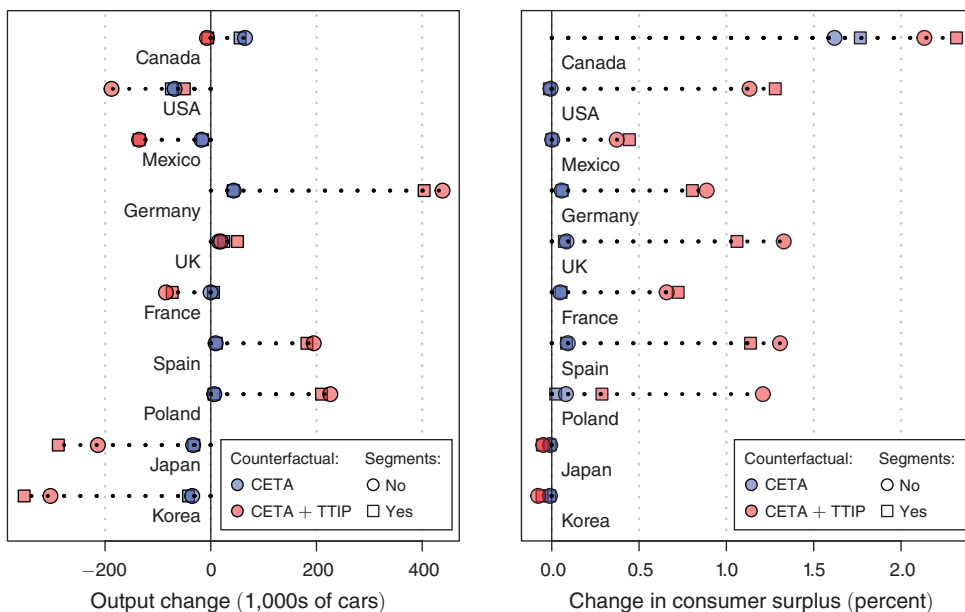


FIGURE 8. CETA (EU-CANADA DEEP RTA)

European Union and the United States, were put to an indefinite halt following the 2016 US elections. We therefore consider the first scenario to be the full implementation of CETA: the abolition of tariffs between the European Union and Canada, combined with deep integration. On top of those, the second scenario applies the same policy changes to the EU28-USA country pairs.

Figure 8 shows that most gains in production under CETA accrue to Canada (56,000 cars, 8 percent production increase), followed by Germany (43,000 cars) and Britain (25,000 cars). The countries losing sales are the United States, Korea, Japan, and Mexico (in that order). Korea and Mexico face erosion of their existing preferences granted by the European Union and Canada as part of preexisting FTAs. The only country where consumers have notable gains is Canada, where surplus rises by 1.8 percent. Canadian production represents too small a share of sales in Europe even after CETA to make much difference for EU customers.

Adding the United States to the transatlantic liberalization of car trade would lead to potentially large output gains for Germany (7 percent), Spain (16 percent), and Poland (74 percent).⁴⁶ These countries increase production by 795,000 cars. Interestingly, their losses at home from lowering tariffs on US-made cars are negligible: Germany loses sales of just 31,000 cars at home and Poland and Spain actually *increase* production for their home market. These gains come from a mix of increasing returns and γ effects (US-owned factories are important in both countries). Meanwhile production falls in the United States as they cede more sales in their home market than they gain in Europe. This asymmetry is initially surprising

⁴⁶Online Appendix H provides evidence that such large production increases are feasible within the medium run contemplated in our policy counterfactuals.

since it is the European Union that lowers its tariffs more under transatlantic free trade (10 percent versus 2.5 percent MFN tariffs). The United States losses can be understood in terms of the tendency of trade liberalization to reallocate output to the low-cost producers. In the prediction of our model, and in the data, the European Union collectively has a cost advantage, exporting about six times more to the United States than what it imports.⁴⁷

Consumers gain widely, but modestly, from transatlantic integration. Gains in Germany, the United States, Spain and Britain all round to 1 percent. Canadian buyers actually gain more (2.3 percent) *with the United States* in the agreement. This is because the German brands VW, Mercedes, and BMW all have manufacturing plants in the United States. Deep integration implies a 6 percent reduction in the production costs of those plants (γ effects), which are also cheaper to distribute in Canada as a result of the agreement (δ effects).

D. Transpacific Integration

Figure 9 displays the predicted impact of the Trans-Pacific Partnership, showing outcomes for the original agreement involving the United States and its successor, the CPTPP, which came into force December 30, 2018. The TPP is mainly of interest as a “might-have-been” policy since the United States exited the agreement in January 2017. Despite claims that TPP was more of a regulatory agreement than a trade agreement, the TPP and CPTPP include substantial tariff cuts for some country pairs. In 2016 Japanese exporters faced a 44 percent tariff when exporting to Vietnam, and 6–7 percent to Canada and New Zealand. US exports faced a 55 percent duty in Vietnam and 23 percent in Malaysia.⁴⁸

Transpacific integration shows how policy impacts on the location of production and the fortunes of individual brands are intertwined. In terms of production, the most affected country is Canada. It is predicted to increase output by 33 percent under TPP and by 42 percent under CPTPP. In terms of brand nationalities, Japan is the primary beneficiary of both TPP and CPTPP. Both policies increase the share of the Canadian market served by Japanese brands by about 10 percentage points. Under TPP, Japanese brand market shares rise by 6 percentage points in the United States and they even rise a small amount (1 percentage point) under CPTPP.

Several forces at work underlie these predicted reallocations in production and market shares. They represent aspects of our framework that do not feature in traditional models. Japanese plants, which account for 45 percent of Canadian production in 2016, are predicted to obtain a 6.4 percent cost reduction as a result of the deep γ_{il} aspect of the agreement. Elimination of the 2.7 percent Canadian tariff on auto parts (an average across the main auto parts HS codes) will further reduce costs at Toyota and Honda’s Ontario plants. Under TPP, similar gains would also accrue to the Japanese plants in the United States, although US parts tariffs are only 1.4 percent on average.

⁴⁷ The model predicts an export/import ratio of 6.2, very close to the 5.6 ratio in the 2016 data.

⁴⁸ Japan signed a free trade agreement with Malaysia that entered into force in July 2006, and stipulated a 9-year phase-out period for cars.

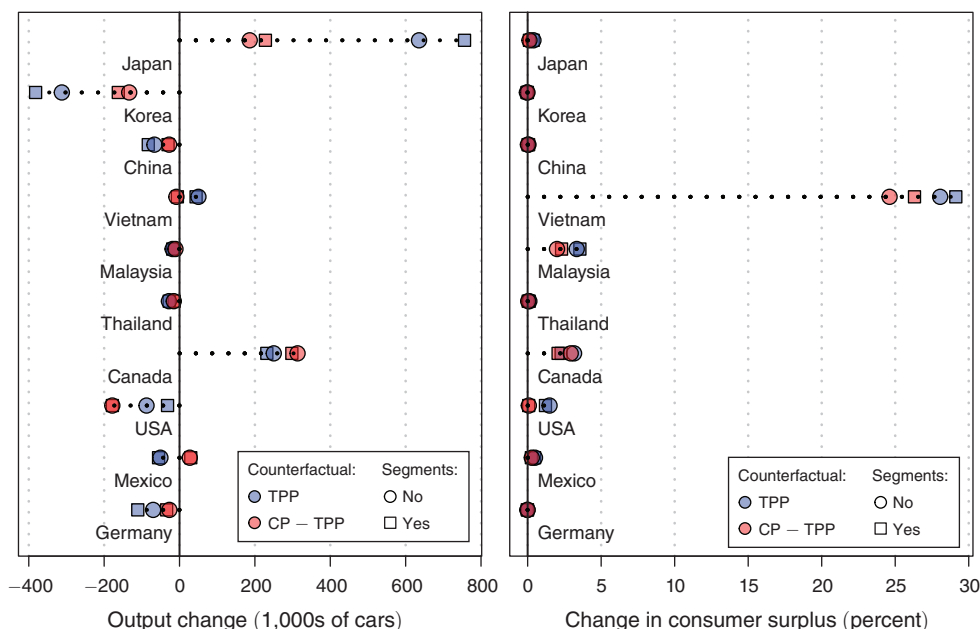


FIGURE 9. TRANS-PACIFIC PARTNERSHIP

Even though the γ effects are mainly balanced, under TPP the reduction in marketing costs from Japan leads to increased market shares for Japanese brands in the United States with much of the additional demand sourced from Canada. Under CPTPP, the situation is different. Although Japanese brand shares do not rise much in the United States, they increase the fraction of models sourced from their Canadian plants. Meanwhile, because Japan-origin cars avoid Canada's 6 percent car tariff, Japan's sourcing share rises in the Canadian market. The total number of Japanese cars sold in Canada rises enough to offset this negative effect.

Lower costs of marketing Japanese cars are predicted to stimulate entry of nearly 20 percent more Japanese car varieties in Canada under the CPTPP, and would stimulate a similar increase in the United States had the TPP been implemented. Since a significant fraction of these new models will be sourced from the Toyota and Honda plants in Canada, model entry stimulates production gains. There are also Japanese brands that would have been unlikely to enter Canada but are more likely to do so under CPTPP: Acura, Scion, and Infiniti increase their entry probability in Canada by a substantial range (19 percent, 18 percent, 8 percent).

The γ gains our model predicts for Canadian production are not present in the conventional trade model used by Global Affairs Canada to predict the consequences of CPTPP.⁴⁹ This study finds close to negligible effects of CPTPP on Canadian output. Another study conducted in 2012 by Van Biesebroeck, Gao, and Verboven

⁴⁹ <http://international.gc.ca/trade-commerce/trade-agreements-accords-commerciaux/agr-acc/cptpp-ptpgp/impact-repercussions.aspx?lang=eng>.

(2012)⁵⁰ looked at the impact of several RTAs for automobile production in Canada. While their use of detailed data on car characteristics allowed the authors to use the whole apparatus of BLP-type demand estimation, they do not account for the γ -type frictions we include here. This is critical in terms of predicted outcomes. When we run our CPTPP counterfactual without any γ liberalization, we find a very small overall impact on Canadian output, confirming the results of the two mentioned studies.

The other country strongly affected by TPP and CPTPP is Vietnam. Vietnam increases production with the TPP scenario. This is because of improved access to the US market and more efficient operations at Ford and Chevrolet plants which are expected to stimulate a big increase in exports to the United States. However, the Japanese brands currently assembling in Vietnam will radically increase their sourcing from Japan. If the United States is not in the agreement, the net effect on production becomes negative for Vietnam. The brand entry margin is also very active in Vietnam under CPTPP. Because of the increased competitiveness of Japanese brands, many competitors are predicted to exit. Land Rover exits Vietnam in 46 percent of the replications, and Chevrolet also faces a large exit probability at 32 percent. Mercedes-Benz, BMW, and Hyundai have a probability of exiting the country between 14 and 19 percent. By contrast, Subaru and Lexus (currently not sold in Vietnam) enter in 17 percent and 11 percent of the replications. The bright side of CPTPP for Vietnam is that the reduction of the 55 percent tariffs on assembled cars leads to an over 26 percent increase in the surplus of Vietnamese car buyers.

VII. Conclusion

We deploy extremely detailed data from the car industry to estimate the structural parameters of an extended version of the double-CES MP model. We use this framework to predict the medium-run consequences of numerous policy proposals circulating in the post-2016 context (after the victories of Leave in the Brexit referendum, and of Donald Trump in the US elections). Our counterfactuals simulate adjustments in firm-level decisions of (i) which markets to enter, (ii) the fraction of varieties to offer, (iii) the quantity of cars to supply in each of those markets, and (iv) which location to source from.

Several insights emerge from our eight counterfactual exercises that have broad applicability. First, the stakes for consumers and producers in the outcomes of trade policy are often sizable. Canada could lose two-thirds of its auto sector in one scenario (Trumpit). The Vietnamese price index of cars falls by one-quarter in another (TPP). A common factor underlying the large effects predicted here are the magnitudes of the two key elasticities: 7.7 for substitution between assembly sites (θ) and 3.87 for substitution between varieties (η). Moreover, the tariff changes proposed in the car industry are far from negligible. Second, the structure of multinational production matters a great deal. The much larger output reductions in Canada when losing duty-free access to the US market as compared to the analogous losses for the United Kingdom upon a Hard Brexit illustrates how the origins and networks

⁵⁰<http://www.international.gc.ca/economist-economiste/analysis-analyse/studies-etudes/studies-etudes-01.aspx?lang=eng>.

of plants for each brand shape counterfactual outcomes. One reason why Canada suffers so much from the end of NAFTA is because 11 of the 12 brands that produce in Canada have plants in the United States to which they can switch production. Third, some striking results, like the boom of Canadian car production under the CPTPP hinge importantly on the γ effects of RTAs. The reduced cost of operating Japanese plants in Canada impacts the American market (lowering output) even though the United States retains the same trade policies. A second important source of interdependencies associated with trade policy changes is external returns to scale. Policies that substantially contract or expand output of important producers have spillover effects on other markets as we saw in the prediction that Chinese car output actually expands in response to Section 232 tariffs.

We view the counterfactual outcomes as the medium-run response because they hold the set of production locations constant. In this time frame, each brand's network of sourcing alternatives strongly shapes the responses to policy changes. The long-run decision of opening and closing production operations in different countries is of course very interesting. Our focus on the medium-run follows from the desire to keep the estimation tractable and the scope of this paper finite. The medium-run is also important in its own right because production networks are strongly persistent. Even over a 17-year period, covering a major disruptive crisis for this industry, 88 percent of OECD countries' car production still takes place within brand-country combinations that existed in 2000.

The large estimated benefits of producing, designing, and selling within a country, within an RTA, or with nearby countries all motivate future research to identify the mechanisms that underlie these frictions. This is particularly the case for the δ effects we have added to the framework. While we specify them as marketing costs, preference-based mechanisms may play important roles. A third topic (along with plant location choices) calling for more research is the decision of where to source the components to be assembled in each car plant. Due to data limitations, we focused on the role of tariffs in raising costs for parts originating from headquarters, but actual sourcing problems are further complicated by rules of origin. These and other aspects of multinational expansion strategies provide a full agenda for future research.

REFERENCES

- Anderson, James E., and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature* 42 (3): 691–751.
- Andrews, Donald W. K., and Moshe Buchinsky. 2000. "A Three-Step Method for Choosing the Number of Bootstrap Repetitions." *Econometrica* 68 (1): 23–51.
- Antràs, Pol, Teresa C. Fort, and Felix Tintelnot. 2017. "The Margins of Global Sourcing: Theory and Evidence from US Firms." *American Economic Review* 107 (9): 2514–64.
- Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review* 102 (1): 94–130.
- Arkolakis, Costas, and Fabian Eckert. 2017. "Combinatorial Discrete Choice." Unpublished.
- Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. 2018. "Innovation and Production in the Global Economy." *American Economic Review* 108 (8): 2128–73.
- Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review* 98 (5): 1998–2031.
- Bartelme, Dominick, Arnaud Costinot, and Andrés Rodríguez-Clare. 2018. "External Economies of Scale and Industrial Policy: A View from Trade." Unpublished.

- Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott.** 2011. "Multiproduct Firms and Trade Liberalization." *Quarterly Journal of Economics* 126 (3): 1271–1318.
- Berry, Steven, James Levinsohn, and Ariel Pakes.** 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 841–90.
- Berry, Steven, James Levinsohn, and Ariel Pakes.** 1999. "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy." *American Economic Review* 89 (3): 400–30.
- Björnerstedt, Jonas, and Frank Verboven.** 2016. "Does Merger Simulation Work? Evidence from the Swedish Analgesics Market." *American Economic Journal: Applied Economics* 8 (3): 125–64.
- Ciliberto, Federico, Charles Murry, and Elie T. Tamer.** 2018. "Market Structure and Competition in Airline Markets." CEPR Discussion Paper 13346.
- Coşar, Kerem A., Paul L. E. Grieco, Shengyu Li, and Felix Tintelnot.** 2018. "What Drives Home Market Advantage?" *Journal of International Economics* 110: 135–50.
- Costinot, Arnaud, and Andrés Rodríguez-Clare.** 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization." In *Handbook of International Economics*, Vol. 4, edited by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, 197–261. Amsterdam: Elsevier.
- de Gortari, Alonso.** 2017. "Disentangling Global Value Chains." Unpublished.
- Dekle, Robert, Jonathan Eaton, and Samuel Kortum.** 2007. "Unbalanced Trade." *American Economic Review* 97 (2): 351–55.
- Eaton, Jonathan, and Samuel Kortum.** 2002. "Technology, Geography, and Trade." *Econometrica* 70 (5): 1741–79.
- Eaton, Jonathon, Samuel Kortum, and Sebastian Sotelo.** 2013. "International Trade: Linking Micro and Macro." In *Advances in Economics and Econometrics Tenth World Congress*, Vol. II: Applied Economics, edited by Daron Acemoglu, Manuel Arellano and Eddie Dekel, 329–70. New York: Cambridge University Press.
- Fajgelbaum, Pablo, Gene M. Grossman, and Elhanan Helpman.** 2011. "Income Distribution, Product Quality, and International Trade." *Journal of Political Economy* 119 (4): 721–65.
- Feenstra, Robert C., and James A. Levinsohn.** 1995. "Estimating Markups and Market Conduct with Multidimensional Product Attributes." *Review of Economic Studies* 62 (1): 19–52.
- Fuss, Melvyn, and Leonard Waverman.** 1990. "The Extent and Sources of Cost and Efficiency Differences between US and Japanese Motor Vehicle Producers." *Journal of the Japanese and International Economics* 4 (3): 219–56.
- Goldberg, Pinelopi Koujianou.** 1995. "Product Differentiation and Oligopoly in International Markets: The Case of the US Automobile Industry." *Econometrica* 63 (4): 891–951.
- Goldberg, Pinelopi Koujianou, and Frank Verboven.** 2001. "The Evolution of Price Dispersion in the European Car Market." *Review of Economic Studies* 68 (4): 811–48.
- Grigolon, Laura, and Frank Verboven.** 2014. "Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation." *Review of Economics and Statistics* 96 (5): 916–35.
- Guimarães, Paulo, Octávio Figueirdo, and Douglas Woodward.** 2003. "A Tractable Approach to the Firm Location Decision Problem." *Review of Economics and Statistics* 85 (1): 201–04.
- Hanemann, W. Michael.** 1984. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52 (3): 541–61.
- Head, Keith, and Thierry Mayer.** 2013. "What Separates Us? Sources of Resistance to Globalization." *Canadian Journal of Economics* 46 (4): 1196–1231.
- Head, Keith, and Thierry Mayer.** 2018. "Poor Substitutes? Counterfactual Methods in IO and Trade Compared." Unpublished.
- Head, Keith, and Thierry Mayer.** 2019. "Brands in Motion: How Frictions Shape Multinational Production: Dataset." *American Economic Review*. <https://doi.org/10.1257/aer.20161345>.
- Irrarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla.** 2013. "The Margins of Multinational Production and the Role of Intrafirm Trade." *Journal of Political Economy* 121 (1): 74–126.
- Klier, Thomas, and Daniel P. McMillen.** 2008. "Evolving Agglomeration in the U.S. Auto Supplier Industry." *Journal of Regional Science* 48 (1): 245–67.
- König, Michael D., Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti.** 2017. "Networks in Conflict: Theory and Evidence from the Great War of Africa." *Econometrica* 85 (4): 1093–1132.
- Kucheryavyy, Konstantin, Gary Lyn, and Andrés Rodríguez-Clare.** 2016. "Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale." NBER Working Paper 22484.
- Melitz, Marc J.** 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.
- Ramondo, Natalia.** 2014. "A Quantitative Approach to Multinational Production." *Journal of International Economics* 93 (1): 108–22.

- Ramondo, Natalia, and Andrés Rodríguez-Clare.** 2013. "Trade, Multinational Production, and the Gains from Openness." *Journal of Political Economy* 121 (2): 273–322.
- Santos Silva, J. M. C., and Silvana Tenreyro.** 2006. "The Log of Gravity." *Review of Economics and Statistics* 88 (4): 641–58.
- Schmitt, Alexander, and Johannes Van Biesebroeck.** 2013. "Proximity Strategies in Outsourcing Relations: The Role of Geographical, Cultural and Relational Proximity in the European Automotive Industry." *Journal of International Business Studies* 44 (5): 475–503.
- Smith, Donald F., and Richard Florida.** 1994. "Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automotive-Related Industries." *Journal of Urban Economics* 36 (1): 23–41.
- Tintelnot, Felix.** 2017. "Global Production with Export Platforms." *Quarterly Journal of Economics* 132 (1): 157–209.
- Van Biesebroeck, Johannes, Hang Gao, and Frank Verboven.** 2012. *Impact of FTAs on Canadian Auto Industry*. Ottawa: Department of Foreign Affairs and International Trade Canada.
- Verboven, Frank.** 1996. "International Price Discrimination in the European Car Market." *RAND Journal of Economics* 27 (2): 240–68.
- Wang, Zi.** 2017. "Headquarters Gravity." Unpublished.