Journal of International Economics 82 (2010) 168–180

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Measuring the benefits of foreign product variety with an accurate variety set ☆

Bruce A. Blonigen [a,*], Anson Soderbery [b]

[a] Department of Economics and NBER, University of Oregon, Eugene, OR 97403-1285, USA
[b] Department of Economics, University of California, 1 Shields Ave., Davis, CA 95616-8578, USA

ARTICLE INFO

Article history:
Received 5 August 2009
Received in revised form 5 August 2010
Accepted 5 August 2010

JEL classification:
F12
F23
L16

Keywords:
Product variety
Trade
Armington
Monopolistic competition
Automobiles

ABSTRACT

Recent studies have used import data to assess the impact of foreign varieties on domestic prices and welfare. We employ a market-based data set on the U.S. automobile market that allows us to define goods varieties at a more precise level, as well as discern location of production and ownership of varieties. Our estimates of price and welfare changes from new varieties in the U.S. automobile sector are twice as large as standard estimates when using our detailed market-based data. We also show that new varieties introduced by foreign-owned affiliates provided an additional 70% welfare gain during our sample.

1. Introduction

The seminal work of Krugman (1979) highlights that the benefits of trade may stem not only from lower prices, but also greater product variety. This has generated considerable subsequent literature focused on the role of product variety in international trade and its welfare effects. A crucial innovation in this literature is Feenstra (1994) which develops a method for adjusting price indexes to account for changes in varieties available to consumers. Using this method one can calculate the impact of changes in product varieties on economy-wide prices, as well as aggregate welfare. Implementing Feenstra's (1994) methodology and using highly-detailed product-level U.S. import data, Broda and Weinstein (2006) estimate that the substantial rise in net new imported varieties (entry of new varieties minus exit of varieties) over the 1972–2001 period suggests a 1.2 percentage point lower effective aggregate import price per year than that estimated by a conventional import price index. This translates into a cumulative U.S. welfare gain from new imported varieties that is equivalent to roughly 2.6% of U.S. GDP.

While understanding the origins and impact of variety on international trade patterns and outcomes has been a recent focus in the literature, empirical analysis of these issues is handicapped by the lack of data on varieties. Empirical studies of the role of product varieties in international trade have exclusively relied on an Armington-type assumption for which each country's exports of a given product code represents one unique variety.[1]

In this paper we explore how severe the measurement bias may be from the previous product variety assumptions by revisiting the effect of net new foreign varieties on prices and welfare for goods where we can very accurately determine varieties: automobiles. Data on automobiles are very carefully documented in well-known publications by make and models, which we use to define unique varieties. We call these data our market data since they rely on market-based categories of automobile varieties. There are obvious examples of how different variety classifications will be in the market data than the standard import data used in previous studies. For instance, whereas the standard import data would classify all makes and models from

☆ This paper has benefited from the comments of the Editor, an anonymous referee, Maggie Chen, David Hummels, Robert Feenstra, Lionel Fontagne, Amit Khandelwahl, Thierry Mayer, Marc Melitz, Rob Nevin, Ankur Patel, Katheryn Russ, Sasha Skiba, Jose de Sousa, Alan Spearot, Deborah Swenson, Farid Toubal, Eric Verhoogen, Jonathan Vogel, Andreas Waldkirch, David Weinstein and seminar participants at Columbia University, an Empirical Investigations in International Trade Conference, George Washington University, the OECD, the Paris School of Economics and the University of California-Davis. Any remaining errors or omissions are our own.
* Corresponding author. Tel.: +1 541 3464680.
E-mail addresses: bruceb@uoregon.edu (B.A. Blonigen), asoderbe@ucdavis.edu (A. Soderbery).

[1] Both Feenstra (1994) and Broda and Weinstein (2006) provide a brief discussion of pitfalls of this restrictive documentation of varieties, but ultimately must settle on measuring varieties by import source due to data limitations.

0022-1996/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jinteco.2010.08.001

Japan as one variety for a given automobile good (say, compact cars), our market data would show numerous varieties, including Honda Civic, Toyota Corolla, Nissan Sentra, Mitsubishi Lancer, etc. A related issue we will show is that the Harmonized System (HS) goods codes do not map very well into accepted market-based definitions of automobile goods. For example, compact cars are classified under numerous HS10-digit codes and often lumped with sports cars.

Standard import data may provide biased estimates of price changes and welfare for other reasons beyond this precision issue with measuring goods and varieties. First, some varieties that are counted as "new" imported varieties may simply be production of exiting varieties that are offshored to foreign production locations and then imported back to the home economy. Such varieties should likely not be considered new.[2] Second, foreign producers may introduce new varieties to an economy through new foreign-affiliate production that is obviously not included in import data. Our data provide locations of production for each variety, as well as ownership of production. Thus, we can accurately measure new imported varieties, as well as estimate the effect of net new varieties not only from imports, but also from foreign-affiliate production.

Our results first show that the market-based data uncover product variety churning (entering and exiting of varieties) that is an order of magnitude larger in the market-based data than what is represented in the standard import data. Thus, the Armington assumption "hides" substantial variety change. Applying the methods of Feenstra (1994) and Broda and Weinstein (2006) to these respective data sets, we find that the estimated impact of net new variety change on price indexes and, hence, consumer welfare is roughly double in magnitude when calculated using our market data versus the import data. Our market data allow us to also examine a number of other issues that one cannot with the import data. First, we show that the introduction of new varieties and their impact on effective consumer prices has been greater from import varieties than domestic varieties over our sample. We also have data on both location and ownership of varieties and therefore can examine the total impact of variety change from foreign sources (both imports and foreign-affiliate activity in the U.S.) We show that the additional introduction of new varieties by foreign affiliates adds gains that are around 70% larger than those calculated from only import sources. In summary, misclassification of goods and varieties by import data, along with the implicit omission of variety change from foreign-affiliate activity, leads to a very large underestimate of net new variety impact for our sample of the U.S. automobile market. Finally, we show that our results are virtually identical whether we make calculations using three-year intervals or annual changes, suggesting that new varieties quickly reach their equilibrium market share when first introduced.

While our exercise applies specifically to automobiles and cannot be easily applied to other products in the absence of similarly detailed market data, we note that automobiles are a significant share of U.S. imports and foreign-affiliate sales in the U.S. Using data from the Center of International Data, we calculate the fraction of automobile imports by value to be 9.4% (by value) of all U.S. imported manufactured products, and 14.7% of all differentiated products,[3] during the period 1990–2005. In addition, BEA statistics show affiliate sales of motor vehicles to be 11.6% (by value) of all manufacturing affiliate sales by firms in the U.S. as of 2002. Thus, the automobile sector comprises a significant share of all differentiated products delivered to the U.S. from "foreign" sources.

The paper proceeds as follows. The next section describes the detailed market-based data we have collected for automobiles, comparing it to the import data with HS10 product classifications for autos that have been used by prior studies. Section 3 briefly reviews the standard methodology developed by Feenstra (1994) and Broda and Weinstein (2006) to account for variety change on price indexes before turning to Section 4 which implements this methodology for our two automobile data sets and directly compares the magnitude of net variety change and welfare. Section 4 also examines further issues, including the effect of foreign-affiliate activity on gains from net variety change and the robustness of results to varying the time interval used to calculate net variety change. Section 5 concludes.

2. Data

The contribution of our paper stems directly from the use of a more accurate data set of automobile varieties. A number of prior industrial organization studies have used automobile data to quantify the effects of trade policy, price discrimination practices, and even technological innovation in this industry using detailed data on specific makes and models of automobiles.[4] Such data for the U.S. market can be gathered from *Ward's Automotive Yearbook* and *Automotive News Market Data Book*, and include information on model characteristics (height, width, horsepower, etc.), as well as units sold and listed retail prices. For our purposes, we gathered data from these sources on units sold, listed retail prices, manufacturer, and locations of production for all passenger vehicles sold in the United States from 1990 through 2005.[5] With these data, we define 'goods" as various categories of automobiles that are commonly used by industry organizations to classify types of vehicles primarily by size; namely "subcompact", "compact", "midsize", etc. We then define "varieties" of these goods as specific makes and models of automobiles. Dodge Neon, Ford Focus, Mazda 3, Toyota Corolla, and Volkswagen Jetta are examples of varieties of "compact" automobiles sold during our sample. We will call this data sample our "market data." Table 1 provides a list of the goods in our market data and counts of varieties for each good.

We contrast this with the type of data used by prior studies to evaluate the effects of new imported varieties on prices and welfare. These studies rely on what we call "import data," which for the U.S. are import data collected by U.S. Customs by 10-digit Harmonized System (HS10) codes. We collect these data from the Center for International Studies at UC-Davis,[6] which provide units of automobiles sold and unit-value prices by country and HS10 combination. Using import data, "goods" are defined as individual HS10 categories and "varieties" are, by an Armington-style assumption, each unique country-level import source. For example, for a given 10-digit HS code, imports from Japan are one unique variety, imports from Germany are another unique variety, etc. Table 2 provides a list of the goods for our import data on passenger automobiles, as well as counts of varieties for each good.

There are a number of important advantages of the market data over the import data. The primary issue is the definition of varieties and the potential for the import data to hide variety change. We know that for many import sources of a good, especially ones that account for significant market share, there are multiple goods that consumers would see as different varieties. It is clear to see this in our market data. For example, there are many Japanese automobile producers that ship their unique model of automobile and which all fall under the same HS10, yet the import data treat these all as one unique variety. This aggregation issue has the potential to hide significant

[2] For example, a Honda Civic produced at time t in Japan and then at $t+1$ in Japan and Mexico will not be considered a gain of a new variety using market level data, but very well may be considered a variety gain using import data.

[3] We use the classification proposed by Rauch (1999) to define differentiated products.

[4] Representative studies in this literature include Goldberg (1995), Berry et al. (1995), Verboven (1996), Berry et al. (1999), Petrin (2002), and Goldberg and Verboven (2005).

[5] Following previous studies using automobile data, our retail prices are for base packages of a model.

[6] Data can be found at, http://cid.econ.ucdavis.edu/.

Table 1
Varieties in market data.

Good	Total distinct models			Average models per year		
	Foreign	Domestic	Pooled[a]	Foreign	Domestic	Pooled[a]
Cars						
Compact	42	33	58	16	15	26
Compact Executive[b]	10	1	10	5	1	5
Electric[c]	1	0	1	1	0	1
Executive	14	0	14	6	0	6
Fullsize	24	24	42	11	12	22
Grand Tourer[b]	7	2	9	4	1	4
Luxury	4	0	4	3	0	3
Midsize	48	38	74	20	17	35
Sport	33	12	43	14	5	18
Sport compact	11	8	18	5	4	9
Subcompact	33	9	36	12	5	14
Supercar[c]	1	0	1	1	0	1
SUVs						
Crossover SUV[b]	19	7	26	10	5	14
Station wagon[c]	0	1	1	0	1	1
SUT[c]	0	4	4	0	3	3
SUV	38	45	74	16	21	35
Trucks						
Medium duty[c]	0	1	1	0	1	1
Pickup	12	22	26	6	12	14
Vans						
Minivan	14	19	28	6	11	14
Van	8	6	12	4	4	6

[a] Models produced both domestically and abroad are not considered distinct, thus total does not necessarily equal the sum of domestic and foreign models.
[b] Insufficient reference variety, however these classes are extremely comparable to other goods.
[c] Cannot satisfy the requirements of a reference model, and do not easily fit in another class, thus not included in our estimation.

changes in varieties over time as automobile firms discontinue or introduce new models.

Indeed, as shown in Table 3, there are significant differences in observed variety change for automobile goods between the two data sets. Columns 1 and 2 show average annual number of new and exiting varieties per good across the two data sets, while columns 3 and 4 show the average annual share of new and exiting varieties in total sales for a given year. The final row displays the weighted average across the goods categorized. As one can see, new annual varieties per good are 1.9 on average for the import data, while significantly higher at 2.5 for the market data set. The average number of exiting varieties is more similar across the data sets, which means that the market data reveal much greater *net* new variety (1.27 versus 0.26). More importantly for measuring the importance of net new variety is the *share* of these new and exiting varieties in total annual sales. Here, the differences are much more stark. The average share of both new and disappearing varieties is an order of magnitude higher in the market data than the import data (5.9% and 2.9% versus 0.4% and 0.3%), with the average share of net new variety in the market at 3.0%, while only 0.1% for the import data. An important factor in these patterns is that the import data show a number of low-volume imports from countries with no known automobile production.[7]

A related advantage of our market data is the ability to determine the ownership of a variety so that we do not confound "domestic" and "foreign" varieties. The following are two stark examples to show how the import data can confound this distinction and lead to serious measurement biases of variety change from foreign sources. Suppose that all imports for a given good are coming from U.S.-owned firms who decided to outsource the production of exiting varieties (formerly produced domestically) to various foreign countries. The import data would then suggest net new variety introduction through imports when, in fact, U.S. consumers would not be enjoying any net new varieties.[8] The opposite example is where foreign firms introduce all new varieties through foreign-affiliate production in the U.S., not through their imports. In this case, the import data show no net new variety change,[9] even though U.S. consumers are enjoying new foreign varieties through foreign-affiliate production and sales in the U.S. The general message is clear. Import data cannot identify outsourced production which creates a bias toward finding *more* import variety change than really exists. Import data also ignore foreign-affiliate production which creates a bias toward finding *less* foreign variety change than really exists. How these two effects net out is obviously an empirical question.

Our market data can directly address this issue because we have automobile production data by location for each make and model, even when it is produced in more than one location. Thus, we can examine the price and welfare effects of variety changes from "foreign-owned" manufacturers as a question distinct from variety changes from imports only.[10] In doing so, we will also not confound outsourced production of exiting varieties by U.S. manufacturers as new varieties to U.S. consumers. The automobile sector is an excellent example to investigate these issues as the value of foreign-affiliate production of automobiles in the U.S. rivals the value of imported automobiles, as mentioned in the introduction. In addition, U.S. manufacturers have significantly increased their outsourcing of automobile models, especially to North American Free Trade Area countries, over our sample period.

Beyond variety definitions, a likely final advantage of the market data over the import data is the classification of goods. While the HS10 classifications are fairly narrowly defined over various car sizes, this does not always correspond well with typical classifications made by automobile industry sources. Table 4 provides a basic concordance between goods in our import data and goods in our market data. Virtually all the HS10 goods span multiple goods in the market data. For example, HS good 8703230046 may contain "compact" automobiles, as well as "sports" and "sport compact" automobiles.[11] This is of concern for the estimation of elasticities of substitution since one would expect that such elasticities would be significantly different for "compact" automobiles versus "sport" automobiles. Lumping such dissimilar automobiles into the same good can then bias these elasticity estimates. There are many other such examples of dissimilar automobiles falling under the same HS10 classification in Table 4.

There is one dimension in which the import data good categories may be preferred to those in our market data. The import data

7 These are presumably transactions where a final consumer purchased a car in the foreign country and arranged to have it personally shipped, rather than a result of manufacturers shipments. This inflates the amount of churning one sees in terms of the number of net new varieties in the import data, but translates to little churning in terms of the market share component of these "new" varieties.

8 This assumes that consumers either are not informed about production location or do not differentiate between automobile models that are identical except for production location. Anecdotally, to "buy American" implies purchasing either a General Motors, Ford, or Chrysler vehicle. However, these companies outsource a significant percentage of their production, including entire model lines, most frequently to Canada and Mexico.

9 The data may even show losses of varieties if production of an entire good is transplanted.

10 Cross-border mergers can blur the distinction between domestic- and foreign-owned varieties. Daimler and Chrysler were merged from 1998 to 2007, which covers a part of our sample years. Since the merged company continued to have models sold separately under their different makes and kept their operations relatively segmented, we continue to classify Chrysler models as domestic-owned throughout our sample. This ensures that our variety change calculations are not artificially affected by this merger.

11 Establishing a concordance between the import data and market data is not always straightforward. We first use the U.S. Customs' definition of the good, then compare the characteristics of the imported goods with the EPA's classification of vehicle classes, and concord goods with overlapping characteristics. Specialty vehicles, such as "sports" and "luxury" autos are the most difficult to concord. For the most part these vehicles do not follow stringent definitions, and we may conceivably observe their imports in seemingly unrelated groups.

Table 2
Varieties in import data.

Good		Interior (ft^3)		Cylinders		Cyl vol (cc)		Varieties	
HS10	Description	Min	Max	Min	Max	Min	Max	Total	Average
8703210000	Pass vehicle spark ignition	NESOI		NESOI			1000	43	24
8703220000	Pass vehicle spark ignition	NESOI		NESOI		1000	1500	32	18
8703230010[a]	Pass vehicle spark ignition	NESOI		NESOI		1500	3000	11	5
8703230022	Wagon or vans: height under 160 cm		99		4	1500	3000	11	5
8703230024	Wagon or vans: height under 160 cm	99	109.5		4	1500	3000	13	4
8703230026	Wagon or vans: height under 160 cm	109.5	120		4	1500	3000	14	4
8703230028	Wagon or vans: height under 160 cm	120			4	1500	3000	15	7
8703230032	Wagon or vans: height above 160 cm		99		4	1500	3000	17	7
8703230034[a]	Wagon or vans: height above 160 cm	99	109.5		4	1500	3000	14	5
8703230036	Wagon or vans: height above 160 cm	109.5	120		4	1500	3000	11	4
8703230038	Wagon or vans: height above 160 cm	120			4	1500	3000	16	9
8703230042	Other: cars and pickups		85		4	1500	3000	23	15
8703230044	Other: cars and pickups	85	99		4	1500	3000	26	19
8703230046	Other: cars and pickups	99	109.5		4	1500	3000	21	18
8703230048	Other: cars and pickups	109.5	120		4	1500	3000	17	13
8703230052	Other: cars and pickups	120			4	1500	3000	17	14
8703230062	Other: cars and pickups		99	5	6	1500	3000	23	20
8703230064	Other: cars and pickups	99	109.5	5	6	1500	3000	19	17
8703230066	Other: cars and pickups	109.5	120	5	6	1500	3000	19	13
8703230068	Other: cars and pickups	120		5	6	1500	3000	26	18
8703230072[a]	Other: cars and pickups		99	7		1500	3000	8	4
8703230074[a]	Other: cars and pickups	99	109.5	7		1500	3000	8	5
8703230076[a]	Other: cars and pickups	109.5	120	7		1500	3000	12	5
8703230078[a]	Other: cars and pickups	120		7		1500	3000	13	7
8703240032[a]	Pass vehicle spark ignition		85		4	3000		10	3
8703240034[a]	Pass vehicle spark ignition	85	99		4	3000		8	3
8703240036[a]	Pass vehicle spark ignition	99	109.5		4	3000		8	4
8703240038[a]	Pass vehicle spark ignition	109.5	120		4	3000		7	4
8703240042[a]	Pass vehicle spark ignition	120			4	3000		12	5
8703240052	Pass vehicle spark ignition		99	5	6	3000		19	12
8703240054	Pass vehicle spark ignition	99	109.5	5	6	3000		21	11
8703240056	Pass vehicle spark ignition	109.5	120	5	6	3000		19	14
8703240058	Pass vehicle spark ignition	120		5	6	3000		25	17
8703240062	Pass vehicle spark ignition		99	7		3000		19	12
8703240064	Pass vehicle spark ignition	99	109.5	7		3000		18	11
8703240066	Pass vehicle spark ignition	109.5	120	7		3000		18	12
8703240068	Pass vehicle spark ignition	120		7		3000		21	16
8703310000[a]	Pass vehicle diesel	NESOI		NESOI			1500	18	8
8703320010	Pass vehicle diesel	NESOI		NESOI		1500	2500	22	11
8703330045	Pass vehicle diesel	NESOI		NESOI		2500		17	8

[a] Goods are not included in our analysis since there is not a variety that satisfies the requirements of a reference variety during our 1990–2005 sample. NESOI stands for Not Elsewhere Specified Or Included.

distinguish automobile goods by cylinders and cylinder volume. Ideally, our market data would have not only sales and prices by make and model, but also by various packages of models (or "trim" levels) offered. For example, in many models one can upgrade from a 4-cylinder version of a model to a 6-cylinder version. Our market data do not provide individual data on sales by these various packages for each variety. However, we believe that the elasticity of substitution differences across these dimensions (and resulting bias of pooling them as one variety) are likely less serious than the pooling across major automobile classifications, such as "compact" versus "sport" automobiles.[12]

3. Methodology

3.1. Nested CES model

In this section, we briefly review the methodology developed by Feenstra (1994) and expanded upon by Broda and Weinstein (2006), to account for product variety change in price indexes which can then be translated into welfare changes for an economy specify the upper level of the nested CES utility function as,

$$U_t = \left(D_t^{\frac{\kappa-1}{\kappa}} + M_t^{\frac{\kappa-1}{\kappa}} \right)^{\frac{\kappa}{\kappa-1}}, \tag{1}$$

where $\kappa > 1$ represents the elasticity of substitution between D_t, the composite domestic good, and M_t, the composite imported good. We define the composite good in the second tier of the utility function as,

$$D_t = \left(\sum_{g \in G} D_{gt}^{\frac{\gamma_D - 1}{\gamma_D}} \right)^{\frac{\gamma_D}{\gamma_D - 1}} \text{ and } M_t = \left(\sum_{g \in G} M_{gt}^{\frac{\gamma_M - 1}{\gamma_M}} \right)^{\frac{\gamma_M}{\gamma_M - 1}}.$$

The subutilities derived from the consumption of the domestic and imported goods from the set of all goods $g \in G$ in time t are denoted by D_{gt} and M_{gt}, respectively. The elasticity of substitution across goods in the set of all domestic and imported goods is denoted by γ_D and γ_M, with each strictly greater than one. As in the recent works of Broda and Weinstein (2006) and Imbs and Mejean (2009), the second tier differentiates across HS10 goods. Alternatively, our market-based data define goods by the categories listed in Table 1.

[12] A related issue is measurement error in our price data if actual transaction prices differ from suggested retail prices that we use as our price variable in the market data. This is assuredly true in many cases, however we expect the differences to be small. We also have data on the prices of different trim levels for all models and have investigated using a simple average of these prices. The results differ insignificantly, and we find using prices of the baseline model most tractable as a simple average likely gives more weight to the expensive trim levels than is appropriate.

Table 3
Annual average variety change across goods.

Good	Number of		Share of value for		Total
	New varieties	Exiting varieties	New varieties	Exiting varieties	Value ($Mil)
Import data (HS10)					
8703210000	3.69	3.00	0.000	0.000	1843
8703220000	2.38	2.13	0.000	0.000	3626
8703230022	1.00	1.00	0.076	0.019	159
8703230024	1.31	1.31	0.020	0.051	31
8703230026	1.00	1.19	0.003	0.005	213
8703230028	1.63	1.38	0.003	0.000	2542
8703230032	2.19	2.13	0.161	0.160	48
8703230036	0.94	1.31	0.138	0.197	39
8703230038	1.69	1.81	0.026	0.061	799
8703230042	2.00	1.88	0.003	0.002	2851
8703230044	2.06	2.00	0.002	0.000	11,065
8703230046	1.94	1.75	0.001	0.000	18,913
8703230048	1.75	1.56	0.027	0.002	4242
8703230052	1.94	2.06	0.000	0.035	8034
8703230062	2.25	1.69	0.009	0.001	6652
8703230064	1.50	1.44	0.001	0.000	3984
8703230066	1.88	1.88	0.001	0.001	6175
8703230068	2.44	2.19	0.001	0.000	9992
8703240052	1.63	1.38	0.015	0.005	3186
8703240054	1.81	1.63	0.006	0.003	2701
8703240056	1.44	1.31	0.001	0.000	18,757
8703240058	2.19	1.81	0.001	0.001	29,402
8703240062	1.38	1.25	0.046	0.000	3743
8703240064	1.13	1.06	0.002	0.001	1691
8703240066	1.69	1.50	0.000	0.000	2719
8703240068	1.50	1.19	0.000	0.000	18,290
8703320010	2.38	1.94	0.053	0.052	291
8703330045	1.63	1.25	0.017	0.028	87
Average[a]	1.90	1.64	0.004	0.003	13,131
Market data					
Compact	2.31	2.00	0.065	0.049	36,377
Compact Executive	0.44	0.25	0.053	0.015	4432
Executive	0.31	0.63	0.046	0.051	4519
Fullsize	1.44	1.75	0.050	0.023	26,283
Grand Tourer	0.38	0.25	0.100	0.013	1639
Luxury	0.19	0.13	0.073	0.002	1333
Midsize	3.00	2.50	0.054	0.023	57,976
Minivan	1.19	0.94	0.052	0.027	21,820
PU	0.88	0.75	0.035	0.016	40,203
Sport	1.88	1.81	0.090	0.096	6380
Sport compact	0.50	1.06	0.061	0.051	8217
Subcompact	1.25	2.06	0.076	0.100	3177
SUV	5.38	2.31	0.081	0.015	62,570
Van	0.44	0.63	0.060	0.062	6400
Average[a]	2.49	1.22	0.059	0.029	40,358

[a] Calculated as a weighted average using the ideal log-change weights.

The focus of Feenstra (1994) is to characterize consumers' choices given they have chosen a product group. Thus, he limits his analysis to the third tier utility function given by,

$$M_{gt} = \left(\sum_{v \in V} b_{gvt}^{\frac{1}{\sigma_g}} m_{gvt}^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g - 1}}, \tag{2}$$

with $\sigma_g > 1 \, \forall g \in G$. Nonsymmetry in our third tier CES function comes from the taste parameter b_{gvt}, which varies across goods and varieties. Consumption of a particular variety of a given import $v \in V$, in period t is m_{gvt}.

A valid concern one may raise from this representative consumer specification for automobiles is that individual consumers are likely not purchasing continuous quantities of each variant. Automobiles are an infrequently purchased durable good that lend themselves to the concept of discrete choice. In fact, the bulk of the industrial organization literature maintains discrete choice individual utility for consumers.[13] However, Anderson et al. (1992) demonstrate that an economy consisting of consumers who choose a variable amount of variety v in time period t and obtain indirect utility

$$u_{gvt}^h = \ln(y) - \alpha_g \ln\left(b_{gvt}^{\frac{1}{\alpha_g}} p_{gvt} \right) + \epsilon_{gvt}^h$$

with $\alpha_g = \sigma_g - 1 > 0$, is theoretically equivalent to the preceding third tier of the CES representative consumer model.[14] Thus, allowing for a nested discrete choice of similar form we can reconcile these seemingly distinct demand systems.

Both types of preferences (Dixit–Stiglitz or a discrete choice logit framework) have a common feature that any variety garnering positive market share implies that there are consumers that love that variety the most. Thus, the entry of new varieties automatically implies welfare gains, giving these frameworks a built-in love of variety. This feature has the potential to overestimate welfare gains if demand-side characteristics, such as product attributes, are not adequately controlled for (e.g., see Petrin (2002), and Ackerberg and Rysman (2005)). While it is relatively straightforward to control for such demand characteristics in the logit demand framework, it has not been done in the Dixit–Stiglitz framework and related estimation routine employed by Feenstra that we follow in this paper.[15] Since we wish to compare our results to prior studies that use Dixit–Stiglitz preferences, we also use these preference assumptions. While this source of bias in the Dixit–Stiglitz framework means that absolute welfare estimates should be treated with appropriate caution, we presume that this source of bias is less of an issue with the relative comparisons we highlight in this common framework.

3.2. Price index

The resulting unit cost requirement of Eq. (2) is readily shown to be,[16]

$$\phi_{gt}^M\left(I_{gt}, \mathbf{b}_{gt}\right) = \left(\sum_{v \in I_{gt}} b_{gvt} p_{gvt}^{1-\sigma_g} \right)^{\frac{1}{1-\sigma_g}},$$

where bold font denotes a vector, such that $\mathbf{b}_{gt}$ is a good specific vector of taste shocks for each variety b_{gvt}. Substituting in each

[13] There is a rich history of work dealing with the estimation of the underlying demand and supply structure in the automobile market. Goldberg (1995) analyzes the market assuming a discrete choice utility model and estimating a nested logit. Using consumer level data, she is able to estimate plausible substitution patterns across available automobile varieties. Berry et al. (1995) develop a random coefficients logit model that effectively estimates the demand system for autos in the absence of consumer specific data.

[14] There is a comparable discussion in Appendix B of Feenstra's (2004) *Advanced International Trade*.

[15] See Sheu (2010) for a study that provides a comparison of price and welfare effects from variety change using Dixit–Stiglitz preferences versus the discrete-choice framework of Berry et al. (1995). Sheu's analysis introduces product attributes in a Dixit–Stiglitz framework, but not in the context of the Feenstra estimation routine we pursue here.

[16] Henceforth, disaggregate expressions are explicitly defined for imported varieties, but can be analogously defined for domestic varieties by replacing M and m_{gvt} with D and d_{gvt}.

Table 4
Estimates using import data and concordance with market data.

Good	$\hat{\sigma}$	End-point ratio			Ideal weights	Cars								Vans		Pickup	SUV
		Conventional price index	Corrected λ ratio	Exact price index		Comp	Exec	Fullsize	Lux[a]	Mid size[b]	Sport	Sport comp	Sub comp	Minivan	Van		
8703210000	9.523	1.096	1.000	1.096	0.009	X						X	X				
8703220000	34.425	2.093	1.000	2.093	0.031	X	X	X	X	X	X	X	X	X			X
8703230022	6.621	2.503	0.393	0.983	0.002	X	X	X	X	X	X	X	X	X			X
8703230024	100.050	0.559	1.038	0.580	0.000									X			X
8703230026	56.100	1.211	1.001	1.212	0.002									X			X
8703230028	6.392	2.362	0.988	2.334	0.013									X			X
8703230032	13.050	1.125	1.061	1.194	0.001									X			X
8703230036	2.341	0.219	73.684	16.104	0.000										X		X
8703230038	18.393	1.077	1.201	1.294	0.011										X		X
8703230042	86.775	2.536	1.000	2.535	0.025										X		X
8703230044	6.375	1.829	0.990	1.810	0.079						X	X	X				
8703230046	17.496	1.946	0.999	1.945	0.104	X					X	X					
8703230048	20.449	1.843	0.965	1.779	0.022					X	X						
8703230052	27.450	0.735	1.044	0.767	0.037		X	X	X								
8703230062	3.000	0.904	0.880	0.795	0.051						X	X	X				
8703230064	3.750	1.252	0.995	1.245	0.043	X					X	X					
8703230066	4.425	2.352	0.992	2.333	0.064					X							
8703230068	6.870	1.354	0.996	1.348	0.061		X	X	X								
8703240052	3.600	0.693	0.882	0.612	0.025						X	X	X				
8703240054	1.575	1.083	0.942	1.020	0.018	X					X	X					
8703240056	9.222	1.711	0.997	1.707	0.090					X							
8703240058	9.097	1.918	0.999	1.916	0.148		X	X	X								
8703240062	1.875	1.846	0.253	0.467	0.027						X	X	X				
8703240064	2.726	4.326	0.981	4.244	0.012	X					X	X					
8703240066	7.327	2.045	1.000	2.044	0.023					X	X						
8703240068	1.875	1.964	0.995	1.954	0.096		X	X	X								
8703320010	5.550	1.629	1.293	2.200	0.002	X	X		X	X			X		X	X	X
8703330045	9.450	1.680	1.003	1.684	0.001		X		X						X	X	X
Average[c]	11.382	1.758	0.977	1.709	0.034												
Median	7.099	1.695	0.998	1.695	0.024												

[a] Includes Grand Tourer class cars.
[b] Includes Compact Executive class cars.
[c] Calculated as a weighted average using the ideal log-change weights. Sources contributing to the concordance of data: http://www.gobiodiesel.org/years.html, http://www.fueleconomy.gov, and http://en.wikipedia.org.

preceding tier of the nested CES utility produces the overall price index,

$$p_t = \left(\left(\phi_t^D\right)^{1-\kappa} + \left(\phi_t^M\right)^{1-\kappa}\right)^{\frac{1}{1-\kappa}}, \text{ where}$$

$$\phi_t^M(G) = \left(\sum_{g\in G}\left(\phi_{gt}^M\left(I_{gt},\mathbf{b}_{gt}\right)\right)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}.$$

In order to quantify the bias from neglecting variety change when calculating these equations, Feenstra (1994) manipulates relative unit cost requirements into an aggregate price index free of taste parameters, assuming $b_{gvt} = b_{gvt-1}$ for $v\in\mathbf{I}_g\subseteq\left(I_{gt}\cap I_{gt-1}\right), \mathbf{I}_g\neq\emptyset$. We define the exact price index, π, as a function of price and cost minimizing quantity and variety vectors such that,

$$\begin{aligned}\frac{\phi_{gt}^M\left(I_{gt},\mathbf{b}_g\right)}{\phi_{gt-1}^M\left(I_{gt-1},\mathbf{b}_g\right)} &= \pi_g^M\left(\mathbf{p}_{gt},\mathbf{p}_{gt-1},\mathbf{m}_{gt},\mathbf{m}_{gt-1},\mathbf{I}_g\right)\\ &= P_g^M\left(\mathbf{p}_{gt},\mathbf{p}_{gt-1},\mathbf{m}_{gt},\mathbf{m}_{gt-1},\mathbf{I}_g\right)\left(\frac{\lambda_{gt}}{\lambda_{gt-1}}\right)^{\frac{1}{\sigma_g-1}}.\end{aligned} \quad (3)$$

Further, by Sato (1976) and Vartia (1976) the "conventional" price index is

$$P_g^M\left(\mathbf{p}_{gt},\mathbf{p}_{gt-1},\mathbf{m}_{gt},\mathbf{m}_{gt-1},\mathbf{I}_g\right) \equiv \prod_{v\in\mathbf{I}_g}\left(\frac{p_{gvt}}{p_{gvt-1}}\right)^{w_{gvt}\left(\mathbf{I}_g\right)}.$$

This is the geometric mean of particular variety price changes weighted by their ideal log-change weight,

$$w_{gvt}\left(\mathbf{I}_g\right) \equiv \frac{\left(\frac{s_{gvt}-s_{gvt-1}}{\ln s_{gvt}-\ln s_{gvt-1}}\right)}{\sum_{v\in\mathbf{I}_g}\left(\frac{s_{gvt}-s_{gvt-1}}{\ln s_{gvt}-\ln s_{gvt-1}}\right)},$$

which is a normalized logarithmic mean of the variety cost shares $s_{gvr} = \frac{p_{gvr}m_{gvr}}{\sum_{v\in\mathbf{I}_g}p_{gvr}m_{gvr}}$ for $r = t, t-1$.

The bias generated by neglecting new and disappearing varieties in Eq. (3) on the conventional price index is quantified by the λ ratio raised to the power of $\frac{1}{\sigma_g-1}$, which we will call the *corrected* λ ratio. The λ ratio captures the net gain or loss of varieties by value:

$$\lambda_{gr} \equiv \frac{\sum_{v\in\mathbf{I}_g}p_{gvr}m_{gvr}}{\sum_{v\in I_{gr}}p_{gvr}m_{gvr}} \quad \text{for } r = t-1, t,$$

where $v\in\mathbf{I}_g = \left(I_{gt}\cap I_{gt-1}\right)$ and $g\in G$. When the expenditure on new varieties in the current period exceeds that of disappearing varieties from the previous period, the λ ratio scales down the exact price index. Intuitively, consumers gain (lose) from an excess value of new (disappearing) varieties. The corrected λ ratio adjusts the period by period net value of new and disappearing varieties by the substitutability of these varieties. For large σ_g varieties are more homogeneous, and increasing or decreasing the set available to consumers has little

Table 5
Estimates for Imported Market Goods.

Class	$\hat{\sigma}$	End-point ratio			Ideal weights
		Conventional price index	Corrected λ ratio	Exact price index	
Compact[a]	14.250	1.630	0.989	1.613	0.162
Executive	3.001	1.414	1.102	1.558	0.028
Fullsize	16.650	1.505	0.941	1.416	0.059
Luxury	6.298	1.597	0.674	1.077	0.020
Midsize	18.750	1.513	0.956	1.447	0.241
Pickup	6.535	1.854	1.015	1.881	0.151
Sport	21.629	1.651	0.954	1.575	0.029
Sport Compact	3.305	1.300	0.985	1.281	0.003
SUV	5.341	1.762	0.780	1.355	0.242
Van[b]	9.525	1.412	0.972	1.373	0.066
Average[c]	11.562	1.639	0.926	1.509	0.100
Median	8.030	1.555	0.964	0.964	0.062

[a] Includes Subcompact.
[b] Includes in Minivan.
[c] Calculated as a weighted average using the ideal log-change weights. The following classes have been combined with classes possessing highly comparable traits due to the unavailability of a well defined reference variety.

effect on the exact price index. This is demonstrated clearly through the corrected λ ratio, which approaches 1 regardless of the net variety change for surviving goods as $\sigma_g \to \infty$.

3.2.1. Aggregate price index

Broda and Weinstein (2006) expand upon the exact price index for the subutility function derived by Feenstra (1994) and show that one can easily aggregate price indexes across many goods into the aggregate exact import price index,

$$\Pi^M = \prod_{g \in G} P_g^M \left(\mathbf{I}_g\right)^{w_{gt}} \left(\frac{\lambda_{gt}}{\lambda_{gt-1}}\right)^{\frac{w_{gt}}{\sigma_g - 1}}. \tag{4}$$

The weights, w_{gt}, are ideal log-change at the goods level. This second tier price index is dual for domestic varieties. Thus, the aggregate market price index is,

$$\Pi = \left(\Pi^D\right)^{w_t^D} \left(\Pi^M\right)^{w_t^M}. \tag{5}$$

The exponents are ideal log-change weights at the good origin level. Due to the lack of domestic data detailing variety change, Broda and Weinstein (2006) cannot calculate an overall price index as we have defined. They instead focus on the effect of varieties on the aggregate import price index.[17] Thus, for direct comparisons we concentrate on the aggregate import price index using various assumptions about our market level data, since our import data are deficient in the same way.

3.3. Estimating σ

To implement the price index calculations in Eq. (4), we need consistent estimates of the elasticity of substitution for each variety.

[17] As a robustness check, Broda and Weinstein (2006) relax the assumption of Krugman (1979), which specifies that the number of domestic varieties is unassailable from competition with new foreign varieties. They adopt a model from Helpman and Krugman (1985) which suggests, for their estimates, it is appropriate to reduce their welfare gains to U.S. consumers from imported varieties by 16%. While we will not be able to estimate if this is the correct scaling factor to account for displaced domestic varieties, we do expand upon the understanding of variety change in the domestic sector.

Following Feenstra (1994), the underlying demand in first differences for each variety can be written as

$$\Delta \ln\left(s_{gvt}\right) = \varphi_{gt} - \left(\sigma_g - 1\right)\Delta \ln\left(p_{gvt}\right) + \epsilon_{gvt}, \text{where } \varphi_{gt} = \left(\sigma_g - 1\right)\ln\left[\frac{\phi_{gt}^M\left(\mathbf{b}_{gt}\right)}{\phi_{gt-1}^M\left(\mathbf{b}_{gt-1}\right)}\right]$$

is a time-good specific random shock driven by the vector of random variety specific taste parameters $\mathbf{b}_{gt}$. The error term of the log of the variety cost share, $\ln(s_{gvt})$, is determined by the random tastes of consumers across varieties such that $\epsilon_{gvt} = \Delta \ln(b_{gvt})$.

We assume producers compete in monopolistically competitive markets for their varieties such that prices are, as in Broda and Weinstein (2006), given by,

$$p_{gvt} = \left(\frac{\sigma_g}{\sigma_g - 1}\right)\exp\left(\nu_{gvt}\right) m_{gvt}^{\omega_g}.$$

Random technology shocks are capture by ν_{gvt}. The inverse supply elasticity for each good is $\omega_g \geq 0$.[18] Taking logs, first differences and using the definition of shares, we obtain

$$\Delta \ln\left(p_{gvt}\right) = \psi_{gt} + \frac{\omega_g}{1+\omega_g}\Delta \ln\left(s_{gvt}\right) + \delta_{gvt}, \text{where } \psi_{gt} = -\omega_g \Delta \ln\left(\frac{\sum_{v \in I_{gt}} p_{gvt} m_{gvt}}{1+\omega_g}\right)$$

represents time-good specific shocks to production. Random technology changes in the production of each variety, ν_{gvt} manifest themselves through $\delta_{gvt} = \Delta \ln\left(\frac{\nu_{gvt}}{1+\omega_g}\right)$. We assume ϵ_{gvt} and δ_{gvt} are independent, which is vital for identification.

Choosing a reference variety $k \in V$ that is present in each period of our sample, we can difference our demand and supply equations, yielding

$$\Delta^k \ln\left(s_{gvt}\right) = -\left(\sigma_g - 1\right)\Delta^k \ln\left(p_{gvt}\right) + \epsilon_{gvt}^k \text{ and} \Delta^k \ln\left(p_{gvt}\right) = \frac{\omega_g}{1+\omega_g}\Delta^k \ln\left(s_{gvt}\right) + \delta_{gvt}^k.$$

Variables denoted with a superscript k are differenced from the reference good; such as $\epsilon_{gvt}^k = \epsilon_{gvt} - \epsilon_{gkt}$. Independence of our errors implies $E[\epsilon_{gvt}^k \delta_{gvt}^k] = 0$, which we utilize to obtain the estimating equation

$$\Delta^k \ln\left(p_{gvt}\right)^2 = \theta_1 \Delta^k \ln\left(s_{gvt}\right)^2 + \theta_2 \Delta^k \ln\left(p_{gvt}\right)\Delta^k \ln\left(s_{gvt}\right) + u_{gvt} \tag{6}$$

$$\text{where } u_{gvt} = \frac{\epsilon_{gvt}^k \delta_{gvt}^k}{\sigma_g - 1}, \theta_1 = \frac{\omega_g}{\left(1+\omega_g\right)\left(\sigma_g - 1\right)^2}$$

$$\text{and} \quad \theta_2 = \frac{1 - \omega_g\left(\sigma_g - 2\right)}{\left(1+\omega_g\right)\left(\sigma_g - 1\right)}$$

by multiplying the differenced error terms and dividing by $\sigma - 1$. Endogeneity is apparent since the error term in our estimating equation is comprised of the error terms of the regressands. Feenstra (1994)

[18] This implicitly assumes that the supply elasticity is identical across varieties.

Table 6
Summary table of aggregate end-point results across different data sets and scenarios.

	Conventional price index	Corrected λ ratio	Exact price index
Comparing estimates between import sample and market sample			
Imported varieties – import sample	1.657	0.959	1.590
Imported varieties – market sample	1.564	0.922	1.442
Market sample – estimates for imported and domestic varieties			
Imported varieties	1.564	0.922	1.442
Domestic varieties	1.666	0.962	1.604
Market sample – estimates for foreign-owned and domestic-owned varieties			
Foreign-owned varieties	1.564	0.871	1.362
Domestic-owned varieties	1.699	0.935	1.589

demonstrates that by taking advantage of the panel nature of the data one can control for this endogeneity by using variety specific dummies as instruments. This will produce consistent and efficient estimates of θ_1 and θ_2 as long as, for some countries $i \neq k$ and $j \neq k$

$$\frac{\sigma^2_{\epsilon,i}+\sigma^2_{\epsilon,k}}{\sigma^2_{\epsilon,j}+\sigma^2_{\epsilon,k}} \neq \frac{\sigma^2_{\delta,i}+\sigma^2_{\delta,k}}{\sigma^2_{\delta,j}+\sigma^2_{\delta,k}}.$$

From the consistent estimates of θ_1 and θ_2, we can calculate

$$\hat{\rho}_g = \begin{cases} \dfrac{1}{2}+\left(\dfrac{1}{4}-\dfrac{1}{4+\frac{\hat{\theta}_2}{\hat{\theta}_1}}\right)^{\frac{1}{2}} & \text{for } \theta_1,\ \theta_2>0 \\ \dfrac{1}{2}-\left(\dfrac{1}{4}-\dfrac{1}{4+\frac{\hat{\theta}_2}{\hat{\theta}_1}}\right)^{\frac{1}{2}} & \text{for } \theta_1>0,\ \theta_2<0, \end{cases}$$

where $\rho_g \equiv \frac{\omega_g(\sigma_g-1)}{1+\sigma_g\omega_g}$ is the correlation between shifts in the demand curve and the resulting change in equilibrium prices. The elasticity of substitution for each good is then

$$\hat{\sigma}_g = 1+\left(\frac{2\hat{\rho}_g-1}{1-\hat{\rho}_g}\right)\frac{1}{\hat{\theta}_2}. \tag{7}$$

In the case where $\theta_1<0$, theoretically feasible values for ρ_g and σ_g generally cannot be computed.[19] In which case, we use a grid search over the set of economically feasible values of σ_g and ρ_g to minimize the GMM objective function implied by the IV estimation when Eq. (7) cannot be calculated. Explicitly, we choose values $\sigma_g \in [1.05, 100.05]$ at equally spaced intervals of .075 and 100 equal intervals of $\rho_g \in [0,(\sigma_g-1)/\sigma_g)$ to minimize $G(\rho_g,\sigma_g)'WG(\rho_g,\sigma_g)$, where $G(\rho_g,\sigma_g)$ is the sample analog of the moment condition, $G(\rho_g,\sigma_g)=E_t[u_{gvt}]=0, \forall v$.[20]

The use of unit values in place of prices is inherent to import data. Feenstra (1994) demonstrates that including a constant when estimating Eq. (6), solves the problem of known measurement error bias from using unit values. We will also include a constant when we estimate using market data, however we expect measurement error to be less severe since we observe *Manufacturer's Suggested Retail Prices* (MSRPs), which are likely closer to actual transaction prices than unit values.[21]

4. Empirical analysis

We begin our empirical analysis with a comparison of estimated new variety effects using the traditional import data versus using the market data. As described in the previous section, this involves first estimating elasticities of substitution separately by good for each data set following Feenstra (1994) and calculating conventional price indexes and exact price indexes (which correct for net new variety change) by good. Then, following Broda and Weinstein (2006), we construct aggregate price index changes for all automobile goods in each data set, part of which separates out the change in the aggregate price index due to net new varieties. This can then also be easily related to welfare effects due to net new varieties. As we will see, there are significant differences across the data sets in the estimated effects of net new varieties.

After these comparisons, we next explore a few issues that one can only address using our market data sample. First, unlike import data, we have data on domestic varieties and can compare the patterns of net variety change on prices across domestic and import varieties. Second, we have information on ownership of varieties in our market data. Therefore, we can examine the effect of new variety change on U.S. welfare from "foreign-owned" varieties, rather than simply from imports, some of which maybe outsourced domestic-owned varieties. Finally, we examine whether calculated variety change and its effects are sensitive to the time interval used.

4.1. Comparison between import and market data

We begin our comparison by first estimating elasticities of substitution for goods for each data sample and then calculating annual price index changes by good which are comprised of pure price changes of continuing models and equivalent price changes due to net variety change. Columns 2 through 5 of Table 4 report relevant estimates and constructed measures by goods when using our import data.

Column 2 reports our estimates of elasticities of substitution by HS10 goods. There is substantial heterogeneity in these estimates across goods with a weighted-average mean elasticity of substitution around 11.4 and a median of 7. There are a couple outliers with large estimated elasticities, but these are for goods with obviously low import activity given the weighted-average elasticity.

We then use these estimates along with our price, quantity, entry, and exit data by variety to calculate annual changes in the exact price index for each good from 1990 through 2005, which can be decomposed into the change in the conventional price index multiplied by the corrected λ ratio that translates net variety change into effective price index changes. We normalize the price indexes and corrected λ ratios to be 1 for each good in our base year, 1990, and cumulate these changes over our sample.

Columns 3, 4, and 5 of Table 4 show the value of the conventional price index, the corrected λ term in the final year, and the exact price index for the final year of our sample, 2005, for each good. For example, a value of 2 in the final year means that price index of the good has doubled over our sample. Also, recall that the corrected λ ratio measures how much the net variety change scales up or down

[19] It is possible to obtain feasible estimates under relatively stringent conditions. Feenstra (1994) discusses these conditions, of which none of our estimates satisfy, and demonstrates a case where feasible estimates are obtained using these equations with $\theta_1<0$.

[20] Soderbery (2010) provides a Monte Carlo analysis of the Feenstra estimation methodology for estimating elasticities of substitution. While there is significant bias in the estimate of some of the base parameters, the bias in the constructed estimate of substitution elasticities are not unreasonably large (10–15% for reasonably sized samples). He also shows that the upward bias in the elasticity of substitution creates a downward bias of about 10% in the welfare calculations of new variety change.

[21] Broda and Weinstein's (2006) propose an optimal weighting scheme assuming that the measurement error in unit values is inversely related to the volume of imports. We expect the same relationship to apply to our market level data for an additional reason. We use the *Manufacturer's Suggested Retail Price* MSRP of the base model as our price measure, which will be biased down for units sold with expensive options, and biased up in the presence of dealer discounts. However, there is anecdotal evidence that high sales vehicles tend to have transaction prices closer to the MSRP of the base model than vehicles with fewer sales.

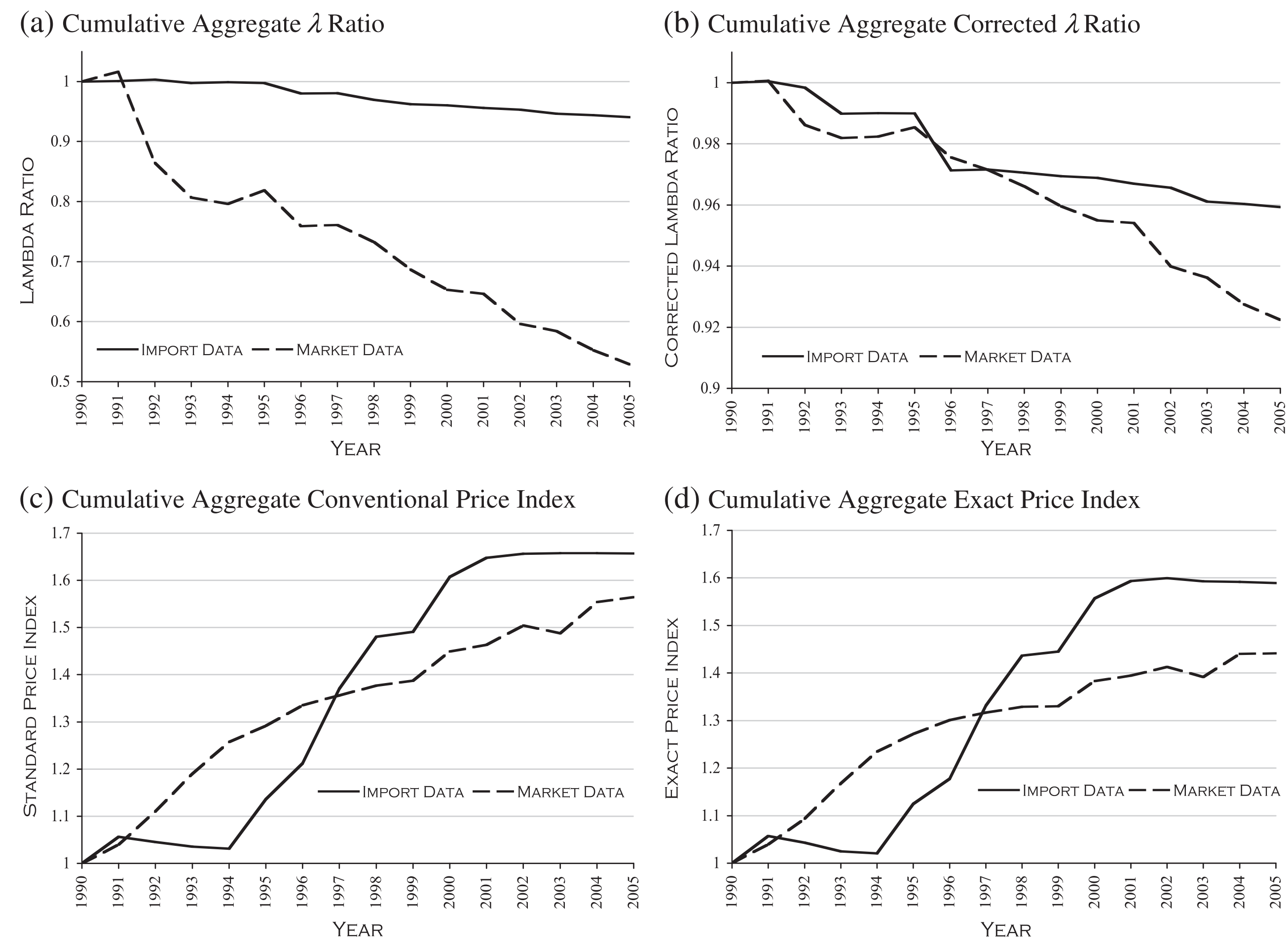

Fig. 1. Comparing aggregate estimates for market and import data.

the exact price index relative to the conventional price index with values below 1.000 indicating net new variety gains.

Overall, there is modest evidence of net variety entry with a weighted-average end-point corrected λ ratio of 0.977 and a median of 0.998 across goods in Table 4. This represents significantly less net variety change than that found by Broda and Weinstein (2006), who report a median value of 0.950 for the same measure across all imported products over the 1990–2001 period. This finding is consistent with the low turnover of automobile varieties we find in the import data, as discussed earlier.[22] As a result of little net new variety change, the differences in the exact price index over the conventional price index are small on average across products.

While these goods-level comparisons are interesting, the true aggregate picture of net variety change effects on prices across the automobile sector is obtained by aggregating across goods following Eq. (4). These calculations suggest more variety change with an aggregate corrected λ ratio of 0.959 over our 1990–2005 sample of automobile goods. Thus, while the conventional price index went up 65.7% over our sample of imported automobiles, the exact price index (controlling for net variety change) only goes up 59.0%. This translates into a ratio of the exact price index to the conventional price index of 0.960 which compares to Broda and Weinstein (2006) finding of 0.917 for the same ratio across all imported products in their 1990–2001 sample. These are our baseline measures using the import data and associated HS10 goods definitions.

We next apply the same methodology to generate analogous results using our market data and market-based goods definitions. The latter columns of Table 4 show the rough concordance between the import (HS10) goods and our market goods categories. It is clear that the concordance is far from a one-to-one mapping and that this will be one source of differences in estimates across the two samples.

Table 5 provides substitution elasticities, the conventional price index changes, the corrected λ ratio changes, and the exact price index changes over our sample of goods using our market data. We cannot make separate estimates for a couple of market goods (subcompact and van) using our market data because there is no single variety that survives our entire sample to serve as a reference variety. Therefore, we combine the data for subcompact automobiles with compact automobiles, and vans with minivans.

From Table 5, estimated elasticities of substitution tend to be slightly higher across our market-based goods (11.562) than for the import-based goods (11.382).[23] These higher substitution elasticities

[22] HS10 good 8703230036 is an outlier with a very large increase in its exact price index and corrected λ ratio. This stems from a very large exit of varieties during one of our sample's earlier years. We conjecture that there may have been an incorrect HS10 classification in this early year, which was right after the U.S. switched to the HS method of categorizing import goods. Such an artificial change in HS10 classification would then show up as a large exit in our data. Since this particular HS10 good accounts for a miniscule share of U.S. automobile imports, it has essentially no impact on our aggregate price index and welfare calculations.

[23] We estimate Eq. (6) many times, so for the sake of brevity we do not present the individual coefficient estimates. However, the presence of measurement error follows our previous predictions; estimates of our constant term suggest measurement error plays a significant role in our import data, but has little distinguishable effect in our market data.

Table 7
Imported versus domestic varieties.

Class	$\hat{\sigma}$		Corrected λ ratio		Exact price index		Ideal log weights	
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign
Compact	5.282	14.250	1.036	0.989	1.693	1.613	0.155	0.177
Fullsize	4.200	16.500	0.924	0.964	1.407	1.420	0.100	0.153
Midsize	24.472	18.750	0.980	0.956	1.497	1.447	0.204	0.260
Pickup	8.132	6.535	0.956	1.015	1.776	1.881	0.180	0.068
Sport	13.475	9.026	1.046	0.871	1.696	1.364	0.036	0.089
SUV	17.965	5.341	0.933	0.780	1.620	1.355	0.222	0.152
Van	4.275	9.525	0.890	0.972	1.484	1.373	0.103	0.102

Notes: Certain classes have been combined with others possessing highly comparable traits due to the unavailability of a well defined reference variety for both foreign and domestically produced goods. Specifically, compact, fullsize midsize, SUV, sport, and van include subcompact, executive, crossover SUV, sport compact, and minivan, respectively.

for the market data sample are even more notable given that market-based goods categories only number ten, while the import goods (HS10) for automobiles numbers 28. This suggests that our market-based goods definitions are better at assigning varieties of the same good than the import-based HS10 good definitions.

The average conventional price change across the goods in Table 5 is about 64% using the market data. This represents a smaller average price increase than for the import goods in Table 4 which was about 75%. While this difference is interesting, the interpretation is not clear since our market prices are at the retail level, while the import prices are border prices and most likely correspond best to wholesale prices.[24]

The corrected λ ratio change is significantly lower on average across goods using the market data than for the import data, suggesting more net variety change. Thus, the market data suggest more significant net variety gains than the import data, and this means a greater average difference between the exact and conventional price indexes.

These same results hold when we properly aggregate across our goods using the method of Broda and Weinstein (2006) and characterized in Eq. (4). The first two rows of Table 6 compare the import and market data sets in their estimates of the relevant measures of aggregate automobile price and net variety changes. There are large differences. The primary difference of interest is that the corrected λ ratio across all automobile products using the market data and goods definitions is 0.922, which represents about 90% more effective variety change than our benchmark using the import data and goods definitions (0.959). Assuming the same functional forms as Broda and Weinstein (2006), this means that the welfare effect from net new variety change is also about twice as large using our market data as when using our import data.[25]

Fig. 1 provides further information on the differences between the import and market data sets estimates of net new variety change and its effects on the exact price index, by showing the annual changes over our sample. Panel A of Fig. 1 shows the raw variety change over time captured by the simple λ ratio, where declines in the ratio indicate net new variety gains. The market data show substantially more net new variety gains over time, though these differences are muted when one views the corrected λ ratio shown in Panel B which accounts for the degree of substitutability between varieties for each good. As one can see, the latter years of our sample are when the difference in net new variety changes between the two data sets has become most pronounced. Panels C and D of Fig. 1 show the annual changes in the conventional and exact price indexes for automobiles for each data set. Two clear differences emerge. First, while the market-based price indexes rise at a fairly consistent rate over the sample, the import-based price indexes show relatively flat price changes in the beginning and end of the sample with a relatively large annual price increases from 1994 to 2000. This observation may be attributable to consumption smoothing practices by retailers that cannot be gleaned from unit values, or a pronounced effect of measurement error in our import data contributing to the relatively drastic price swings calculated from these data. Second, and of more interest for this paper, the exact price index shows much lower price increases relative to the conventional price index (due to net new variety change) for the market data set than for the import data set.

In summary, there appear to be two opposing biases from the import-based goods (mis)classifications. First, the Armington assumption on varieties hides significant churning of varieties, which in this case also means it misses a significant amount of net new variety change and biases one toward finding lower gains from new imported varieties. On the other hand, to the extent that import goods classifications (HS10) deviate from (true) market-based classifications, elasticities of substitution will be biased downward, which will bias one toward finding a greater effect of net imported variety change on prices and welfare in the import data. In our case, we find that the former effect outweighs the latter effect in significant fashion, creating a bias that underestimates the price and welfare effects of imported variety change by about half.

4.2. Comparing effects across domestic and import varieties

Past studies of the effects of product variety do not have data on domestic varieties. This means that they cannot estimate and compare the impact of net variety change on prices stemming from domestic sources versus import sources. In contrast, our market data allow us to easily calculate these measures for domestic varieties, which we do and report in Row 3 of Table 7. As one might expect given the declining market shares of U.S. manufacturers in the U.S. automobile market, price increases are larger and net variety effects are smaller for domestic varieties over our sample as compared to what is found for import varieties. We estimate that the effect of net variety change on prices is twice as large for import varieties as from domestic varieties. (0.922 is 0.078 points below a scale factor of 1.000, whereas 0.962 is only 0.038 points below 1.000).

4.3. Accounting for foreign-affiliate production

A country likely cares most about the gains they get from net new varieties from foreign sources, not just those foreign varieties that are

[24] An obvious possibility is that retail profit margins have fallen over our sample. However, there are also measurement issues in both price terms where systematic changes in those errors over time could account for the difference. Since this is not the focus of this paper, we do not pursue the issue further.

[25] As Broda and Weinstein (2006) shows, welfare gains from foreign varieties in this model can be calculated as a compensating variation that is equal to a scaling factor multiplied with the inverse of the conventional price index. This scaling factor is equal to the inverse of the weighted product of the corrected λ ratios across goods, raised to the share of foreign varieties in total consumption. When corrected λ ratios and price indexes are reasonably close to one and the share of foreign varieties in total consumption is greater than zero, then the compensating variation will roughly correlate directly to how far the corrected λ ratios are from one. For example, the inverse of 0.922 is 1.085, which is a scaling factor that will roughly lead to twice the increase from a scaling factor of 1.043 — the inverse of 0.959.

Table 8
Foreign-owned versus domestic-owned varieties.

Class	$\hat{\sigma}$		Corrected λ ratio		Exact price index		Ideal log weights	
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign
Compact	5.629	10.065	0.890	1.032	1.437	1.680	0.114	0.272
Fullsize	20.775	12.620	0.982	0.970	1.529	1.414	0.120	0.113
Midsize	8.081	5.550	0.956	0.781	1.539	1.137	0.193	0.275
Pickup	7.681	13.944	0.952	0.980	1.793	1.596	0.186	0.052
Sport	9.146	12.353	1.048	0.906	1.699	1.414	0.040	0.083
SUV	6.946	11.135	0.844	0.862	1.479	1.483	0.218	0.164
Van	5.100	2.546	0.969	0.187	1.640	0.342	0.130	0.041

Notes: Certain classes have been combined with others possessing highly comparable traits due to the unavailability of a well defined reference variety for both foreign and domestically produced goods. Specifically, compact, fullsize midsize, SUV, sport, and van include subcompact, executive, crossover SUV, sport compact, and minivan, respectively.

imported. As a result, accounting for foreign-affiliate activity is very important in understanding the gains to an economy from foreign sources. Since the early 1980s foreign automobile manufacturers have located substantial production in the United States, often introducing new varieties through foreign-affiliate production, not imports. In this way, a focus on import data alone could substantially underestimate the gains to U.S. consumers from foreign varieties. Relatedly, U.S. automobile manufacturers have outsourced some production to foreign affiliates (most often in NAFTA partner countries) over the past decades as well. These U.S. varieties may then get erroneously assigned as “foreign” because they are part of U.S. imports, and may lead estimates of new variety gains from foreign imports to be overestimated.

Our market data allow us to easily disentangle foreign from U.S. varieties because we have information on both the ownership and production location of an automobile variety. Thus, as an alternative to the import/domestic distinction that is determined by production location, we split our data by ownership into domestic-owned versus foreign-owned varieties. Domestic-owned varieties include any varieties produced by a U.S. auto manufacturer, regardless of whether it was produced in the U.S. or outsourced to an affiliate abroad and imported back into the U.S. Foreign-owned varieties include both those imported into the U.S. and those produced by foreign affiliates in the U.S.

Table 8 provides estimates of substitution elasticities, price indexes, and corrected lambda ratios by goods for our samples of domestic and foreign-owned varieties, while the last rows of Table 6 provide the price indexes and corrected lambda ratio for automobiles when we appropriately aggregate up over these goods. The results contrast significantly with our previous estimates separating varieties into “imports” and “domestic”. While the estimated elasticities of substitution are generally in the same range and conventional price indexes almost identical, the corrected lambda ratios suggest significantly more net variety change when separating varieties into domestic- and foreign-owned. The corrected lambda ratio for foreign-owned is 0.871, which translates into about a 70% larger impact of net variety change on the conventional price index and, thus, similar differences in welfare gains, than from imported varieties alone. Interestingly, the estimated effect of net variety change from domestic-owned varieties is also significantly larger relative to the previously defined “domestic” (or non-imported) category as well.

In summary, it is clear that defining varieties by ownership, rather than by production location, has a significant effect on the estimates. To the extent that ownership is the correct way to classify foreign versus domestic varieties, classification by location (i.e., imports versus non-imports) significantly underestimates the actual welfare gains experienced by U.S. consumers from foreign varieties of automobiles over our sample. This certainly means that the bias from not including foreign-owned, domestic-produced varieties as foreign substantially outweighs the bias from including domestic-owned foreign-produced (or outsourced) varieties as foreign.[26]

4.4. Robustness of results to the time interval of calculations

Given annually recorded data, it is natural to make calculations of net variety changes on an annual basis. However, this implicitly assumes that new varieties achieve their equilibrium market share within the year of introduction. Alternatively, first-year market shares of new varieties may be quite different from their ultimate equilibrium market share. For example, it may take consumers significant time to become fully informed about a new variety. This may be particularly true in the automobile sector with sophisticated products for which product reviews and reliability figures take time to gather and observe. If consumers are cautious in their purchase of new varieties for these reasons, first-year market shares may be below their eventual equilibrium market share and calculation of the share of net new variety consumption relative to exiting varieties may very well underestimate the true impact of variety change on the marketplace.[27] There are also reasons why market shares of exiting varieties may not be true representations of their equilibrium market share. For example, an automobile manufacturer may simply close down production and sales of an exiting variety in the middle of its final year, leading to a much lower market share than if the variety had been available for the entire year.

As a simple way to examine the sensitivity of our estimates to these considerations, we alternatively calculate price indexes and net variety change measures (corrected λ indexes) using three-year intervals of our data. For example, rather than calculate net new variety and price index changes for each year from 1990 to 1993, we only calculate the three-year change over this period. As a result, new varieties that occur in the early part of the three-year period will have had time to reach their equilibrium market share. Obviously this is not a perfect solution, as new varieties in the last years of the interval may not have had time to reach their equilibrium market share, but it will at least give some indication of the potential bias. This procedure would also miss significant variety change if varieties often turned over in less than three years. This is not an issue with the automobile data as new models almost always remain in the market for at least three years.

Table 9 provides calculations when using three-year intervals of our data for the same scenarios reported in Table 6 using annual intervals to make calculations. Our calculations of price indexes and

[26] We have done separate estimates where we only control for one of these channels of bias at a time and find that the bias from not taking into net variety change from U.S. outsourcing is very small, whereas the bias from not taking into account variety change introduced by foreign affiliates is therefore very large.

[27] It is certainly possible that first-year market shares could be *larger* than equilibrium market shares causing the effects of net variety change to be overestimated. For example, automobile manufacturers may devote much more advertising expenditures to new varieties distorting their “true” market share.

Table 9
Summary table of aggregate end-point results across different data sets and scenarios using three-year intervals.

	Conventional price index	Corrected λ ratio	Exact price index
Comparing estimates between import sample and market sample			
Imported varieties — import sample	1.749	0.991	1.730
Imported varieties — market sample	1.561	0.915	1.428
Market sample — estimates for imported and domestic varieties			
Imported varieties	1.561	0.915	1.428
Domestic varieties	1.687	0.962	1.624
Market sample — estimates for foreign-owned and domestic-owned varieties			
Foreign-owned varieties	1.546	0.876	1.355
Domestic-owned varieties	1.694	0.932	1.579

net variety changes using our market-based data set are virtually identical across the two tables. Fig. 2 repeats the exercise undertaken in Fig. 1, and yields qualitatively similar outcomes. This speaks to the robustness of our results and suggests that first-year new variety market shares and last-year variety market shares are representative of their equilibrium market shares. We also find in the raw data that the average market share of new varieties in their first year is not significantly different from their average market share in subsequent years. On a final note, we also find very little differences in our estimates when we calculate price and variety changes using a simple long-difference over the entire period of our sample.

5. Conclusion

The effects of product variety have been important to international trade economists for decades, yet measuring varieties remains problematic. This paper revisits important recent work estimating the impact of variety change from foreign sources on domestic prices and welfare. Using detailed market-based data on the U.S. automobile sector, we show significant biases resulting from the use of HS product codes to define goods and the Armington assumption used to assign varieties. Compared to our market-based automobile data set, HS codes often lump quite dissimilar products into the same good classification, which biases elasticities of substitution downward. On the other hand, the Armington assumption hides substantial net variety change. On net, our market-based estimates suggest the effect of net variety change from automobile imports on U.S. welfare for the 1990–2006 period was almost twice as large as that estimated by the typical trade (HS) data and Armington-defined varieties. Taking into account net variety change by affiliates of foreign firms (something the import data cannot identify) further increases these welfare effects by about an additional 70%.

While our paper examines the automobile market, we have examined a number of issues that likely affect the estimation of new variety effects for all (differentiated) products. First, HS10 goods codes may not correspond very well with market-based goods definitions, which will then likely bias substitution elasticities downward and

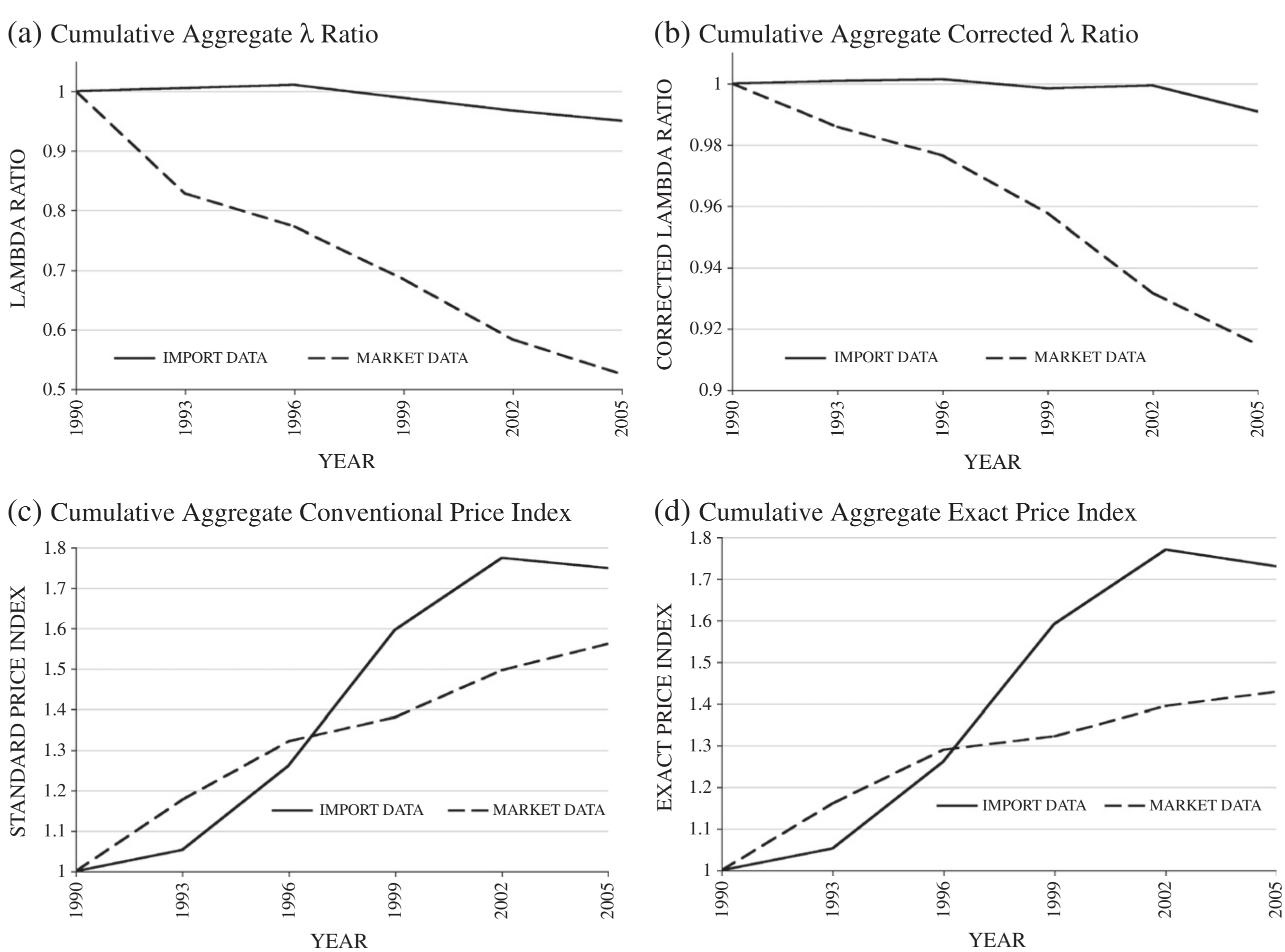

Fig. 2. Comparing aggregate estimates for market and import data in 3 year intervals.

inflate the estimated impact of new variety change. Second, the Armington assumption can hide true variety churning (as measured by market share), but may also artificially inflate churning of varieties measured as simple counts because import data reflect consumer's purchases of products foreign country that is not the location of production. Finally, using only import data ignores the effects of variety change from foreign-affiliate activity that we find to be a significant source of net new variety change in our data. Recent statistics show foreign-affiliate sales in the U.S. to be roughly similar in size to U.S. imports for not only the motor vehicle sector, but also for total manufactured goods.

References

Ackerberg, D., Rysman, M., 2005. Unobserved product differentiation in discrete-choice models: estimating price elasticities and welfare effects. RAND Journal of Economics 36, 771–788.

Anderson, S.P., dePalma, A., Thisse, J.F., 1992. Discrete Choice Theory of Product Differentiation. The MIT Press.

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica 63, 841–890.

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

Broda, C., Weinstein, D., 2006. Globalization and the gains from variety. Quarterly Journal of Economics 121, 541–585.

Feenstra, R.C., 1994. New product varieties and the measurement of international prices. The American Economic Review 84, 157–177.

Feenstra, R.C., 2004. Advanced International Trade : Theory and Evidence. Princeton University Press, Princeton, N.J.

Goldberg, P.K., 1995. Product differentiation and oligopoly in international markets: the case of the U.S. automobile industry. Econometrica 63, 891–951.

Goldberg, P.K., Verboven, F., 2005. Market integration and convergence to the law of one price: evidence from the European car market. Jounal of International Economics 65, 49–73.

Helpman, E., Krugman, P.R., 1985. Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. MIT Press, Cambridge, Mass.

Imbs, J., Mejean, I., 2009. Elasticity Optimism. IMF Working Papers.

Krugman, P.R., 1979. Increasing returns, monopolistic competition, and international trade. Journal of International Economics 9, 469–479.

Petrin, A., 2002. Quantifying the benefits of new products: the case of the minivan. Journal of Political Economy 110, 705–729.

Rauch, J., 1999. Neworks versus markets in international trade. Jounal of International Economics 48, 7–37.

Sato, K., 1976. The ideal log-change index number. Review of Economics and Statistics 58, 223–228.

Sheu, G., 2010. Price, quality, and variety: Measuring the gains from trade in differentiated products. mimeo Harvard University

Soderbery, A., 2010. Investigating the asymptotic properties of import elasticity estimates. Forthcoming in Economics Letters.

Vartia, Y.O., 1976. Ideal log-change index numbers. Scandinavian Journal of Statistics 3, 121–126.

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