# Quantifying the Sources of Firm Heterogeneity *

Colin Hottman
Stephen J. Redding†
David E. Weinstein

November 13, 2015

### Abstract

We develop and structurally estimate a model of heterogeneous multiproduct firms that can be used to decompose the firm-size distribution into the contributions of costs, "appeal" (quality or taste), markups, and product scope. Using Nielsen bar-code data on prices and sales, we find that variation in firm appeal and product scope explains at least four fifths of the variation in firm sales. We show that the imperfect substitutability of products within firms, and the fact that larger firms supply more products than smaller firms, implies that standard productivity measures are highly dependent on implicit demand system assumptions and probably dramatically understate the relative productivity of the largest firms. Although most firms are well approximated by the monopolistic competition benchmark of constant markups, we find that the largest firms that account for most of aggregate sales depart substantially from this benchmark, and exhibit both variable markups and substantial cannibalization effects.

JEL CLASSIFICATION: L11, L21, L25, L60

KEYWORDS: firm heterogeneity, multiproduct firms, cannibalization effects

# I.  Introduction

Why are some firms larger than others? Some companies, such as the Coca Cola Corporation, generate billions of dollars of sales and dominate the markets in which they operate. Other companies account for only a small fraction of the sales of their larger competitors. What explains these vast differences in firm performance? Answering this question is important for quantifying equilibrium models of firm heterogeneity developed in the recent trade and macro literatures and for understanding the relationship between microeconomic firm performance and macroeconomic outcomes.

Recent research on firm heterogeneity in trade and macroeconomics (*e.g.*, Melitz [2003]; Manova and Zhang [2012]; Feenstra [2014]) points to four components of firm heterogeneity: marginal cost, "appeal" (quality or taste), markups and product scope (*i.e.,* the number of products produced by firms).[1] We develop a structural model of heterogeneous multiproduct firms that can be used to decompose the firm-size distribution into the relative contributions of each component. We first use this framework to structurally estimate elasticities of substitution across varieties between and within multiproduct firms. We next implement our model-based decomposition for the around 20,000 firms that supply goods with bar codes in the Nielsen HomeScan Database in a typical quarter. This decomposition uses the structure of the model to isolate different margins in the data without making assumptions about how those margins are related to one another (as in the business cycle decomposition of Chari, Kehoe and McGratten 2007 in the macroeconomics literature). Our framework requires only price and expenditure data and hence is widely applicable. We separate out the contributions of cost and appeal using the exclusion restriction that cost only affects sales through price. In contrast, appeal affects sales conditional on price.

Our results point to demand differences (which could arise from quality or taste variation) as being the principal reason why some firms are successful in the marketplace and others are not. Depending on the specification considered, we find that 50-70 percent of the variance in firm size can be attributed to differences in firm appeal, about 20-25 percent to differences in product scope, and less than 25 percent to cost. When we turn to examine time-series evidence, the results become even more stark. Virtually all firm growth can be attributed to firm appeal with most of the remainder due to product scope. These results suggest that most of what economists call differences in revenue productivity reflects differences in appeal (*e.g.,* quality or taste) rather than cost.

Our framework uses a nested constant elasticity of substitution (CES) utility system that allows the elasticity of substitution between varieties within a firm to differ from the elasticity of substitution between varieties supplied by different firms. Our choice of this CES demand structure is guided by its prominence, tractability, and empirical feasibility. Across international trade, economic geography, and macroeconomics, there is little doubt that this framework is the preferred approach to modeling product variety. Since our approach nests the standard CES-monopolistic competition model as a special case, we can compare our results with those that

---

[1]Much of the existing empirical literature refers to any shifter of demand conditional on price as "quality," as in Shaked and Sutton (1983); Berry (1994); Khandelwal (2010); Broda and Weinstein (2010); Hallak and Schott (2011); and Feenstra and Romalis (2014). We use the term "appeal" to avoid taking a stand as to whether the shift in demand arises from vertical quality differentiation or subjective differences in consumer taste. Empirically, we find that higher appeal products are on average more costly to produce, which is consistent with a quality interpretation.

would be obtained in a standard trade, macro or economic geography model. However, we generalize this standard model to allow firms to supply multiple products (a pervasive feature of our data) and to have market power (since the largest firms in our data are far from being measure zero).

Incorporating these features into our structural model yields a number of additional insights. Our model makes clear a conceptual problem in the estimation of firm productivity that is likely to bias existing estimates. Most productivity estimates rely on the concept of real output, which is calculated by dividing nominal output by a price index. However, the formula for any economically motivated price index, which is the same as a unit expenditure function, is dependent on implicit assumptions about how the output of firms enters utility. *Thus, one cannot move from nominal output to real output without imposing assumptions about the structure of the demand system.*[2] Our results show theoretically and empirically that the CES measure of a multiproduct firm's price is highly sensitive to how differentiated its output is and how many products it supplies. The sensitivity of the CES price index to demand parameters, such as the elasticity of substitution and whether multiproduct firms exist, implies that estimates of real output are equally sensitive to these demand parameters. We show that if demand has a nested CES structure, conventional measures of real output will have a downward bias that rises with firm size with an elasticity of around one third. In other words, real output variation is substantially greater than nominal output variation. This bias also implies that true productivity differences are much larger than conventionally measured ones.

The bias is driven by two features of reality that are typically ignored in most analyses. First, most analyses treat producers as single-product firms so they can avoid complications arising from the challenges of measuring the real output of multiproduct firms. However, our results indicate that multiproduct firms are the norm. For example, we show that 69 percent of firms supply more than one bar code and these firms account for more than 99 percent of output in their sectors.[3] Therefore truly single product firms account for a negligible share of sales in our data. Second, we show that if the output of multiproduct firms is differentiated, the common assumption that total firm output is simply the sum of the output of each good understates real output for multiproduct firms, and the degree of this downward bias rises in the number of products supplied.

Our framework also provides a new metric for quantifying the extent to which a firm's products are differentiated from those of its rivals. If a firm's products are perfectly substitutable with each other but not with those of other firms, then 100 percent of a new product's sales will come from the firm's existing sales, which implies a cannibalization rate of one. However, if a firm's market share is negligible (as it is for most firms) and its products are as differentiated from each other as they are from products supplied by other firms, none of a new product's sales will come at the expense of the firm's other products, which implies a cannibalization rate of zero. We estimate that the cannibalization rate for the typical firm is about 50 percent, indicating that although products supplied by the same firm are more substitutable with each other than with those of other firms, it is not correct to think of them as perfect substitutes.

---

[2]This point was stressed by Aristotle in his *Nicomachean Ethics* (Book V, Section 5): "Demand holds things together as a single unit.... In truth it is impossible that things differing so much should become commensurate, but with reference to demand they may become so." (see http://classics.mit.edu/Aristotle/nicomachaen.5.v.html)

[3]By contrast, U.S. Census of Manufactures data indicates that producers of multiple five-digit Standard Industrial Classification (SIC) products account for 37 percent of firms and 87 percent of shipments (Bernard, Redding and Schott [2010]). The difference comes from defining single product firms as one-industry firms as opposed to one-bar-code firms.

We find that the typical sector is characterized by a few large firms with substantial market shares and a competitive fringe of firms with trivial market shares. Firms from this competitive fringe choose prices close to the benchmark of monopolistic competition, because they have market shares too small to exploit their market power. However, the very largest firms that account for disproportionate shares of sales within product groups have substantially higher markups. This variation in markups is greater under quantity competition than under price competition. In most sectors, the largest firm has a market share above 20 percent, which enables it to charge a markup that is 24 percent higher than that of the median firm under price competition and double that of the median firm under quantity competition. We use the model to undertake counterfactuals, in which we show that these variable markups for the largest firms have quantitatively relevant effects on the firm-size distribution and aggregate consumer price indices. Our counterfactual estimates indicate that the multiple varieties supplied by multiproduct firms reduce the consumer price index by around one third.

The remainder of the paper is structured as follows. Section II. reviews the related literature. Section III. discusses the data. Section IV. introduces the model. Section V. uses the structure of the model to derive moment conditions to estimate elasticities of substitution and undertake our decomposition of firm sales. Section VI. presents our estimation results. Section VII. reports the results of a number of robustness tests. Section VIII. undertakes counterfactuals. Section IX. concludes.

## II. Related Literature

Over the last decade, the fields of international trade and macroeconomics have undergone a transformation as the dissemination of micro datasets and the development of new theories has led to a shift in attention towards firm heterogeneity. Existing research has suggested a number of candidate explanations for differences in firm performance, including differences in production efficiency (*e.g.,* Melitz [2003]), demand in the form of product quality (*e.g.*, Sutton [1991]; Schott [2004]; Khandelwal [2010]; Hallak and Schott [2011]; Johnson [2012]; Di Comite, Thisse and Vandenbussche [2014]; Eslava, Fieler and Xu [2014]; Feenstra and Romalis [2015]), markups (*e.g.,* De Loecker and Warzynski [2012]; De Loecker et al. [2015]), fixed costs (*e.g.*, Das, Roberts and Tybout [2007]), and the ability to supply multiple products (*e.g.*, Arkolakis and Muendler [2010]; Bernard, Redding and Schott [2010, 2011]; Eckel and Neary [2010]; Mayer, Melitz and Ottaviano [2014]; Eckel, et al. [2015]). While existing research typically focuses on one or more of these candidate explanations, they are all likely to operate to some degree in the data. We develop a general theoretical model that incorporates each of these candidate explanations and estimate that model structurally using disaggregated data on prices and sales by firm and product. We use the estimated model to provide evidence on the quantitative importance of each source of firm heterogeneity and the ways in which they interact with one another.

In much of the literature on firm heterogeneity following Melitz (2003), marginal cost and appeal (quality or taste) are isomorphic. Under the assumption of CES preferences and monopolistic competition, cost and appeal enter equilibrium firm revenue in exactly the same way. However, these different sources of firm heterogeneity have different implications for firm revenue conditional on prices (*e.g.*, Berry [1994]; Khandelwal [2010]). While marginal cost affects firm revenue *through* prices, appeal affects firm revenue *conditional* on prices.

Thus, if two firms charge the same price, but one firm has a more appealing (*e.g.,* higher quality) product than the other, that firm will generate more sales. An advantage of our approach is that we observe prices and sales in our data, and hence we are able to separate out cost and appeal as sources of dispersion in firm sales.[4]

Most of the existing research on firm heterogeneity in trade and macroeconomics has assumed that firms are measure zero and compete under conditions of monopolistic competition (*e.g.,* Melitz [2003] and Melitz and Ottaviano [2008]). In contrast, a small number of papers have allowed firms to be large relative to the markets in which they operate (*e.g.,* Atkeson and Burstein [2008]; Amiti, Itskhoki and Konings [2014]; and Edmond, Midrigan and Xu [2015]). When firms internalize the effects of their decisions on market aggregates, they behave systematically differently from measure zero firms. Even under CES demand, firms charge variable mark-ups, because each firm internalizes the effects of its pricing decisions on market price indices and these effects are greater for larger firms.[5] Furthermore, since firms are of positive measure, idiosyncratic shocks to these "granular" firms can affect aggregate outcomes, as in Gabaix (2011) and di Giovanni and Levchenko (2012). In contrast to these papers, we structurally estimate a model of heterogeneous firms, and show how it can be used to recover the determinants of firm sales dispersion and undertake counterfactuals.

To the extent that such large firms supply multiple products, they take into account the effect of introducing new varieties on the sales of existing varieties. Most of the existing theoretical research on multiproduct firms in trade and macroeconomics has abstracted from these cannibalization effects by again assuming measure zero firms (*e.g.*, Allanson and Montagna [2005]; Agur [2010]; Arkolakis and Muendler [2010]; Bernard, Redding and Schott [2010, 2011]; Mayer, Melitz and Ottaviano [2014]; and Nocke and Yeaple [2014]). Important exceptions that explore cannibalization effects theoretically are Feenstra and Ma (2008); Eckel and Neary (2010); and Dhingra (2013). In contrast to these theoretical studies, we develop a structural model that can be used to provide quantitative evidence on how important it is to introduce such cannibalization effects into models of firm behavior.

Most existing empirical research on multiproduct firms has measured products using production classification codes (*e.g.*, around 1,500 five-digit Standard Industrial Classification (SIC) categories in Bernard, Redding and Schott [2010]) or trade classification codes (*e.g.*, around 10,000 Harmonized System codes in Bernard, Jensen and Schott [2009]). In contrast, we measure products at a much finer level of resolution using what we term "bar codes"—either 12-digit Universal Product Codes (UPCs) or 13-digit European Article Numbers (EANs)—in scanner data.[6] This measure corresponds closely to the level at which product choice decisions are made by firms, because it is rare for an observable change in product attributes to occur without the introduction of a new bar code.

Our econometric approach builds on the literature estimating elasticities of substitution and quantifying the contribution of new varieties to welfare following Feenstra (1994) and Broda and Weinstein (2006, 2010). We

---

[4]Our CES formulation of market demand can be derived from a discrete choice model of the demands of individual consumers, as shown in Anderson, de Palma and Thisse (1992).

[5]Our model generates variable mark-ups without a "choke price" above which demand is zero. Therefore this model lies outside the classes considered by Arkolakis, Costinot and Rodriguez-Clare (2012) and Arkolakis, Costinot, Donaldson and Rodriguez-Clare (2015), in which aggregate statistics such as the trade share and trade elasticity are sufficient statistics for welfare.

[6]Recently, the 12-digit UPCs have been upgraded to 13-digit EAN-13s (European Article Numbers). The extra digit of the EAN-13 allows for more products and firms.

extend this estimation approach to allow firms to be of positive measure relative to the market and to supply multiple products, which introduces variable markups and cannibalization effects. We show how this extended approach can be used to recover demand heterogeneity, marginal cost heterogeneity, variable markups, and cannibalization effects from the data. More generally, although other studies have used scanner data such as Chevalier et al. (2003), so far these data have not been used to estimate a structural model of heterogeneous multiproduct firms and quantify the sources of dispersion in firm sales.

## III. Data

Our data source is the Nielsen HomeScan database which enables us to observe price and sales information for millions of products with a bar code.[7] Bar-code data have a number of advantages for the purpose of our analysis. First, since bar codes are inexpensive but provide sellers access to stores with scanners as well as internet sales, producers have a strong incentive to purchase bar codes for all products that have more than a trivial amount of sales.[8] This feature of the data means that it is likely we observe all products supplied by firms. Second, since assigning more than one product to a single bar code can interfere with a store's inventory system and pricing policy, firms have a strong incentive not to reuse bar codes. This second feature of the data ensures that identical goods do not have different bar codes.[9] Thus, a bar code is the closest thing we have empirically to the theoretical concept of a good. Finally, since the cutoff size for a firm is to make a sale rather than an arbitrary number of workers, we actually observe something close to the full distribution of firms. We match bar codes to firms using the first few digits of the bar code, which identify the company that owns the brand, for both domestically-produced and imported bar codes, and for both domestic and foreign companies.[10]

Nielsen collects its bar-code data by providing handheld scanners to on average 55,000 households per year to scan each good purchased that has a bar code.[11] Prices are either downloaded from the store in which the good was purchased, or they are hand entered, and the household records any deals used that may affect the price. These households represent a demographically balanced sample of households in 42 cities in the United States. Overall, the database covers around 30 percent of all expenditure on goods in the CPI.[12] We collapse the household dimension in the data and collapse the weekly purchase frequency to construct a national quarterly database by bar code on the total value sold, total quantity sold, and average price.[13]

---

[7]Our results are calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business. Information on availability and access to the data is available at http://research.chicagobooth.edu/nielsen

[8]GS1 provides a company with up to 10 bar codes for a $250 initial membership fee and a $50 annual fee. There are deep discounts in the per bar code cost for firms purchasing larger numbers of them (see http://www.gs1us.org/get-started/im-new-to-gs1-us)

[9]For example, Broda and Weinstein (2010) compared bar codes with their product descriptions and found "no two identical UPCs with different product descriptions, or two different UPCs with the same product description." [p. 695]

[10]Firm prefixes are usually seven digits but they can be as long as eleven digits, so we use GS1 data to map the bar codes into firm identifiers. We could not obtain a firm identifier for just under 5 percent of the bar codes, and so we dropped these products. As an example of a firm identifier, for an imported spark plug, GS1 identifies the (potentially foreign) company that owns the brand (*e.g.*, Bosch). If a bar code is made under license (*e.g.*, Whole Foods contracts with the Coca Cola Corporation to manufacture a 12-ounce bottle of Cola that is marketed under the Whole Foods brand), GS1 again identifies the company that owns the brand (in this example Whole Foods).

[11]The data for 2004 through 2006 come from a sample of 40,000 households, while the data for 2007 through 2011 come from a sample of 60,000 households.

[12]For further discussion of the Nielsen data, see Broda and Weinstein (2010).

[13]One question that naturally arises is whether it is appropriate to think of US firms competing in a national market or in different

Defining products as goods with bar codes has a number of advantages over defining goods by industry classifications. While industry classifications aggregate products produced by a firm within an industry, our data reveals that large firms typically sell hundreds of different products within even a narrowly-defined sector. In other words, while prior work equates multiproduct firms with multi-industry firms, our data hews extremely closely to what an economist would call a multiproduct firm. In principle, it could be appropriate to aggregate all output of a firm within an industry into a single good (Is whole milk the same as skim milk? Is a six-pack of soda the same as a two-liter bottle?). However, as we will show theoretically, the validity of this procedure depends on the elasticity of substitution between the products supplied by a firm, which is an empirical question.

Instead of relying on product data for a single industry, we observe virtually the entire universe of goods purchased by households in the sectors that we examine. Our database covers approximately 1.6 million goods purchased at some point by households in our sample. The data were weighted by Nielsen to correct for sampling error. For example, if the response rate for a particular demographic category is low relative to the census, Nielsen re-weights the averages so that the price paid and the quantity purchased is representative of the United States as a whole.[14]

Nielsen organizes the bar codes into product groups according to where they would likely be stocked in a store. The five largest of our 98 product groups are carbonated beverages, pet food, paper products, bread and baked goods, and tobacco. We report a full list of the product groups in the online supplement. Output units are common within a product group: typically volume, weight, area, length, or counts. Importantly, we deflate by the number of units in the bar code, so prices are expressed in price per unit (*e.g.,* price per ounce). When the units are in counts, we also deflate by the number of goods in a multipack, so, for instance, we would measure price per battery for batteries sold in multipacks. While about two thirds of these bar-coded items correspond to food items, the data also contains significant amounts of information about non-food items like medications, housewares, detergents and electronics.

Table I presents descriptive statistics for our sample of firms and bar codes in the 98 product groups. We weight the data by the sales of the product group in each quarter and average across quarters because sectors like carbonated beverages are much larger economically than sectors like "feminine hygiene." There are on average 512 firms in each product group with 90 percent of the product groups having more than 200 firms. We see enormous range in firm product scope. The median number of products supplied by a firm is 4 and the average is 13. On average, around 440,000 different UPCs were sold each quarter.

One of the most striking facts displayed in this table is the degree of firm heterogeneity. This is manifest in the skewness of the size and bar-code distributions. The largest firm in an industry typically sells 2,500 times

---

regional markets. Handbury and Weinstein (2015) argue there is not that much heterogeneity in the price levels of bar-coded goods across cities. Similarly, we find that national and regional expenditure shares and prices are strongly correlated in the data. National brands and firms (brands and firms that sell in 10 or more cities) account for 97 and 99 percent of expenditure respectively in cities with more than 1,000 households in the AC Nielsen data. Regressing the level of prices in these cities on national prices, quarter-year dummies, and bar-code fixed effects, we find a partial $R^2$ (*i.e.,* an $R^2$ not inclusive of the fixed effects) of 0.98. These results suggest that most of the variation in U.S. consumer prices reflects national shocks not local ones.

[14]Although the bar-code data include mostly final-consumption products, available at retail stores, Section S9 of the online supplement reports the results of a robustness test in which we replicate our empirical procedure using Chilean international trade transactions data (which includes both final-consumption and intermediate products). We find a similar pattern of results using these international trade transactions data.

more than the median firm. We see similar patterns in terms of product scope and sales per product. The firm with the most products typically has 97 times more products than the firm with the median number of bar codes, and the bar code with the most sales on average generates almost 900 times more revenue than the revenue of the median bar code.

We see in Table II that almost 90 percent of sales in a product group was produced by firms with sales in the top decile of sales. Table III provides a more detailed description of this firm heterogeneity by focusing on the ten largest firms in each product group (where we weight the averages by the sales of the product group). Table III reveals an almost fractal nature of firm sales. Around two-thirds of all of the sales of firms in the top decile is produced by the ten largest firms (which on average only account for 1.9 percent of firms in this decile). While on average half of all output in a product group is produced by just five firms, 98 percent of firms have market shares of less than 2 percent. Thus, the typical sector is characterized by a few large firms and a competitive fringe comprised of firms with trivial market shares. A second striking feature of the data is that even the largest firms are not close to being monopolists. The largest firm in a product group on average only has a market share of 22 percent. Finally, the data reveal that firms in the top decile of sales are all multiproduct firms, supplying on average 68 different goods with the largest firms supplying hundreds of goods.

The extent of multiproduct firms can be seen more clearly in Table IV, which shows the results of splitting the data by the number of UPCs supplied by a firm. While single-product firms constitute about one third of all firms on average, these firms account for less than one percent of all output. In other words virtually all output is supplied by multiproduct firms. Moreover, the fact that over ninety percent of output is sold by firms selling eleven or more varieties and nearly two-thirds of all output is supplied by firms selling more than fifty varieties suggests that single product firms are more the exception than the rule.

Large firms not only sell more products but they also sell a lot more of each product. The penultimate column of Table IV documents that while the typical bar code sold by a single-product firm only brings in $63,871 in revenue, the typical bar code sold by a firm selling over one hundred bar codes brings almost twice as much ($122,045). In other words, large firms not only supply more products but they sell more of each product. If firms differed only in the fixed cost of adding new varieties, one would not expect to see large firms sell more of each variety. The fact that they do strongly suggests that large firms must also differ in the marginal cost or demand for their output.

Finally, although the largest firms have non-trivial shares of particular product groups, these firms are small compared to the U.S. economy. Ninety-nine percent of firms in our sample have aggregate market shares (across all product groups) of less than 0.1 percent of total bar-code sales. Even the largest firm only sells 3 percent of total bar-code sales in our sample. Given that sales of packaged goods is only a fraction of total U.S. sales in all sectors, it is reasonable to conclude that no individual firm has the capacity to affect aggregate U.S. prices, expenditure or welfare.

In sum, our overview of the data reveals some key features that we model in our empirical exercise. First, the vast majority of firms have trivial market shares, which means that if we believe that all firms may have some market power, we need to work with demand systems that do not imply that firms with trivial market shares have trivial markups. Second, the fact that most economic output is produced by multiproduct firms

impels us to build this feature directly into the estimation system and allow for both differences in the fixed cost of developing new products as well as differences in marginal cost and demand across products. Third, the fact that there are firms with non-trivial market shares implies that at least some firms are likely to have market power. To the extent that these large firms internalize the effects of their price choices on market aggregates, this concentration of market shares will induce departures from the monopolistic competition benchmark. Finally, the fact that all the firms in our sample are small compared to the U.S. economy means that it is reasonable to assume that firms do not consider the implications of their pricing behavior on the aggregate economy.

## IV.    Theoretical Framework

Our choice of functional forms is motivated not only by the data issues we identified above but also by some theoretical concerns. Since much of the theoretical literature in international trade and economic geography has worked with CES models, we want our results to nest this case (at least within product groups), so that our results easily can be compared with existing work in trade, macro and regional economics. However, we also need a framework that allows the elasticity of substitution for products supplied by the same firm to be different than that for products supplied by different firms, thereby to allow for the possibility of cannibalization effects. Finally, we need a setup that can be applied to firm-level data without imposing implausible assumptions or empirical predictions. This last requirement rules out two common demand systems: linear demand and the symmetric translog. A linear demand system would be problematic in our setting because one needs to impose the assumption that income elasticities are equal to zero and the estimation of marginal costs derived from a linear demand system can often result in negative values. While the symmetric translog demand system improves on the linear demand system in this regard, it is a difficult system to implement at the firm level because it has the undesirable result that firms with negligible market shares have negligible markups.

Our estimation strategy therefore is based on an upper level Cobb-Douglas demand system across product groups with CES nests below it. The upper-level Cobb-Douglas assumption implies that no firm has an incentive to try to manipulate prices in one product group to influence behavior in another product group. The reason is that each firm is assumed to be small relative to the aggregate economy (and hence cannot affect aggregate expenditure) and product group expenditure shares are determined by parameters alone. Therefore the firm problem becomes separable by product group.[15] However, the nested CES structure within broadly defined "product groups" allows for strategic interactions among firms supplying similar products. Within this structure the real consumption of each product group is composed of the real consumption of each firm's output, which itself is made up of the consumption of each of the varieties supplied by the firm.

---

[15] In Section S5 of the online supplement, we show how our approach can be extended to allow for a CES upper tier of utility, in which case firms have an incentive to price strategically across product groups. We focus on the Cobb-Douglas case in our baseline specification, because we find that product group expenditure shares are relatively constant over time despite changes in product group price indices, which suggests that this specification provides a reasonable approximation to the data.

## IV.A.   Demand

In order to implement this approach, we assume that utility, $\mathbb{U}_t$, at time $t$ is a Cobb-Douglas aggregate of real consumption , $C_{gt}^G$, of a continuum of product groups:[16]

$$\ln \mathbb{U}_t = \int_{g \in \Omega^G} \varphi_{gt}^G \ln C_{gt}^G dg, \qquad \int_{g \in \Omega^G} \varphi_{gt}^G dg = 1,$$

where $g$ denotes each product group; $\varphi_{gt}^G$ is the share of expenditure on product group $g$ at time $t$; and $\Omega^G$ is the set of product groups. Within product groups, we assume two CES nests for firms and UPCs. The first nest for firms enables us to connect with the existing literature on measuring firm productivity. The second nest for UPCs enables us to incorporate multi-product firms. Therefore the consumption indices for product groups $(C_{gt}^G)$ and firms $(C_{fgt}^F)$ can be written as respectively:

$$C_{gt}^G = \left[ \sum_{f \in \Omega_{gt}^F} \left( \varphi_{fgt}^F C_{fgt}^F \right)^{\frac{\sigma_g^F - 1}{\sigma_g^F}} \right]^{\frac{\sigma_g^F}{\sigma_g^F - 1}}, \qquad C_{fgt}^F = \left[ \sum_{u \in \Omega_{fgt}^U} \left( \varphi_{ut}^U C_{ut}^U \right)^{\frac{\sigma_g^U - 1}{\sigma_g^U}} \right]^{\frac{\sigma_g^U}{\sigma_g^U - 1}}, \qquad (1)$$

In other words, the real consumption in any product group, $g$, is a function of the consumption of each firm $f$'s output, $C_{fgt}^F$, weighted by the consumer appeal of that firm's output, $\varphi_{fgt}^F > 0$, and adjusted for the substitutability of the output of each firm, $\sigma_g^F > 1$, where the set of active firms is $\Omega_{gt}^F$.[17] Similarly, the sub-utility derived from the consumption of a firm $f$'s output within product group $g$, $C_{fgt}^F$, is a function of the consumption of each UPC (i.e. bar code) $u$, $C_{ut}^U$, multiplied by the consumer appeal of that bar code's output, $\varphi_{ut}^U > 0$, and adjusted by the substitutability between the various UPCs supplied by the firm, $\sigma_g^U$, where the set of these UPCs is $\Omega_{fgt}^U$.[18]

There are a few features of this specification that are worth noting. First, if the elasticity of substitution across varieties supplied by a firm, $\sigma_g^U$, is finite, then the real output of a multiproduct firm is not equal to the sum of the outputs of each product. For example, if the only reason firms differ in size is that larger firms supply more varieties than smaller firms, then assuming firm real output is the sum of the output of each variety will tend to understate the relative size of larger firms. This size bias is a topic that we will explore in much more detail later. Second, for much of what follows, we are focused on the sales decomposition within a given product group ($g$), so for notational simplicity we can suppress the $g$ subscript on $\sigma_g^F$ and $\sigma_g^U$ until we need it again to pool results across sectors in Section IV.F.. Third, we would expect (but do not impose) that

---

[16]We use the superscripts $G$, $F$ and $U$ to denote product group variables, firm variables, and UPC variables respectively. Time is indexed by the subscript $t$. We use the subscripts $g$, $f$ and $u$ to index individual product groups, firms and UPCs respectively, which belong to the sets $\Omega^G$, $\Omega_{gt}^F$ and $\Omega_{fgt}^U$ respectively, where the set of product groups $\Omega^G$ is constant over time. For example, $C_{gt}^G$ is the product group consumption index for product group $g \in \Omega^G$, and $\sigma_g^U$ is the elasticity of substitution across UPCs within product group $g \in \Omega^G$.

[17]We allow firms to be active in multiple product groups and firm appeal ($\varphi_{fgt}^F$) to vary across product groups $g$. Empirically, we find firm appeal to be strongly positively correlated across product groups within firms.

[18]Our definition of appeal is utility per common physical unit (*e.g.*, utility per ounce of a firm's output). However, variation in appeal could be interpreted either as a difference in utility per physical unit or as variation in the number of identical unobservable sub-units within a physical unit. For example, it is isomorphic to say that Firm A produces products with twice the utility per ounce as Firm B and to say that 1/2 an ounce of Firm A's product generates as much utility as an ounce of Firm B's product. In the latter case, an ounce of Firm A's product would contain two "1/2 ounce sub-units". We think our interpretation of utility per physical unit is the most natural for our data.

the elasticity of substitution across varieties is larger within firms than across firms, $i.e.$, $\sigma^U \geq \sigma^F$. When the two elasticities are equal, our system will collapse to a standard CES at the product-group level, and if the inequality is strict, we will show that our setup features cannibalization effects. While our nested CES specification provides a parsimonious and natural approach to modeling multi-product firms, we consider the robustness of our results to alternative specifications in Section VII..

Fourth, we allow firms to be large relative to product groups (and hence internalize their effects on the consumption and price index for the product group). But we assume a continuum of product groups so that each firm is of measure zero relative to the economy as a whole (and hence takes aggregate expenditure $E_t$ as given). When we take the model to the data, we approximate the continuum of product groups in the model with a large number of product groups in the data, in which each firm is small relative to aggregate expenditure, as discussed in Section III. above. Finally, since the utility function is homogeneous of degree one in firm appeal it is impossible to define firm appeal, $\varphi^F_{fgt}$, independently of the product appeal of the varieties produced by that firm, $\varphi^U_{ut}$. We therefore need to choose a normalization. It will prove convenient to normalize the geometric means of the $\varphi^F_{fgt}$ for each product group and the $\varphi^U_{ut}$ for each firm to equal one:

$$\tilde{\varphi}^F_{gt} = \left( \prod_{f \in \Omega^F_{gt}} \varphi^F_{fgt} \right)^{\frac{1}{N^F_{gt}}} = 1, \qquad \tilde{\varphi}^U_{fgt} = \left( \prod_{u \in \Omega^U_{fgt}} \varphi^U_{ut} \right)^{\frac{1}{N^U_{fgt}}} = 1, \qquad (2)$$

where $N^F_{gt}$ is the number of firms in product group $g$ at time $t$ (the number of elements in $\Omega^F_{gt}$) and $N^U_{fgt}$ is the number of UPCs supplied by firm $f$ within product group $g$ at time $t$ (the number of elements in $\Omega^U_{fgt}$). Under this normalization, firm appeal ($\varphi^F_{fgt}$) affects the sales of all products supplied by firm $f$ within product group $g$ *proportionately*, while product appeal ($\varphi^U_{ut}$) determines the *relative* sales of individual products $u$ within firm $f$ and product group $g$.[19]

We can gain some intuition for this framework by using the product group "carbonated beverages" as an example. Aggregate utility depends on the expenditure share (given by $\varphi^G_{gt}$) and amount of consumption of goods in the carbonated-beverages product group, (given by $C^G_{gt}$). The utility derived from the consumption of carbonated beverages depends on the appeal of Coke versus Perrier ($\varphi^F_{fgt}$), the amounts of each firm's real output consumed ($C^F_{fgt}$), and the degree of substitutability between Coke and Perrier ($\sigma^F$). Finally, the real amount of Coke or Perrier consumed ($C^F_{fgt}$) depends on the number of different types of soda produced by each company ($N^U_{fgt}$), the demand for each of these types of soda ($\varphi^U_{ut}$), the consumption of each variety of soda ($C^U_{ut}$), and how similar varieties of Coke (or Perrier) products are with other varieties offered by the same company ($\sigma^U$).

The corresponding exact price indices for consumption are:

$$P_{gt}^G = \left[ \sum_{f \in \Omega_{gt}^F} \left( \frac{P_{fgt}^F}{\varphi_{fgt}^F} \right)^{1-\sigma^F} \right]^{\frac{1}{1-\sigma^F}}, \qquad P_{fgt}^F = \left[ \sum_{u \in \Omega_{fgt}^U} \left( \frac{P_{ut}^U}{\varphi_{ut}^U} \right)^{1-\sigma^U} \right]^{\frac{1}{1-\sigma^U}}. \tag{3}$$

where $P_{gt}^G$ is the product group price index for product group $g$ at time $t$; $P_{fgt}^F$ is the firm price index for firm $f$ and product group $g$ at time $t$; and $P_{ut}^U$ is the price of UPC $u$ at time $t$.

Using the properties of CES demand, the expenditure share of firm $f$ within product group $g$ ($S_{fgt}^F$ for $f \in \Omega_{gt}^F$) equals the elasticity of the product group price index with respect to the price index for firm $f$. Similarly, the expenditure share of product $u$ within firm $f$ ($S_{ut}^U$ for $u \in \Omega_{fgt}^U$) equals the elasticity of the firm price index with respect to the price of product $u$:

$$S_{fgt}^F = \frac{\left( P_{fgt}^F/\varphi_{fgt}^F \right)^{1-\sigma^F}}{\sum_{k \in \Omega_{gt}^F} \left( P_{kgt}^F/\varphi_{kgt}^F \right)^{1-\sigma^F}} = \frac{\partial P_{gt}^G}{\partial P_{fgt}^F} \frac{P_{fgt}^F}{P_{gt}^G}, \qquad S_{ut}^U = \frac{\left( P_{ut}^U/\varphi_{ut}^U \right)^{1-\sigma^U}}{\sum_{k \in \Omega_{fgt}^U} \left( P_{ut}^U/\varphi_{ut}^U \right)^{1-\sigma^U}} = \frac{\partial P_{fgt}^F}{\partial P_{ut}^U} \frac{P_{ut}^U}{P_{fgt}^F}, \tag{4}$$

Equation (4) makes clear exactly how we conceive of firm and product appeal in this setup. Holding fixed prices, a UPC with more appeal has a larger market share. Similarly, holding fixed prices, firms that supply goods with more appeal have larger market shares.

The role of both firm and product appeal also can be seen by writing down the demand for the output of UPC $u$ supplied by firm $f$ within product group $g$:

$$C_{ut}^U = \left( \varphi_{fgt}^F \right)^{\sigma^F - 1} \left( \varphi_{ut}^U \right)^{\sigma^U - 1} E_{gt}^G \left( P_{gt}^G \right)^{\sigma^F - 1} \left( P_{fgt}^F \right)^{\sigma^U - \sigma^F} \left( P_{ut}^U \right)^{-\sigma^U}, \tag{5}$$

where $E_{gt}^G$ denotes product group expenditure for product group $g \in \Omega^G$. Equation (5) is critical in determining how we can use this framework for understanding the different roles played by cost and appeal for the sales of a firm. While appeal affects consumption independently of price, cost only affects consumption through price.

## IV.B. Technology

We allow the costs of supplying products to the market to vary across UPCs and firms. This specification encompasses both heterogeneity in productivity across firms (as in Melitz [2003]) and heterogeneity in productivity within firms (as in Bernard, Redding and Schott [2011]). All costs are incurred in terms of a composite factor input that is chosen as our numéraire. We assume that the variable cost function is separable across UPCs and that supplying $Y_{ut}^U$ units of output of UPC $u$ incurs a total variable cost of $A_{ut} \left( Y_{ut}^U \right) = a_{ut} \left( Y_{ut}^U \right)^{1+\delta_g}$, where $a_{ut}$ is a cost shifter and $\delta_g$ parameterizes the elasticity of marginal costs with respect to output. In addition, each firm $f$ faces a fixed market entry cost for each product group of $H_{gt}^F > 0$ (e.g., the fixed costs of supplying the market) and a fixed market entry cost for each UPC supplied of $H_{gt}^U > 0$ (e.g., the fixed costs of product development and distribution). We allow for entry and exit of both firms and products, where for the data to be an equilibrium of the model, it must be the case that no firm can profitably enter or exit and no product can be profitably added or dropped. To simplify notation, we again suppress the product group subscript $g$ on $\delta_g$ until Section IV.F..

## IV.C.  Profit Maximization

In our baseline specification, we assume that firms choose prices under Bertrand competition, though we also report results in which firms instead choose quantities under Cournot competition. Since each firm is small relative to the aggregate economy and the upper tier of utility is Cobb-Douglas, the firm's problem is separable by product group.[20] With CES preferences within product groups, the decisions of any one firm only affect the decisions of other firms within that product group through the product group price index ($P_{gt}^G$). Each firm $f$ within product group $g$ chooses its set of UPCs $u \in \{\underline{u}_{fgt}, \ldots, \overline{u}_{fgt}\}$ and their prices $\{P_{ut}^U\}$ for the final consumer to maximize its profits, taking into account these effects on the product group price index:[21]

$$\max_{\{\underline{u}_{fgt},\ldots,\overline{u}_{fgt}\},\{P_{ut}^U\}} \Pi_{fgt}^F = \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \left[ P_{kt}^U Y_{kt}^U - A_{kt}\left(Y_{kt}^U\right) \right] - N_{fgt}^U H_{gt}^U - H_{gt}^F, \tag{6}$$

subject to the constraint that in equilibrium output of each UPC equals consumption ($Y_{kt}^U = C_{kt}^U$, where $C_{kt}^U$ is determined by (5)); we index the UPCs supplied by the firm within the product group from the largest ($\underline{u}_{fgt}$) to the smallest ($\overline{u}_{fgt}$) in sales; and the total number of goods supplied by the firm within the product group is denoted by $N_{fgt}^U$, where $\overline{u}_{fgt} = \underline{u}_{fgt} + N_{fgt}^U$.

Multiproduct firms that are large relative to the market internalize the effects of their decisions for any one variety on the sales of their other varieties. From the first-order conditions for profit maximization, we can derive the firm markup for each UPC, as shown in Appendix AA.:

$$\mu_{fgt}^F = \frac{\varepsilon_{fgt}^F}{\varepsilon_{fgt}^F - 1}, \tag{7}$$

where we define the firm's perceived elasticity of demand as

$$\varepsilon_{fgt}^F = \sigma^F - \left(\sigma^F - 1\right) S_{fgt}^F = \sigma^F \left(1 - S_{fgt}^F\right) + S_{fgt}^F, \tag{8}$$

and the firm's pricing rule as

$$P_{ut}^U = \mu_{fgt}^F \gamma_{ut}, \qquad \gamma_{ut} = (1 + \delta) a_{ut} \left(Y_{ut}^U\right)^\delta, \tag{9}$$

where $\gamma_{ut}$ denotes marginal cost.

One of the surprising features of this setup is that *markups only vary at the firm level within product groups*. This is a generic property of nested demand systems, including nested translog, as shown in Sections S2 and S3 of the online supplement. The intuition is that the firm internalizes that it is the monopoly supplier of its real output within the product group, which in our model equals real consumption of the firm's bundle of goods, $C_{fgt}^F$. Hence its profit maximization problem can be thought of in two stages. First, the firm chooses the price index ($P_{fgt}^F$) to maximize the profits from supplying real consumption ($C_{fgt}^F$), which implies a markup at the

firm level within product groups over the cost of supplying real output. Second, the firm chooses the price of each UPC to minimize the cost of supplying real output within the product group ($C_{fgt}^F$), which requires setting the relative prices of UPCs equal to their relative marginal costs. Together these two results ensure the same markup across all UPCs supplied by the firm within a given product group. Nonetheless, markups vary *across product groups* within firms ($\mu_{fgt}^F \neq \mu_{fmt}^F$ for $g \neq m$).[22]

Although consumers have constant elasticity of substitution preferences ($\sigma^F$), each firm internalizes the effect of its pricing decisions on market price indices, and hence perceives a variable elasticity of demand for UPC $u$ ($\varepsilon_{fgt}^F$) that is decreasing in the expenditure share of the firm $f$ supplying that UPC within product group $g$ ($S_{fgt}^F$), as in Atkeson and Burstein (2008) and Edmond, Midrigan and Xu (2015). As a result, the firm's equilibrium pricing rule (9) involves a variable markup ($\mu_{fgt}^F$) that is increasing in its expenditure share within the product group ($S_{fgt}^F$). For a positive equilibrium price (9), we require that the perceived elasticity of demand ($\varepsilon_{fgt}^F$) is greater than one (firms produce substitutes), which requires that the elasticity of substitution between firms ($\sigma^F$) is sufficiently large. As a firm's sales become small relative to the product group ($S_{fgt}^F \to 0$), the mark-up (7) collapses to the standard constant mark-up of price over marginal cost under monopolistic competition and nested CES preferences (see for example Allanson and Montagna [2005] and Arkolakis and Muendler [2010]).

Using the above equilibrium pricing rule, profits for UPC $u$ supplied by firm $f$ ($\Pi_{ut}^U$) are equal to variable profits ($\pi_{ut}^U \left( N_{fgt}^U \right)$) minus fixed costs. UPC variable profits in turn can be written in terms of UPC revenues ($P_{ut}^U Y_{ut}^U$), the markup ($\mu_{fgt}^F$), and the elasticity of costs with respect to output ($\zeta_{ut}$):

$$\Pi_{ut}^U = \pi_{ut}^U \left( N_{fgt}^U \right) - H_{gt}^U, \tag{10}$$

$$\pi_{ut}^U \left( N_{fgt}^U \right) = P_{ut}^U Y_{ut}^U - A_{ut} \left( Y_{ut}^U \right) = \left( \frac{\zeta_{ut} \mu_{fgt}^F - 1}{\zeta_{ut} \mu_{fgt}^F} \right) P_{ut}^U Y_{ut}^U, \text{ where } \zeta_{ut} = \frac{dA_{ut} \left( Y_{ut}^U \right)}{dY_{ut}^U} \frac{Y_{ut}^U}{A_{ut} \left( Y_{ut}^U \right)} = 1 + \delta. \tag{11}$$

That is, $\pi_{ut}^U \left( N_{fgt}^U \right)$ denotes the variable profits from UPC $u$ when firm $f$ supplies $N_{fgt}^U$ UPCs within product group $g$.

## IV.D. Cannibalization Effects

The number of UPCs supplied by each firm $f$ within product group $g$, $N_{fgt}^U$, is determined by the requirement that the increase in profits from introducing an additional UPC, $\overline{u}_{fgt} + 1$, minus the reduction in profits from reduced sales of existing UPCs $u \in \{ \underline{u}_{fgt}, \ldots, \overline{u}_{fgt} \}$ is less than the fixed cost of introducing the new UPC, $H_{gt}^U$. If a firm supplies $N_{fgt}^U$ products in equilibrium, then it must be the case that if it were to introduce a new

---

[22]Section S4 of the online supplement shows how the model can be extended to incorporate another CES nest (product groups, firms, brands and UPCs). Under the assumption that prices are chosen separately by brand, this extension can generate variation in markups across brands within a given firm and product group.

good, its profits would fall, *i.e.*,

$$\sum_{u=\underline{u}_{fgt}}^{\overline{u}_{fgt}+1} \pi_{ut}^U \left( N_{fgt}^U + 1 \right) - \left( N_{fgt}^U + 1 \right) H_{gt}^U \quad < \quad \sum_{u=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \pi_{ut}^U \left( N_{fgt}^U \right) - N_{fgt}^U H_{gt}^U$$

$$\Leftrightarrow \pi_{\overline{u}_{fgt}+1,t}^U \left( N_{fgt}^U + 1 \right) - \sum_{u=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \left\{ \pi_{ut}^U \left( N_{fgt}^U \right) - \pi_{ut}^U \left( N_{fgt}^U + 1 \right) \right\} \quad < \quad H_{gt}^U. \tag{12}$$

In the case where the number of UPCs is large and can be approximated by a continuous variable, we obtain after some manipulation (see Appendix AB.) the following measure of cannibalization:

$$-\frac{\partial Y_{ut}^U}{\partial N_{fgt}^U} \frac{N_{fgt}^U}{Y_{ut}^U} \quad = \quad \left[ \left( \frac{\sigma^U - \sigma^F}{\sigma^U - 1} \right) + \left( \frac{\sigma^F - 1}{\sigma^U - 1} \right) S_{fgt}^F \right] S_{N_{fgt}^U}^U N_{fgt}^U > 0, \; for \; \sigma^U \geq \sigma^F > 1. \tag{13}$$

where $S_{N_{fgt}^U}^U$ is the share of the new UPC in expenditure on the firm within the product group and $S_{fgt}^F$ is the share of the firm in expenditure within the product group. This measure of cannibalization is the partial elasticity of the sales of existing products with respect to the number of products. This partial elasticity captures the direct effect of the introduction of a new product on the sales of existing products, through the firm and product group price indices, holding constant the prices and marginal costs of these existing products. This partial elasticity with respect to number of products is expressed in terms of observed expenditure shares, and takes the same value under both price and quantity competition, as shown in the appendix.

The first term on the right hand-side captures cannibalization within firms: The introduction of a UPC reduces the firm price index ($P_{fgt}^F$), which reduces the revenue of existing UPCs if varieties are more substitutable within firms than across firms ($\sigma^U > \sigma^F$). The second term captures cannibalization across firms: The introduction of the new UPC reduces the product-group price index ($P_{gt}^G$), which reduces the revenue of existing UPCs if varieties are more substitutable within product-groups than across product-groups ($\sigma^F > 1$).

There are two useful benchmarks for understanding the magnitude of cannibalization. In both cases, it useful to think of the introduction of a "standardized" product that has a market share equal to the average market share of the firm's other goods within the product group ($S_{N_{fgt}^U}^U N_{fgt}^U = 1$). We define the "cannibalization rate" as the value of the partial elasticity (13) for such a standardized product. At one extreme, suppose that firms are monopolistic competitors, $S_{fgt}^F \approx 0$, and all products supplied by a firm are as differentiated among themselves as they are with the output of other firms, *i.e.*, $\sigma^U = \sigma^F$. In this case, the cannibalization rate is zero, because all sales revenue arising from introducing a new product comes from the sales of goods supplied by other firms. Thus, a world with monopolistic competition and equal product differentiation is a world with no cannibalization. Clearly, the cannibalization rate will rise if firms cease being small $S_{fgt}^F > 0$ or if goods supplied by the same firm are more substitutable with each other than with goods supplied by different firms: $\sigma^U > \sigma^F$. At the other extreme, suppose that goods are "perfect substitutes" within firms, *i.e.*, $\sigma^U = \infty > \sigma^F$, so that varieties are differentiated across firms but there is no difference between varieties supplied by the same firm. In this case, the cannibalization rate is 1, because any sales of a new product are exactly offset by a reduction in the sales of existing products. Thus the cannibalization rate provides a measure of where in the spectrum ranging from perfect substitutes to equal differentiation the output of a firm lies.

## IV.E. The Sources of Firm Heterogeneity

In this section, we use the model to quantify the contribution of the different sources of firm heterogeneity to the dispersion in sales across firms. Nominal sales of firm $f$ within product group $g$, $E^F_{fgt}$, is the sum of sales across UPCs supplied by the firm:

$$E^F_{fgt} \equiv \sum_{u \in \Omega^U_{fgt}} P^U_{ut} C^U_{ut}.$$

Using CES demand (5), firm sales can be re-written as:

$$E^F_{fgt} = \left(\varphi^F_{fgt}\right)^{\sigma^F - 1} E^G_{gt} \left(P^G_{gt}\right)^{\sigma^F - 1} \left(P^F_{fgt}\right)^{\sigma^U - \sigma^F} \sum_{u \in \Omega^U_{fgt}} \left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1 - \sigma^U}. \tag{14}$$

Using the firm price index (3) to substitute for $P^{\sigma^U - \sigma^F}_{fgt}$, dividing and multiplying by $\left(N^U_{fgt}\right)^{\frac{\sigma^F - 1}{\sigma^U - 1}}$, and taking logarithms, we obtain:

$$\ln E^F_{fgt} = \ln E^G_{gt} + (\sigma_F - 1) \ln P^G_{gt} + (\sigma^F - 1) \ln \varphi^F_{fgt} + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right) \ln N^U_{fgt}$$

$$+ \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right) \ln \left(\frac{1}{N^U_{fgt}} \sum_{u \in \Omega^U_{fgt}} \left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1 - \sigma^U}\right), \tag{15}$$

which decomposes the variation in firm sales into contributions of product group variables ($E^G_{gt}$ and $P^G_{gt}$), firm appeal ($\varphi^F_{fgt}$), product scope ($N^U_{fgt}$) and average product-appeal-adjusted prices (final term). This decomposition holds regardless of assumptions made about firm pricing behavior and markups (*e.g.*, Bertrand versus Cournot pricing or the introduction of retail markups), because it uses the observed prices paid by consumers ($P^U_{ut}$).

Using the equilibrium pricing rule (9), the final term can be further decomposed into the contributions of the geometric mean of marginal costs ($\widetilde{\gamma}_{fgt}$), relative appeal-adjusted marginal costs across UPCs ($(\gamma_{ut}/\widetilde{\gamma}_{fgt})/\varphi^U_{ut}$), and the firm markup ($\mu^F_{fgt}$):

$$\ln E^F_{fgt} = \left\{\ln E^G_{gt} + (\sigma^F - 1) \ln P^G_{gt}\right\} \tag{16}$$

$$+ \left\{(\sigma^F - 1) \ln \varphi^F_{fgt} + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right) \ln N^U_{fgt}\right\}$$

$$+ \left\{[-(\sigma^F - 1) \ln \tilde{\gamma}_{fgt}] + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right) \ln \left(\frac{1}{N^U_{fgt}} \sum_{u \in \Omega^U_{fgt}} \left(\frac{\gamma_{ut}/\tilde{\gamma}_{fgt}}{\varphi^U_{ut}}\right)^{1 - \sigma^U}\right)\right\} - (\sigma^F - 1) \ln \mu^F_{fgt},$$

where $\tilde{\gamma}_{fgt} = \left(\prod_{u \in \Omega^U_{fgt}} \gamma_{ut}\right)^{\frac{1}{N^U_{fgt}}}$ and recall that $\gamma_{ut} = (1 + \delta) a_{ut} \left(Y^U_{ut}\right)^\delta$ denotes marginal cost.

Equation (16) decomposes firm sales into seven terms that capture the various margins through which firms can differ in sales. Clearly these margins are related to one another, since both the number of products and the markup are endogenous to firm appeal, product appeal and marginal cost. Nonetheless, the decomposition (16) isolates the direct effects of firm appeal, product appeal and marginal cost from their indirect effects through the number of products and the markup.

Although our decomposition is exact for any co-movement in variables, it is easiest to obtain intuition for this equation if we consider one-at-a-time, small movements in each variable so that we can safely ignore interactions between different variables. We therefore will explain the intuition for this equation in terms of small movements in each variable, and note that in general and in our empirical implementation we will allow all variables to move simultaneously.

The seven terms in equation (16) can be grouped into four main elements. The first element, in the first set of braces, captures market size and relative pricing. Our demand system is homogeneous of degree one in product group expenditures, so firm sales rise one to one with aggregate expenditures. The second term in the first set of braces captures the impact of the product group price index that summarizes the prices of competing varieties. Holding fixed a firm's characteristics, an increase in the product group price level of one percent will cause the firm's sales to rise by $\left(\sigma^F - 1\right)$ percent. Here, the elasticity of substitution between the firm's output and the output of other firms, $\sigma^F$, plays the crucial role of determining how much a relative price movement affects firm sales.

"Total firm appeal" is captured by the second two terms in braces, which capture the impacts of "firm appeal," $\varphi_{fgt}^F$, and the number of products the firm offers, "product scope," $N_{fgt}^U$. Consider two firms that supply the same number of products, but one firm supplies varieties that consumers prefer (measured in utility per unit), meaning that $\varphi_{fgt}^F > \varphi_{f'gt}^F$ and $N_{fgt}^U = N_{f'gt}^U$. Firm $f$ will then have higher sales; how much depends on the elasticity of substitution between firm output, $\sigma^F$. For a larger value of this elasticity, a given difference in firm appeal will translate into a larger difference in firm sales.

Now consider two firms that supply products with identical appeal, but one firm supplies more UPCs than another ($N_{fgt}^U > N_{f'gt}^U$). Here, it is easiest to think about this term in a symmetric world in which all goods and firms have identical appeal ($\varphi_{ut}^U = \varphi_{ft}^F = 1$ for all $u \in \Omega_{fgt}^U$ and $f \in \Omega_{gt}^F$) and identical marginal cost ($\gamma_{ut} = \gamma_t$ for all $u \in \Omega_{fgt}^U$ and $f \in \Omega_{gt}^F$), so we can just focus on the role played by product scope. Although firms have identical appeal, they do not have identical market shares because they differ in the number of products they offer, i.e., $\ln N_{fgt}^U > \ln N_{f'gt}^U$. For example, if consumers treated all UPCs identically regardless of which firm supplied them, i.e., $\sigma^U = \sigma^F$, firm $f$ would sell $\ln\left(N_{fgt}^U / N_{f'gt}^U\right)$ percent more output than firm $f'$. More generally, if the products supplied by a firm are more substitutable with each other than with those of other firms, $\sigma^U > \sigma^F$, the percentage gain in sales accruing to a firm that adds a product will be less than one, reflecting the fact that a new product will cannibalize the sales of its existing products. Indeed the degree of cannibalization will depend on the magnitude of $\sigma^U$; as this elasticity approaches infinity, the cannibalization rate will approach one, and all sales of new products will come from the sales of the firm's existing products. Hence, in this limiting case of $\sigma^U \to \infty$, adding product scope will have no impact on sales.

The terms in the third set of braces capture the role played by marginal costs. These costs can be divided into average marginal cost ($\tilde{\gamma}_{fgt}$) and "cost dispersion". Average marginal costs ($\tilde{\gamma}_{fgt}$) is the more conventional measure which captures the fact that high-cost firms have lower sales in equilibrium. The second term captures the fact that a firm sells more in equilibrium as the *dispersion* in the cost-to-appeal ratio across its products increases. The intuition is straightforward. The term in logs is a form of Theil index of dispersion with the numerator being the cost of production of each variety relative to the average and the denominator being the

product appeal relative to the average (which is always normalized to one). Consider a firm that supplies $N_{fgt}^U$ identical varieties at identical cost so that $\varphi_{ut} = \gamma_{ut} = \widetilde{\gamma}_{fgt} = 1$ for all $u \in \Omega_{fgt}^U$. In this case, the cost dispersion term will be zero because there is no dispersion in appeal or marginal cost among the firm's products. Now consider a comparative static in which we multiply $\gamma_{ut}$ by $\tau > 1$ and divide $\gamma_{kt}$ for all $k \neq u$ by $\tau^{1/(N_{fgt}^U - 1)}$. This change has the property of keeping geometric mean marginal costs, $\widetilde{\gamma}_{fgt}$, equal to one while increasing the dispersion in the cost of providing each element of the production bundle. Even though geometric mean marginal costs are unaffected by construction, firm sales will rise because the increase in dispersion allows the firm to supply its production bundle more cheaply by shifting its output towards the sales of cheaper varieties. This cost-dispersion term indicates that a firm sells more if its cost-to-appeal ratio is less evenly distributed across its varieties. Here, we also see the first instance of the insidiousness of demand for understanding multiproduct firms—in the nested CES case, one cannot express firm-level marginal costs without reference to the demand parameters $\sigma^F$ and $\sigma^U$.

Finally, the last term captures the role played by firm markups ($\mu_{fgt}^F$), which are themselves a function of the elasticity of substitution between firms ($\sigma^F$) and firm market share ($S_{fgt}^F$).

## IV.F. Firm Sales Decompositions

While the decomposition given in equation (16) is exact for each sector, it is difficult to use that specification to understand the general determinants of firm size because aggregate product group expenditures are a major determinant of firm sales. Therefore we decompose firm sales in each product group relative to the geometric mean for that product group:

$$
\Delta^g \ln E_{fgt}^F = \left\{ (\sigma_g^F - 1)\Delta^g \ln \varphi_{fgt}^F + \left( \frac{\sigma_g^F - 1}{\sigma_g^U - 1} \right) \Delta^g \ln N_{fgt}^U \right\} \tag{17}
$$
$$
+ \left\{ \left[ -\left( \sigma_g^F - 1 \right) \Delta^g \ln \tilde{\gamma}_{fgt} \right] + \left( \frac{\sigma_g^F - 1}{\sigma_g^U - 1} \right) \Delta^g \ln \left( \frac{1}{N_{fgt}^U} \sum_{u \in \Omega_{fgt}^U} \left( \frac{\gamma_{ut}/\tilde{\gamma}_{fgt}}{\varphi_{ut}^U} \right)^{1-\sigma_g^U} \right) \right\}
$$
$$
- \left( \sigma_g^F - 1 \right) \Delta^g \ln \mu_{fgt}^F,
$$

where we have reintroduced the product group subscript $g$ for all the relevant variables, and $\Delta^g$ is the difference operator relative to the geometric mean for product group $g$, such that $\Delta^g \ln E_{fgt}^F = \left[ \ln E_{fgt}^F - \frac{1}{N_{gt}^F} \sum_{k \in \Omega_{gt}^F} \ln E_{kgt}^F \right]$.

The left-hand side of equation (17) is the sales of firm $f$ relative to average firm sales in the sector, and the five terms on the right-hand side tell us the importance of relative firm appeal, scope, average marginal costs, cost dispersion, and markups in understanding cross-sectional differences in firm size. Moreover, we can also undertake this decomposition for a firm's sales growth rate by taking the first difference of equation (17) over time, which in our notation simply involves replacing $\Delta^g$ in (17) with $\Delta^{g,t}$, where $\Delta^{g,t}$ is the double difference operator relative to the geometric mean for product group $g$ and over time $t$, so that $\Delta^{g,t} \ln E_{fgt}^F = \left[ \ln E_{fgt}^F - \frac{1}{N_{gt}^F} \sum_{k \in \Omega_{gt}^F} \ln E_{kgt}^F \right] - \left[ \ln E_{fgt-1}^F - \frac{1}{N_{gt-1}^F} \sum_{k \in \Omega_{gt-1}^F} \ln E_{kgt-1}^F \right]$.

We now can decompose the cross-sectional variation in firm sales using a procedure analogous to Eaton, Kortum, and Kramarz's (2004) variance decomposition commonly used in the international trade literature. In

particular, we regress each of the components of log firm sales in the decomposition (17) on log firm sales as follows:

$$(\sigma_g^F - 1)\Delta^g \ln \varphi_{fgt}^F = \alpha_g^\varphi \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^\varphi, \tag{18}$$

$$-\left(\sigma_g^F - 1\right)\Delta^g \ln \tilde{\gamma}_{fgt} = \alpha_g^\gamma \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^\gamma, \tag{19}$$

$$\left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\Delta^g \ln \left[\frac{1}{N_{fgt}^U} \sum_{u \in \Omega_{fgt}^U} \left(\frac{\gamma_{ut}/\tilde{\gamma}_{fgt}}{\varphi_{ut}^U}\right)^{1-\sigma^U}\right] = \alpha_g^D \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^a, \tag{20}$$

$$\left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\Delta^g \ln N_{fgt}^U = \alpha_g^N \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^N, \tag{21}$$

$$-\left(\sigma_g^F - 1\right)\Delta^g \ln \mu_{fgt}^F = \alpha_g^\mu \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^\mu, \tag{22}$$

where we have again differenced relative to the geometric mean for the product group. We allow the coefficients $\{\alpha_g^\varphi, \alpha_g^\gamma, \alpha_g^D, \alpha_g^N, \alpha_g^\mu\}$ to differ across product groups. By the properties of OLS, this decomposition allocates the covariance terms between the components of firm sales equally across those components, and implies $\alpha_g^\varphi + \alpha_g^\gamma + \alpha_g^D + \alpha_g^N + \alpha_g^\mu = 1$. The values for each of the $\alpha$'s provide us with a measure of how much of the variation in firm sales can be attributed to each component. If we replace $\Delta^g$ with $\Delta^{g,t}$ in the above equations, we can determine how much of the variance in firm sales growth rates can be attributed to each component. Thus, estimation of these five equations provides us with a simple way to decompose firm sales in the cross-section and in the time-series.

## IV.G.   Decomposing Changes in Firm Average Appeal

When decomposing firm growth, we can go one step further and understand how much of a firm's change in appeal is due to the introduction of new products and how much is due to changes in the appeal of existing products. Note that our normalization for UPC appeal implies that the geometric mean of the $\varphi_{ut}^U$ equals one for each firm $f$ within each product group $g$ in each period $t$, *i.e.*,

$$\varphi_{fgt}^F = \varphi_{fgt}^F \left(\prod_{u \in \Omega_{fgt}^U} \varphi_{ut}^U\right)^{\frac{1}{N_{fgt}^U}}. \tag{23}$$

We can then can take log differences of this equation relative to the mean within a product group and over time to yield:

$$\Delta^{g,t} \ln \varphi_{fgt}^F \equiv \Delta^t \ln \varphi_{fgt}^F - \Delta^t \overline{\ln \varphi_{gt}^F} = \Delta^t \ln \varphi_{fgt}^F + \frac{1}{N_{fgt}^U} \sum_{u \in \Omega_{fgt}^U} \ln \varphi_{ut}^U - \frac{1}{N_{fg,t-1}^U} \sum_{u \in \Omega_{fg,t-1}^U} \ln \varphi_{u,t-1}^U - \Delta^t \overline{\ln \varphi_{gt}^F}, \tag{24}$$

where $\Delta^t$ denotes the first difference operator over time such that $\Delta^t \ln \varphi_{fgt}^F = \ln \varphi_{fgt}^F - \ln \varphi_{fgt-1}^F$; a bar above a variable denotes a mean such that $\overline{\ln \varphi_{gt}^F} = \frac{1}{N_{gt}^F} \sum_{f \in \Omega_{gt}^F} \ln \varphi_{fgt}^F$; and our normalization of product appeal to have a geometric mean of one implies that the second and third terms are both equal to zero.

To isolate the sources of changes in firm average appeal, we distinguish between UPCs that are supplied in both periods versus those that are supplied in only one of the two periods. Let $I_{fgt}^U = \Omega_{fgt}^U \cap \Omega_{fgt-1}^U$ denote the

set of UPC's that are supplied by firm $f$ within product group $g$ in both periods $t$ and $t-1$. Similarly, define $I_{fgt}^{U+}$ to be the set of newly introduced UPCs, *i.e.,* the set of UPCs in $\Omega_{fgt}^{U}$ but not in $\Omega_{fgt-1}^{U}$, and $I_{fgt}^{U-}$ to be the set of disappearing UPCs, *i.e.,* the set of UPCs in $\Omega_{fgt-1}^{U}$ but not in $\Omega_{fgt}^{U}$.

Noting that $\left|\Omega_{fgt}^{U}\right| = \left|I_{fgt}^{U}\right| + \left|I_{fgt}^{U+}\right|$ and $\left|\Omega_{fgt-1}^{U}\right| = \left|I_{fgt}^{U}\right| + \left|I_{fgt}^{U-}\right|$, we can decompose the change in firm average appeal into changes in appeal for a constant set of products and changes in appeal from the adding and dropping of UPCs:

$$
\Delta^{g,t} \ln \varphi_{fgt}^{F} = \left\{ \frac{\left|I_{fgt}^{U}\right|}{\left|\Omega_{fgt}^{U}\right|} \left[ \frac{1}{\left|I_{fgt}^{U}\right|} \sum_{u \in I_{fgt}^{U}} \ln \varphi_{ut}^{U} \right] - \frac{\left|I_{fgt}^{U}\right|}{\left|\Omega_{fgt-1}^{U}\right|} \left[ \frac{1}{\left|I_{fgt}^{U}\right|} \sum_{u \in I_{fgt}^{U}} \ln \varphi_{ut-1}^{U} \right] + \left( \Delta^{t} \ln \varphi_{fgt}^{F} - \Delta^{t} \overline{\ln \varphi_{gt}^{F}} \right) \right\}
$$

(25)

$$
+ \left\{ \frac{\left|I_{fgt}^{U+}\right|}{\left|\Omega_{fgt}^{U}\right|} \left[ \frac{1}{\left|I_{fgt}^{U+}\right|} \sum_{u \in I_{fgt}^{U+}} \ln \varphi_{ut}^{U} \right] - \frac{\left|I_{fgt}^{U-}\right|}{\left|\Omega_{fgt-1}^{U}\right|} \left[ \frac{1}{\left|I_{fgt}^{U-}\right|} \sum_{u \in I_{fgt}^{U-}} \ln \varphi_{ut-1}^{U} \right] \right\}.
$$

The first term in braces captures changes in appeal for a constant set of products, and includes the change in firm average appeal ($\Delta^{t} \ln \varphi_{fgt}^{F}$), as well as the difference between average product appeal for the common set of products (in square brackets) weighted by their relative importance in the set of products in each quarter. This difference adjusts changes in our measure of firm appeal for changes in our normalization due to the entry and exit of new goods. We refer to the first term in braces in equation (25) as the "demand effect". It represents the change in the appeal of the UPCs present in both periods. If the firm does not add or eliminate any products, the summation terms will be zero since our normalization means that average log product appeal must be zero in all periods.

However, if the set of products changes, the summation terms will not necessarily be zero. Product turnover can influence sales either through product upgrading, which is measured in the second term in braces in equation (25), as well as through changes in the number of products, which enters separately into the decomposition of log firms sales in (17). The "product-upgrading effect" (the second term in braces in (25)) captures changes in appeal from the adding and dropping of products. Product upgrading depends on average product appeal for the entering and exiting products (in square brackets) weighted by their relative importance in the set of products in each quarter. For example, suppose a firm introduces a new high-appeal product and retires a low-appeal product leaving the number of products unchanged. In this case, the second term in braces will be positive, reflecting the fact that the firm upgraded the average appeal of its product mix.

## V. Structural Estimation

Our structural estimation of the model has two components. First, given data on expenditure shares and prices $\left\{ S_{ut}^{U}, \; S_{fgt}^{F}, \; P_{ut}^{U} \right\}$ and known values of the model parameters $\left\{ \sigma_{g}^{U}, \; \sigma_{g}^{F}, \delta_{g} \right\}$, we show how the model can be used to determine unique values of firm appeal ($\varphi_{fgt}^{F}$), product appeal ($\varphi_{ut}^{U}$), and marginal cost shocks ($a_{ut}$) up to our normalization of appeal. These correspond to structural residuals of the model that are functions of the observed data and parameters and ensure that the model exactly replicates the observed data. We use these

structural residuals to implement our decomposition of firm sales from Section IV.E. with our observed data on expenditure shares and prices.

Second, we estimate the model parameters $\{\sigma_g^U, \sigma_g^F, \delta_g\}$ using a generalization of Feenstra (1994) and Broda and Weinstein (2006, 2010) to allow firms to be large relative to the markets in which they operate (which introduces variable markups) and to incorporate multiproduct firms (so that firm pricing decisions are made jointly for all varieties). This estimation uses moment conditions in the double-differenced structural residuals $\{\varphi_{ft}^F, \varphi_{ut}^U, a_{ut}\}$ and also has a recursive structure. In a first step, we estimate the elasticity of substitution across UPCs within firms and the marginal cost elasticity for each product group $\{\sigma_g^U, \delta_g\}$. In a second step, we use these estimates for UPCs to estimate the elasticity of substitution across firms for each product group $\{\sigma_g^F\}$.

## V.A.  Structural Residuals

We begin by showing that there is a one-to-one mapping from the observed data on expenditure shares and prices $\left\{S_{ut}^U, S_{fgt}^F, P_{ut}^U\right\}$ and the model's parameters $\{\sigma_g^U, \sigma_g^F, \delta_g\}$ to the unobserved structural residuals $\left\{\varphi_{ut}^U, \varphi_{fgt}^F, a_{ut}\right\}$.

Given known values for the model's parameters $\{\sigma_g^U, \sigma_g^F, \delta_g\}$ and the observed UPC expenditure shares and prices, we can use the expression for the expenditure share given in equation (4) to determine UPC appeal $\left\{\varphi_{ut}^U\right\}$ up to our normalization that the geometric mean of UPC appeal is equal to one. These solutions for UPC appeal and observed UPC prices can be substituted into the CES price index (3) to compute firm price indices $\left\{P_{fgt}^F\right\}$. These solutions for firm price indices and observed firm expenditure shares can be combined with the CES expenditure share (4) to determine firm appeal $\left\{\varphi_{fgt}^F\right\}$ up to our normalization that the geometric mean of firm appeal is equal to one. Furthermore observed firm expenditure shares and the CES markup (7) are sufficient to recover firm markups $\left\{\mu_{fgt}^F\right\}$. These solutions for markups and observed UPC prices and expenditures can be substituted into the CES pricing rule (9) to determine the marginal cost shock ($a_{ut}$).

Finally, our solutions for markups and observed UPC expenditures can be combined with CES variable profits (11) to obtain upper bounds to the fixed costs of supplying UPCs ($H_{gt}^U$) and the fixed costs of supplying the market ($H_{gt}^F$). The upper bound for UPC fixed costs for each product group ($H_{gt}^U$) is defined by the requirement that variable profits for the least profitable UPC within a product group must be greater than this fixed cost. Similarly, the upper bound for firm fixed costs for each product group ($H_{gt}^F$) is defined by the requirement that variable profits for the least profitable firm must be greater than this fixed cost.

This mapping from the observed data on expenditure shares and prices $\left\{S_{ut}^U, S_{fgt}^F, P_{ut}^U\right\}$ and the model's parameters $\{\sigma_g^U, \sigma_g^F, \delta_g\}$ to the unobserved structural residuals $\left\{\varphi_{ut}^U, \varphi_{fgt}^F, a_{ut}\right\}$ does not impose assumptions about the functional forms of the distributions for the structural residuals or about their correlation with one another. When we estimate the model's parameters $\{\sigma_g^U, \sigma_g^F, \delta_g\}$ below, we impose some identifying assumptions on the double-differenced values of these structural results but not upon their levels. Therefore, having recovered these structural residuals, we can examine the relationship between them.

## V.B. UPC Moment Conditions

We now discuss our methodology for estimating the elasticities of substitution $\{\sigma_g^U, \sigma_g^F\}$ and the elasticity of marginal costs with respect to output $\{\delta_g\}$, which uses moment conditions in double-differenced values of the structural residuals $\{\varphi_{fgt}^F, \varphi_{ut}^U, a_{ut}\}$. This estimation again has a recursive structure. In a first step, we estimate the elasticity of substitution across UPCs within firms for each product group $\{\sigma_g^U\}$ and the marginal cost elasticity $\{\delta_g\}$. In a second step, we use these estimates for UPCs to estimate the elasticity of substitution across firms for each product group $\{\sigma_g^F\}$.

In the first step, we double-difference log UPC expenditure shares (4) over time and relative to the largest UPC within each firm to obtain the following equation for relative UPC demand:

$$\Delta^{\underline{u},t} \ln S_{ut}^U = (1 - \sigma_g^U)\Delta^{\underline{u},t} \ln P_{ut}^U + \omega_{ut}, \tag{26}$$

where $u$ is a UPC supplied by the firm; $\underline{u}$ corresponds to the largest UPC supplied by the same firm (as measured by the sum of expenditure across the two quarters); $\Delta^{\underline{u},t}$ is the double-difference operator across UPCs and over time such that $\Delta^{\underline{u},t} \ln S_{ut}^U = \Delta^t S_{ut}^U - \Delta^t S_{\underline{u}t}^U$; $\Delta^t$ is the first-difference operator over time such that $\Delta^t \ln S_{ut}^U = S_{ut}^U - S_{ut-1}^U$; and $\omega_{ut} = (\sigma_g^U - 1)\left[\Delta^t \ln \varphi_{ut}^U - \Delta^t \ln \varphi_{\underline{u}t}^U\right]$ is a stochastic error. We can compute this double difference for all firms that have two or more UPCs and are present for two or more consecutive time periods.

Double-differencing the UPC pricing rule (9) enables us to obtain an equation for relative UPC supply. Using the cost function $(A_{ut}(Y_{ut}^U) = a_{ut}(Y_{ut}^U)^{1+\delta})$ and noting that $Y_{ut}^U = E_{ut}^U/P_{ut}^U = E_{fgt}^F S_{ut}^U/P_{ut}^U$, the UPC pricing rule given in equation (9) can be re-written as:

$$P_{ut}^U = \mu_{fgt}^F \gamma_{ut} = \left(\mu_{fgt}^F\right)^{\frac{1}{1+\delta_g}} (1+\delta_g)^{\frac{1}{1+\delta_g}} a_{ut}^{\frac{1}{1+\delta_g}} \left(E_{fgt}^F\right)^{\frac{\delta_g}{1+\delta_g}} \left(S_{ut}^U\right)^{\frac{\delta_g}{1+\delta_g}}.$$

Taking logs and double-differencing, we obtain the following equation for relative UPC supply:

$$\Delta^{\underline{u},t} \ln P_{ut}^U = \frac{\delta_g}{1+\delta_g}\Delta^{\underline{u},t} \ln S_{ut}^U + \kappa_{ut}, \tag{27}$$

where the markup $(\mu_{fgt}^F)$ and total firm expenditure $(E_{fgt}^F)$ have differenced out because they take the same value across UPCs within the firm; and $\kappa_{ut} = \frac{1}{1+\delta_g}\left[\Delta^t \ln a_{ut} - \Delta^t \ln a_{\underline{u}t}\right]$ is a stochastic error. Since the firm markup differences out and we observe prices, our estimation approach is the same under either price or quantity competition, and hence is robust across these different forms of competition.

Following Broda and Weinstein (2006), the orthogonality of the double-differenced demand and supply shocks defines a set of moment conditions (one for each UPC):

$$G(\beta_g) = \mathbb{E}_{\mathbb{T}}[v_{ut}(\beta_g)] = 0, \tag{28}$$

where $\beta_g = \begin{pmatrix} \sigma_g^U \\ \delta_g \end{pmatrix}$; $v_{ut} = \omega_{ut}\kappa_{ut}$; and $\mathbb{E}_{\mathbb{T}}$ is the expectations operator over time. For each product group, we stack all the moment conditions to form the GMM objective function and obtain:

$$\hat{\beta}_g = \arg\min_{\beta_g} \left\{ G^*(\beta_g)'WG^*(\beta_g) \right\} \qquad \forall g, \tag{29}$$

where $G^*(\beta_g)$ is the sample analog of $G(\beta_g)$ stacked over all UPCs in a product group and $W$ is a positive definite weighting matrix. As in Broda and Weinstein (2010), we weight the data for each UPC by the number of raw buyers for that UPC to ensure that our objective function is more sensitive to UPCs purchased by larger numbers of consumers. We impose the theoretical restriction that $\sigma_g^U, \sigma_g^F > 0$, which is a necessary (though not sufficient) condition for positive markups. We do not impose the theoretical restrictions that $\sigma_g^U > \sigma_g^F > 1$ and instead check whether these inequalities are satisfied by our estimates. We also do not restrict $\delta_g$.

The moment condition (28) for each UPC involves the expectation of the product of the double-differenced demand and supply shocks: $\upsilon = \omega_{ut}\kappa_{ut}$. From relative demand (26) and relative supply (27), this expectation depends on the variance of prices, the variance of expenditure shares, the covariance of prices and expenditure shares, and parameters. Our identifying assumption that this expectation is equal to zero defines a rectangular hyperbola in ($\sigma_g^U$, $\delta_g$) space for each UPC within a product group $g$, along which a higher value of $\sigma_g^U$ has to be offset by a lower value of $\delta_g$ in order for the expectation to be equal to zero (Leontief [1929]). Therefore, this rectangular hyperbola places bounds on the demand and supply elasticities for each UPC within that product group, even in the absence of instruments for demand and supply. Furthermore, if the variances for the double-differenced demand and supply shocks are heteroskedastic across UPCs, the rectangular hyperbolas are different for each pair of UPCs within the product group, and their intersection can be used to separately identify the demand and supply elasticities for that product group (Feenstra [1994]). Consistent with these identifying assumptions, we find that the double-differenced demand and supply shocks are in general heteroskedastic.[23]

Our assumption that the double-differenced demand and supply shocks are orthogonal follows a standard approach in the international trade and macroeconomics literatures (see in particular Feenstra [1994]; Broda and Weinstein [2006, 2010]). This double differencing is important in addressing most standard endogeneity concerns. For example, a standard concern is that there may be *firm-level* shocks (*e.g.*, changes in management) that affect both costs and appeal (quality or taste) across all products. Differencing across products within firms eliminates such common firm-level shocks. Another standard concern is that products may have different production technologies, which could affect both costs and appeal (quality or taste) in all time periods. Differencing over time within products eliminates such time-invariant heterogeneity between products. Therefore our double differencing nets out both these types of shocks. In other words, our identification is based only on *relative* differences in demand and supply of individual bar codes.

The second main potential threat to identification is a change in observable *product* characteristics that affects both relative costs and relative appeal, but this endogeneity concern is not present in bar-code data. Fortunately, any substantive change in product characteristics is accompanied by the introduction of a new bar code.[24] Therefore, by using variation within bar codes over time, we hold constant observable product characteristics. As a result, it is much harder to think of reasons why double-differenced changes in production costs that leave the observable characteristics of a product constant should affect double-differenced consumer demand for that product conditional on price. We provide further evidence in support of these identifying

---

[23]In a White test for heteroskedasticity, we are able to reject the null hypothesis of homoskedasticity at conventional significance levels for 94 percent of product groups.

[24]Minor changes like the style of product label do not result in new bar codes, but retailers and manufacturers do not use the same bar codes for different products because doing so would interfere with their inventory management.

assumptions and the robustness of our results in Section VII. below.

## V.C.  Firm Moment Conditions

We use our estimates of the UPC elasticities of substitution $\left\{\sigma_g^U\right\}$ from the first step to solve for UPC appeal $\left\{\varphi_{ut}^U\right\}$ and compute the firm price indices $\left\{P_{fgt}^F\right\}$ using equations (3) and (4). In our second step, we double difference log firm expenditure shares (4) over time and relative to the largest firm within each product group, $\underline{f}$, to obtain the following equation for relative firm market share:

$$\Delta^{\underline{f},t} \ln S_{fgt}^F = \left(1 - \sigma_g^F\right) \Delta^{\underline{f},t} \ln P_{fgt}^F + \omega_{fgt}, \tag{30}$$

where $\Delta^{\underline{f},t}$ is the double difference operator across firms within a product group and over time such that $\Delta^{\underline{f},t} \ln S_{fgt}^F = \Delta^t \ln S_{fgt}^F - \Delta^t \ln S_{\underline{f}gt}^F$ and the stochastic error is $\omega_{fgt} \equiv \left(\sigma_g^F - 1\right) \Delta^{\underline{f},t} \ln \varphi_{fgt}^F$.

Estimating equation (30) using ordinary least squares could be problematic because changes in firm price indices could be correlated with changes in firm appeal: $\mathrm{Cov}\left(\Delta^{\underline{f},t} \ln P_{fgt}^F, \Delta^{\underline{f},t} \ln \varphi_{fgt}^F\right) \neq 0$. To find a suitable instrument for changes in firm price indices, we use the structure of the model to write changes in firm price indices in terms of the underlying UPC characteristics of the firm. Using the CES expenditure shares (4), we can write relative UPC expenditures in terms of relative UPC prices and relative UPC demand shifters:

$$\frac{S_{ut}^U}{\widetilde{S}_{fgt}^U} = \left(\frac{P_{ut}^U/\varphi_{ut}^U}{\widetilde{P}_{fgt}^U/\widetilde{\varphi}_{fgt}^U}\right)^{1-\sigma_g^U}, \qquad u \in \Omega_{fgt}^U, \tag{31}$$

where here we compare each UPC to the geometric mean of UPCs within the firm and product group, which is denoted by a tilde above a variable so that for example $\tilde{S}_{fgt}^U = \exp\left\{\frac{1}{N_{fgt}^U} \sum_{u \in \Omega_{fgt}^U} \ln S_{ut}^U\right\}$. Using this expression for relative expenditure shares to substitute for product appeal $(\varphi_{ut}^U)$ in the CES price index (3), we can write the firm price index solely in terms of observed relative expenditures and the geometric mean of UPC prices:

$$\ln P_{fgt}^F = \ln \tilde{P}_{fgt}^U + \frac{1}{1-\sigma_g^U} \ln\left[\sum_{u \in \Omega_{fgt}^U} \frac{S_{ut}^U}{\widetilde{S}_{fgt}^U}\right], \tag{32}$$

where we have used our normalization that $\tilde{\varphi}_{fgt}^U = 1$.

Equation (32) decomposes the firm price index into two terms. The first term is entirely conventional: the geometric mean of the prices of all goods supplied by the firm. When researchers approximate this firm price index using firm-level unit values or the prices of representative goods supplied by firms they are essentially capturing this component of the firm price index. The second term is novel and arises because the conventional price indexes are not appropriate for multiproduct firms. The log component of the second term is a variant of the Theil index of dispersion.[25] If the shares of all products are equal, the Theil index will equal $\ln N_{ft}^U$, which is increasing in the number of products supplied by the firm, $N_{fgt}^U$. Since the Theil index is multiplied by $\left(1 - \sigma_g^U\right)^{-1} \leq 0$, the firm's price index falls as the number of goods supplied by the firm rises. This Theil index also increases as the dispersion of the market shares across UPCs within the firm rises.

---

[25] The standard Theil index uses shares relative to simple average shares, while ours expresses shares relative to the geometric mean.

One can obtain some intuition for this formula by comparing it to the conventional price indexes commonly used in economics. Firm price indexes are typically constructed by making at least one of two critical assumptions—firms only supply one product ($N_{fgt}^U = 1$), or the goods supplied by firms are perfect substitutes (*i.e.*, $\sigma_g^U = \infty$). Either of these assumptions is sufficient to guarantee that the firm's price level equals its average price level. However, both of these assumptions are likely violated in reality. The average price of a firm's output overstates the price level for multiproduct firms because consumers derive more utility per dollar spent on a firm's production if that production bundle contains more products. *Thus, while the (geometric) average firm price is a theoretically rigorous way to measure the price level of firms that produce perfect substitutes, it overstates the prices of firms that produce differentiated goods, and this bias will tend to rise as the number of products supplied by the firm increases.*

The structure of the model implies that the dispersion of the shares of UPCs in firm expenditure ($S_{ut}^U$) only affects the shares of firms in product group expenditure ($S_{fgt}^F$) through the firm price indices ($P_{fgt}^F$). Double differencing equation (32) for the log firm price index over time and relative to the largest firm within each product-group, we obtain:

$$\Delta^{\underline{f},t} \ln P_{fgt}^F = \Delta^{\underline{f},t} \ln \widetilde{P}_{fgt}^U + \frac{1}{1-\sigma_g^U} \Delta^{\underline{f},t} \ln \left[ \sum_{u \in \Omega_{ft}^U} \frac{S_{ut}^U}{\widetilde{S}_{fgt}^U} \right], \tag{33}$$

where the model implies that the second term on the right-hand side containing the shares of UPCs in firm expenditure is a suitable instrument for the double-differenced firm price index in equation (30). We find that this instrument is powerful in the first-stage regression (33), with a first-stage $F$-statistic for the statistical significance of the excluded exogenous variable that is substantially above the recommended threshold of 10 from Stock, Wright and Yogo (2002).

# VI. Estimation Results

We present our results in several stages. First, we present our elasticity and appeal estimates and show that they are reasonable. Second we use these estimates to examine cannibalization, and finally we present our results on the sources of firm heterogeneity.

## VI.A. Estimated Elasticities of Substitution

Because we estimate 100 $\sigma_g^U$'s and $\sigma_g^F$'s, it would needlessly clutter the paper to present all of them individually. Table V shows that goods supplied by the same firm are imperfect substitutes. For UPCs, the estimated elasticity of substitution ranges from 4.7 at the 95th percentile to 17.6 at the 5th percentile with a median elasticity of 6.9. These numbers are large compared with trade elasticities, reflecting the fact that products supplied by the same firm are closer substitutes than products supplied by different firms. The median elasticity implies that a one percent price cut causes the sales of that UPC to rise by 6.9 percent.[26]

---

[26]The UPC own price elasticity derived in equation (38) is given by

A second striking feature of our results is that the elasticity of substitution among varieties supplied by a firm is always larger than that between firms (*i.e.*, $\sigma_g^U > \sigma_g^F$). It is important to remember that this is not a result that we imposed on the data. Moreover, most of the elasticities are precisely estimated—in 82 percent of the cases, we can statistically reject the hypothesis that $\sigma_g^U = \sigma_g^F$ at the 5% level. The higher elasticities of substitution across UPCs than across firms imply that varieties are more substitutable within firms than across firms, which implies cannibalization effects from the introduction of new varieties by firms. Moreover, the fact that the estimated elasticities of substitution are always greater than one implies that firms' varieties are substitutes, which is required for positive markups of price over marginal costs (*c.f.*, equation (9)).

In order to assess whether our elasticities are plausible, it is useful to compare our estimates with those of other papers. In order to do this, we restricted ourselves to comparing our results with studies that used US scanner data and estimated elasticities for the same product groups as ours. We do this because, as Broda and Weinstein (2006) show, elasticity estimates for aggregate data can look quite different than those for disaggregate data. Unfortunately, we did not find studies estimating the elasticity of substitution within firms, but we did find studies that examined elasticities that can be compared with our cross-firm elasticity, $\sigma_g^F$. Gordon et al. (2013) and the literature review therein presents results for comparable product groups using quite different estimation methodologies. For the twelve overlapping product groups, we find similar results with an average estimated elasticity of 3.65 compared to 3.14 in Gordon (a difference of around 14 percent), as reported in further detail in Section S10 of the online supplement.[27] Therefore, while our empirical approach has a number of novel features in modeling multiproduct firms that are of positive measure relative to the markets in which they operate, the empirical estimates generated by our procedure are reasonable compared to the benchmark of findings from other empirical studies.

## VI.B. Firm and Product Appeal

Firm and product appeal are a second key component of our decomposition. Our estimation procedure allows us to estimate a different $\varphi_{ut}^U$ for every quarter in our dataset. We have strong priors that product appeal should be fairly stable across time. One way to gauge this stability is to regress these variables on UPC fixed effects in order to determine how much of a product's appeal is common across time periods. When we do this and include time fixed effects to control for inflation and other common demand shocks, we find that the $R^2$ is 0.84, which implies that very little of the variation in our measures of appeal arises from changes in appeal within UPCs over time. We should also expect firm appeal, $\varphi_{fgt}^F$, to exhibit a strong firm component. Running an analogous regression for firm appeal including firm and time fixed effects, we find that the $R^2$ is 0.83.[28]

We can gain further intuition for the reasonableness of our estimates and the role played by the various sources of firm heterogeneity by computing some simple correlations. Our procedure for estimating the

$$\frac{\partial C_{ut}^U}{\partial P_{ut}^U} \frac{P_{ut}^U}{C_{ut}^U} = \left(\sigma_g^F - 1\right) S_{ut}^F S_{ut}^U + \left(\sigma_g^U - \sigma_g^F\right) S_{ut}^U - \sigma_g^U.$$

---

[27]The product categories are Butter and Margarine, Carbonated Soft Drinks, Coffee, Deodorant, Ketchup, Laundry Detergent, Mayonnaise, Mustard, Peanut Butter, Spaghetti Sauce, Toilet Paper, and Yogurt.

[28]Most of the explanatory power in these regressions is contributed by the UPC and firm fixed effects. Running the same regressions without the time fixed effects, we continue to find $R^2$ of 0.84 and 0.83 respectively.

model's parameters ($\sigma_g^U$, $\sigma_g^F$, $\delta_g$) assumes that the double-differenced product appeal ($\Delta^{\underline{u},t} \ln \varphi_{ut}^U$) and cost shocks ($\Delta^{\underline{u},t} \ln a_{ut}$) are orthogonal to one another, where $\Delta^{\underline{u},t}$ denotes a double difference relative to the largest product within a firm and over time. But we do not impose restrictions on the relationship between the relative levels of sales ($\Delta^f \ln S_{ut}^U$), prices ($\Delta^f \ln P_{ut}^U$), appeal ($\Delta^f \ln \varphi_{ut}^U$) and marginal costs ($\Delta^f \ln \gamma_{ut} = \Delta^f \ln \left( (1 + \delta_g) a_{ut} Y_{ut}^{\delta_g} \right)$), where $\Delta^f$ denotes a first difference relative to the geometric mean for the firm. In Table VI, we present the correlations between the relative levels of these variables across bar codes. There are several important insights from this table. First, relative bar-code appeal ($\Delta^f \ln \varphi_{ut}^U$) and relative marginal cost ($\Delta^f \ln \gamma_{ut}^U$) are positively correlated. The fact that the correlation is 0.94 means that on average it is more costly for firms to produce more appealing products. This high correlation is not that surprising because much of the variation in appeal amongst a single firm's products probably reflects manufacturing differences (*e.g.,* size and quality) that are reflected in marginal costs.

Table VI is also useful in understanding the determinants of bar-code sales. While relative bar-code appeal ($\Delta^f \ln \varphi_{ut}^U$) is strongly correlated with relative bar-code sales ($\Delta^f \ln E_{ut}^U$), relative marginal cost ($\Delta^f \ln \gamma_{ut}^U$) is not. How can it be that marginal costs are strongly associated with bar-code appeal, which in turn is positively correlated with sales, but marginal costs are not correlated with sales? The answer can be garnered from rewriting our expenditure share equation (4) after log differencing:

$$\Delta^f \ln E_{ut}^U = \left( 1 - \sigma_g^F \right) \Delta^f \ln \gamma_{ut} - \left( 1 - \sigma_g^F \right) \Delta^f \ln \varphi_{ut}^U \tag{34}$$

where we have used the fact that $\Delta^f \ln E_{ut}^U = \Delta^f \ln S_{ut}^U$, $P_{ut}^U = \mu_{fgt}^F \gamma_{ut}$, and the markup $\mu_{fgt}^F$ is common across UPCs within firms and hence differences out.

From the above relationship (34), there are two forces at work in these correlations. First, there is the direct effect of marginal costs on sales through the price channel, which is captured by the first term: higher marginal costs lead to higher prices and *lower* sales. However, we also saw that the correlation between appeal and marginal cost is positive, which arises from the fact that higher marginal costs are associated with higher appeal products, which tend to have *higher* sales. The fact that these forces tend to cancel explains why the correlation coefficients between sales and either marginal costs or appeal are much less than one. The reason that this attenuation matters more for marginal cost than appeal stems from the variance of the two variables. The variance of relative appeal is 16 percent larger than the variance of relative marginal cost.[29] In other words, appeal varies more than marginal costs so the high covariance between the two variables cancels out the tendency of high cost products to have low market shares, but it does not fully negate the tendency of high appeal products to have high market shares.

Table VII shows that this pattern is even stronger at the firm level. Again we express each variable in logs and now take differences across firms relative to the geometric mean for the product group (denoted by

---

[29]One can think of this issue in terms of the standard formula for omitted variable bias. If we estimate equation (34) and omit appeal, the regression coefficient on marginal cost $\left( 1 - \sigma_g^F \right)$ will be biased upwards by $\left( \sigma_g^F - 1 \right) \text{cov} \left( \Delta^f \ln (\gamma_{ut}), \Delta^f \ln (\varphi_{ut}) \right) / \text{var} \left( \Delta^f \ln (\gamma_{ut}) \right)$. On the other hand, if we estimate the equation omitting marginal cost, the coefficient on appeal $\left( \sigma_g^F - 1 \right)$ will be biased downward by $\left( \sigma_g^F - 1 \right) \text{cov} \left( \Delta^f \ln (\varphi_{ut}), \Delta^f \ln (\gamma_{ut}) \right) / \text{var} \left( \Delta^f \ln (\varphi_{ut}) \right)$. The important point is that the biases in each regression just differ by the magnitude of the variances, but the true coefficients are of the same magnitude (but opposite sign). Thus, if the variance of appeal exceeds the variance of marginal cost, we are more likely to find no correlation between sales and marginal cost than between sales and appeal.

the $\Delta^g$ operator). We find that firm appeal is still strongly positively correlated with firm average marginal costs ($\Delta^g \ln(\tilde{\gamma}_{fgt})$), presumably because firm appeal is closely associated with manufacturing costs, but the magnitude of this correlation is lower than at the bar-code level. This lower correlation may reflect the fact that appeal differences across firms are more reflective of taste differences than products made by the same firm. Here too, firm appeal is positively correlated with firm sales, but there is no correlation between firm average marginal costs and firm sales. In the firm-level results, the stronger correlation between sales and appeal is driven by the fact that the variance of firm appeal is 60 percent larger than that of firm average marginal cost.

Table VII also indicates strong positive correlations between scope and sales and between cost dispersion and sales, which reflects the fact that larger firms tend to produce more goods.[30] Taking Tables VI and VII together, we find strong but imperfect correlations between appeal and prices at both the bar code and firm level. These correlations give some support to the work that uses price as a proxy for appeal or quality. While there is a very strong link between price and quality for products produced by the same firm, there is a weaker, but still substantial, correlation between price and quality across firms. These results highlight the additional insights one can obtain by measuring appeal through the lens of our model. Finally, although markups are positively correlated with sales, they are, if anything, negatively correlated with firm price indices. The reason is that, although large firms have higher markups for each bar code, they produce so many more bar codes than small firms that their price index is lower.[31]

In sum, these correlations point to a number of ways in which our measures appear reasonable: cost, price, and appeal are all positively correlated at the bar-code and firm level. In addition, appeal is less correlated with marginal costs across firms than across products within firms. Most importantly for what follows, the fact that appeal and scope are highly correlated with firm sales but average marginal cost is almost uncorrelated with it provides the intuition for why our decompositions are going to point to appeal and scope as being the principle drivers of firm heterogeneity.

## VI.C. Markups

We use our estimated elasticities of substitution to compute implied firm markups, as summarized in equations (7) and (8). These markups vary along two dimensions. The first is cross-sector variation which captures the fact that $\sigma_g^F$ varies systematically across product groups, and the second is within-sector variation which captures the fact that larger firms have higher markups within a sector than smaller firms. As discussed above, our procedure for estimating the elasticities $\{\sigma_g^U, \sigma_g^F\}$ is robust to the assumption of either price or quantity competition, because the firm markup differences out in equation (26) and we observe prices. However, the markup formula, and hence the decomposition of prices into markups and marginal cost, depends in an important way on the form of competition. In our baseline specification, we assume Bertrand competition in prices, but we

---

[30]Across products within firms, we find that large firms are characterized by more dispersion in product appeal. Regressing the coefficient of variation of log product appeal within firms on log firm sales, and controlling for the log number of products, we find a positive and statistically significant coefficient on log firm sales.

[31]Our estimates of the components of firm sales can be related to the literature on the firm-size distribution (*e.g.*, Rossi-Hansberg and Wright [2007]). In line with existing results in that literature, we reject the null hypothesis of an untruncated Pareto distribution for firm sales (the relationship between log rank and log size is concave). Consistent with our finding that firm appeal is the most important determinant of firm size, we also reject the null hypothesis of an untruncated Pareto distribution for firm appeal (the relationship between log rank and log size is again concave).

report the derivation of the markup under Cournot competition in quantities in the appendix. A key takeaway from this comparison is that Cournot competition generates substantially higher markups for larger firms than Bertrand competition if goods are substitutes. Additionally, as shown in the appendix, under Bertrand competition, the relative markups of firms depend solely on the distribution of market shares (and hence are robust to different estimates of the elasticity of substitution between firms). In contrast, under Cournot competition, the relative markups of firms depend on both the distribution of market shares and the elasticity of substitution between firms.

Table VIII reports the distribution of markups across firms in our 100 product groups. The first column reports the percentile of the markup distribution; the second column reports the Cournot markup; the third column reports the Bertrand markup; the fourth column reports the markup implied by our parameter estimates under monopolistic competition across firms (based on $\sigma^F$); and the fifth column reports the markup implied by our parameter estimates under monopolistic competition across UPCs (based on $\sigma^U$). The median markup (under either Bertrand or Cournot) is 31 percent, which is slightly lower than the Domowitz et al. (1988) estimate of 36 percent for U.S. consumer goods, and slightly above the median markups estimated by De Loecker and Warzynski (2012) for Slovenian data which range from 17 to 28 percent. This suggests that our markup estimates are reasonable in the sense that they do not differ greatly from those found in prior work.

Within product groups, the markups of most firms are essentially the same regardless of whether we assume firms compete through prices or quantities, because most firms have trivial market shares.[32] In contrast, the markups of the largest firms are substantially different. Table IX shows the distribution of the markups for the three largest firms by product group. In each case, we report the markup relative to the average markup in the product group, so that one can see how different the markup of the largest firm is. The median largest firm within product groups has a markup that is 24 percent larger than average if we assume competition is Bertrand and almost double the average if we assume competition is Cournot. However, as one can also see from this table, these markups drop off quite rapidly for the second and third-largest firms. Therefore, both modes of competition suggest that very few firms can exploit their market power.

## VI.D. Cannibalization Rate

Our estimated elasticities of substitution also determine the magnitude of the "cannibalization rate," defined as the partial elasticity of the sales of existing products with respect to the introduction of a standardized new product with a market share equal to the average market share of the firm's other goods ($S_{N_{fgt}^U}^U N_{fgt}^U = 1$ in (13)). As discussed above, a cannibalization rate of 0 corresponds to equal differentiation of a firm's products from its rivals and a rate of 1 corresponds to a world in which the products of firms are perfect substitutes. On average across product groups, we find a cannibalization rate for the median firm of around 0.50, which lies almost exactly between the benchmarks of perfect substitutes and equal differentiation. This cannibalization level implies that about half of the sales of a new product introduced by a firm comes from the sales of existing products and half from the sales of other firms. The fact that these cannibalization rates are much less than

---

[32]From the first to the tenth decile of the distribution of sales within product groups, the average Bertrand (Cournot) markup varies from 0.38 to 0.37 (0.40 to 0.37).

one underscores the fact that it is not appropriate to treat the goods supplied by a multiproduct firm as perfect substitutes nor is it appropriate to treat multiproduct firms as if the introduction of a new good had no impact on the sales of existing goods.

Cannibalization rates rise with firm size reflecting the fact that large firms take into account that more of their sales of new products comes at the expense of their existing products. However, we find an average cannibalization rate across all deciles of the firm size distribution that ranges between 0.50 and 0.51, which reflects the fact that even in the upper decile of the size distribution most firms have trivial market shares. In contrast, when the largest firm within a sector (which has an average market share of 22 percent) introduces a new product, we estimate that 61 percent of the sales of that product comes from the sales of its existing products. Since we saw from Table III that the largest firms typically have the most products per firm, this implies that cannibalization is likely to be a first-order issue for them.[33]

## VI.E.   Decomposing Firm Sales

The first panel of Table X presents the decompositions described in equations (18) to (22) for the full sample of firms and the second presents the results for the largest firms (those with a market share in excess of 0.5 percent). We run the results over both sets of firms to see if there are any differences between them. Overall, the results suggest a powerful role for firm appeal in accounting for the size of firms. The variance decomposition indicates that 76 percent of the overall size distribution can be attributed to firm appeal and 21 percent is attributed to scope. Thus, "total firm appeal" accounts for 97 percent of variation in firm sales. If we restrict ourselves to looking at the data for the fifty largest firms we still find that firm appeal explains on average 55 percent of the variation with scope accounting for 26 percent. In other words, even if we restrict our sample to the largest firms, variation in firm appeal accounts for around 80 percent of the variation across firms.

What is perhaps most surprising is the small role played by cost in the determination of firm size. There are two important features of this decomposition to bear in mind. First, a positive coefficient means that larger firms have *lower* costs: *i.e.,* cost differentials help explain size differentials. Second, our decomposition allows us to split firm-cost differentials into two components: average marginal cost differentials and the dispersion of costs across products ("cost dispersion"). In all cross-sectional decompositions, the cost dispersion term is positively correlated with firm size. The intuition is simple. Larger firms tend to sell more products and therefore are better positioned to benefit from cost differentials in supplying their consumption index. The results for average marginal costs are more mixed. In the full sample, we find almost no association between firm size and average costs, indicating that larger firms have essentially the same average marginal costs as smaller firms. Their total costs are lower because they can take advantage of the cost dispersion in their output. Interestingly, we find a stronger association with average costs when we restrict our sample to only the largest fifty firms. Here, we find that 17 percent of the firm-size distribution can be attributed to lower average costs

---

[33]In Section S11 of the online supplement, we compare our estimates of cannibalization effects to those in the marketing literature. In making this comparison, we note that the marketing literature uses different measures of cannibalization effects, and our model emphasizes that cannibalization effects vary across product groups (with elasticities of substitution) and across firms (with expenditure shares). Despite these caveats, our findings of substantial cannibalization effects are broadly in line with the consensus in the existing empirical marketing literature.

of large firms and a quarter of the overall distribution is explained by these cost differences. As discussed in Section VI.B., this ambiguous pattern of results reflects two offsetting effects. On the one hand, higher marginal costs directly reduce firm sales through a higher price. On the other hand, higher marginal costs are correlated with higher firm appeal, which raises firm sales. Finally markup variation, which depends on both appeal and cost through expenditure shares, plays a relatively small role in explaining firm sales dispersion.[34]

Figure I presents the results of the cross-sectional decomposition for the largest 50 firms in each product group. Each point represents the average contribution of a particular factor towards understanding the relative level of sales in a sector. For example, the column of points at the right of the figure indicate that the largest firm in a product group is on average almost nine log units larger than the average firm. Of this nine-log-unit difference the diamond-shaped point indicates that on average about 4.9 log units (54 percent) can be attributed to firm appeal differences. The circular point reveals that about 2.3 log units (25 percent) is on average attributable to scope differences, while the triangular and square points indicate that lower average costs and greater cost dispersion account for about 1.3 (14 percent) and 0.6 log units (7 percent) of sales. These percentages are extremely similar to the percentages for larger firms that we saw in Table X. However, the fact that each set of points is relatively tightly arrayed along a line indicates that the results presented in Table X are not the result of any outlier but are a robust feature of the data.[35]

There are two other features of the data that this plot also makes clear. First, we can see that while there is a tendency for average marginal costs to fall with size for the four largest firms, which accounts for the upward trend in the sequence of triangular points, this upward trend is not a robust feature of the data. This pattern helps explain why we found in Table X that average marginal costs did not fall with firm size when we looked at the full sample of firms, but did fall when we restricted our attention to the very largest firms. Moreover, we can see that the gains from increasing cost dispersion are a robust feature of the data.

The ability of the largest firms to exploit market power is also clearly seen in Figure I. We saw in Table X that while markup variation played almost no role in understanding the firm-size distribution when we focused on the full sample of firms, it was negatively associated with firm sales when we examined the largest firms. Figure I shows that this negative association comes from the fact that if it were not for the fact that the very largest two or three firms had higher markups (a result we also saw in Table IX), their output levels would be higher. The implied inefficiently low levels of output of the largest firms suggest that the monopolistic competition model breaks down for these firms and motivates our attempt to quantify the implications of this pattern for aggregate consumer price indices in Section VIII..

As discussed in Section IV.F., we can also examine the time-series determinants of firm growth by differencing with respect to time (*i.e.*, using $\Delta^{g,t}$ instead of $\Delta^g$ in the decompositions). As shown in Table X, appeal matters even more for firm growth. Virtually all of a firm's growth can be understood in terms of increases in appeal and scope, with cost changes playing almost no role. Interestingly, product upgrading—*i.e.*, the addition

---

[34]Consistent with these findings, Roberts et al. (2011) find substantial heterogeneity in demand across Chinese footwear producers, while Foster, Haltiwanger and Syverson (2015) emphasize the role of demand in the growth of entering firms.

[35]In principle, large firms could supply disproportionately more products than small firms for strategic reasons outside the model (*e.g.*, to gain shelf space). If this were the case, one would expect the scatter of points corresponding to scope in Figure I to be convex. However, the relationship between size and scope is remarkably linear, which suggests that we do not observe this type of strategic behavior on average, although this does not rule out it mattering in some sectors.

of new higher appeal products—plays little role in the typical firm's expansion. However, one must be cautious in the interpretation of this result for a number of reasons. First, this result does not include the important role played by scope as opposed to upgrading. Thus, if a firm innovates by adding new products of roughly the same appeal, we will measure this as an expansion in scope, not appeal. Since we do find that changes in product scope play an important role in understanding firm growth, our findings should not be interpreted as suggesting that firms do not grow by adding new products. Indeed, over a quarter of the growth of large firms comes from this route.

Second, the result masks some important cross-sectoral heterogeneity. Some of our product groups, like frozen juice, flour, and eggs, exhibit low rates of bar-code turnover reflecting the fact that these are mature sectors with little potential for innovation. In these sectors, there may be little scope for product upgrading. However, in the three highest turnover sectors (disposable diapers, photographic supplies, and electronics), we might expect to see product upgrading play a more important role. In order to investigate this, we created a measure of innovation in each product group. Following Broda and Weinstein (2010), we defined the product turnover rate as the total number of bar codes in 2004 that did not survive until 2010 plus the number of new bar codes in 2010 divided by the total number of bar codes that existed in 2004 and 2010. By construction, this index will equal zero if there is no turnover of bar codes and one if all bar codes in 2010 are new. The index ranges from 0.3 for flour and 0.96 for photographic supplies, which corresponds to our priors about the importance of product innovation in these two sectors.

In order to see how important this heterogeneity is for understanding product upgrading, we regressed the importance of product upgrading for firm growth against how much product turnover occurs in a product group. We obtained a coefficient of 0.24 (s.e. 0.08). Similarly, we obtained a coefficient of 0.49 (s.e. 0.10) when we regressed the contribution of scope towards firm growth on product turnover. In other words, there is a clear statistical association between how important developing new and improved products are for understanding firm growth and the overall innovation rate in the sector. Firms in sectors with high rates of product turnover need to grow by developing new and improved products even more rapidly than their competitors, while this channel is almost absent in low turnover sectors. These differences are economically significant. New and improved products (scope and turnover) account for 64 percent of firm growth in our three highest turnover sectors, but only 10 percent of firm growth in the three lowest turnover sectors. While product upgrading alone accounts for 15 percent of firm growth in the three highest turnover sectors, it only accounts for 0.2 percent of firm growth in our three lowest turnover sectors.

Taken together, these decomposition results are quite striking in their consistency about the sources of firm heterogeneity. Regardless of whether we examine firms in the cross-section or in the time-series, improvements in firm appeal followed by product scope are the most important drivers of firm sales for both small and large firms alike. The role played by average marginal cost differences tends to be small and depend on the sample and decomposition method used. But we consistently find a positive contribution towards differences in firm size from the cost dispersion term. Finally, we see that although markup differences are unimportant determinants of firm sales for most firms, they do appear to be important for the very largest firms, something that suggests aggregate implications from the exploitation of market power by these firms.

# VII. Robustness

In this section, we demonstrate the robustness of our main finding that firm appeal accounts for most of the variation in firm sales to three sets of potential concerns: (i) theoretical specification, (ii) econometric estimation, and (iii) data and measurement.

## VII.A. Theoretical Specification

The first class of concerns focuses on our nesting structure. One limitation of our approach is that to structurally estimate the model we need to assume a particular nesting structure for demand consisting of firms and products within firms. We choose this nesting structure to connect with the existing literature on measuring firm productivity (which implicitly or explicitly treats firms as a nest) and as a natural approach to modeling multi-product firms. But it is reasonable to wonder whether our results are robust to other nesting structures.

In the online supplement, we consider a number of generalizations of our approach, including other nested demand structures (Sections S2 and S3), the introduction of an additional nest for brands (Section S4), and a CES upper tier of utility (Section S5). In this section, we provide further evidence that our main empirical finding about the importance of appeal for sales variation is not sensitive to our assumed nesting structure. To do so, we use the CES functional form to derive a reduced-form equation that can be estimated for alternative nesting structures. Although this reduced-form equation cannot be used to recover structural parameters, it can be used to provide a lower bound on the importance of variation in appeal for alternative nesting structures.

We combine the following two results for CES demand. First, from (30), double-differenced expenditure shares depend on double-differenced price indices for any tier of utility $J$. Second, from (32), the price index for this tier of utility $J$ can be expressed in terms of expenditure shares and the geometric mean of prices for the lower tier of utility $K$. Together these two results yield the following reduced-form equation that can be estimated using OLS:

$$\Delta^{j,t} \ln S_{jgt}^J = \alpha_g + \beta_{1g} \Delta^{j,t} \ln \tilde{P}_{jgt}^K + \beta_{2g} \Delta^{j,t} \ln \left[ \sum_{k \in \Omega_{jgt}^K} \frac{S_{kt}^K}{\tilde{S}_{jgt}^K} \right] + \epsilon_{jgt}, \tag{35}$$

where $j \in \Omega_{gt}^J$ indexes varieties in utility tier $J$ and $k \in \Omega_{jgt}^K$ indexes varieties in the lower utility tier $K$; $\Delta^{j,t}$ is the double-difference operator across varieties $j$ and over time $t$; we allow the parameters $\{\alpha_g, \beta_{1g}, \beta_{2g}\}$ to vary across product groups $g$; $\tilde{P}_{jgt}^K = \exp\left\{ \frac{1}{N_{jgt}^K} \sum_{k=1}^{N_{jgt}^K} \ln P_{kt}^K \right\}$ and the error $\epsilon_{jgt}$ captures double-differenced appeal ($\Delta^{j,t} \ln \varphi_{jgt}^J$).

In our structural model, $\beta_{2g}$ can be related to structural parameters of the model ($\beta_{2g} = 1/\left(\sigma_g^K - 1\right)$). In the reduced-form regression (35), this coefficient does not have a structural interpretation, because the term in appeal for tier $j$ ($\epsilon_{jgt}$) can be correlated with the geometric mean of prices in the lower tier $k$ ($\tilde{P}_{jgt}^K$). Therefore, if the reduced-form regression (35) is estimated using OLS, the coefficient $\beta_{2g}$ is subject to omitted variable bias and is not consistently estimated. For this reason, we use our structural approach developed above to estimate the model's parameters for our assumed nesting structure.

Nonetheless, the reduced-form specification (35) can be used to provide a lower bound on the importance of appeal for alternative nesting structures. The reason that this specification provides a lower bound is that changes in appeal are captured in the estimated residual ($\widehat{\epsilon}_{jgt}$). Therefore, the component of appeal that is correlated with the explanatory variables is attributed to the coefficients $\{\beta_{1g}, \beta_{2g}\}$ on these explanatory variables, leaving only the orthogonal component of appeal in the estimated residual ($\widehat{\epsilon}_{jgt}$).

In Table XI, we report the average $R^2$ across all product groups from estimating this reduced-form regression for alternative nesting structures, where one minus this $R^2$ corresponds to the variance in double-differenced expenditure shares that is unexplained by the explanatory variables (and is instead explained by appeal). We estimate the regression separately for each product group and quarter, and summarize the mean and standard deviation of the $R^2$, weighting each product group and quarter by their sales.

In the first row of the table, we report results for our assumed nesting structure. In the remaining rows of the table, we report results for a wide range of alternative nesting structures. While there are 98 product groups, Nielsen subdivides these into approximately 1,000 "modules," which are finer categories (*e.g.,* "Nutritional Supplements" is a product group, and Multi-Vitamins" is a module"). "Firm-modules" correspond to all bar codes in a module sold by a firm (*e.g.,* "Wyeth") and "Brand-modules" are even more narrow nests consisting of all bar codes sold by a firm in a module under a particular brand (*e.g.,* "Centrum" and "Centrum Performance"). Some of these alternative nesting structures disaggregate firms relative to our baseline specification (*e.g.,* Specification II uses Product Group and Brand-Module instead of Product Group and Firm). Other alternative nesting structures omit firms as a nest altogether (*e.g.,* Specification IV uses Module and Brand-Module). Across this wide range of specifications, we find that appeal accounts for at least 43 percent of the variation in double-differenced sales. Thus, regardless of one's choice of nesting structure, it is hard to escape the conclusion that around half of sales variation is attributable to variation in appeal. Since this lower bound also leaves out the important role played by scope, this further reinforces the importance of total appeal (including scope) in explaining sales variation.

## VII.B. Econometric Estimation

Our firm sales decomposition depends on the model's parameters and hence a second class of concerns relates to the estimation of these parameters. At the simplest level, one might worry about whether our estimation procedure in subsections V.B.-V.C. can successfully recover the model's parameters. To address this concern, we undertake a Monte Carlo simulation (section S8 of the online supplement), in which we show that our estimation procedure consistently recovers the true parameters when the data are generated from the model. Therefore our estimation procedure works when our identifying assumptions are satisfied.

However, what if our identifying assumptions are not satisfied? As a first step to addressing this concern, we report a number of additional specification checks. First, we use the fact that we have data on different brands supplied by firms within product groups. Therefore, instead of double differencing relative to a firm's largest product, we double difference relative to a firm's largest product within the same brand. We continue to find a similar pattern of results as in our baseline specification, which suggests that our estimates are not

driven by shocks to both demand and supply for brands.[36] Second, under our identifying assumption that double-differenced demand and supply shocks are orthogonal to one another, the time interval over which we difference should not affect the results. In contrast, if demand and supply shocks are correlated, the time interval over which we difference could matter, because for example of seasonal fluctuations in demand. Estimating over time intervals ranging from a quarter to a year, we continue to find a similar pattern of results, consistent with our identifying assumption.[37]

Nonetheless, there could be further econometric concerns or data issues that are not fully addressed in these robustness checks. Fortunately, it is easy to deal with concerns about whether our particular elasticity estimates are driving our firm sales decomposition results. Put simply, if our decomposition results are influenced by an estimation problem, then we should obtain different results from this decomposition using different values for the model's parameters. We therefore can dispose of this concern by examining the sensitively of our decomposition of firm sales (15) to *all* reasonable estimates of the model's parameters.

In Section S7 of the online supplement, we report the results of a grid search over a wide range of values for the elasticities of substitution $\{\sigma_g^U, \sigma_g^F\}$. In particular, we consider a range of possible values for $\sigma^F$ at intervals of 0.5 from 2 (half our median estimate for $\sigma^F$) to 12 (four times our median estimate for $\sigma^F$) and $\sigma^U$ ranging from 2.5 (just under half the median estimate) to 20 (almost triple the median estimate). Consistent with our empirical estimates above, we restrict attention to values on this parameter grid for which $\sigma_g^U > \sigma_g^F > 1$. For each parameter configuration, we replicate our firm sales decomposition (15), which holds regardless of the value for $\delta_g$ because it is based on observed prices. Naturally, the precise relative contribution of each component varies depending on parameter values. But we find that firm appeal and scope account for more than 80 percent of firm sales variation (and firm appeal alone accounts for at least 50 percent of firm sales variation) across the wide range of parameter values on the grid. *In other words, the result that firm appeal and scope are the main determinants of firm size is a deep feature of looking at the world through a CES setup and is not dependent on our particular elasticity estimates.* The underlying feature of the data driving this result is the substantial variation in firm sales conditional on price. For plausible values of the elasticity of substitution, the model cannot explain this sales variation by price variation, and hence it is attributed to appeal. While we derive our results for a CES setup, we conjecture that this underlying feature of the data would generate a substantial role for demand shifters in explaining firm sales variation for a range of plausible demand systems.

## VII.C.   Data and Measurement

A third class of concerns relates to data and measurement. We use data on bar codes because they correspond closely to the economic concept of a product and hold observable product characteristics constant. But one might worry that marketing, taste and other non-technical dimensions of appeal might be more important for consumer non-durables than for traded goods (which include intermediate inputs) or consumer durables. Alternatively, one might worry that because we do not have data on wholesale prices, we are missing important

---

[36]Differencing within brands (rather than within firms), we find median estimates for $\sigma^U$ of 6.49 (compared to 6.93 in our baseline specification) and for $\sigma^F$ of 3.74 (compared to 3.92 in our baseline specification).

[37]Differencing over years rather than over quarters, we find median estimates for $\sigma^U$ of 7.06 (compared to 6.93 in our baseline specification) and for $\sigma^F$ of 4.37 (compared to 3.92 in our baseline specification).

variation in retail markups or marketing. To address these concerns, Section S9 of the online supplement re-estimates our econometric specification using Chilean international trade transactions data. Despite the use of different types of data and different definitions of products (*e.g.*, bar codes versus 8-digit Harmonized System (HS) products), we continue to find that most of firm sales variation is explained by firm appeal and scope. Therefore, our findings are not dependent on our data and apply to other types of goods.

## VIII.    Firm and Aggregate Implications

Thus far, we have been focused on using our structural model to quantify the sources of firm heterogeneity. However, our results also have implications for the measurement of firm productivity and the aggregate consequences of departures from the benchmark of monopolistic competition. We now consider each of these implications in turn.

### VIII.A.    Firm Output and Productivity

All productivity estimates are based on a concept of real output, which in theory equals nominal output divided by a price index, but the choice of price index is not arbitrary: it is determined by the utility function. In other words, for multiproduct firms, the concept of real output is not independent of the demand system, so all attempts to measure productivity based on a real output concept contain an implicit assumption about the structure of the demand system.[38]

We can see how much it matters whether a researcher takes into account the fact that virtually all firms produce more than one product by considering two approaches to measuring firm output. The first is a "common sense" or conventional price index, $\mathbb{P}_{fgt}^{F}$, which is a quantity-weighted average price of a firm's output. The second is the CES price index, which equals nominal output divided by the minimum expenditure necessary to generate a unit of utility. According to equation (1) the quality-adjusted flow of consumption from a firm's output is $\varphi_{fgt}^{F} C_{fgt}^{F}$, so the corresponding expenditure function for a unit of consumption is $P_{fgt}^{F}/\varphi_{fgt}^{F}$.

Obviously, different price index formulas will yield different measures of prices, but there are two issues: how big is the measurement error and is the common-sense approach biased? It's easy to see that the answer to the second question must be "yes." Equation (32) indicates that, in the CES system, the use of average goods prices to measure firm-level prices will overstate the price level (and understate real output) more for large, multiproduct firms than for small, single-product ones. Therefore, if the true model is CES but a researcher uses a conventional price index, the results are likely to underestimate the output of large firms relative to small firms.

We can understand the magnitudes of the measurement error and bias of conventional approaches through a simple exercise. If we denote the consumer's expenditure on a firm's output by $E_{fgt}^{F}$, we can write the conventional measure of real output, $\mathbb{Q}_{fgt}^{F}$, as $E_{fgt}^{F}/\mathbb{P}_{fgt}^{F}$, and the CES measure of real output, $Q_{fgt}^{F}$, as $E_{fgt}^{F}/\left(P_{fgt}^{F}/\varphi_{fgt}^{F}\right)$. These measures will vary across sectors due to the units, so in order to compare size vari-

---

[38]Since firms supply differentiated products, the relevant measure of productivity is TFPR (Foster, Haltiwanger and Syverson [2008]), which necessarily depends on prices and hence the specification of demand.

ation across sectors we will work with unitless shares of each firm in sectoral output (*i.e.*, $Q_{fgt}^F / \sum_{f \in \Omega_{gt}^F} Q_{fgt}^F$ and $\mathbb{Q}_{fgt}^F / \sum_{f \in \Omega_{gt}^F} \mathbb{Q}_{fgt}^F$).

Figure II plots these two measures of real output. We place the CES market shares on the horizontal axis because the conventional index is only correct in the special case in which the elasticity of substitution is infinite and therefore output measures computed using the conventional index are likely to be measured with error. As one can see in Figure II, the choice of price index matters enormously for the computation of real output. If one regresses $\mathbb{Q}_{fgt}^F / \sum_{f \in \Omega_{gt}^F} \mathbb{Q}_{fgt}^F$ on $Q_{fgt}^F / \sum_{f \in \Omega_{gt}^F} Q_{fgt}^F$, one only obtains an $R^2$ of 0.77 with a coefficient of 0.65 (s.e. 0.004). These results imply that about a quarter of the variation of firm-level real output one would obtain by using a conventional price index is simply due to the fact the conventional price index assumes that firms produce homogeneous output: an assumption that easily can be rejected. Moreover, studies based on conventional measures of firm prices understate the real output of large firms by a third relative to small firms, with implications for estimates of returns to scale and productivity. Even if one does not believe that the CES framework is the right one to measure demand, the results mean that the implicit choice of demand system contained in a price index is of first-order importance in the measurement of real output.

## VIII.B. Aggregate Consumer Price Indices

We now turn to the implications of our estimates for aggregate consumer price indices. We undertake two counterfactuals to highlight the quantitative relevance of our departures from the benchmark of single-product, monopolistically competitive firms. First, we consider the implications of a regulator preventing large firms from exercising their market power and thereby requiring each firm to compete under monopolistic competition. Second, we consider the implications of allowing firms to supply a number of differentiated products. Thus, this second counterfactual examines the effect on the consumer price index from the different varieties supplied by each firm.

In each counterfactual, we report results for both Bertrand and Cournot competition. In each case, we hold constant aggregate expenditure, factor prices, the sets of active UPCs and firms $\left\{\Omega_{fgt}^U, \, \Omega_{gt}^F\right\}$, firm and product appeal $\left\{\varphi_{fgt}^F, \, \varphi_{ut}^U\right\}$, and the cost shifters $\{a_{ut}\}$. Therefore we abstract from the effect of changes in profits on aggregate expenditure. We also abstract from general equilibrium effects through firm entry and exit. Our goal is not to capture the full general equilibrium effects of each counterfactual, but rather to show that our firm-level estimates are of quantitative relevance for aggregate consumer price indices.

In the first set of counterfactuals (Counterfactuals I-B and I-C), we compare the actual equilibrium in which firms exploit their market power with a counterfactual equilibrium in which a price regulator requires all firms to charge the same constant markup over marginal cost equal to the monopolistically competitive firm markup:

$$\mu_{fgt}^F = \frac{\sigma_g^F}{\sigma_g^F - 1}. \tag{36}$$

Moving from the actual to the counterfactual markup reduces the relative prices of successful firms with high market shares and hence redistributes market shares towards these firms. Therefore, we expect these counterfactuals to reduce the consumer price index and increase the dispersion of firm sales, but the magnitude

of these effects depends on the estimated parameters $\left\{ \sigma_g^U,\ \sigma_g^F,\ \delta_g \right\}$, firm and product appeal $\left\{ \varphi_{fgt}^F,\ \varphi_{ut}^U \right\}$, and the cost shifters $\left\{ a_{ut}^U \right\}$.

Given the assumed constant markup in equation (36), we solve for the counterfactual equilibrium using an iterative procedure to solve for a fixed point in a system of five equations for expenditure shares and price indices $\left\{ S_{ut}^U, P_{ut}^U,\ S_{fgt}^F, P_{fgt}^F,\ P_{gt}^G \right\}$. These five equations are the UPC expenditure share (given by equation (4) for UPCs), the UPC pricing rule (equation (9)), the firm expenditure share (given by equation (4) for firms), the firm price index (given by (3) for firms) and the product group price index (given by equation (3) for product groups). We use the resulting solutions for product group price indices to compute the aggregate consumer price index: $\prod_{g=1}^G \left( P_{gt}^G \right)^{\varphi_{gt}}$.

In Counterfactuals II-B and II-C, we examine the quantitative relevance of multiproduct firms. We compare the actual equilibrium in which firms supply multiple products to a counterfactual equilibrium in which firms are restricted to supply a single UPC (their largest). Counterfactual II-B undertakes this comparison assuming Bertrand competition, while Counterfactual II-C undertakes the same comparison assuming Cournot competition. We again solve for the counterfactual equilibrium by solving for a fixed point in the system of five equations for expenditure shares and price indices discussed above.

Table XII reports the results from both sets of counterfactuals. Each row of the table corresponds to a different counterfactual. Each column reports the value of a variable in the counterfactual relative to that in the actual data. The second column reports the relative coefficient of variation of firm sales. The third column reports the relative consumer price index.

In Counterfactual I-B, we find that moving from Bertrand to monopolistic competition reduces the consumer price index by around 4 percent, which is quantitatively relevant compared to standard estimates of the welfare gains from trade for an economy such as the United States. In Counterfactual I-C, we find much larger effects from moving from Cournot to monopolistic competition, which reduces the consumer price index by around 13 percent. These much larger effects reflect the fact that large firms charge much higher markups under Cournot than under Bertrand competition as we saw in Table IX. This result implies larger price reductions from moving to monopolistically competitive markups, and hence larger reductions in the consumer price index. In both Counterfactuals I-B and I-C, the relative prices of the largest firms fall and demand is elastic. Therefore, the coefficient of variation of firms sales increases, by around 11 percent for Bertrand competition and around 39 percent for Cournot competition. Again, the larger effects for Cournot reflect the greater variation in markups under this mode of competition.

In both Counterfactuals II-B and II-C, we find quantitatively large effects of multiproduct firms. Restricting firms to supply a single product under Bertrand competition increases the consumer price index by around 49 percent and reduces the coefficient of variation of firm sales by around 17 percent. The corresponding increase in the consumer price index and reduction in the coefficient of variation of firm sales under Cournot competition are similar: 47 percent and 14 percent respectively. These results suggest that consumers benefit enormously from the range of products supplied within a single firm.

All of these counterfactuals are subject to a number of caveats. In particular, they do not allow for firm entry or exit, because they hold the set of firms constant, and they abstract from the effect of changes in profits on

aggregate expenditure. Nonetheless, these counterfactuals establish the quantitative relevance of our departures from the case of monopolistic competition with single product firms.

# IX.  Conclusions

We develop and structurally estimate a model of heterogeneous multiproduct firms that can be used to decompose the firm-size distribution into the contributions of costs, appeal, scope and markups. Our framework requires only price and expenditure data and hence is widely applicable. We use these price and expenditure data to estimate the key parameters of the model, namely the elasticities of substitution between and within firms, and the elasticities of marginal cost with respect to output. We use the resulting parameter estimates and the structure of the model to recover overall firm and product appeal, marginal costs, the number of products a firm supplies, and the firm's markup.

Our results point to differences in firm appeal as being the principal reason why some firms are large and others are not. Depending on the specification considered, we find that 50-75 percent of the variance in firm size can be attributed to differences in appeal, about 20-25 percent to differences in product scope, and less than 20 percent to average marginal cost differences. If we use a broad measure of total firm appeal, which encompasses both firm appeal as well as scope, we find that total firm appeal accounts for almost all of firm size differences.

We estimate substantially higher elasticities of substitution between varieties within firms than between firms (median elasticities of 6.9 and 3.9 respectively), implying that a firm's introduction of new product varieties cannibalizes the sales of existing varieties. We estimate that the cannibalization rate for the typical firm is 0.50, roughly half-way in-between the extreme of no cannibalization (equal elasticities of substitution within and between monopolistically competitive firms), and the extreme of complete cannibalization (varieties perfectly substitutable within firms).

We find that the typical sector comprises a few large firms with substantial market shares and a competitive fringe of firms with trivial market shares. Therefore most firms charge markups close to the monopolistic competition benchmark of constant markups, because they have trivial market shares and hence are unable to exploit their market power. However, the largest firms that account for up to around 25 percent of sales within sectors charge markups between 24 and 100 percent higher than the median firm. Using the estimated model to undertake counterfactuals, we find that these departures from the monopolistically competitive benchmark raise aggregate consumer price indices by between 4 and 13 percent.

Our findings that firms supply multiple imperfectly substitutable varieties have important implications for the measurement of firm productivity, highlighting the role of assumptions about demand in the measurement of productivity for multiproduct firms. Conventional price indices based on a weighted average of firm prices do not take into account that the theoretical price index for the firm depends on the number of products it supplies whenever consumers care about variety. Indeed, our counterfactual exercise indicates that the multiple varieties supplied by multiproduct firms reduce the aggregate consumer price index by around one third. Moreover, ignoring this effect introduces a systematic bias into the measurement of firm productivity, because larger firms

supply more products than smaller firms. We find this bias to be quantitatively large. An increase in firm sales is associated with around a one third larger increase in true real output (using the ideal price index) than in measured real output (using the conventional price index).

Federal Reserve Board of Governors


Princeton University, NBER and CEPR


Columbia University and NBER

# A   Appendix A: Derivations

## AA.   Derivation of Equations (7)-(9)

The first-order condition with respect to the price of an individual UPC implies:

$$Y_{ut}^U + \sum_{k=\underline{u}_{fgt}}^{\underline{u}_{fgt}+N_{fgt}^U} \left( P_{kt}^U \frac{dY_{kt}^U}{dP_{ut}^U} - \frac{dA_{kt}\left(Y_{kt}^U\right)}{dY_{kt}^U} \frac{dY_{kt}^U}{dP_{ut}^U} \right) = 0. \tag{37}$$

Using equation (5) and setting UPC supply equal to demand we have

$$\frac{\partial Y_{kt}^U}{\partial P_{ut}^U} = \left(\sigma^F - 1\right) \frac{Y_{kt}^U}{P_{gt}^G} \frac{\partial P_{gt}^G}{\partial P_{ut}^U} + \left(\sigma^U - \sigma^F\right) \frac{Y_{kt}^U}{P_{fgt}^F} \frac{\partial P_{fgt}^F}{\partial P_{ut}^U} - \sigma_u \frac{Y_{kt}^U}{P_{ut}^U} \frac{\partial P_{ut}^U}{\partial P_{ut}^U}.$$

We now can use equation (4) to solve for the elasticities and rewrite $\frac{\partial Y_k}{\partial P_u}$ as

$$
\begin{aligned}
\frac{\partial Y_{kt}^U}{\partial P_{ut}^U} &= \left(\sigma^F - 1\right) \left(\frac{\partial P_{gt}^G}{\partial P_{fgt}^F} \frac{P_{fgt}^F}{P_{gt}^G}\right) \left(\frac{\partial P_{fgt}^F}{\partial P_{ut}^U} \frac{P_{ut}^U}{P_{fgt}^F}\right) \frac{Y_{kt}^U}{P_{ut}^U} + \left(\sigma^U - \sigma^F\right) \left(\frac{\partial P_{fgt}^F}{\partial P_{ut}^U} \frac{P_{ut}^U}{P_{fgt}^F}\right) \frac{Y_{kt}^U}{P_{ut}^U} - \sigma^U \frac{Y_{kt}^U}{P_{ut}^U} 1_{\{u=k\}} \\
&= \left(\sigma^F - 1\right) S_{fgt}^F S_{ut}^U \frac{Y_{kt}^U}{P_{ut}^U} + \left(\sigma^U - \sigma^F\right) S_{ut}^U \frac{Y_{kt}^U}{P_{ut}^U} - \sigma^U \frac{Y_{kt}^U}{P_{ut}^U} 1_{\{u=k\}}. \tag{38}
\end{aligned}
$$

If we now substitute equation (38) into equation (37) and divide both sides by $Y_u^U$, we get

$$1 + \sum_k \left(\sigma^F - 1\right) S_{fgt}^F S_{ut}^U \frac{P_{kt}^U Y_{kt}^U}{P_{ut}^U Y_{ut}^U} + \sum_k \left(\sigma^U - \sigma^F\right) S_{ut}^U \frac{P_{kt}^U Y_{kt}^U}{P_{ut}^U Y_{ut}^U} - \sigma^U$$

$$- \sum_k \left(\sigma^F - 1\right) S_{fgt}^F S_{ut}^U \frac{\frac{\partial A_{kt}\left(Y_{kt}^U\right)}{\partial Y_{kt}^U} Y_{kt}^U}{P_{ut}^U Y_{ut}^U} - \sum_k \left(\sigma^U - \sigma^F\right) S_{ut}^U \frac{\frac{\partial A_{kt}\left(Y_{kt}^U\right)}{\partial Y_{kt}^U} Y_{kt}^U}{P_{ut}^U Y_{ut}^U} + \sigma_u \frac{\frac{\partial A_{kt}\left(Y_{kt}^U\right)}{\partial Y_{kt}^U}}{P_{ut}^U} = 0. \tag{39}$$

We define the markup as $\mu_{kt} \equiv P_{kt}^U \big/ \frac{\partial A_{kt}\left(Y_{kt}^U\right)}{\partial Y_{kt}^U}$. Since $S_{ut}^U \frac{1}{P_{ut}^U Y_{ut}^U} = \frac{1}{\sum_k P_{kt}^U Y_{kt}^U}$ and therefore $\sum_k S_{ut}^U \frac{P_{kt}^U Y_{kt}^U}{P_{ut}^U Y_{ut}^U} = 1$, we can rewrite equation (39) as

$$1 + \left(\sigma^F - 1\right) S_{fgt}^F + \left(\sigma^U - \sigma^F\right) - \sigma^U - \left(\sigma^F - 1\right) S_{fgt}^F \frac{\sum_k \frac{\partial A_{kt}\left(Y_{kt}^U\right)}{\partial Y_{kt}^U} Y_{kt}^U}{\sum_k P_{kt}^U Y_{kt}^U} - \left(\sigma^U - \sigma^F\right) \frac{\sum_k \frac{\partial A_{kt}\left(Y_{kt}^U\right)}{\partial Y_{kt}^U} Y_{kt}^U}{\sum_k P_{kt}^U Y_{kt}^U} + \sigma^U \frac{1}{\mu_{ut}} = 0.$$

Because we assume that $\sigma^U$ is the same for all $u$ produced by a firm, $\mu_{ut}$ is the only $u$-specific term in this expression. Hence, $\mu_{ut}$ must be constant for all $u$ produced by firm $f$ in time $t$; in other words, *markups only vary at the firm level*. Together these two results ensure the same markup across all UPCs supplied by the firm.

We can now solve for $\mu_{ft}^F$ by

$$
\begin{aligned}
1 + \left(\sigma^F - 1\right) S_{fgt}^F + \left(\sigma^U - \sigma^F\right) - \sigma^U - \left(\sigma^F - 1\right) S_{fgt}^F \frac{1}{\mu_{ft}^F} - \left(\sigma^U - \sigma^F\right) \frac{1}{\mu_{ft}^F} + \sigma^U \frac{1}{\mu_{ft}^F} &= 0 \\
\Rightarrow \mu_{ft}^F &= \frac{\sigma^F - \left(\sigma^F - 1\right) S_{fgt}^F}{\sigma^F - \left(\sigma^F - 1\right) S_{fgt}^F - 1}.
\end{aligned}
$$

## AB. Derivation of Equation (13)

Equilibrium price indices $P_{gt}^G$ and $P_{fgt}^F$ are functions of the number of products $N_{fgt}^U$. Recall that demand for UPC $u$ supplied by firm $f$ within product group $g$ is:

$$Y_{ut}^U = \left(\varphi_{fgt}^F\right)^{\sigma^F-1} \left(\varphi_{ut}^U\right)^{\sigma^U-1} E_{gt}^G \left(P_{gt}^G\right)^{\sigma^F-1} \left(P_{fgt}^F\right)^{\sigma^U-\sigma^F} \left(P_{ut}^U\right)^{-\sigma^U}.$$

Now suppose that the number of UPCs supplied by the firm $(N_{fgt}^U)$ can be approximated by a continuous variable and consider the partial elasticity of $Y_{ut}^U$ with respect to this measure of UPCs:

$$
\begin{aligned}
\frac{\partial Y_{ut}^U}{\partial N_{fgt}^U} &= \left(\sigma^F - 1\right) \frac{Y_{ut}^U}{P_{gt}^G} \frac{\partial P_{gt}^G}{\partial N_{fgt}^U} + \left(\sigma^U - \sigma^F\right) \frac{Y_{kt}^U}{P_{fgt}^F} \frac{\partial P_{fgt}^F}{\partial N_{fgt}^U}, \\
&= \left(\sigma^F - 1\right) \frac{Y_{ut}^U}{P_{gt}^G} \frac{N_{fgt}^U}{P_{fgt}^F} \frac{P_{fgt}^F}{N_{fgt}^U} \frac{\partial P_{gt}^G}{\partial P_{fgt}^F} \frac{\partial P_{fgt}^F}{\partial N_{fgt}^U} + \left(\sigma^U - \sigma^F\right) \frac{Y_{kt}^U}{P_{fgt}^F} \frac{N_{fgt}^U}{N_{fgt}^U} \frac{\partial P_{fgt}^F}{\partial N_{fgt}^U}, \\
&= \left(\sigma^F - 1\right) \left(\frac{\partial P_{gt}^G}{\partial P_{fgt}^F} \frac{P_{fgt}^F}{P_{gt}^G}\right) \left(\frac{\partial P_{fgt}^F}{\partial N_{fgt}^U} \frac{N_{fgt}^U}{P_{fgt}^F}\right) \frac{Y_{ut}^U}{N_{fgt}^U} + \left(\sigma^U - \sigma^F\right) \left(\frac{\partial P_{fgt}^F}{\partial N_{fgt}^U} \frac{N_{fgt}^U}{P_{fgt}^F}\right) \frac{Y_{kt}^U}{N_{fgt}^U},
\end{aligned}
$$

where this partial elasticity holds constant the prices of existing UPCs. Using equation (3) for the firm price index and (4) for the UPC expenditure share, we have:

$$
\frac{\partial P_{fgt}^F}{\partial N_{fgt}^U} \frac{N_{fgt}^U}{P_{fgt}^F} = \frac{\left(\frac{P_{N_{fgt}^U}^U}{\varphi_{N_{fgt}^U}^U}\right)^{1-\sigma^U}}{1-\sigma^U} \left(P_{fgt}^F\right)^{\sigma^U} \frac{N_{fgt}^U}{P_{fgt}^F} = \frac{N_{fgt}^U}{1-\sigma^U} \frac{\left(\frac{P_{N_{fgt}^U}^U}{\varphi_{N_{fgt}^U}^U}\right)^{1-\sigma^U}}{\sum_{n=1}^{N_{fgt}^U} \left(\frac{P_{ut}^U}{\varphi_{ut}^U}\right)^{1-\sigma^U}} = \frac{N_{fgt}^U}{1-\sigma^U} S_{N_{fgt}^U}^U.
$$

Recalling that expenditure shares are also the elasticities of the price index with respect to prices, we obtain:

$$
\begin{aligned}
\text{Cannibalization} \equiv -\frac{\partial Y_{ut}^U}{\partial N_{fgt}^U} \frac{N_{fgt}^U}{Y_{ut}^U} &= -\left(\sigma^F - 1\right) S_{fgt}^F \frac{N_{fgt}^U}{1-\sigma^U} S_{N_{fgt}^U}^U - \left(\sigma^U - \sigma^F\right) \frac{N_{fgt}^U}{1-\sigma^U} S_{N_{fgt}^U}^U \\
&= \left[\left(\frac{\sigma^U - \sigma^F}{\sigma^U - 1}\right) + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right) S_{fgt}^F\right] S_{N_{fgt}^U}^U N_{fgt}^U > 0, \qquad \text{for } \sigma^U \geq \sigma^F > 1.
\end{aligned}
$$

Note that this derivation is based on solely on the derivatives of the CES price index with respect to the measure of UPCs and observed prices. Since the assumption of Bertrand or Cournot competition merely affects the decomposition of observed prices into markups and marginal costs, these derivatives are the same under both forms of competition. Therefore the derivation of the cannibalization rate is the same under both forms of competition.

## AC. Cournot Quantity Competition

In our baseline specification in the main text above, we assume that firms choose prices under Bertrand competition. In this appendix, we discuss a robustness test in which we assume instead that firms choose quantities under Cournot competition. Each firm chooses the number of UPCs $(N_{fgt}^U)$ and their quantities $(Y_{ut}^U)$ to maximize its profits:

$$\max_{\{\underline{u}_{fgt},\ldots,\overline{u}_{fgt}\},\{Y_{ut}^U\}} \Pi_{fgt}^F = \sum_{u=\underline{u}_{fgt}}^{\underline{u}_{fgt}+N_{fgt}^U} P_{ut}^U Y_{ut}^U - A_{ut}\left(Y_{ut}^U\right) - N_{fgt}^U H_{gt}^U - H_{gt}^F, \tag{40}$$

where we again index the UPCs supplied by the firm from the largest to the smallest in sales. From the first-order conditions for profit maximization, we obtain the equilibrium markup:

$$\mu_{fgt}^F = \frac{\varepsilon_{fgt}^F}{\varepsilon_{fgt}^F - 1}, \tag{41}$$

where the firm's perceived elasticity of demand is now:

$$\varepsilon_{fgt}^F = \frac{1}{\frac{1}{\sigma^F} - \left(\frac{1}{\sigma^F} - 1\right) S_{fgt}^F}, \tag{42}$$

and the firm's pricing rule is:

$$P_{ut}^U = \mu_{fgt}^F \frac{dA_{ut}\left(Y_{ut}^U\right)}{dY_{ut}^U} = \mu_{fgt}^F \left[(1 + \delta) a_{ut} \left(Y_{ut}^U\right)^\delta\right]. \tag{43}$$

Therefore the analysis with Cournot competition in quantities is similar to that with Bertrand competition in prices, except that firms' perceived elasticities of demand ($\varepsilon_{fgt}^F$) and hence their markups differ between these two cases. Since firms internalize their effects on product group aggregates, markups are again variable. Furthermore, markups are common across UPCs within a given firm and product group, but vary across product groups within firms for the reasons discussed in the main text above.

Our estimation procedure for $\{\sigma^U, \sigma^F, \delta\}$ remains entirely unchanged because the firm markup differences out when we take double differences between a pair of UPCs within the firm over time. Therefore our estimates of the parameters $\{\sigma^U, \sigma^F, \delta\}$ are robust to the assumption of either Bertrand or Cournot competition. Our solutions for firm and product quality $\{\varphi^U, \varphi^F, \delta\}$ are also completely unchanged, because they depend solely on observed expenditure shares and prices $\left\{S_{ut}^U, S_{fgt}^F, P_{ut}^U\right\}$ and our estimates of the parameters $\{\varphi^U, \varphi^F, \delta\}$. Therefore whether we assume Cournot or Bertrand competition is only consequential for the decomposition of observed UPC prices ($P_{ut}^U$) into markups ($\mu_{fgt}^F$) and marginal costs ($a_{ut}$). Markups are somewhat more variable in the case of Cournot competition, which strengthens our finding of substantial departures from the monopolistically competitive markup for the largest firms within each product group.

## AD. Relative Firm Markups Under Bertrand Competition

For firms with non-negligible shares of expenditure within a product group ($0 < S_{fgt}^F < 1$):

$$\mu_{fgt}^F - 1 = \left(\frac{\sigma^F - (\sigma^F - 1)S_{fgt}^F}{\sigma^F - (\sigma^F - 1)S_{fgt}^F - 1}\right) - 1$$

For a firm with a negligible share of expenditure within a product group ($S_{fgt}^F \approx 0$):

$$\mu_{fgt}^F - 1 = \left(\frac{\sigma^F}{\sigma^F - 1}\right) - 1.$$

The ratio of the two markups (for $\sigma^F$ finite and $S_{fgt}^F < 1$) is:

$$\frac{1}{1 - S_{fgt}^F}.$$

Therefore the dispersion of relative markups under Bertrand competition depends solely on the dispersion of expenditure shares.

## AE. Relative Firm Markups Under Cournot Competition

For firms with non-negligible shares of expenditure within a product group ($0 < S^F_{fgt} < 1$):

$$\mu^F_{fgt} - 1 = \left( \frac{\frac{1}{\sigma^F} - \left( \frac{1}{\sigma^F} - 1 \right) S^F_{fgt}}{\frac{1}{\sigma^F} - \left( \frac{1}{\sigma^F} - 1 \right) S^F_{fgt} - 1} \right) - 1.$$

For a firm with a negligible share of expenditure within a product group ($S^F_{fgt} \approx 0$):

$$\mu^F_{fgt} - 1 = \left( \frac{\sigma^F}{\sigma^F - 1} \right) - 1.$$

The ratio of the two markups (for $\sigma^F$ finite and $S^F_{fgt} < 1$) is:

$$\frac{1 + S^F_{fgt}(\sigma^F - 1)}{1 - S^F_{fgt}}.$$

Therefore, the dispersion of relative markups under Cournot competition is increasing in the elasticity of substitution ($\sigma^F$), and depends on both this elasticity and the dispersion of expenditure shares.

# References

[1] Agur, Itai, "Trade Liberalization, Firm Selection and Variety Growth," *Review of International Economics*, 18 (2010), 582-594.

[2] Allanson, Paul and Catia Montagna, "Multiproduct Firms and Market Structure: An Explorative Application to the Product Life Cycle," *International Journal of Industrial Organization*, 23 (2005), 587-597.

[3] Amiti, Mary, Oleg Itskhoki, and Jozef Konings, "Importers, Exporters and Exchange Rate Disconnect," *American Economic Review*, 104 (2014), 1942-78.

[4] Anderson, Simon P., André de Palma, and Jacques-François Thisse, *Discrete Choice Theory of Product Differentiation*, (Cambridge, MA: MIT Press, 1992).

[5] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodriguez-Clare, "New Trade Models, Same Old Gains?" *American Economic Review*, 102 (2012), 94-130.

[6] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodriguez-Clare, "The Elusive Pro-Competitive Effects of Trade," NBER Working Paper No., w21370, 2015.

[7] Arkolakis, Costas and Marc-Andreas Muendler (2010) "The Extensive Margin of Exporting Products: A Firm-Level Analysis," *NBER Working Paper*, 16641.

[8] Atkeson, Andrew and Ariel Burstein, "Pricing to Market, Trade Costs and International Relative Prices," *American Economic Review*, 98 (2008), 1998-2031.

[9] Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott, "Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods," in *Producer dynamics: New Evidence from Micro Data*, T. Dunne, J. B. Jensen and M. J, Roberts, ed. (Chicago: University of Chicago Press, 2009).

[10] Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott, "Multiproduct Firms and Product Switching," *American Economic Review*, 100 (2010), 70-97.

[11] Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott, "Multiproduct Firms and Trade Liberalization," *Quarterly Journal of Economics*, 126 (2011), 1271-1318.

[12] Berry, Steven T., "Estimating Discrete-choice Models of Product Differentiation," *RAND Journal of Economics*, 25 (1994), 242-262.

[13] Broda, Christian and David E. Weinstein, "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121 (2006), 541-86.

[14] Broda, Christian and David E. Weinstein, "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 100 (2010), 691-723.

[15] Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan, "Business Cycle Accounting," *Econometrica*, 75 (2007), 781-836.

[16] Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi, "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93 (2003), 15-37.

[17] Das M, Mark J. Roberts, and James R. Tybout, "Market Entry Costs, Producer Heterogeneity and Export Dynamics," *Econometrica*, 75 (2007), 837-73.

[18] De Loecker, Jan and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 102 (2012), 2437-2471.

[19] De Loecker, Jan, Penny Goldberg, Amit Khandelwal, and Nina Pavcnik, "Prices, Markups and Trade Reform," *Econometrica*, forthcoming, 2015.

[20] Di Comite, F., J. Thisse, and H. Vandenbussche, "Verti-zontal Differentiation in Export Markets", *Journal of International Economics*, 93 (2014), 50-66.

[21] Dhingra, Swati, "Trading Away Wide Brands for Cheap Brands," *American Economic Review*, 103 (2013), 2554-84.

[22] Di Giovanni, Julian and Andrei Levchenko, "Country Size, International Trade, and Aggregate Fluctuations in Granular Economies," *Journal of Political Economy*, 120 (2012), 1083-1132.

[23] Domowitz, Ian, R. Glenn Hubbard, and Bruce C. Petersen, "Business Cycles and the Relationship between Concentration and Price-Cost Margins," *The RAND Journal of Economics*, 17 (1988), 1-17.

[24] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, "Dissecting Trade: Firms, Industries, and Export Destinations," *The American Economic Review*. Papers and Proceedings, , 94 (2004), 150-154.

[25] Eckel, Carsten, Leonardo Iacavone, Beata Javorcik, and J. Peter Neary, "Multiproduct Firms at Home and Away: Cost Versus Quality-Based Competence," *Journal of International Economics*, 95 (2015), 216-232.

[26] Eckel, Carsten and J. Peter Neary, "Multiproduct Firms and Flexible Manufacturing in the Global Economy," *Review of Economic Studies*, 77 (2010), 188-217.

[27] Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "Competition, Markups, and the Gains from International Trade," *American Economic Review*, forthcoming, 2015.

[28] Eslava, Marcela, Cecilia Fieler, and Daniel Yi Xu, "Trade, Skills and Quality Upgrading: A Theory with Evidence from Colombia," NBER Working Paper No., w19992, 2014.

[29] Feenstra, Robert C., "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84 (1994), 157-177.

[30] Feenstra, Robert C., "Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity," NBER Working Paper No., 19833, 2014.

[31] Feenstra, Rob and Hong Ma, "Optimal Choice of Product Scope for Multiproduct Firms," in *The Organization of Firms in a Global Economy*, Elhanan Helpman, Dalia Marin and Thierry Verdier, ed. (Cambridge MA: Harvard University Press, 2008).

[32] Feenstra, Robert C. and John Romalis, "International Prices and Endogenous Quality,"*Quarterly Journal of Economics*, forthcoming, 2015.

[33] Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?"*American Economic Review*, 98 (2008), 394-425.

[34] Foster, Lucia, John Haltiwanger, and Chad Syverson, "The Slow Growth of New Plants: Learning about Demand," *Economica*, forthcoming, 2015.

[35] Gabaix, Xavier, "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79 (2011), 733-772.

[36] Gordon, Brett R., Avi Goldfarb, and Yang Li, "Does Price Elasticity Vary with Economic Growth? A Cross-Category Analysis," *Journal of Marketing Research*, L (2013), , 4-23.

[37] Hallak, Juan Carlos and Peter K. Schott, "Estimating Cross-Country Differences in Product Quality,"*Quarterly Journal of Economics*, 126 (2011), 417-474.

[38] Handbury, Jessie and David E. Weinstein, "Goods Prices and Availability in Cities," *Review of Economic Studies*, 82 (2015), 258-296.

[39] Johnson, Robert C., "Trade and Prices with Heterogeneous Firms," *Journal of International Economics*, 86 (2012), 43-56.

[40] Khandelwal, Amit K., "The Long and Short (of) Quality Ladders," *Review of Economic Studies*, 77 (2010), 1450-1476.

[41] Leontief, Wassily, "Ein Versuch Zur Statistichen Analyse von Angebot und Nachfrage," *Weltwirtschaftliches Archiv*, 30 (1929), 1-53.

[42] Manova, Kalina and Zhiwei Zhang, "Export Prices across Firms and Destinations," *Quarterly Journal of Economics*, 127 (2012), 379-436.

[43] Mayer, Thierry, Marc Melitz, and Gianmarco I. P. Ottaviano, "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 104 (2014), 495-536.

[44] Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71 (2003), 1695-725.

[45] Melitz, Marc J. and Gianmarco I. P. Ottaviano, "Market Size, Trade, and Productivity," *Review of Economic Studies*, 75 (2008), 295-316.

[46] Nocke, Volker and Stephen R. Yeaple, "Globalization and Endogenous Firm Scope," *International Economic Review*, 55 (2014), 993-1018.

[47] Roberts, Mark J., Xiaoyan Fan, Daniel Xu, and Shengxing Zhang, "A Structural Model of Demand, Cost, and Export Market Selection for Chinese Footwear Producers," NBER Working Paper No., 17725, 2011.

[48] Rossi-Hansberg, Esteban and Mark L. J. Wright, "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, 97 (2007), 1639-66.

[49] Sato, Kazuo, "The Ideal Log-Change Index Number," *Review of Economics and Statistics*, 58 (1976), 223-228.

[50] Shaked, Avner and John Sutton, "Natural Oligopolies," *Econometrica*, 51 (1983), 1469-1484.

[51] Schott, Peter K., "Across-product Versus Within-product Specialization in International Trade," *Quarterly Journal of Economics*, 119 (2004), 647-678.

[52] Stock, James H., Jonathan H. Wright, and Motohiro Yogo, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20 (2002), 518-529.

[53] Sutton, John, *Sunk Costs and Market Structure* (Cambridge MA: MIT Press, 1991).

[54] Vartia, Yrjo, "Ideal Log-Change Index Numbers," *Scandinavian Journal of Statistics*, 3 (1976), 121-126.

## TABLE I: Sample Statistics

| | Mean | Median | Std. Dev. | 10th Percentile | 90th Percentile | Max Value |
|---|---|---|---|---|---|---|
| No. Firms per Product Group | 512 | 410 | 309 | 200 | 947 | 1434 |
| Firm Sales | 4176 | 153 | 25386 | 5 | 4661 | 380108 |
| Log Firm Sales | 12 | 12 | 3 | 8 | 15 | 19 |
| No. of UPCs per Firm | 13 | 4 | 32 | 1 | 33 | 388 |
| UPC Sales | 309 | 40 | 1186 | 2 | 645 | 34839 |

*Note:* Weighted by product group–quarter sales. Firm and UPC sales in thousands. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

## TABLE II: Size Distribution by Decile

| Ranked Decile | Decile Market Share | Mean Firm Market Share | Mean Log Firm Sales | Avg. Std. Dev. log UPC Sales | Mean No. UPCs per Firm | Median No. UPCs per Firm |
|---|---|---|---|---|---|---|
| 1 | 87.72 | 2.69 | 16.4 | 1.8 | 68.1 | 44.9 |
| 2 | 7.09 | 0.22 | 14.5 | 1.6 | 26.0 | 21.0 |
| 3 | 2.69 | 0.08 | 13.5 | 1.6 | 14.7 | 11.8 |
| 4 | 1.26 | 0.04 | 12.7 | 1.5 | 9.0 | 7.2 |
| 5 | 0.63 | 0.02 | 12.0 | 1.4 | 5.8 | 4.6 |
| 6 | 0.32 | 0.01 | 11.4 | 1.3 | 4.0 | 3.2 |
| 7 | 0.17 | 0.00 | 10.7 | 1.2 | 2.8 | 2.2 |
| 8 | 0.08 | 0.00 | 9.9 | 1.1 | 2.1 | 1.5 |
| 9 | 0.03 | 0.00 | 8.9 | 0.9 | 1.5 | 1.1 |
| 10 | 0.01 | 0.00 | 7.3 | 0.6 | 1.2 | 1.0 |

*Note:* Largest decile is ranked first. Weighted by product group–quarter sales. UPC counts and mean firm sales are across firm within each decile. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE III: Size Distribution by Firm Rank

| Firm Rank | Firm Market Share (%) | Log Firm Sales | No. UPCs per Firm |
|-----------|----------------------|----------------|-------------------|
| 1 | 22.0 | 19.4 | 276.7 |
| 2 | 12.7 | 18.8 | 198.6 |
| 3 | 8.0 | 18.4 | 158.1 |
| 4 | 5.5 | 18.1 | 121.5 |
| 5 | 4.2 | 17.8 | 117.4 |
| 6 | 3.4 | 17.6 | 113.7 |
| 7 | 2.8 | 17.4 | 107.6 |
| 8 | 2.4 | 17.2 | 94.4 |
| 9 | 2.1 | 17.1 | 87.5 |
| 10 | 1.9 | 17.0 | 74.4 |

*Note:* Largest firm is ranked first. Weighted by product group–quarter sales. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

## TABLE IV: Size Distribution by Number of UPCs

| No. of UPCs | No. of Firms | Share of Value (%) | Mean Sales (Thousands) | Median Sales (Thousands) | Avg. St. Dev. UPC Sales |
|---|---|---|---|---|---|
| 1 | 160 | 0.8 | 72 | 63871 | |
| 2–5 | 169 | 3.6 | 322 | 63945 | 1.29 |
| 6–10 | 63 | 4.8 | 1180 | 56272 | 1.50 |
| 11–20 | 48 | 8.8 | 2876 | 63067 | 1.56 |
| 21–50 | 44 | 19.0 | 7626 | 62678 | 1.65 |
| 51–100 | 18 | 19.1 | 23328 | 116088 | 1.72 |
| > 100 | 10 | 43.8 | 109999 | 122045 | 1.79 |

*Note:* Weighted by product group–quarter sales. To get last column, calculate standard deviation over log UPC sales by product group–quarter, then take weighted average across product group–quarters. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE V: Distribution of 100 GMM Estimates

| Ranked Percentile | $\sigma_g^U$ | $\sigma_g^F$ | $\sigma_g^U - \sigma_g^F$ | $\delta_g$ |
|---|---|---|---|---|
| 1 | 24.4 | 12.8 | 16.8 | 0.62 |
| 5 | 17.6 | 8.5 | 9.4 | 0.42 |
| 10 | 14.1 | 7.3 | 6.5 | 0.30 |
| 25 | 8.6 | 5.1 | 4.0 | 0.20 |
| 50 | 6.9 | 3.9 | 2.6 | 0.16 |
| 75 | 5.4 | 3.1 | 2.2 | 0.09 |
| 90 | 5.0 | 2.6 | 1.8 | 0.06 |
| 95 | 4.7 | 2.3 | 1.7 | 0.04 |
| 99 | 4.4 | 2.1 | 1.2 | 0.02 |

*Note:* Percentiles are decreasing: largest estimate is ranked first. $\sigma_g^U$ is the elasticity of substitution between bar codes sold by the same firm in a product group, $\sigma_g^F$ is the elasticity of substitution between firms within a product group, and $\delta_g$ is the elasticity of marginal cost with respect to output for firms in a product group. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE VI: Sales Component Correlations (UPC Level)

| | $\Delta^f \ln E_{ut}^U$ | $\Delta^f \ln P_{ut}^U$ | $\Delta^f \ln \varphi_{ut}^U$ | $\Delta^f \ln \gamma_{ut}$ |
|---|---|---|---|---|
| $\Delta^f \ln E_{ut}^U$ | 1.00 | | | |
| $\Delta^f \ln P_{ut}^U$ | 0.03 | 1.00 | | |
| $\Delta^f \ln \varphi_{ut}^U$ | 0.36 | 0.94 | 1.00 | |
| $\Delta^f \ln \gamma_{ut}$ | 0.03 | 1.00 | 0.94 | 1.00 |

*Note:* All correlations significant at the 0.1 percent level. $\Delta^f \ln E_{ut}^U$ is the log sales of each bar code less the log geometric mean sales of all bar codes sold by the firm. $\Delta^f \ln P_{ut}^U$ is the log price of each bar code less the log geometric mean price of all bar codes sold by the firm. $\Delta^f \ln \varphi_{ut}^U$ is the log appeal of each bar code. $\Delta^f \ln \gamma_{ut}^U$ is the log marginal cost of each bar code less the log geometric mean marginal cost of all bar codes sold by the firm. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

#### TABLE VII: Sales Component Correlations (Firm Level)

| | $\Delta^g \ln E_{ft}^F$ | $\Delta^g \ln P_{ft}^F$ | $\Delta^g \ln \varphi_{ft}^F$ | $\Delta^g \ln N_{ft}^U$ | $\Delta^g \ln \tilde{\gamma}_{ft}$ | $\Delta^g \ln CD_{ft}$ | $\Delta^g \ln \mu_{ft}^F$ |
|---|---|---|---|---|---|---|---|
| $\Delta^g \ln E_{ft}^F$ | 1.00 | | | | | | |
| $\Delta^g \ln P_{ft}^F$ | -0.20 | 1.00 | | | | | |
| $\Delta^g \ln \varphi_{ft}^F$ | 0.51 | 0.68 | 1.00 | | | | |
| $\Delta^g \ln N_{ft}^U$ | 0.79 | -0.35 | 0.24 | 1.00 | | | |
| $\Delta^g \ln \tilde{\gamma}_{ft}$ | 0.01 | 0.96 | 0.80 | -0.10 | 1.00 | | |
| $\Delta^g \ln CD_{ft}$ | 0.64 | -0.28 | 0.19 | 0.68 | -0.07 | 1.00 | |
| $\Delta^g \ln \mu_{ft}^F$ | 0.26 | -0.07 | 0.15 | 0.23 | -0.01 | 0.16 | 1.00 |

*Note:* All correlations significant at the 0.1 percent level. The difference operator $\Delta^g \ln x_{ft}$ denotes the value of log $x$ for firm $f$ in quarter $t$, less the log geometric mean value of $x$ among all firms in product group $g$ that quarter. The variable $x$ is one of the following: $E_{ft}^F$ is firm sales, $P_{ft}^F$ is the firm price index, $\varphi_{ft}^F$ is firm appeal, $N_{ft}^U$ is the number of products sold by a firm, $\tilde{\gamma}_{ft}^F$ is the average marginal cost among a firm's products,

$CD_{ft} = \frac{1}{N_{ft}^U} \sum_{u \in \Omega_{ft}^U} \left( \frac{\gamma_{ut}/\tilde{\gamma}_{ut}}{\varphi_{ut}^U} \right)^{1-\sigma_g^U}$ is the firm's cost dispersion, and $\mu_{ft}^F$ is the firm's markup.

## TABLE VIII: Distribution of Markups

| Ranked Percentile | Cournot | Bertrand | Monopolistic Competition ($\sigma_g^F$) | Monopolistic Competition ($\sigma_g^U$) |
|---|---|---|---|---|
| 10 | 0.57 | 0.57 | 0.52 | 0.24 |
| 25 | 0.45 | 0.45 | 0.45 | 0.23 |
| 50 | 0.31 | 0.31 | 0.31 | 0.16 |
| 75 | 0.24 | 0.24 | 0.24 | 0.14 |
| 90 | 0.16 | 0.16 | 0.16 | 0.09 |

*Note:* Percentiles are decreasing: largest is ranked first. Markup=(Price-Marginal Cost)/Marginal Cost. The "Cournot" column presents markups computed assuming firms compete as Cournot competitors. The "Bertrand" column presents markups computed assuming firms compete as Bertrand competitors. "Monopolistic Competition ($\sigma_g^F$)" and "Monopolistic Competition ($\sigma_g^U$)" presents markups constructed under the assumption that firms are monopolistic competitors with an elasticity of substitution of $\sigma_g^F$ or $\sigma_g^U$, respectively. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE IX: Distribution of Largest Firms' Markups Relative to Product Group Average

| Ranked Percentile | Cournot | | | Bertrand | | |
|---|---|---|---|---|---|---|
| | Largest | Second | Third | Largest | Second | Third |
| 10 | 4.36 | 2.61 | 1.81 | 1.72 | 1.29 | 1.14 |
| 25 | 2.97 | 1.73 | 1.46 | 1.41 | 1.18 | 1.11 |
| 50 | 1.94 | 1.50 | 1.30 | 1.24 | 1.13 | 1.08 |
| 75 | 1.55 | 1.33 | 1.23 | 1.16 | 1.09 | 1.06 |
| 90 | 1.36 | 1.23 | 1.16 | 1.10 | 1.07 | 1.05 |

*Note:* Percentiles are decreasing: largest is ranked first. Markup=(Price-Marginal Cost)/Marginal Cost. The "Cournot" column presents markups computed assuming firms compete as Cournot competitors. The "Bertrand" column presents markups computed assuming firms compete as Bertrand competitors. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE X: Variance Decompositions

| PANEL A: ALL FIRMS | | | | | | |
|---|---|---|---|---|---|---|
| Decomposition | Appeal | Scope | Average MC | Cost Dispersion | Markup | Upgrading |
| *Cross-Sectional* | 0.758 | 0.2103 | -0.037 | 0.0706 | -0.00195 | |
| | (0.002) | (0.0002) | (0.002) | (0.0001) | (0.00001) | |
| *Firm Growth* | 0.909 | 0.1581 | -0.135 | 0.0685 | -0.00060 | 0.0004 |
| | (0.009) | (0.0013) | (0.009) | (0.0010) | (0.00004) | (0.0014) |
| PANEL B: FIRMS WITH > 0.5% PRODUCT-GROUP SHARE | | | | | | |
| Decomposition | Appeal | Scope | Average MC | Cost Dispersion | Markup | Upgrading |
| *Cross-Sectional* | 0.547 | 0.2638 | 0.172 | 0.0614 | -0.04396 | |
| | (0.013) | (0.0016) | (0.013) | (0.0008) | (0.00016) | |
| *Firm Growth* | 0.906 | 0.2627 | -0.169 | 0.0270 | -0.02620 | 0.0226 |
| | (0.053) | (0.0080) | (0.052) | (0.0083) | (0.00129) | (0.0140) |

*Note: MC* is marginal cost  Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE XI: Within Nest $R^2$

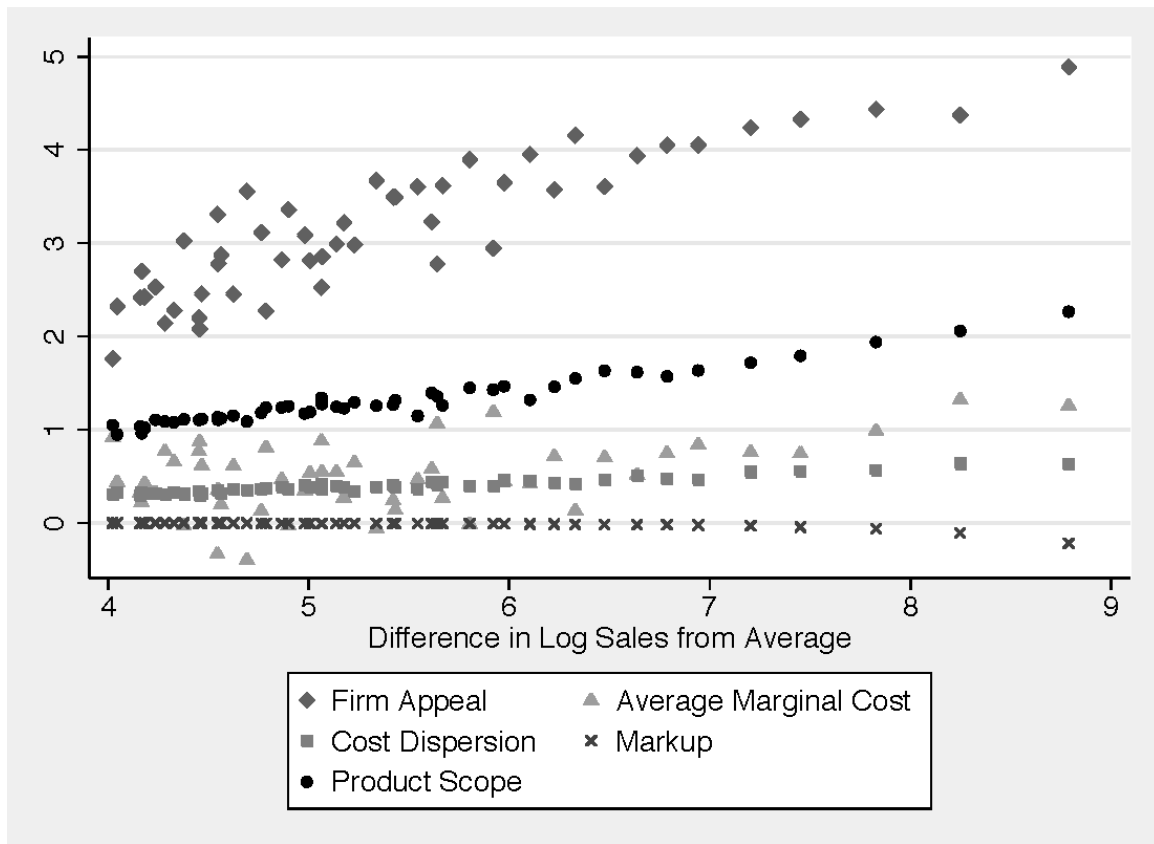| Specification | Nest 1 | Nest 2 | Nest 3 | Nest 4 | Mean | Std. Dev. |
|---|---|---|---|---|---|---|
| I | Product Group | Firm | UPC | N/A | 0.53 | 0.05 |
| II | Product Group | Brand-Module | UPC | N/A | 0.46 | 0.06 |
| III | Product Group | Firm-Module | UPC | N/A | 0.47 | 0.07 |
| IV | Module | Brand-Module | UPC | N/A | 0.48 | 0.06 |
| V | Module | Firm-Module | UPC | N/A | 0.49 | 0.05 |
| VI | Module | Firm-Module | Brand | UPC | 0.57 | 0.07 |

*Note:*  Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

TABLE XII: Counterfactual Exercises (Relative to Observed)

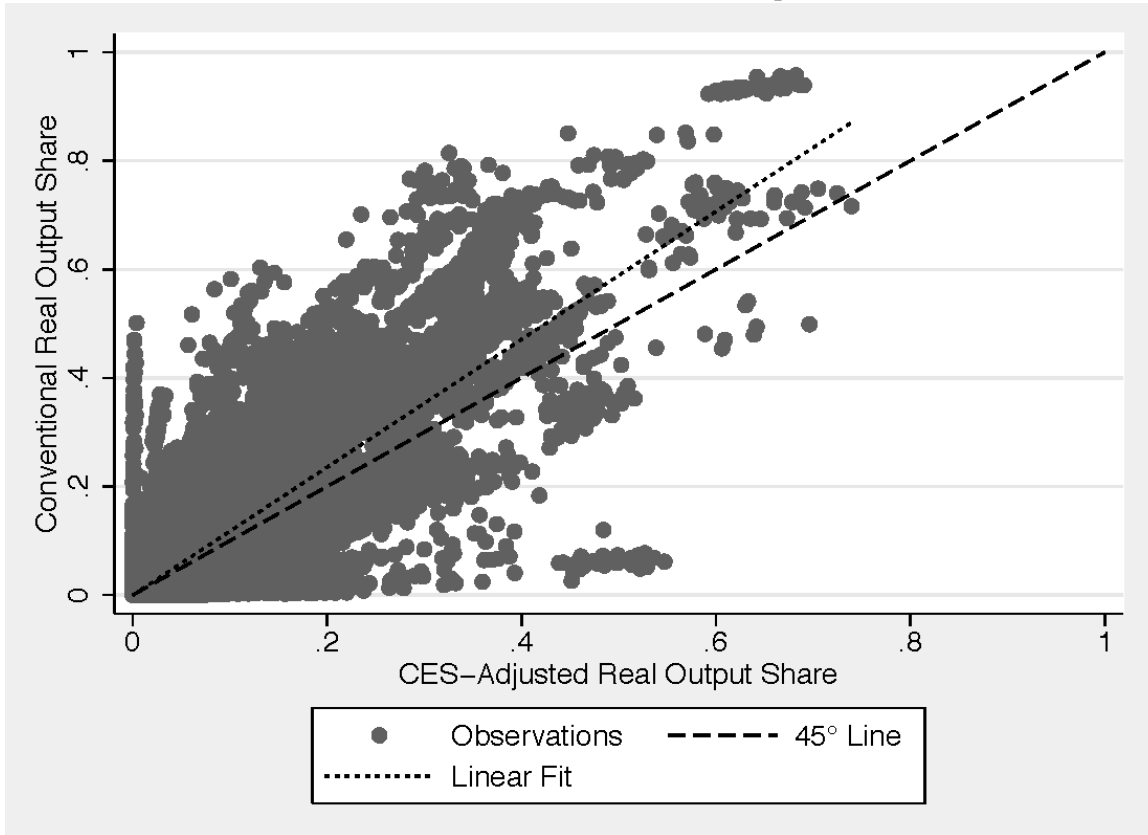| | Coef. Var. Firm Sales | Consumer Price Index |
|---|---|---|
| *I-B:* Bertrand to Monopolistic Competition | 1.11 | 0.96 |
| *I-C:* Cournot to Monopolistic Competition | 1.39 | 0.87 |
| *II-B:* Multiproduct to Single-Product Bertrand | 0.83 | 1.49 |
| *II-C:* Multiproduct to Single-Product Cournot | 0.86 | 1.47 |

*Note:* Each statistic is expressed as the value in the counterfactual divided by the value in the observed data. Calculations on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

FIGURE I: Sales Decomposition by Firm Rank



*Note:* The graph contains the weighted average across product groups (by quarter) of the decomposition of relative firm sales into the contributions of the firm appeal, marginal cost, markup and product scope for the fifty firms with the largest average market share in 2004 and 2011. The vertical axis measures the contribution of each factor given by the left-hand sides of equations (18) to (22). The horizontal axis is the log difference between the firm's sales and that of the average firm defined by the right-hand side of these equations. Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

FIGURE II: Bias in Conventional Real Output Share



Note: Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.