Online Supplement for Quantifying the Sources of Firm Heterogeneity (Not for Publication)*

Colin Hottman
Federal Reserve Board of Governors†

Stephen J. Redding
Princeton University and NBER‡

David E. Weinstein
Columbia University and NBER§

November 8, 2015

Contents

*Thanks to Ildiko Magyari and Scott Marchi for providing excellent research assistance. David Weinstein would like to thank the NSF (Award 1127493) for generous financial support. We also would like to thank GS1 for providing us with a concordance between the bar code data and firm identifiers. The views expressed are solely those of the authors and do not represent the views of the Board of Governors or other members of the Federal Reserve System. All 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.

†20th Street and Constitution Avenue N.W. Washington, D.C. 20551. Email: colin.j.hottman@frb.gov.

‡Fisher Hall, Princeton, NJ 08544. Email: reddings@princeton.edu.

§420 W. 118th Street, MC 3308, New York, NY 10027. Email: dew35@columbia.edu.

S1 Overview

This technical appendix contains additional supplementary material for the paper, "Quantifying the Sources of Firm Heterogeneity." Here, we consider a number of extensions and robustness checks:

- Section S2 analyzes the properties of nested demand systems in general. We show that the prediction that markups are the same across products within a firm-sector (but vary within firms across sectors) is a generic prediction of nested demand systems.

- Section S3 examines a symmetric nested translog demand system based on Feenstra and Weinstein (2015) as an alternative to the nested CES demand system considered in the paper. We show that the moment conditions used to estimate the model's parameters take a similar form as in the nested CES specification in the paper, except that the dependent variable in the demand equation is the change in the expenditure share rather than the change in the log expenditure share.
- Section S4 considers an extension of our nested CES demand system to introduce brands as an additional nest within firms. We show how this extension can generate variation in markups across brands within firm-sectors.
- Section S5 presents a robustness test in which we assume that the upper tier of utility is CES rather than Cobb-Douglas. In this robustness test, there is an incentive for firms to price strategically across product groups. We show that the firm pricing rule again takes the form of a common markup across UPCs within a given product group and a given firm. These common markups difference out from our moment conditions and hence leave our parameter estimates unchanged. Given these unchanged parameters, our main decomposition result for the importance of firm appeal remains unaffected (this extension merely changes the decomposition of observed prices into markups and costs).
- Section S6 shows that our empirical approach can allow for separate retail and production markups that are common across UPCs for a given product group and a given firm. These common markups again difference out from our moment conditions, leaving our parameter estimates and our main decomposition result for the importance of firm appeal unaffected (this extension again changes the decomposition of observed prices into markups and costs). This section also shows how our model of the production sector can be embedded in a simple explicit model of the retail sector.
- Section S7 demonstrates the robustness of our results to alternative values for the model's parameters. We undertake a grid search over the parameter space and show that we find that firm appeal and product scope account for most of the observed variation in firms sales for a wide range of plausible values for the model's parameters.
- Section S8 reports a Monte Carlo Simulation for our estimation procedure. We show that this procedure correctly recovers the true values of the model's parameters when the data are generated according to the model.
- Section S9 examines a robustness test in which we replicate our estimation procedure using Chilean international trade transactions data. Although these data include different categories of goods (including intermediates) and also use different definitions of sector and product (based on Harmonized System categories rather than bar codes), we again find that firm appeal and product scope account for most of the observed variation in firm sales.

- Section S10 compares our elasticity estimates to those of other studies using scanner data but different methodologies.

- Section S11 relates our estimates of cannibalization effects to those in the marketing literature.

- Section S12 reports the results of our firm sales decomposition for the ten largest product groups.

- Section S13 lists the individual product groups, the number of firms in each product group, and the share of each product group in total sales in our data.

S2 Nested Demand Systems

In this section, we show that the prediction that markups are the same across products within a firm-sector (but vary within firms across sectors) is a general feature of all homogeneous-of-degree-one nested utility functions when firms are small relative to the aggregate economy.

S2.1 Preferences

Utility is assumed to be a function of the consumption indices (C_g^G) of a number of sectors indexed by $g \in \Omega^G$:

$$\mathbb{U} = F\left(\left\{C_g^G\right\}\right), \tag{1}$$

where $F(\cdot)$ is homogeneous of degree one. The consumption index for each sector g is a function of the consumption indices (C_{fg}^F) of a number of firms indexed by $f \in \Omega_g^F$:

$$C_g^G = F_g\left(\left\{C_{fg}^F\right\}\right), \tag{2}$$

where $F_g(\cdot)$ is also homogeneous of degree one. The consumption index for each firm f within sector g is a function of consumption (C_u^U) of a number of products indexed by $u \in \Omega_{fg}^U$:

$$C_{fg}^F = F_{fg}\left(\left\{C_u^U\right\}\right), \tag{3}$$

where $F_{fg}(\cdot)$ is again homogeneous of degree one.

S2.2 Firm Problem

We assume that each firm is small relative to the aggregate economy (so firm pricing does not affect aggregate expenditure across all sectors) and that pricing and product introduction decisions are made at the sector level by each firm. The firm internalizes that it is the monopoly supplier of the consumption index C_{fg}^F in each sector and chooses the price for that consumption index (P_{fg}^F) to maximize its profits. The firm's problem is:

$$\max_{P_{fg}^F} P_{fg}^F C_{fg}^F - \gamma_{fg}^F C_{fg}^F,$$

where $\gamma_{fg}^{F} = \gamma_{fg}\left(\left\{\gamma_u^U\right\}\right)$ is the marginal cost of supplying the consumption index C_{fg}^{F}, which depends on the marginal cost for each product (γ_u^U), and is the dual to (3). Recall that $F_{fg}\left(\left\{C_u^U\right\}\right)$ is homogeneous of degree one in $\left\{C_u^U\right\}$. Therefore $\gamma_{fg}\left(\left\{\gamma_u^U\right\}\right)$ is homogeneous of degree one in $\left\{\gamma_u^U\right\}$. The first-order condition for profit maximization implies:

$$P_{fg}^{F} = \frac{\epsilon_{fg}^{F}}{\epsilon_{fg}^{F} - 1}\gamma_{fg}\left(\left\{\gamma_u^U\right\}\right), \qquad \epsilon_{fg}^{F} = -\frac{\partial C_{fg}^{F}}{\partial P_{fg}^{F}}\frac{P_{fg}^{F}}{C_{fg}^{F}}. \tag{4}$$

Hence, the firm chooses a mark-up over the cost of supplying the consumption index (C_{fg}^{F}) in each sector, where the size of this markup depends on the elasticity of demand (ϵ_{fg}^{F}), which only varies by firm and sector.

Now consider the firm's choice of the consumption of each product within the sector. Given the marginal cost for each product (γ_u^U), the firm chooses the set of products $u \in \left\{\underline{u}_{fg}, \ldots, \overline{u}_{fg}\right\}$ and their consumption $\{C_u^U\}$ to minimize the cost of supplying the consumption index:

$$\min_{\left\{\underline{u}_{fgt}, \ldots, \overline{u}_{fgt}\right\}, \left\{C_u^U\right\}} \sum_{u=\overline{u}_{fg}}^{\underline{u}_{fg}+N_{fg}^{U}} \gamma_u^U C_u^U - \lambda\left[F_{fg}\left(\left\{C_u^U\right\}\right) - C_f^F\right] - N_{fg}^{U} H_g^U - H_g^F,$$

where we index the products supplied by the firm within the sector 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 sector is denoted by N_{fg}^{U}, where $\overline{u}_{fg} = \underline{u}_{fg} + N_{fg}^{U}$.

The first-order conditions for the consumption of any two products u and k are:

$$\gamma_u^U - \lambda\frac{\partial F_{fg}\left(\{C_u^U\}\right)}{\partial C_u^U} = 0,$$

$$\gamma_k^U - \lambda\frac{\partial F_{fg}\left(\{C_k^U\}\right)}{\partial C_k^U} = 0,$$

which implies:

$$\frac{\gamma_u^U}{\gamma_k^U} = \frac{\partial F_{fg}\left(\{C_u^U\}\right)/\partial C_u^U}{\partial F_{fg}\left(\{C_k^U\}\right)/\partial C_k^U}, \tag{5}$$

where $\gamma_u^U = dA_u\left(C_u^U\right)/dC_u^U$. Cost minimization therefore implies that the firm equates relative marginal costs (γ_u^U/γ_k^U) and relative marginal utilities $\left(\partial F_{fg}\left(\{C_u^U\}\right)/\partial C_u^U\right) / \left(\partial F_{fg}\left(\{C_k^U\}\right)/\partial C_k^U\right)$. But utility maximization implies that consumers equate relative prices and relative marginal utilities:

$$\frac{P_u^U}{P_k^U} = \frac{\mu_u^U\gamma_u^U}{\mu_k^U\gamma_k^U} = \frac{\partial F_{fg}\left(\{C_u^U\}\right)/\partial C_u^U}{\partial F_{fg}\left(\{C_k^U\}\right)/\partial C_k^U}, \tag{6}$$

where the only way that (5) and (6) can be both satisfied is if the firm sets a common markup across products within the sector ($\mu_u^U = \mu_k^U = \mu_{fg}^F$), so that relative prices in (6) equal relative marginal costs in (5). Combining these results in (4), (5) and (6), we obtain:

$$P_u^U = \mu_{fg}^F\gamma_u^U, \qquad \mu_{fg}^F = \frac{\epsilon_{fg}^F}{\epsilon_{fg}^F - 1} \tag{7}$$

$$P_{fg}^{F} = \gamma_{fg}\left(\left\{P_{u}^{U}\right\}\right) = \frac{\epsilon_{fg}^{F}}{\epsilon_{fg}^{F}-1}\gamma_{fg}\left(\left\{\gamma_{u}^{U}\right\}\right),$$

where the second equation uses the fact that $\gamma_{fg}\left(\left\{P_{u}^{U}\right\}\right)$ is homogeneous of degree one in $\left\{P_{u}^{U}\right\}$.

Therefore, the prediction that a firm's markups are the same across products within a sector is a general prediction of a nested demand system and is not limited to a nested CES demand system. However, markups vary across sectors within the firm.

S3 Nested Translog Demand System

In this section, we consider a nested translog demand system based on the symmetric translog specification of Feenstra and Weinstein (2015). We show that the moment conditions used to estimate the model's parameters take a similar form as in the nested CES specification in the paper, except that the dependent variable in the demand equation is the change in the expenditure share rather than the change in the log expenditure share.

S3.1 Preferences

The upper tier of utility is defined over consumption indices for each sector g and takes the Cobb-Douglas form as in the baseline specification in the paper. The dual price index for each sector g (P_{gt}^{G}) is defined over the appeal-adjusted price indices for firms f within that sector ($P_{fgt}^{F}/\varphi_{fgt}^{F}$) and takes the translog form:

$$\ln P_{gt}^{G} = \alpha_{0g}^{F} + \sum_{f=1}^{\tilde{N}_{g}^{F}} \alpha_{fg}^{F} \ln\left(\frac{P_{fgt}^{F}}{\varphi_{fgt}^{F}}\right) + \frac{1}{2}\sum_{f=1}^{\tilde{N}_{g}^{F}}\sum_{k=1}^{\tilde{N}_{g}^{F}} \gamma_{fkg}^{F} \ln\left(\frac{P_{fgt}^{F}}{\varphi_{fgt}^{F}}\right)\ln\left(\frac{P_{kgt}^{F}}{\varphi_{kgt}^{F}}\right), \quad \text{with} \quad \gamma_{fkg}^{F} = \gamma_{kfg}^{F}, \tag{8}$$

where $\tilde{N}_{g}^{F}$ is the universe of possible firms within sector g and we impose the following symmetry restriction:

$$\gamma_{ffg}^{F} = -\gamma_{g}^{F}\left(\frac{\tilde{N}_{g}^{F}-1}{\tilde{N}_{g}^{F}}\right) < 0, \qquad \gamma_{fkg}^{F} = \frac{\gamma_{g}^{F}}{\tilde{N}_{g}^{F}} > 0 \qquad \text{for} \qquad f \neq k, \qquad f,k \in \left\{1,\ldots,\tilde{N}_{g}^{F}\right\}.$$

The dual price index for each firm f within sector g (P_{fgt}^{F}) is defined over the appeal-adjusted prices for products ($P_{ut}^{U}/\varphi_{ut}^{U}$) and also takes the translog form:

$$\ln P_{fgt}^{F} = \beta_{0g}^{U} + \sum_{u=1}^{\tilde{N}_{fg}^{U}} \beta_{ug}^{U} \ln\left(\frac{P_{ut}}{\varphi_{ut}}\right) + \frac{1}{2}\sum_{u=1}^{\tilde{N}_{fg}^{U}}\sum_{r=1}^{\tilde{N}_{fg}^{U}} \delta_{urg}^{U} \ln\left(\frac{P_{ut}^{U}}{\varphi_{ut}^{U}}\right)\ln\left(\frac{P_{rt}^{U}}{\varphi_{rt}^{U}}\right), \quad \text{with} \quad \delta_{urg}^{U} = \delta_{rug}^{U}, \tag{9}$$

where $\tilde{N}_{fg}^{U}$ is the universe of possible products for firm f within sector g and we impose the following symmetry restriction:

$$\delta_{uug}^{U} = -\delta_{g}^{U}\left(\frac{\tilde{N}_{fg}^{U}-1}{\tilde{N}_{fg}^{U}}\right) < 0, \qquad \delta_{urg}^{U} = \frac{\delta_{g}^{U}}{\tilde{N}_{fg}^{U}} > 0, \qquad \text{for} \qquad u \neq r, \qquad u,r \in \left\{1,\ldots,\tilde{N}_{fg}^{U}\right\}.$$

S3.2 Firm Expenditure shares

From the sector price index (8), we obtain the following expression for the expenditure share (S^F_{fgt}) of firm f within sector g at time t (which corresponds to equation (9) in Feenstra and Weinstein 2015):

$$S^F_{fgt} = \left(\alpha^F_{fg} + \alpha^F_{gt}\right) - \gamma^F_g \left(\ln\left(\frac{P^F_{fgt}}{\varphi^F_{fgt}}\right) - \overline{\ln P^F_{gt}}\right), \tag{10}$$

where:

$$\alpha^F_{gt} = \frac{1}{N^F_{gt}}\left(1 - \sum_{f\in\Omega^F_{gt}} \alpha^F_{fg}\right),$$

$$\overline{\ln P^F_{gt}} = \frac{1}{N^F_{gt}} \sum_{f\in\Omega^F_{gt}} \ln\left(\frac{P^F_{fgt}}{\varphi^F_{fgt}}\right),$$

where Ω^F_{gt} is the set of firms active in sector g at time t and N^F_{gt} is the number of elements in this set. Differencing relative to the largest firm in the sector (denoted by $\underline{f}$), we obtain:

$$\Delta^{\underline{f}} S^F_{fgt} = \left(\alpha^F_{fg} - \alpha^F_{\underline{f}g}\right) - \gamma^F_g \left(\Delta^{\underline{f}} \ln P^F_{fgt} - \Delta^{\underline{f}} \ln \varphi^F_{fgt}\right),$$

where $\Delta^{\underline{f}}$ denotes the difference operator relative to the largest firm in the sector (such that $\Delta^{\underline{f}} S_{fgt} = S_{fgt} - S_{\underline{f}gt}$). Differencing again over time, we obtain:

$$\Delta^{\underline{f},t} S^F_{fgt} = -\gamma^F_g \Delta^{\underline{f},t} \ln P^F_{fgt} + w^F_{fgt},$$

where $\Delta^{\underline{f},t}$ denotes the double-difference operator relative to the largest firm and over time (such that $\Delta^{\underline{f},t} S^F_{fgt} = \left(S^F_{fgt} - S^F_{\underline{f}gt}\right) - \left(S^F_{fgt-1} - S^F_{\underline{f}gt-1}\right)$); and $w^F_{fgt} = \gamma^F_g \Delta^{\underline{f},t} \ln \varphi^F_{fgt}$ is a stochastic error that captures double-differenced firm appeal.

Note that this demand equation takes a similar form as in our nested CES specification, except that the dependent variable is the change in the expenditure share rather than the change in the log expenditure share.

S3.3 Product Expenditure Shares

From the firm price index (9), we obtain the following expression for the expenditure share (S^U_{ut}) of product u within sector f and sector g at time t (which again corresponds to equation (9) in Feenstra and Weinstein 2015):

$$S^U_{ut} = \left(\beta^U_{ug} + \beta^U_{gt}\right) - \delta^U_g \left(\ln\left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right) - \overline{\ln P^U_{fgt}}\right), \tag{11}$$

where:

$$\beta^U_{gt} = \frac{1}{N^U_{fgt}}\left(1 - \sum_{u\in\Omega^U_{fgt}} \beta^U_u\right),$$

$$\overline{\ln P^U_{ft}} = \frac{1}{N^U_{fgt}} \sum_{u\in\Omega^U_{fgt}} \ln\left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right),$$

where Ω^U_{fgt} is the set of products supplied by firm f within sector g at time t and N^U_{fgt} is the number of elements in this set. Differencing relative to the largest product supplied by the firm (denoted by $\underline{u}$), we obtain:

$$\Delta^{\underline{u}} S^U_{ut} = \left(\beta^U_{ug} - \beta^U_{\underline{u}g}\right) - \delta^U_g \left(\Delta^{\underline{u}} \ln P^U_{ut} - \Delta^{\underline{u}} \ln \varphi^U_{ut}\right),$$

where $\Delta^{\underline{u}}$ denotes the difference operator relative to the largest firm in the sector (such that $\Delta^{\underline{u}} S^U_{ut} = S^U_{ut} - S^U_{\underline{u}t}$). Differencing again over time, we obtain:

$$\Delta^{\underline{u},t} S^U_{ut} = -\delta^U_g \Delta^{\underline{u},t} \ln P^U_{ut} + e^U_{ut},$$

where $\Delta^{u,t}$ denotes the double-difference operator relative to the largest product and over time (such that $\Delta^{\underline{u},t} S^U_{ut} = \left(S^U_{ut} - S^U_{\underline{u}t}\right) - \left(S^U_{ut-1} - S^U_{\underline{u}t-1}\right)$); and $e^U_{ut} = \delta^U_g \Delta^{\underline{u},t} \ln \varphi^U_{ut}$ is a stochastic error that captures double-differenced product appeal.

Note that this demand equation again takes a similar form as in our nested CES specification, except that the dependent variable is the change in the expenditure share rather than the change in the log expenditure share.

S3.4 Pricing Rule

Given the nested demand structure, the firm will again choose the monopoly price for supplying the consumption index as a whole (as shown in Section S2 above), which implies a common (variable) markup across all products within the nest. Therefore, our supply-side moment condition from the firm's pricing rule will take the same form as in our nested CES specification, because the common markup across products within firms will again difference out.

S4 Robustness Test with Additional Nest

In this section, we show how our framework can be extended to introduce an additional nest into the CES demand system, so that it includes product groups, firms, brands and UPCs. We show that this extension can generate variation in markups across brands within firm-sectors (in addition to the variation in markups across sectors within firms in our baseline specification).

S4.1 Preferences

Utility is a Cobb-Douglas aggregate of real consumption of each product group:

$$\ln \mathbb{U}_t = \int_{g \in \Omega^G} \varphi^G_{gt} \ln C^G_{gt} dg, \qquad \int_{g \in \Omega^G} \varphi^G_{gt} dg = 1. \tag{12}$$

Product group consumption indices are a CES function of real consumption of each firm:

$$C^G_{gt} = \left[\sum_{f \in \Omega^F_{gt}} \left(\varphi^F_{fgt} C^F_{fgt}\right)^{\frac{\sigma^F_g - 1}{\sigma^F_g}}\right]^{\frac{\sigma^F_g}{\sigma^F_g - 1}}, \qquad \sigma^F_g > 1,\ \varphi^F_{fgt} > 0. \tag{13}$$

Firm consumption indices are a CES function of real consumption for each brand:

$$C_{fgt}^{F} = \left[\sum_{b \in \Omega_{fgt}^{B}} \left(\varphi_{bt}^{B} C_{bt}^{B} \right)^{\frac{\sigma_g^B - 1}{\sigma_g^B}} \right]^{\frac{\sigma_g^B}{\sigma_g^B - 1}}, \qquad \sigma_g^B > 1, \, \varphi_{bt}^{B} > 0. \tag{14}$$

Brand consumption indices are a CES function of real consumption for each UPC:

$$C_{bt}^{B} = \left[\sum_{u \in \Omega_{bt}^{U}} \left(\varphi_{ut}^{U} C_{ut}^{U} \right)^{\frac{\sigma_g^U - 1}{\sigma_g^U}} \right]^{\frac{\sigma_g^U}{\sigma_g^U - 1}}, \qquad \sigma_g^U > 1, \, \varphi_{ut}^{U} > 0, \tag{15}$$

with dual price indices defined analogously. From now on, we suppress the product group subscript g on $\{\sigma_g^F, \sigma_g^B, \sigma_g^U\}$ to simplify notation. The system of expenditure shares is homogeneous of degree zero in prices and appeal. Therefore we impose the following normalization on appeal:

$$\tilde{\varphi}_{gt}^{F} = \left(\prod_{f \in \Omega_{gt}^{F}} \varphi_{fgt}^{F} \right)^{\frac{1}{N_{gt}^{F}}} = 1, \qquad \tilde{\varphi}_{ft}^{B} = \left(\prod_{b \in \Omega_{fgt}^{B}} \varphi_{bt}^{B} \right)^{\frac{1}{N_{fgt}^{B}}} = 1, \qquad \tilde{\varphi}_{bt}^{U} = \left(\prod_{u \in \Omega_{bt}^{U}} \varphi_{ut}^{U} \right)^{\frac{1}{N_{bt}^{U}}} = 1, \tag{16}$$

where N_{gt}^{F} is the number of firms in product group g at time t (the number of elements in Ω_{gt}^{F}); N_{fgt}^{B} is the number of brands in firm f in product group g at time t (the number of elements in Ω_{fgt}^{B}); and N_{bt}^{U} is the number of UPCs supplied by brand b at time t (the number of elements in Ω_{bt}^{U}). Demand for each product u can be written as:

$$C_{ut}^{U} = \left(\varphi_{fgt}^{F} \right)^{\sigma^F - 1} \left(\varphi_{bt}^{B} \right)^{\sigma^B - 1} \left(\varphi_{ut}^{U} \right)^{\sigma^U - 1} E_{gt}^{G} \left(P_{gt}^{G} \right)^{\sigma^F - 1} \left(P_{fgt}^{F} \right)^{\sigma^B - \sigma^F} \left(P_{bt}^{B} \right)^{\sigma^U - \sigma^B} \left(P_{ut}^{U} \right)^{-\sigma^U}. \tag{17}$$

S4.2 Technology

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 $\delta_g > 0$ parameterizes the elasticity of marginal costs to output. In addition, each firm faces a fixed market-entry cost for each product group of $H_{gt}^{F} > 0$ (e.g. fixed costs of supplying the market), a fixed market-entry cost for each brand of $H_{gt}^{B} > 0$ (e.g. fixed costs of marketing and branding), and a fixed market-entry cost for each UPC supplied of $H_{gt}^{U} > 0$ (e.g. fixed costs of product development and distribution). We allow for firm, brand and UPC entry and exit, where for the data to be an equilibrium of the model, it must be the case that no firm can profitably enter or exit, no brand can be profitably created or destroyed, and no product can be profitably added or dropped. To simplify notation, we again suppress the product group subscript g on δ_g, unless otherwise indicated.

S4.3 Profit Maximization

We assume that each firm is small relative to the aggregate economy and that pricing and product introduction decisions are made at the brand level. We assume that brands choose prices under

Bertrand competition, though it is straightforward to also consider the case in which brands choose quantities under Cournot competition. Each brand chooses the set of UPCs $u \in \{\underline{u}_{bt}, \ldots, \overline{u}_{bt}\}$ to supply and their prices $\{P^U_{ut}\}$ to maximize its profits:

$$\max_{\{\underline{u}_{bt},\ldots,\overline{u}_{bt}\},\{P^U_{ut}\}} \Pi^B_{bt} = \sum_{u=\underline{u}_{bt}}^{\overline{u}_{bt}} \left[P^U_{ut} Y^U_{ut} - A_{ut} \left(Y^U_{ut} \right) \right] - N^U_{bt} H^U_{gt} - H^B_{gt}, \tag{18}$$

where we index the UPCs supplied by the brand from the largest ($\underline{u}_{bt}$) to the smallest ($\bar{u}_{bt}$) in sales, and the total number of UPCs supplied by the brand is denoted by N^U_{bt}, where $\overline{u}_{bt} = \underline{u}_{bt} + N^U_{bt}$. From the first-order conditions for profit maximization, the equilibrium pricing rule implies the same markup across all UPCs supplied by a brand, but a different markup across brands within firms (both across brands within the same sector and across sectors):

$$P^U_{ut} = \mu^B_{bt} \gamma_{ut}, \qquad \gamma_{ut} = (1+\delta)\, a_{ut} \left(Y^U_{ut} \right)^{\delta}, \tag{19}$$

where γ_{ut} denotes marginal cost. The markup depends on the perceived elasticity of demand for the brand:

$$\mu^B_{bt} = \frac{\varepsilon^B_{bt}}{\varepsilon^B_{bt} - 1}, \tag{20}$$

where this perceived elasticity of demand for the brand depends on its market share within the firm and the firm's market share within the sector:

$$\varepsilon^B_{bt} = \sigma^B - \left(\sigma^F - 1 \right) S^F_{fgt} S^B_{bt} - \left(\sigma^B - \sigma^F \right) S^B_{bt}. \tag{21}$$

S4.4 Sources of Firm Heterogeneity

Nominal sales for firm f within sector g, E^F_{fgt}, is the sum of sales across brands and UPCs supplied by the firm:

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

Using CES demand (17), firm sales can be re-written as:

$$E^F_{fgt} = \left(\varphi^F_{fgt} \right)^{\sigma^F - 1} E^G_{gt} \left(P^G_{gt} \right)^{\sigma^F - 1} \left(P^F_{fgt} \right)^{\sigma^B - \sigma^F} \sum_{b \in \Omega^B_{fgt}} \left(\varphi^B_{bt} \right)^{\sigma^B - 1} \left(P^B_{bt} \right)^{\sigma^U - \sigma^B} \sum_{u \in \Omega^U_{bt}} \left(\frac{P^U_{ut}}{\varphi^U_{ut}} \right)^{1-\sigma^U}. \tag{22}$$

Using the brand price index to substitute for $\left(P^B_b \right)^{\sigma^U - \sigma^B}$, we obtain:

$$E^F_{fgt} = \left(\varphi^F_{fgt} \right)^{\sigma^F - 1} E^G_{gt} \left(P^G_{gt} \right)^{\sigma^F - 1} \left(P^F_{fgt} \right)^{\sigma^B - \sigma^F} \sum_{b \in \Omega^B_{fgt}} \left(\varphi^B_{bt} \right)^{\sigma^B - 1} \left[\sum_{u \in \Omega^U_{bt}} \left(\frac{P^U_{ut}}{\varphi^U_{ut}} \right)^{1-\sigma^U} \right]^{\frac{1-\sigma^B}{1-\sigma^U}}. \tag{23}$$

Using the firm price index to substitute for $\left(P^F_{fgt} \right)^{\sigma^B - \sigma^F}$, we obtain:

$$E^F_{fgt} = \left(\varphi^F_{fgt} \right)^{\sigma^F - 1} E^G_{gt} \left(P^G_{gt} \right)^{\sigma^F - 1} \left[\sum_{b \in \Omega^B_{fgt}} \left(\varphi^B_{bt} \right)^{\sigma^B - 1} \left[\sum_{u \in \Omega^U_{bt}} \left(\frac{P^U_{ut}}{\varphi^U_{ut}} \right)^{1-\sigma^U} \right]^{\frac{1-\sigma^B}{1-\sigma^U}} \right]^{\frac{1-\sigma^F}{1-\sigma^B}}.$$

Dividing and multiplying by $\left(N_{fgt}^{B}\right)^{\frac{\sigma^F-1}{\sigma^B-1}}$, and taking logarithms, this expression for firm sales can be re-written as:

$$\ln E_{fgt}^{F} = \ln E_{gt}^{G} + (\sigma^F - 1)\ln P_{gt}^{G} + (\sigma^F - 1)\ln \varphi_{fgt}^{F} + \frac{\sigma^F - 1}{\sigma^B - 1}\ln N_{fgt}^{B}$$
$$+ \frac{\sigma^F - 1}{\sigma^B - 1}\ln\left[\frac{1}{N_{fgt}^{B}}\sum_{b\in\Omega_{fgt}^{B}}\left(\varphi_{bt}^{B}\right)^{\sigma^B-1}\left[\sum_{u\in\Omega_{bt}^{U}}\left(\frac{P_{ut}^{U}}{\varphi_{ut}^{U}}\right)^{1-\sigma^U}\right]^{\frac{1-\sigma^B}{1-\sigma^U}}\right]. \tag{24}$$

This expression permits a decomposition of relative firm sales within product groups into the contributions of (a) firm appeal, (b) the number of brands, (c) all other components (including markups, product appeal and costs).

S4.5 Moment Conditions

S4.5.1 UPC Moment Conditions

Our UPC moment conditions are similar to those in our baseline specification in the paper, because the markup is common across UPCs within brands, and therefore differences out when we difference relative to the largest UPC within a brand. On the demand-side, we double-difference log UPC expenditure shares over time and relative to the largest UPC within each brand to obtain:

$$\Delta^{\underline{u},t}\ln S_{ut}^{U} = (1-\sigma^U)\Delta^{\underline{u},t}\ln P_{ut}^{U} + \omega_{ut}^{U}, \tag{25}$$

where u is a UPC supplied by the brand; $\underline{u}$ corresponds to the largest UPC supplied by the same brand (as measured by the sum of expenditure across the two years); $\Delta^{\underline{u},t}$ is the double-difference operator across brands and over time such that $\Delta^{\underline{u},t}\ln S_{ut}^{U} = \Delta^{t}\ln S_{ut}^{U} - \Delta^{t}\ln S_{\underline{u}t}^{U}$; Δ^{t} is the first-difference operator over time t such that $\Delta^{t}\ln S_{ut}^{U} = S_{ut}^{U} - S_{ut-1}^{U}$; $\omega_{ut}^{U} = (\sigma^U - 1)\left[\Delta^{t}\ln\varphi_{ut}^{U} - \Delta^{t}\ln\varphi_{\underline{u}t}^{U}\right]$ is a stochastic error.

On the supply-side, using the cost function ($A_{ut}\left(Y_{u}^{U}\right) = a_{ut}\left(Y_{ut}^{U}\right)^{1+\delta}$) and noting that $Y_{ut}^{U} = E_{ut}^{U}/P_{ut}^{U} = E_{bt}^{B}S_{ut}^{U}/P_{ut}^{U}$, the UPC pricing rule given in equation (19) can be re-written as:

$$P_{ut}^{U} = \mu_{bt}^{B}\gamma_{ut} = \left(\mu_{bt}^{B}\right)^{\frac{1}{1+\delta}}(1+\delta)^{\frac{1}{1+\delta}}a_{ut}^{\frac{1}{1+\delta}}\left(E_{bt}^{B}\right)^{\frac{\delta}{1+\delta}}\left(S_{ut}^{U}\right)^{\frac{\delta}{1+\delta}}.$$

Taking logs and double-differencing, we obtain:

$$\Delta^{\underline{u},t}\ln P_{ut}^{U} = \frac{\delta}{1+\delta}\Delta^{\underline{u},t}\ln S_{ut}^{U} + \kappa_{ut}, \tag{26}$$

where the markup (μ_{ut}^{B}) and total brand expenditure (E_{bt}^{B}) have differenced out because they take the same value across UPCs within the brand; $\kappa_{ut} = \frac{1}{1+\delta}\left[\Delta^{t}\ln a_{ut} - \Delta^{t}\ln a_{\underline{u}t}\right]$ is a stochastic error.

Following Broda and Weinstein (2006, 2010), the orthogonality of double-differenced demand and supply shocks defines a set of moment conditions (one for each UPC):

$$G(\beta_g) = \mathbb{E}_{\mathbb{T}}\left[v_{ut}(\beta_g)\right] = 0, \tag{27}$$

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

$$\hat{\beta}_g = \underset{\beta_g}{\arg\min} \left\{ G^*(\beta_g)' W G^*(\beta_g) \right\} \qquad \forall g, \tag{28}$$

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.

S4.5.2 Brand Moment Conditions

Our brand moment conditions take an analogous form to the firm moment conditions in our baseline specification in the paper. We double difference log brand expenditure shares over time and relative to the largest brand within each firm to obtain:

$$\Delta^{\underline{b},t} \ln S_{bt}^B = \left(1 - \sigma^B\right) \Delta^{\underline{b},t} \ln P_{bt}^B + \omega_{bt}^B, \tag{29}$$

where b is a brand supplied by the firm; $\underline{b}$ is the largest brand supplied by the firm (as measured by the sum of expenditure across the two years); $\Delta^{\underline{b},t}$ is the double-difference operator across brands and over time such that $\Delta^{\underline{b},t} \ln S_{bt}^B = \Delta^t \ln S_{bt}^B - \Delta^t \ln S_{\underline{b}t}^B$; Δ^t is the first-difference operator over time t such that $\Delta^t \ln S_{bt}^B = S_{bt}^B - S_{bt-1}^B$; $\omega_{bt}^B \equiv \left(\sigma^B - 1\right) \Delta^{\underline{b},t} \ln \varphi_{bt}^B$ is a stochastic error.

Estimating equation (29) using ordinary least squares could be problematic because changes in brand price indices could be correlated with changes in brand average appeal: $\mathrm{Cov}\left(\Delta^{\underline{b},t} \ln P_{bt}^B, \Delta^{\underline{b},t} \ln \varphi_{bt}^B\right) \neq 0$. To find a suitable instrument for changes in brand price indices, we use the properties of CES demand, which imply that we can write the brand price index solely in terms of observed relative expenditures and the geometric mean of UPC prices:

$$\ln P_{bt}^B = \ln \tilde{P}_{bt}^U + \frac{1}{1-\sigma^U} \ln \left[\sum_{u \in \Omega_{bt}^U} \frac{S_{ut}^U}{\tilde{S}_{bt}^U} \right], \tag{30}$$

where a tilde above a variable denotes a geometric mean such that $\tilde{S}_{bt}^U = \exp\left\{ \frac{1}{N_{bt}^U} \sum_{u \in \Omega_{bt}^U} \ln S_{ut}^U \right\}$; and we have used our normalization that $\tilde{\varphi}_{bt}^U = 1$. Double differencing the log brand price index (30) over time and relative to the largest brand within each firm, we obtain:

$$\Delta^{\underline{b},t} \ln P_{bt}^B = \Delta^{\underline{b},t} \ln \widetilde{P}_{bt}^U + \frac{1}{1-\sigma^U} \Delta^{\underline{b},t} \ln \left[\sum_{u \in \Omega_{bt}^U} \frac{S_{ut}^U}{\widetilde{S}_{bt}^U} \right], \tag{31}$$

where the model implies that the second term on the right-hand side containing the shares of UPCs in brand expenditure is a suitable instrument for the double-differenced brand price index in (29).

S4.5.3 Firm Moment Conditions

Our firm moment conditions take the same form as in our baseline specification in the paper. We double difference log firm expenditure shares over time and relative to the largest firm within each product group to obtain:

$$\Delta^{\underline{f},t} \ln S^F_{fgt} = \left(1 - \sigma^F\right) \Delta^{\underline{f},t} \ln P^F_{fgt} + \omega^F_{fgt}, \tag{32}$$

where f is a firm; $\underline{f}$ is the largest firm within the product group (as measured by the sum of expenditure across the two years); $\Delta^{\underline{f},t}$ is the double-difference operator across firms and over time such that $\Delta^{\underline{f},t} \ln S^F_{fgt} = \Delta^t \ln S^F_{fgt} - \Delta^t \ln S^F_{\underline{f}gt}$; Δ^t is the first difference operator over time such that $\Delta^t \ln S^F_{fgt} = S^F_{fgt} - S^F_{fgt-1}$; $\omega^F_{fgt} \equiv (\sigma_F - 1)\, \Delta^{\underline{f},t} \ln \varphi^F_{fgt}$ is a stochastic error.

Estimating equation (32) using ordinary least squares could be problematic because changes in firm price indices could be correlated with changes in firm average appeal: $\text{Cov}\left(\Delta^{\underline{f},t} \ln P^F_{fgt},\, \Delta^{\underline{f},t} \ln \varphi^F_{fgt}\right) \neq 0$. To find a suitable instrument for changes in firm price indices, we use the properties of CES demand, which imply that we can write the firm price index solely in terms of observed relative expenditures and the geometric mean of brand prices:

$$\ln P^F_{fgt} = \ln \tilde{P}^B_{fgt} + \frac{1}{1 - \sigma^B} \ln \left[\sum_{b \in \Omega^B_{fgt}} \frac{S^B_{bt}}{\tilde{S}^B_{fgt}} \right], \tag{33}$$

where a tilde above a variable denotes a geometric mean such that $\tilde{S}^B_{fgt} = \exp\left\{ \frac{1}{N^B_{fgt}} \sum_{b \in \Omega^B_{fgt}} \ln S^B_{bt} \right\}$; and we have used our normalization that $\tilde{\varphi}^B_{fgt} = 1$. Double differencing the log firm price index (33) over time and relative to the largest firm within each product-group, we obtain:

$$\Delta^{\underline{f},t} \ln P^F_{fgt} = \Delta^{\underline{f},t} \ln \widetilde{P}^F_{fgt} + \frac{1}{1 - \sigma^B} \Delta^{\underline{f},t} \ln \left[\sum_{b \in \Omega^B_{fgt}} \frac{S^B_{bt}}{\widetilde{S}^B_{fgt}} \right], \tag{34}$$

where the model implies that the second term on the right-hand side containing the shares of brands in firm expenditure is a suitable instrument for the double-differenced firm price index in (32).

Therefore, our framework can be extended to introduce brands as an additional nest within firms. Under the assumption that pricing and product introduction decisions are made at the brand level, markups vary across brands within firms within product groups. We do not pursue this extension further, because firms in general have an incentive to price strategically across brands, and we do not know of a simple theory of brand price strategy that differs from firm price strategy. Therefore, we focus in our baseline specification on the case where pricing and product introduction decisions are made at the firm level. Under our assumption of a Cobb-Douglas upper tier of utility, no firm has an incentive to price strategically across product groups, because product group expenditure shares are determined by parameters alone, and the firm problem becomes separable by product group. Hence, in this baseline specification, markups vary across product groups within firms.

S5 Robustness Test with CES Upper Tier

In this section, we consider a robustness test in which we assume that the upper tier of utility takes the constant elasticity of substitution (CES) form rather than the Cobb-Douglas form. In this extension, product group expenditure shares become endogenous, which introduces an incentive for firms to price strategically across product groups. We show that although the formula for the firm markup changes, the firm markup remains common across UPCs within each product group and hence differences out from our UPC moment conditions. This property, together with the recursive nature of our estimation strategy, implies that our estimates at the UPC and firm level $\{\sigma^U, \sigma^F, \delta\}$ are unchanged (only the elasticity of substitution across product groups σ^G changes). Given these unchanged UPC and firm parameter estimates, our decomposition of variation in firm sales within product groups into firm appeal, product scope, and average product appeal adjusted prices also remains unchanged, because prices are observed in the data. Therefore this extension does not change our main result that firm appeal and scope account for most of the variation in observed firm sales within product groups. Instead, this extension merely changes the decomposition of product-appeal adjusted prices into mark-ups, average cost and cost dispersion, because the formula for the markup changes.

S5.1 Demand

Utility at time t is a CES function of the consumption (C^G_{gt}) of a number of product groups $g \in \Omega^G$:

$$U_t = \left[\sum_{g \in \Omega^G} \left(\varphi^G_{gt} C^G_{gt} \right)^{\frac{\sigma^G - 1}{\sigma^G}} \right]^{\frac{\sigma^G}{\sigma^G - 1}}, \tag{35}$$

where φ^G_{gt} is the appeal of product group g at time t; Ω^G is the set of product groups which is constant over time; and σ^G is the elasticity of substitution across product groups. We assume a finite number of product groups and allow each firm to be active in multiple product groups. Consumption for each product group is a CES function of the consumption (C^F_{fgt}) of a number of firms $f \in \Omega^F_{gt}$:

$$C^G_{gt} = \left[\sum_{f \in \Omega^F_{gt}} \left(\varphi^F_{fgt} C^F_{fgt} \right)^{\frac{\sigma^F_g - 1}{\sigma^F_g}} \right]^{\frac{\sigma^F_g}{\sigma^F_g - 1}}, \tag{36}$$

where φ^F_{fgt} is the appeal of firm f within product group g at time t; Ω^F_{gt} is the set of firms within product group g at time t; and σ^F_g is the elasticity of substitution across firms within product group g. Consumption for each firm within each product group is a CES function of the consumption (C^U_{ut}) of a number of UPCs:

$$C^F_{fgt} = \left[\sum_{u \in \Omega^U_{fgt}} \left(\varphi^U_{ut} C^U_{ut} \right)^{\frac{\sigma^U_g - 1}{\sigma^U_g}} \right]^{\frac{\sigma^U_g}{\sigma^U_g - 1}}, \tag{37}$$

where φ^U_{ut} is the appeal of UPC u at time t; Ω^U_{fgt} is the set of UPCs within firm f and product group g at time t; and σ^U_g is the elasticity of substitution across UPCs within firm f and product group g.

We allow firms to be active in multiple product groups. We denote the set of product groups in which a firm is active by Ω^G_{ft} and the set of products supplied by a firm by $\Omega^U_{ft} = \left\{\Omega^U_{fgt} : \Omega^U_{fgt} \neq \varnothing\right\}$. Using the properties of CES demand, equilibrium consumption of UPC u supplied by firm f in product group g is:

$$C^U_{ut} = \left(\varphi^G_{gt}\right)^{\sigma^G-1} \left(\varphi^F_{fgt}\right)^{\sigma^F_g-1} \left(\varphi^U_{ut}\right)^{\sigma^U_g-1} E_t\left(P_t\right)^{\sigma^G-1} \left(P^G_{gt}\right)^{\sigma^F_g-\sigma^G} \left(P^F_{fgt}\right)^{\sigma^U_g-\sigma^F_g} \left(P^U_{ut}\right)^{-\sigma^U_g}, \tag{38}$$

where E_t is aggregate expenditure; P_t, P^G_{gt} and P^F_{fgt} are the price indices dual to (35), (36) and (37) respectively. We assume that labor is the sole primary factor of production and choose labor as the numeraire. Therefore the wage is $w = 1$ and aggregate expenditure is pinned down by the aggregate supply of labor $E_t = L$.

S5.2 Technology

We assume that the variable cost function is separable across UPCs and that supplying Y^U_{ut} units of output of UPC u incurs a total variable cost in terms of labor of $A_{ut}\left(Y^U_{ut}\right) = a_{ut}\left(Y^U_{ut}\right)^{1+\delta_g}$, where a_{ut} is a cost shifter and δ_g is 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 $H^F_{gt} > 0$ (e.g. fixed costs of supplying the market) and a fixed market entry cost for each UPC supplied of $H^U_{gt} > 0$ (e.g. fixed costs of product development and distribution).

S5.3 Profit Maximization

We assume that pricing and product introduction decisions are made at the level of firm, internalizing the effects of choices for one product group on sales across all product groups in which the firm is active. In our baseline specification, we assume that firms choose prices under Bertrand competition. However, it is straightforward to also consider the case in which firms choose quantities under Cournot competition. Each firm f chooses the set of UPCs $u \in \left\{\underline{u}_{ft}, \ldots, \overline{u}_{ft}\right\}$ to supply and their prices $\{P^U_{ut}\}$ to maximize its profits:

$$\max_{\left\{\underline{u}_{fgt},\ldots,\overline{u}_{fgt}\right\},\{P_{ut}\}} \Pi^F_{ft} = \sum_{g\in\Omega^G} \left\{ \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \left[P^U_{kt}Y^U_{kt} - A_{kt}\left(Y^U_{kt}\right)\right] - N^U_{fgt}H^U_{gt} - \mathbb{I}^G_{gt}H^F_{gt} \right\}, \tag{39}$$

where $\left\{\underline{u}_{fgt}, \ldots, \overline{u}_{fgt}\right\}$ is the subset of products that the firm supplies within product group g; $\mathbb{I}^G_{gt}$ is an indicator variable that equals one if a firm is active in product group g; $Y^U_{ut} \geq 0$; $N^U_{fgt} \geq 0$; and in equilibrium we require that output (Y^U_{ut}) equals consumption (C^U_{ut}) in (38):

$$Y^U_{kt} = C^U_{kt} = \left(\varphi^G_{gt}\right)^{\sigma^G-1} \left(\varphi^F_{fgt}\right)^{\sigma^F_g-1} \left(\varphi^U_{kt}\right)^{\sigma^U_g-1} E_t\left(P_t\right)^{\sigma^G-1} \left(P^G_{gt}\right)^{\sigma^F_g-\sigma^G} \left(P^F_{fgt}\right)^{\sigma^U_g-\sigma^F_g} \left(P^U_{kt}\right)^{-\sigma^U_g}. \tag{40}$$

We index the UPCs supplied by the firm within each 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 each product group is denoted by N^U_{fgt}, where $\overline{u}_{fgt} = \underline{u}_{fgt} + N^U_{fgt}$.

Given the nested CES structure of demand in (40), a firm's choice of the price for a UPC in product group g (P^U_{ut}) affects the demand for its other UPCs in that product group through the aggregate, product group, and firm price indices (P_t, P^G_{gt} and P^F_{fgt} respectively). In contrast, the firm's choice of the price for that UPC in product group g (P^U_{ut}) only affects the demand for its other UPCs in other product groups $m \neq g$ through the aggregate price index (P_t). The first-order condition for profit maximization with respect to the price of UPC u (P^U_{ut}) for firm f within product group g is:

$$\begin{aligned}
Y^U_{ut} + \sum_{m \in \Omega^G} \sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}} P^U_{kt} \left(\sigma^G - 1\right) \frac{\partial P_t}{\partial P^G_{gt}} \frac{\partial P^G_{gt}}{\partial P^F_{fgt}} \frac{\partial P^F_{fgt}}{\partial P^U_{ut}} \frac{Y^U_{kt}}{P_t} \\
+ \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P^U_{kt} \left(\sigma^F_g - \sigma^G\right) \frac{\partial P^G_{gt}}{\partial P^F_{fgt}} \frac{\partial P^F_{fgt}}{\partial P^U_{ut}} \frac{Y^U_{kt}}{P^G_{gt}} \\
+ \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P^U_{kt} \left(\sigma^U_g - \sigma^F_g\right) \frac{\partial P}{\partial P^U_{ut}} \frac{Y^U_{kt}}{P^F_{fgt}} \\
- P^U_{ut} \sigma^U_g \frac{Y^U_{ut}}{P^U_{ut}} \\
- \sum_{m \in \Omega^G} \sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}} \frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}} \left(\sigma^G - 1\right) \frac{\partial P_t}{\partial P^G_{gt}} \frac{\partial P^G_{gt}}{\partial P^F_{fgt}} \frac{\partial P^F_{fgt}}{\partial P^U_{ut}} \frac{Y^U_{kt}}{P_t} \\
- \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}} \left(\sigma^F_g - \sigma^G\right) \frac{\partial P^G_{gt}}{\partial P^F_{fgt}} \frac{\partial P^F_{fgt}}{\partial P^U_{ut}} \frac{Y^U_{kt}}{P^G_{gt}} \\
- \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}} \left(\sigma^U_g - \sigma^F_g\right) \frac{\partial P^F_{fgt}}{\partial P^U_{ut}} \frac{Y^U_{kt}}{P^F_{fgt}} \\
+ \frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}} \sigma^U_g \frac{Y^U_{ut}}{P^U_{ut}} = 0.
\end{aligned}$$

This first-order condition can be written as:

$$
\begin{aligned}
Y_{ut}^{U} + \sum_{m\in\Omega^{G}} \sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}} P_{kt}^{U}\left(\sigma^{G}-1\right)\left(\frac{\partial P_{t}}{\partial P_{gt}^{G}}\frac{P_{gt}^{G}}{P_{t}}\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}} \\
+\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P_{kt}^{U}\left(\sigma_{g}^{F}-\sigma^{G}\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}} \\
+\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P_{kt}^{U}\left(\sigma_{g}^{U}-\sigma_{g}^{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}} \\
-P_{ut}^{U}\sigma_{g}^{U}\frac{Y_{ut}^{U}}{P_{ut}^{U}} \\
-\sum_{m\in\Omega^{G}} \sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}} \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\left(\sigma^{G}-1\right)\left(\frac{\partial P_{t}}{\partial P_{gt}^{G}}\frac{P_{gt}^{G}}{P_{t}}\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}} \\
-\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\left(\sigma_{g}^{F}-\sigma^{G}\right)\left(\frac{\partial P_{fgt}^{G}}{\partial P_{ft}^{F}}\frac{P_{ft}^{F}}{P_{fgt}^{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}} \\
-\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\left(\sigma_{g}^{U}-\sigma_{g}^{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}} \\
+\frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\sigma_{g}^{U}\frac{Y_{ut}^{U}}{P_{ut}^{U}} = 0,
\end{aligned}
$$

which can be further re-written as:

$$
\begin{aligned}
Y_{ut}^{U} + \sum_{m\in\Omega^{G}} \sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}} P_{kt}^{U}\left(\sigma^{G}-1\right)S_{gt}^{G}S_{fgt}^{F}S_{ut}^{U}\frac{Y_{kt}^{U}}{P_{ut}^{U}} + \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P_{kt}^{U}\left(\sigma_{g}^{F}-\sigma^{G}\right)S_{fgt}^{F}S_{ut}^{U}\frac{Y_{kt}^{U}}{P_{ut}^{U}} \\
+\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P_{kt}^{U}\left(\sigma_{g}^{U}-\sigma_{g}^{F}\right)S_{ut}^{U}\frac{Y_{kt}^{U}}{P_{ut}^{U}} - P_{ut}^{U}\sigma_{g}^{U}\frac{Y_{ut}^{U}}{P_{ut}^{U}} \\
-\sum_{m\in\Omega^{G}} \sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}} \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\left(\sigma^{G}-1\right)S_{gt}^{G}S_{fgt}^{F}S_{ut}^{U}\frac{Y_{kt}^{U}}{P_{ut}^{U}} - \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\left(\sigma_{g}^{F}-\sigma^{G}\right)S_{fgt}^{F}S_{ut}^{U}\frac{Y_{kt}^{U}}{P_{ut}^{U}} \\
-\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\left(\sigma_{g}^{U}-\sigma_{g}^{F}\right)S_{ut}^{U}\frac{Y_{kt}^{U}}{P_{ut}^{U}} + \frac{dA_{kt}\left(Y_{ut}^{U}\right)}{dY_{ut}^{U}}\sigma_{g}^{U}\frac{Y_{ut}^{U}}{P_{ut}^{U}} = 0.
\end{aligned}
$$

Dividing through by Y_{ut}^{U} and using the definition:

$$
S_{ut}^{U} = \frac{P_{ut}^{U}C_{ut}^{U}}{\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P_{kt}^{U}C_{kt}^{U}} = \frac{P_{ut}^{U}Y_{ut}^{U}}{\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} P_{kt}^{U}Y_{kt}^{U}},
$$

we obtain:

$$1+\sum_{m\in\Omega^G}\sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}}\left(\sigma^G-1\right)S^G_{gt}S^F_{fgt}\frac{P^U_{kt}Y^U_{kt}}{\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}}P^U_{kt}Y^U_{kt}}+\left(\sigma^F_g-\sigma^G\right)S^F_{fgt}+\left(\sigma^U_g-\sigma^F_g\right)-\sigma^U_g$$

$$-\sum_{m\in\Omega^G}\sum_{k=\underline{u}_{fmt}}^{\overline{u}_{fmt}}\left(\sigma^G-1\right)S^G_{gt}S^F_{fgt}\frac{\frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}}Y^U_{kt}}{\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}}\mu_{kt}\frac{dA_k\left(Y^U_{ut}\right)}{dY^U_{ut}}Y^U_{kt}}-\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}}\left(\sigma^F_g-\sigma^G\right)S^F_{fgt}\frac{\frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}}Y^U_{kt}}{\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}}\mu_{kt}\frac{dA_k\left(Y^U_{ut}\right)}{dY^U_{ut}}Y^U_{kt}}$$
$$-\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}}\left(\sigma^U_g-\sigma^F_g\right)\frac{\frac{dA_{kt}\left(Y^U_{ut}\right)}{dY^U_{ut}}Y^U_{kt}}{\sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}}\mu_{kt}\frac{dA_k\left(Y^U_{ut}\right)}{dY^U_{ut}}Y^U_{kt}}+\frac{\sigma^U_g}{\mu_{ut}}=0. \tag{41}$$

Note that this first-order condition is identical for each UPC u supplied by firm f within product group g. Therefore the firm f charges the same markup for all UPCs u within product group g:

$$\mu^F_{ut}=\mu^F_{kt}\qquad\text{for all}\qquad u,k\in\Omega^U_{fgt},$$

but this mark-up differs across UPCs supplied by firm f in two different product groups g and m:

$$\mu^F_{ut}\neq\mu^F_{kt}\qquad\text{for}\qquad u\in\Omega^U_{fgt},\qquad k\in\Omega^U_{fmt},\qquad g\neq m.$$

Re-arranging the first-order condition (41), the common markup for firm f across all UPCs u within product group g is:

$$\mu^F_{fgt}=\frac{\varepsilon^F_{fgt}}{\varepsilon^F_{fgt}-1}, \tag{42}$$

$$\varepsilon^F_{fgt}=\sigma^F_g-\left(\sigma^F_g-\sigma^G\right)S^F_{gt}-\left(\sigma^G-1\right)\frac{S^G_{gt}S^F_{fgt}}{Z^F_{fgt}}, \tag{43}$$

where S^F_{fgt} is the expenditure share of firm f in product group g at time t; S^G_{gt} is the expenditure share of product group g at time t; and Z^F_{fgt} is the share of product group g in overall expenditure on firm f at time t.

Intuitively, the firm charges a high mark-up across all UPCs within a product group when (i) it accounts for a large share of expenditure in that product group, (ii) when that product group accounts for a large share of aggregate expenditure, and (iii) when that product group accounts for a small share of overall expenditure on the firm. First, when the firm accounts for a larger share of expenditure within the product group, an increase in the price of UPCs within that product group leads to a smaller reduction in expenditure on these UPCs, because of a larger rise in the product group price index (increasing expenditure on these UPCs). Second, when the product group accounts for a larger share of aggregate expenditure, the rise in the product group price index caused by the increase in the price of these UPCs leads to a smaller reduction in expenditure on the product group, because of a larger rise in the aggregate price index (increasing expenditure on the product group).

Finally, when the product group accounts for a smaller share of overall expenditure on the firm, the reduction in sales from the increase in the price of these UPCs within a given product group is offset to a larger degree by an increase in sales in other product groups in which the firm is active.

S5.4 UPC Moment Conditions

In the first step of our estimation, we double difference log UPC expenditure shares over time and relative to the largest UPC within the same firm and product group:

$$\Delta^{\underline{u},t} \ln S_{ut}^{U} = (1-\sigma_g^U)\Delta^{\underline{u},t} \ln P_{ut}^{U} + \omega_{ut}, \tag{44}$$

$$\Delta^{\underline{u},t} \ln P_{ut}^{U} = \frac{\delta_g}{1+\delta_g}\Delta^{\underline{u},t} \ln S_{ut}^{U} + \kappa_{ut}, \tag{45}$$

where u is a UPC supplied by the firm within a given product group; $\underline{u}$ corresponds to the largest UPC supplied by the same firm within the same product group (as measured by the sum of expenditure across the two years); $\Delta^{\underline{u},t}$ is the double-difference operator across UPCs within a given firm and product group and over time such that $\Delta^{\underline{u},t} \ln S_{ut}^{U} = \Delta^{t} S_{ut}^{U} - \Delta^{t} S_{\underline{u}t}^{U}$; Δ^{t} is the first difference operator over time such that $\Delta^{t} \ln S_{ut}^{U} = S_{ut}^{U} - S_{ut-1}^{U}$; $\omega_{ut} = \left(\sigma_g^U - 1\right)\left[\Delta^{t} \ln \varphi_{ut}^{U} - \Delta^{t} \ln \varphi_{\underline{u}t}^{U}\right]$ and $\kappa_{ut} = \frac{1}{1+\delta_g}\left[\Delta^{t} \ln a_{ut} - \Delta^{t} \ln a_{\underline{u}t}\right]$ are stochastic errors; the markup (μ_{fgt}^{F}) has differenced out because it is common across UPCs within the same product group within a firm. Note that these moment conditions take the same form as in our baseline specification in the paper and hence leave our parameter estimates $\{\sigma_g^U, \delta_g\}$ unchanged.

S5.5 Firm Moment Conditions

In the second step of our estimation, we double difference log firm expenditure shares over time and relative to the largest firm within the same product group:

$$\Delta^{\underline{f},t} \ln S_{fgt}^{F} = \left(1-\sigma_g^F\right)\Delta^{\underline{f},t} \ln P_{fgt}^{F} + \omega_{fgt}, \tag{46}$$

where f is a firm within product group g; $\underline{f}$ corresponds to the largest firm within the same product group (as measured by the sum of expenditure across the two years);$\Delta^{\underline{f},t}$ is the double difference operator across firms within the 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 $\omega_{fgt} \equiv \left(\sigma_g^F - 1\right)\Delta^{\underline{f},t} \ln \varphi_{fgt}^{F}$. The double differenced log firm price indices within the product group can be expressed as:

$$\Delta^{\underline{f},t} \ln P_{fgt}^{F} = \Delta^{\underline{f},t} \ln \widetilde{P}_{fgt}^{U} + \frac{1}{1-\sigma_g^U}\Delta^{\underline{f},t} \ln\left[\sum_{u\in\Omega_{fgt}^{U}} \frac{S_{ut}^{U}}{\widetilde{S}_{fgt}^{U}}\right], \tag{47}$$

where a tilde above a variable denotes a geometric mean such that $\tilde{S}_{fgt}^{U} = \exp\left\{\frac{1}{N_{fgt}^{U}}\sum_{u\in\Omega_{fgt}^{U}} \ln S_{ut}^{U}\right\}$; and we have used our normalization that $\tilde{\varphi}_{fgt}^{U} = 1$. The model implies that the second term on the right-hand side containing the shares of UPCs in firm expenditure within the product group is a suitable instrument for the double-differenced log firm price index within the product group in (46). Again these relationships take the same form as in our baseline specification in the paper and hence leave our parameter estimates $\{\sigma_g^F\}$ unchanged.

S5.6 Firm Sales Decomposition

Given these unchanged UPC and firm parameter estimates $\{\sigma_g^U, \sigma_g^F, \delta_g\}$, our decomposition of firm sales variation within product groups into the contributions of firm appeal, product scope, and average product appeal adjusted prices also remains unchanged:

$$\begin{aligned}\ln E_{fgt}^F &= (\sigma_g^F - 1)\ln \varphi_{fgt}^F + \ln E_{gt}^G + (\sigma_g^F - 1)\ln P_{gt}^G + \left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\ln N_{fgt}^U \\ &\quad + \left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\ln\left(\frac{1}{N_{fgt}^U}\sum_{u\in\Omega_{fgt}^U}\left(\frac{P_{ut}^U}{\varphi_{ut}^U}\right)^{1-\sigma_g^U}\right),\end{aligned}$$

and hence:

$$\Delta^g \ln E_{fgt}^F = \left\{\underbrace{(\sigma_g^F - 1)\Delta^g \ln \varphi_{fgt}^F}_{\text{firm appeal}} + \underbrace{\left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\Delta^g \ln N_{fgt}^U}_{\text{firm scope}} + \underbrace{\left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\Delta^g \ln\left(\frac{1}{N_{fgt}^U}\sum_{u\in\Omega_{fgt}^U}\left(\frac{P_{ut}^U}{\varphi_{ut}^U}\right)^{1-\sigma_g^U}\right)}_{\text{average product appeal-adjusted prices residual}}\right\}, \tag{48}$$

where Δ^g is the difference operator relative to the geometric mean for product group g and prices are observed in the data. Therefore this extension merely changes the decomposition of average product appeal adjusted prices into mark-ups, average cost and cost dispersion (because the formula for the markup changes).

S6 Retail Markups

In subsection S6.1, we first show that our estimation procedure can allow for separate markups for retailers and producers. If these markups are common across UPCs for a given producer and product group, they difference out from our moment conditions, which leaves our estimation procedure and parameter estimates unchanged. In this case, our decomposition of producer sales into firm appeal, product scope and product-appeal-adjusted prices also remains unchanged, because we observe the prices paid by the consumer in our data. Therefore the introduction of retail markups that are common across UPCs for a given producer and product group merely changes the decomposition of product-appeal-adjusted prices into the contributions of mark-ups and costs (because the formula for the markup changes). In subsection S6.2, we next show that our theoretical model of the production sector can be embedded in a simple model of the retail sector.

S6.1 Retailer and Producer Markups

In our baseline specification in the paper, we assume that each producer chooses its set of UPCs and their prices for the final consumer to maximize its profits, taking into account the effects of these choices on consumer price indices. The resulting equilibrium pricing rule features a variable markup (μ_{fgt}^{F}) of price over marginal cost that is common across UPCs for a given producer and product group (equation (9) in the paper reproduced below):

$$P_{ut}^{U} = \mu_{fgt}^{F} \gamma_{ut}, \tag{49}$$

where γ_{ut} denotes marginal cost. Each producer internalizes that it is a monopoly supplier of a consumption index within the product group and hence charges a common markup across all UPCs within that product group ($\mu_{ut}^{F} = \mu_{kt}^{F} = \mu_{fgt}^{F}$ for all $u, k \in \Omega_{fgt}^{U}$). This common markup depends on the producer's share of final consumption expenditure within the product group.

We now consider a generalization of this pricing rule to allow for separate markups for the producer f (μ_{fgt}^{F}) and for a retailer (μ_{fgt}^{R}) that supplies the UPCs of the producer f to the consumer:

$$P_{ut}^{U} = \mu_{fgt}^{F} \mu_{fgt}^{R} \gamma_{ut}, \tag{50}$$

where γ_{ut} again denotes marginal cost. We assume that both markups are common across all of the producer's UPCs within a given product group ($\mu_{ut}^{F} = \mu_{kt}^{F} = \mu_{fgt}^{F}$ for all $u, k \in \Omega_{fgt}^{U}$ and $\mu_{ut}^{R} = \mu_{kt}^{R} = \mu_{fgt}^{R}$ for all $u, k \in \Omega_{fgt}^{U}$), as for the producer's markup in our baseline specification.

Our UPC moment conditions are based on double differences across UPCs within a given producer and product group and over time. Therefore the common values of the producer and retailer mark-ups difference out and our UPC moment conditions remain the same as in the baseline specification in the paper:

$$\Delta^{\underline{u},t} \ln S_{ut}^{U} = (1 - \sigma^{U}) \Delta^{\underline{u},t} \ln P_{ut}^{U} + \omega_{ut}, \tag{51}$$

$$\Delta^{\underline{u},t} \ln P_{ut}^{U} = \frac{\delta}{1+\delta} \Delta^{\underline{u},t} \ln S_{ut}^{U} + \kappa_{ut}, \tag{52}$$

where we suppress the product group subscript g on $\{\sigma_{g}^{U}, \sigma_{g}^{F}, \delta_{g}\}$ to simplify notation; $\Delta^{\underline{u},t}$ is the double-difference operator across UPCs within a given producer and product group and over time such that $\Delta^{\underline{u},t} \ln S_{ut}^{U} = \Delta^{t} \ln S_{ut}^{U} - \Delta^{t} \ln S_{\underline{u}t}^{U}$. Therefore our estimates of $\{\sigma^{U}, \delta\}$ are unchanged. Given these unchanged parameter estimates and the observed UPC expenditure and prices paid by the final consumer $\{E_{ut}^{U}, P_{ut}^{U}\}$, the implied UPC appeal (φ_{ut}^{U}) and firm consumption price index (P_{fgt}^{F}) are also unchanged. Therefore our producer moment conditions remain identical to those in the paper:

$$\Delta^{\underline{f},t} \ln S_{fgt}^{F} = \left(1 - \sigma^{F}\right) \Delta^{\underline{f},t} \ln P_{fgt}^{F} + \omega_{fgt}, \tag{53}$$

where $\Delta^{\underline{f},t}$ is the double-difference operator across producers within a given product group and over time such that $\Delta^{\underline{f},t} \ln S_{fgt}^{F} = \Delta^{t} \ln S_{fgt}^{F} - \Delta^{t} \ln S_{\underline{f}gt}^{F}$. We again instrument for the firm consumption price index using the following index of sales dispersion across UPCs within a given producer and product group:

$$\mathbb{X}_{fgt}^{F} = \ln \left[\sum_{u \in \Omega_{fgt}^{U}} \frac{S_{ut}^{U}}{\tilde{S}_{fgt}^{U}} \right].$$

These identical producer moment conditions in turn imply that our estimates of $\{\sigma^F\}$ are unchanged.

Given these unchanged parameter estimates $\{\sigma^U, \delta, \sigma^F\}$, our decomposition of firm sales variation into (a) firm appeal, (b) product scope and (c) an average product-appeal-adjusted prices residual remains the same as in our baseline specification in the paper because we observe the prices paid by the final consumer:

$$\begin{aligned} \ln E_{fgt}^{F} &= (\sigma^F - 1) \ln \varphi_{fgt}^{F} + \ln E_{gt}^{G} + (\sigma_F - 1) \ln P_{gt}^{G} + \left(\frac{\sigma^F - 1}{\sigma^U - 1} \right) \ln N_{fgt}^{U} \\ &\quad + \left(\frac{\sigma^F - 1}{\sigma^U - 1} \right) \ln \left(\frac{1}{N_{fgt}^{U}} \sum_{u \in \Omega_{fgt}^{U}} \left(\frac{P_{ut}^{U}}{\varphi_{ut}^{U}} \right)^{1-\sigma^U} \right), \end{aligned}$$

which implies:

$$\Delta^g \ln E_{fgt}^{F} = \left\{ \underbrace{(\sigma^F - 1)\Delta^g \ln \varphi_{fgt}^{F}}_{\text{firm demand}} + \underbrace{\left(\frac{\sigma^F - 1}{\sigma^U - 1} \right) \Delta^g \ln N_{fgt}^{U}}_{\text{firm scope}} + \underbrace{\left(\frac{\sigma^F - 1}{\sigma^U - 1} \right) \Delta^g \ln \left(\frac{1}{N_{fgt}^{U}} \sum_{u \in \Omega_{fgt}^{U}} \left(\frac{P_{ut}^{U}}{\varphi_{ut}^{U}} \right)^{1-\sigma^U} \right)}_{\text{average product-appeal-adjusted prices residual}} \right\}, \tag{54}$$

where Δ^g is the difference operator relative to the geometric mean for product group g.

Therefore our finding that most of observed firm sales variation is explained by (a) firm appeal and (b) product scope is robust to the introduction of separate common markups for producers and retailers ((50) compared to (49)), because the contribution of these two terms in (54) is unchanged. The sole implication of the different markup rule is to alter the decomposition of the average product-appeal-adjusted prices residual (the third term in (54)) into the components of (i) markups, (ii) average marginal cost, and (iii) cost dispersion.

S6.2 Retail Sector

We now show that our theoretical model of the production sector can be embedded in an explicit model of the retail sector. In the interests of simplicity, we assume a perfectly competitive retail sector, but we discuss below the implications of alternative possible market structure assumptions for the retail sector.

Demand Our specification of consumer preferences remains the same as in our baseline specification in the paper. Utility at time t, U_t, is a Cobb-Douglas function of consumption, C_{gt}^{G}, of a continuum of product groups:

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

where g denotes each product group and φ_{gt}^G is the share of expenditure on product group g at time t. The consumption indices for product groups (C_{gt}^G) and firms (C_{fgt}^F) can be written as respectively:

$$C_{gt}^G = \left[\sum_{f\in\Omega_{gt}^F} \left(\varphi_{fgt}^F C_{ft}^F \right)^{\frac{\sigma_g^F - 1}{\sigma_g^F}} \right]^{\frac{\sigma_g^F}{\sigma_g^F - 1}}, \qquad C_{fgt}^F = \left[\sum_{u\in\Omega_{fgt}^U} \left(\varphi_{ut}^U C_{ut}^U \right)^{\frac{\sigma_g^U - 1}{\sigma_g^U}} \right]^{\frac{\sigma_g^U}{\sigma_g^U - 1}}, \tag{55}$$

The corresponding dual consumption price indices for product groups (P_{gt}^G) and firms (P_{fgt}^F) are:

$$P_{gt}^G = \left[\sum_{f\in\Omega_{gt}^F} \left(\frac{P_{fgt}^F}{\varphi_{fgt}^F} \right)^{1-\sigma_g^F} \right]^{\frac{1}{1-\sigma_g^F}} \qquad P_{fgt}^F = \left[\sum_{u\in\Omega_{fgt}^U} \left(\frac{P_{ut}^U}{\varphi_{ut}^U} \right)^{1-\sigma_g^U} \right]^{\frac{1}{1-\sigma_g^U}}, \tag{56}$$

where P_{ut}^U is the consumption goods price (the retail price) for a UPC. Using the properties of CES demand, the share of firm f in final consumption expenditure within product group g (S_{fgt}^F for $f \in \Omega_{gt}^F$) equals the elasticity of the product-group consumption price index with respect to the consumption price index for firm f. Similarly, the share of product u in final consumption expenditure within firm f (S_{ut}^U for $u \in \Omega_{fgt}^U$) is equal to the elasticity of the firm consumption price index with respect to the retail price of product u:

$$S_{fgt}^F = \frac{\left(P_{fgt}^F / \varphi_{fgt}^F \right)^{1-\sigma^F}}{\sum_{k\in\Omega_{gt}^F} \left(P_{kgt}^F / \varphi_{kgt}^F \right)^{1-\sigma^F}} = \frac{\partial P_{gt}^G}{\partial P_{fgt}^F} \frac{P_{fgt}^F}{P_{gt}^G}, \qquad S_{ut}^U = \frac{\left(P_{ut}^U / \varphi_{ut}^U \right)^{1-\sigma^U}}{\sum_{k\in\Omega_{fgt}^U} \left(P_{ut}^U / \varphi_{ut}^U \right)^{1-\sigma^U}} = \frac{\partial P_{ft}^F}{\partial P_{ut}^U} \frac{P_{ut}^U}{P_{ft}^F}, \tag{57}$$

where we suppress the product group subscript g on σ_g^U and σ_g^F to simplify notation. The corresponding final consumption demand for UPC u of firm f within product group g is:

$$C_{ut}^U = \left(\varphi_{fgt}^F \right)^{\sigma^F - 1} \left(\varphi_{ut}^U \right)^{\sigma^U - 1} E_{gt}^G \left(P_{gt}^G \right)^{\sigma^F - 1} \left(P_{fgt}^F \right)^{\sigma^U - \sigma^F} \left(P_{ut}^U \right)^{-\sigma^U}, \tag{58}$$

where E_{gt}^G denotes final consumption expenditure on product group g.

Retail Sector We assume a perfectly competitive retail sector that demands production (wholesale) output (X_{ut}^U) from the production sector and supplies final consumption (retail) output (Y_{ut}^U) according to a linear technology:

$$Y_{ut}^U = \frac{1}{\tau} X_{ut}^U, \tag{59}$$

where retailing costs are assumed to take the iceberg form, such that the retailer requires $\tau > 1$ units of a good in order to supply one unit to the final consumer in all product groups. Retailers choose final consumption output (Y_{ut}^U) to maximize profits taking as given final consumption (retail) prices (P_{ut}^U) and production (wholesale) prices (Q_{ut}^U):

$$\max_{Y_{ut}^U} \left[P_{ut}^U Y_{ut}^U - Q_{ut}^U \tau Y_{ut}^U \right],$$

where we have used the retailing technology (59). The first-order condition for profit maximization implies:

$$P_{ut}^{U} = \tau Q_{ut}^{U}. \tag{60}$$

Therefore profit maximization and zero profits imply that final consumption (retail) prices are a constant multiple of production (wholesale) prices ($P_{ut}^{U} = \tau Q_{ut}^{U}$), where the size of this multiple is determined by the size of retailing costs (τ). We use this result as a natural benchmark for embedding our specification of the production sector into an explicit model of the retail sector. It is straightforward to instead consider the case of a monopolistically competitive retail sector, in which retail prices are a constant markup over wholesale prices adjusted for retailing costs. In principle, the analysis also can be generalized to consider the case of an oligopolistically competitive retail sector, although at the cost of additional complexity.

Production Sector Our specification of the production sector remains the same as in our baseline specification in the paper. We assume that the variable cost function is separable across UPCs and that supplying X_{ut}^{U} units of output of UPC u incurs a total variable cost of $A_{ut}\left(X_{ut}^{U}\right) = a_{ut}\left(X_{ut}^{U}\right)^{1+\delta_g}$, where a_{ut} is a cost shifter and δ_g parameterizes the elasticity of marginal costs with respect to output. We denote the production (wholesale) price charged by a producer by Q_{ut}^{U}. Each producer f also faces a fixed market entry cost for each product group of $H_{gt}^{F} > 0$ (e.g. fixed costs of supplying the market) and a fixed market entry cost for each UPC supplied of $H_{gt}^{U} > 0$ (e.g. fixed costs of product development and distribution). To simplify notation, we again suppress the product group subscript g on δ_g, unless otherwise indicated.

As in the baseline specification in the paper, we assume that producers choose prices under Bertrand competition, though it is straightforward to also consider the case in which they choose quantities under Cournot competition. We assume that each producer is small relative to the aggregate economy and the upper tier of utility is Cobb-Douglas, which implies that the producer's problem is separable by product group. With CES preferences within product groups, the decisions of any one producer in a product group only affect the decisions of other producers within that product group through the product-group consumption price index (P_{gt}^{G}). Each producer f within product group g chooses its set of UPCs $u \in \left\{\underline{u}_{fgt}, \ldots, \overline{u}_{fgt}\right\}$ and their prices $\{Q_{ut}^{U}\}$ to maximize its profits, taking into account these effects on the product-group consumption price index:

$$\max_{\left\{\underline{u}_{fgt},\ldots,\overline{u}_{fgt}\right\},\{Q_{ut}\}} \Pi_{fgt}^{F} = \sum_{k=\underline{u}_{fgt}}^{\overline{u}_{fgt}} \left[Q_{kt}^{U} X_{kt}^{U} - A_{kt}\left(X_{kt}^{U}\right)\right] - N_{fgt}^{U} H_{gt}^{U} - H_{gt}^{F}, \tag{61}$$

subject to the constraints that in equilibrium:

$$Q_{ut}^{U} = \frac{P_{ut}^{U}}{\tau},$$

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

Therefore each producer chooses production goods (wholesale) prices (Q^U_{ut}) to maximize its profits taking into account that wholesale prices are a multiple of consumption (retail) prices ($Q^U_{ut} = P^U_{ut}/\tau$), wholesale output (X^U_{ut}) is a multiple of retail output (Y^U_{ut}) which equals consumption ($X^U_{ut} = \tau Y^U_{ut} = \tau C^U_{ut}$), and taking into account the effect of its choice of wholesale price ($Q^U_{ut} = P^U_{ut}/\tau$) on the firm and product-group consumption goods price indices (P^F_{fgt} and P^G_{gt} respectively). From the first-order conditions for profit maximization, we obtain the same common markup across UPCs within a given producer and product group as in our baseline specification in the paper:

$$\mu^F_{fgt} = \frac{\varepsilon^F_{fgt}}{\varepsilon^F_{fgt} - 1}, \tag{62}$$

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

$$\varepsilon^F_{fgt} = \sigma^F - \left(\sigma^F - 1\right) S^F_{fgt} = \sigma^F \left(1 - S^F_{fgt}\right) + S^F_{fgt}, \tag{63}$$

and the producer's pricing rule takes the same form as in our baseline specification in the paper:

$$Q^U_{ut} = \mu^F_{fgt} \gamma_{ut} \tag{64}$$

where $\gamma_{ut} = (1+\delta)\, a_{ut} \left(X^U_{ut}\right)^{\delta}$ denotes marginal cost. Therefore retail prices are given by:

$$P^U_{ut} = \tau Q^U_{ut} = \tau \mu^F_{fgt} \gamma_{ut}, \tag{65}$$

where $\gamma_{ut} = (1+\delta)\, a_{ut} \left(X^U_{ut}\right)^{\delta} = (1+\delta)\, a_{ut} \left(\tau Y^U_{ut}\right)^{\delta}$.

As in our baseline specification in the paper, markups are the same across all UPCs for a given producer and product group, because each producer internalizes that it is the monopoly supplier of its consumption index within that product group, and retailing costs are the same for all UPCs. To the extent that producers are active in multiple product groups, markups again vary across product groups within producers, as in our baseline specification. Therefore we have shown that our model of the production sector can be embedded within a simple model of the retail sector in way consistent with our baseline specification in the paper.

S7 Parameter Grid Robustness Test

In our baseline specification in the paper, we estimate the parameters of the model (σ^U, σ^F, δ) using a generalization of the estimation procedure of Broda and Weinstein (2006, 2010) and Feenstra (1994) to incorporate multiproduct firms that are large relative to the markets in which they operate. In this section of the online supplement, we demonstrate the robustness of our findings that firm appeal and scope explain most of observed firm sales variation to alternative estimates of the model's parameters. In particular, we consider a range of possible values for σ^F at intervals of 0.5 from 2 (half our median estimate for σ^F) to 12 (four times our median estimate for σ^F) and σ^U ranging from 2.5 (just under half the median estimate) to 20 (almost triple the median estimate). Consistent with our empirical estimates in the paper, we restrict attention to values on this parameter grid for which $\sigma^U > \sigma^F$. Since setting $\sigma^F = 12$ exceeds all estimates of this parameter that we were aware of, and a value of $\sigma^U = 20$ means that a firm's products are highly substitutable, our parameter ranges cover the plausible values used in other studies, and thus can serve as a test of whether any plausible values would likely yield different results.

For each value (σ^U, σ^F) on the resulting parameter grid, we decompose firm sales variation into the contributions of (a) firm appeal, (b) product scope and (c) an average product appeal-adjusted prices residual as follows:

$$\begin{aligned} \ln E^F_{fgt} &= (\sigma^F - 1)\ln \varphi^F_{fgt} + \ln E^G_{gt} + (\sigma_F - 1)\ln P^G_{gt} + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\ln N^U_{fgt} \\ &\quad + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\ln\left(\frac{1}{N^U_{fgt}}\sum_{u \in \Omega^U_{fgt}}\left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1-\sigma^U}\right), \end{aligned}$$

and hence:

$$\Delta^g \ln E^F_{fgt} = \left\{ \underbrace{(\sigma^F - 1)\Delta^g \ln \varphi^F_{fgt}}_{\text{firm demand}} \underbrace{+ \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\Delta^g \ln N^U_{fgt}}_{\text{firm scope}} \underbrace{+ \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\Delta^g \ln\left(\frac{1}{N^U_{fgt}}\sum_{u \in \Omega^U_{fgt}}\left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1-\sigma^U}\right)}_{\text{average product appeal-adjusted prices residual}} \right\}, \tag{66}$$

where Δ^g is the difference operator relative to the geometric mean for product group g.

The decomposition (66) can be undertaken regardless of the value of the marginal cost elasticity (δ), because it is based on observed prices. The average product appeal-adjusted prices residual includes the net contribution of (i) product appeal, (ii) the mark-up, (iii) average marginal cost, (iv) cost dispersion. The value of δ does not affect the relative importance of (a) firm appeal, (b) product scope and (c) the average product appeal-adjusted prices residual. Instead this parameter merely affects the decomposition of the residual into its constituent components (because it affects marginal costs). Therefore we are not required to make any assumptions about the value of δ in order to undertake the decomposition (66).

In Figures S1-S2, we show the shares of (a) firm appeal ($(\sigma^F - 1)\Delta^g \ln \varphi^F_{fgt}$), (b) product scope ($\left(\frac{\sigma^F-1}{\sigma^U-1}\right)\Delta^g \ln N^U_{fgt}$) and (c) average product appeal-adjusted prices residual in firm sales variation for each value of (σ^U, σ^F) on the parameter grid. As in the paper, we report the results as the sales-weighted average of the results for each product group and quarter. As apparent from the figures, across the range of values for (σ^U, σ^F) on the parameter grid, we find that most of the observed variation in firm sales is explained by firm appeal and scope. Appeal accounts for over half of the variation in all cases. Scope remains important, too, except in cases where σ^U is very large, which renders within-firm product variety unimportant. Even so, appeal and scope account for over 80 percent of the variation in all specifications. In contrast, the residual (that includes all of the other components of our decomposition in the paper) accounts for a relatively small share of the observed variation in firm sales for all parameter values.

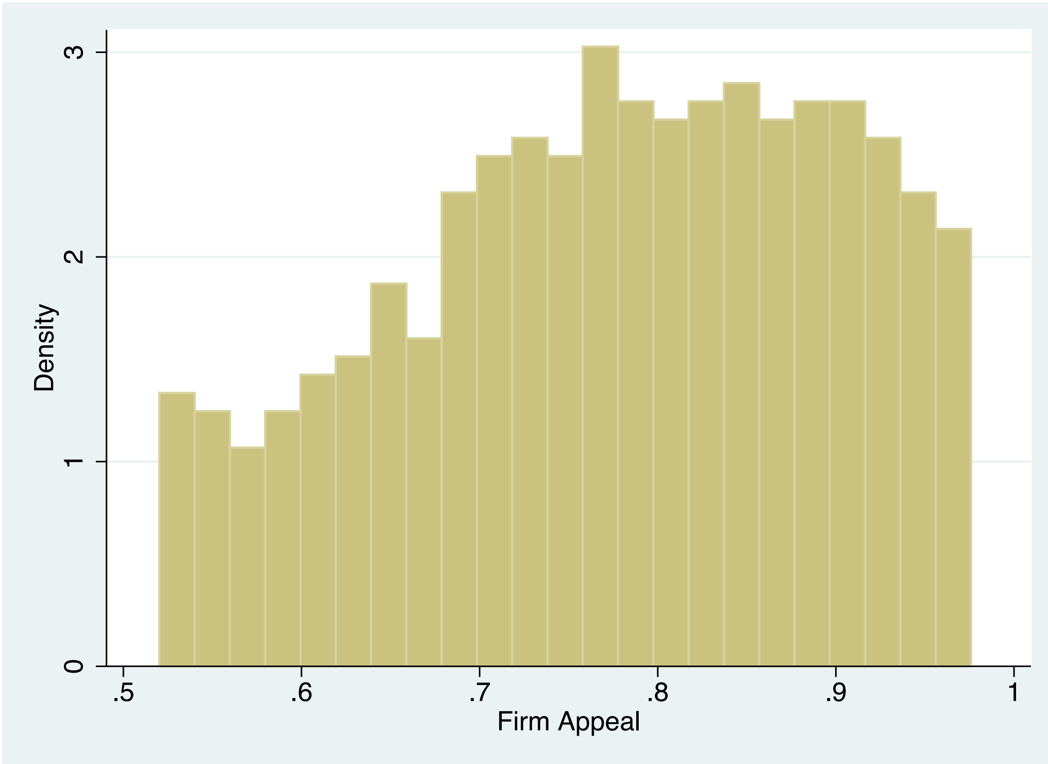

Figure S1: Histogram of Share of Firm Demand in Firm Sales Variation Across the Parameter Grid

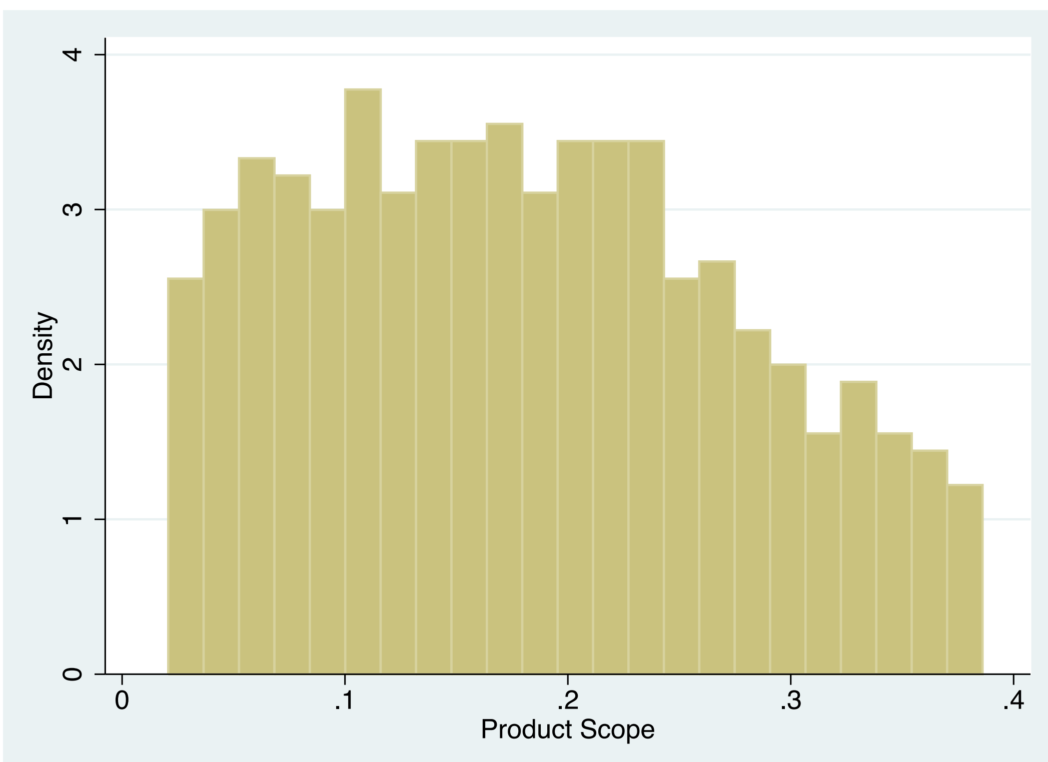

Figure S2: Histogram of Share of Product Scope in Firm Sales Variation Across the Parameter Grid

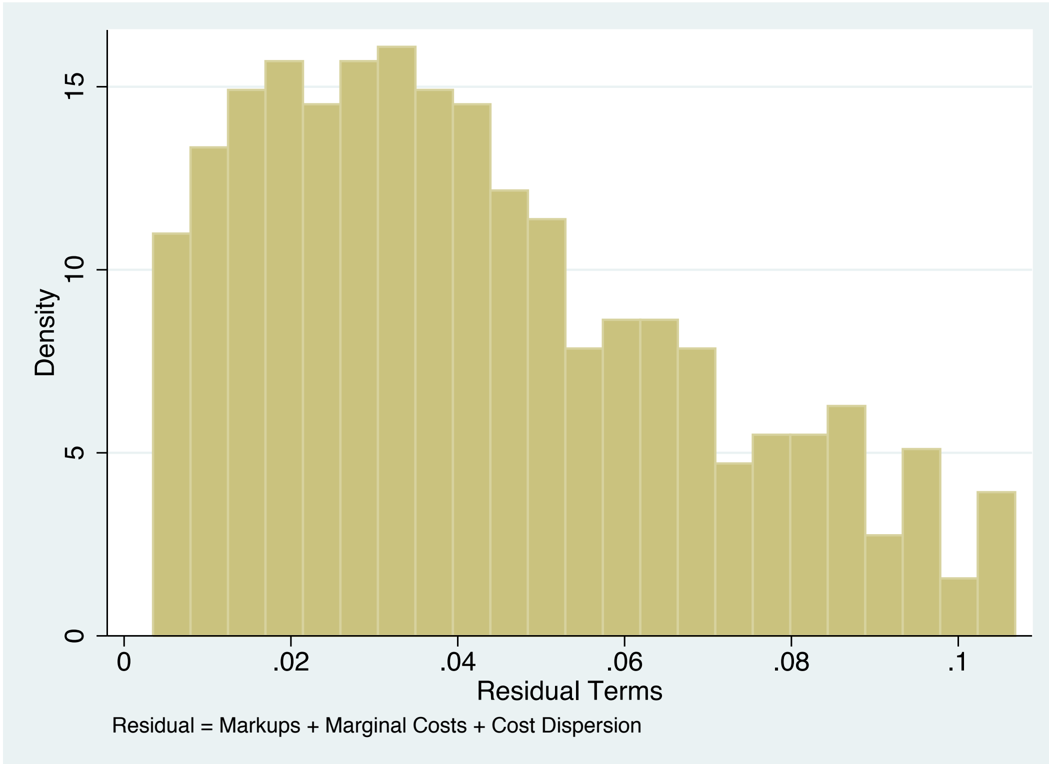

Figure S3: Histogram of Share of Average Product Appeal-Adjusted Prices Residual in Firm Sales Variation Across the Parameter Grid

Therefore *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.

S8 Monte Carlo Simulation

In this section, we report the results of a Monte Carlo simulation for our estimation procedure. We first assume parameter values and a data generating process for firm appeal (φ^F_{ft}), product appeal (φ^U_{ut}) and the cost shock (a_{ut}) for a hypothetical product group of firms and UPCs.[1] We next apply our estimation procedure to this hypothetical product group and show that it correctly recovers the assumed parameter values.

In particular, we consider a hypothetical product group comprised of 500 firms, each of which supplies 10 UPCs for a period of 200 quarters. We assume parameter values of $\{\sigma^U, \sigma^F, \delta\}$=$\{7, 4, 0.15\}$. We consider 50 replications of the model using realizations of random draws for firm appeal

[1]Throughout this section, we suppress the product group subscript g, because we consider a single product group.

(φ^F_{ft}), product appeal (φ^U_{ut}) and the cost shock (a_{ut}). For each replication, we undertake the following procedure:

- Step 1: We draw realizations for firm appeal (φ^F_{ft}), product appeal (φ^U_{ut}) and the cost shock (a_{ut}) from heteroskedastic log normal distributions. In particular, for each UPC-quarter observation, we draw product appeal (φ^U_{ut}) and the cost shock (a_{ut}) from an independent log normal distribution with a mean of zero and a standard deviation equal to one plus the realization of a uniform (0,1) random variable. Having generated these realizations for product appeal (φ^U_{ut}) and the cost shock (a_{ut}), we normalize their values such that they have a geometric mean of one within each firm. For each firm-quarter, we draw a realization for firm appeal (φ^F_{ft}) from the same independent log normal distribution with a mean of zero and a standard deviation of one plus the realization of a uniform (0,1) random variable. Having generated these realizations for firm appeal (φ^F_{ft}), we normalize their values such that they have a geometric mean of one within the product group.

- Step 2: Using the assumed values for the model's parameters $\{\sigma^U, \sigma^F, \delta\}$=$\{7, 4, 0.15\}$, the realizations of the random draws for $\{\varphi^F_{ft}, \varphi^U_{ut}, a_{ut}\}$ for each firm, UPC and quarter, and an assumed value of total expenditure on the product group ($E^G_{gt} = 100$), we solve for equilibrium expenditure shares and prices $\{P^U_{ut}, S^U_{ut}, P^F_{ft}, S^F_{ft}\}$ using the system of equations for equilibrium in the model:
$$P^U_{ut} = \left(\mu^F_{ft}\right)^{\frac{1}{1+\delta}} (1+\delta)^{\frac{1}{1+\delta}} a_{ut}^{\frac{1}{1+\delta}} \left(S^U_{ut}\right)^{\frac{\delta}{1+\delta}} \left(E^F_{ft}\right)^{\frac{\delta}{1+\delta}}, \tag{67}$$
$$S^U_{ut} = \frac{\left(P^U_{ut}/\varphi^U_{ut}\right)^{1-\sigma^U}}{\sum_{k\in\Omega^U_{ft}} \left(P^U_{ut}/\varphi^U_{ut}\right)^{1-\sigma^U}}, \tag{68}$$
$$P^F_{ft} = \left[\sum_{u\in\Omega^U_{ft}} \left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1-\sigma^U}\right]^{\frac{1}{1-\sigma^U}}, \tag{69}$$
$$S^F_{ft} = \frac{\left(P^F_{ft}/\varphi^F_{ft}\right)^{1-\sigma^F}}{\sum_{k\in\Omega^F_{gt}} \left(P^F_{kt}/\varphi^F_{kt}\right)^{1-\sigma^F}}, \tag{70}$$
where $E^F_{ft} = S^F_{ft}E^G_{gt}$.

- Step 3: We test our estimation procedure as outlined in Section 5 of the paper. We assume that we only observe UPC prices and expenditures $\{P^U_{ut}, E^U_{ut}\}$. We use these data and our UPC moment conditions to estimate $\{\sigma^U, \delta\}$ and recover unobserved UPC product appeal $\{\varphi^U_{ut}\}$. We next construct the firm price index $\{P^F_{ft}\}$ and use our firm moment conditions to estimate $\{\sigma^F\}$ and recover unobserved firm appeal $\{\varphi^F_{ft}\}$ and the cost shock $\{a_{ut}\}$.

In Figures S4-S6, we display histograms of the estimated values of the parameters across 50 replications, where the true values of the parameters are shown by the vertical red lines. Figure S4 shows the results for σ^U; Figure S5 displays the results for δ; and Figure S6 presents the results for σ^F. Across all three figures, we find that the estimated parameter values are tightly clustered around the true parameter values. In Table S1, we report the mean and standard error of the parameter estimates across the 50 replications. For all three of the model's parameters, we are unable to reject the null hypothesis that the estimated parameter is equal to the true parameter at conventional levels of significance. Therefore, when the data are generated according to the model, our estimation procedure is successful in recovering the correct values of the model's parameters.

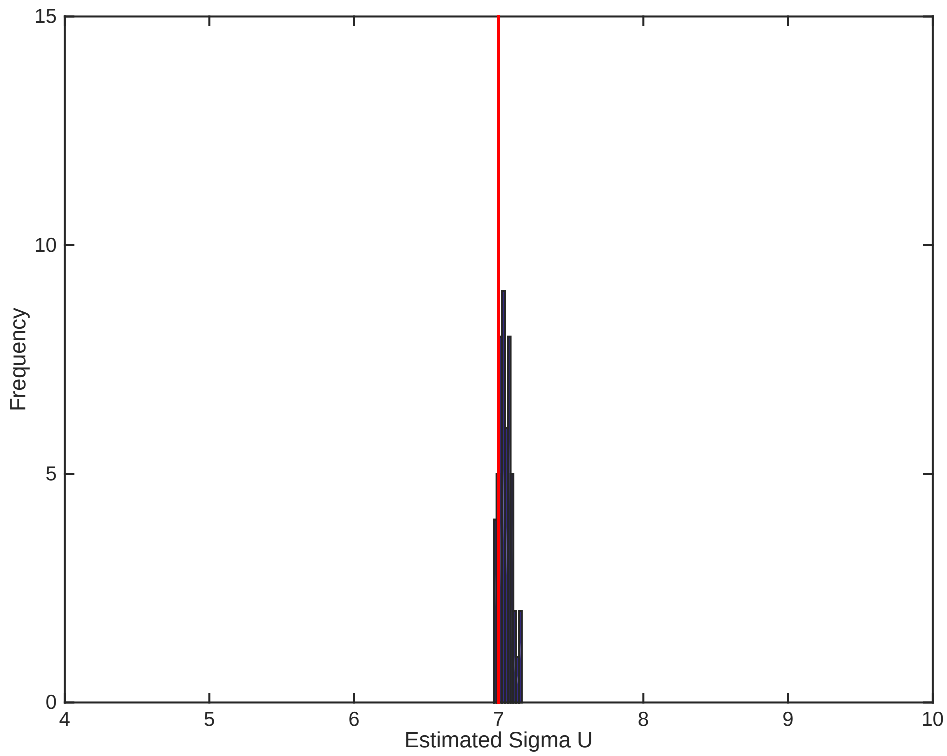

Figure S4: Monte Carlo Results for σ^U (50 Replications)

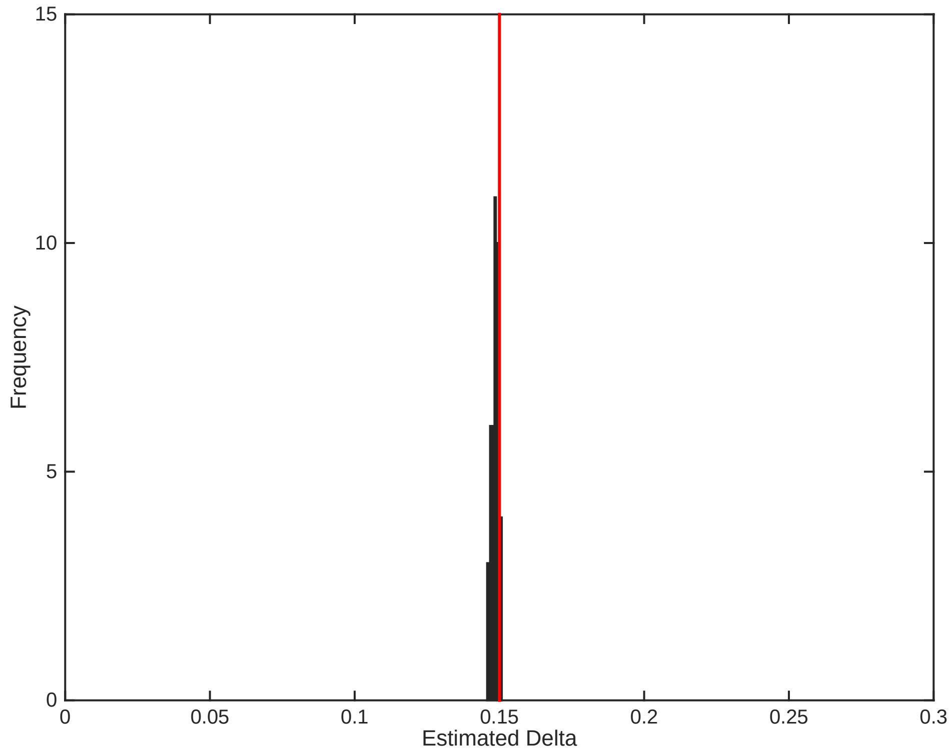

Figure S5: Monte Carlo Results for δ (50 Replications)

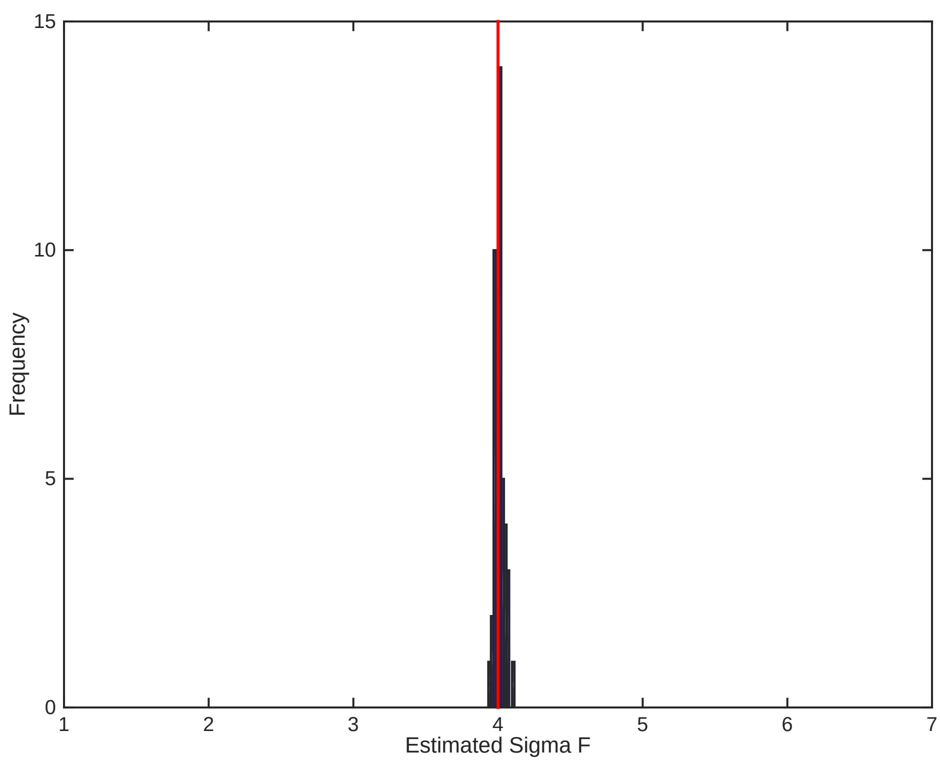

Figure S6: Monte Carlo Results for σ^F (50 Replications)

Table S1: Mean Parameter Estimates (50 Replications)

Parameter	Estimate
Product elasticity of substitution (σ^U)	7.045
	(0.046)
Firm elasticity of substitution (σ^F)	4.009
	(0.033)
Elasticity marginal costs (δ)	0.148
	(0.001)

Note: Standard errors in parentheses.

S9 Robustness Test using International Trade Transactions Data

In our baseline specification in the paper, we use bar-code data that consist mainly of consumer products. These consumer products account for around one third of the goods included in the Consumer Price Index (CPI)). Furthermore, these consumer products span a broad spectrum of goods, ranging from Carbonated Beverages to Car Parts and Accessories (many of which are traded internationally). Nonetheless, one concern is whether our findings using bar-code data would also hold for other categories of goods in international trade data (which includes both intermediate and consumer goods). In this section, we switch countries and product categories to show that even if we use Chilean international trade transactions data, we again find that firm appeal and scope account for most the observed variation in firm sales, confirming our results using bar-code data in the paper.

The Chilean trade transactions data take a similar form as other transactions datasets used in the international trade literature.[2] For each import shipment, we observe the product classification of the shipment (Harmonized System (HS) 8-digit), the value and quantity shipped, the date of the shipment, the source country, the shipping port, the importer's name, the exporter's brand, and a variety of other information. We define product groups (g) as HS 2-digit categories, firms (f) as the exporter's brand (*e.g.,* Nestle, Glade) and products (u) as an HS 8-digit category exported under a particular exporter's brand. We aggregate the import transactions data to the quarterly level (from 2007-14) and implement our estimation procedure from Section 5 of the paper. We first use our product moment conditions to estimate $\{\sigma^U, \delta\}$ and recover unobserved product appeal $\{\varphi^U_{ut}\}$. We next construct the firm price index $\{P^F_{fgt}\}$ and use our firm moment conditions to estimate $\{\sigma^F\}$ and recover unobserved firm appeal $\{\varphi^F_{fgt}\}$ and the cost shock $\{a_{ut}\}$. Finally, we use the observed data $\{P^U_{ut}, E^U_{ut}\}$ and the recovered unobservables $\{\varphi^U_{ut}, a_{ut}, \varphi^F_{fgt}\}$ to implement our decomposition of firm sales variation from Section 4.6 of the paper.

Table S2 reports the median parameter estimates $\{\sigma^U, \sigma^F, \delta\}$ across the HS 2-digit sectors and their standard errors. We find an elasticity of substitution across products (σ^U) of around 4.35, an elasticity of substitution across firms (σ^F) of around 2.26, and an elasticity of marginal cost with respect to output (δ) of around 0.54. These parameter estimates are within the range of our results

[2] See Bernard, Jensen and Schott (2009) and Bernard, Jensen, Redding and Schott (2007) for general discussions of international trade transactions data.

across product groups using bar-code data in the paper. The estimated values of σ^U and σ^F are slightly smaller using HS categories than using bar codes, which is consistent with the Broda and Weinstein (2006) finding of lower estimated elasticities of substitution using more aggregated data. Table S3 uses these estimates to decompose firm sales variation into its components. As for the results using bar-code data in the paper, we again find that most of the observed variation in firm sales is driven by firm appeal and scope, with a relatively small contribution from the other components of the decomposition. We find a somewhat larger contribution from firm appeal and a somewhat smaller contribution from scope than using bar-code data in the paper. This pattern of results is consistent with the idea that HS 8-digit categories are substantially more aggregated than bar codes (so that the bar-code extensive margin within HS 8-digit categories is captured in firm appeal).

Taken together, these results confirm that our finding that firm appeal is the most important source of firm sales variation is not driven by the use of bar-code data. Using international trade transactions data for Chile (including both intermediate and consumer goods) and a different definition of product (HS 8-digit categories rather than bar codes), we continue to find a similar pattern of results as in our baseline specification in the paper.

Table S2: GMM Parameter Estimates Using Chilean Trade Transactions Data

Parameter	Median Estimate
σ^U	4.354
	(0.254)
σ^F	2.257
	(0.166)
δ	0.541
	(0.053)

Note: Median parameter estimates across 50 Harmonized System (HS) 2-digit sectors. Firm corresponds to exporter brand (e.g. Nestle, Glade). Product corresponds to HS 8-digit category exported under a particular exporter's brand. Standard errors in parentheses.

Table S3: Firm Sales Decomposition Using Chilean Trade Transactions Data

Component	Percentage Contribution in Cross-Section	Percentage Contribution in Growth
Firm Appeal	97.26	96.31
Product Scope	5.91	10.76
Average MC	-6.92	-14.81
Markup	-0.02	-0.02
Cost Dispersion	3.78	7.77

Note: MC is marginal cost First column is the variance decomposition of the level of log sales across firms (equations (19)–(23) in the paper). Second column is the variance decomposition of the change in log sales. Results are the sales-weighted average of the results for each Harmonized System (HS) 2-digit sector and quarter.

S10 Comparison to Other Elasticity Estimates

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, σ^F. Gordon et al. (2013) and the literature review therein presents results for comparable product groups using quite different estimation methodologies. In Table S4, we report our estimate of σ^F for each product group and the corresponding estimate from Gordon et al. (2013). As apparent from the table, we find similar estimated elasticities of substitution, with an average value across the twelve overlapping product groups of 3.65 compared to 3.14 in Gordon (a difference of around 16 percent). 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. Therefore there is no reason to expect our estimated elasticities of substitution to be exactly the same as these values from the existing literature. Nonetheless, despite this caveat, the empirical estimates generated by our procedure are reasonable compared to the benchmark of findings from other empirical studies.

Table S4: Comparison of Elasticities to Literature

Product Category	σ_F	Prior Literature
Butter & Margarine	2.41	1.73†
Carbonated Soft Drinks	5.18	3.09*
Coffee	3.89	4.50*
Deodorant	3.71	2.92†
Ketchup	3.28	2.92*
Laundry Detergent	3.91	1.96*
Mayo	3.39	3.84†
Mustard	3.28	2.23†
Peanut Butter	3.88	2.70*
Spaghetti Sauce	3.28	3.80*
Toilet Tissue	2.92	4.09*
Yogurt	4.68	3.87*
Average	3.65	3.14

Note: * taken from average of prior studies in Gordon [2013] Web Appendix. † taken from Gordon [2013]. 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.

S11 Marketing Literature on Cannibalization

As reported in Section 6 of the paper, our parameter estimates imply substantial cannibalization effects, with around 50 percent of the sales of new varieties coming at the expense of existing varieties. In this section, we compare our estimates of cannibalization effects to existing estimates in the marketing literature.

In making this comparison, it is important to note that there is no standard way to measure cannibalization effects in the marketing literature (which often defines "cannibalization" non-structurally as something akin to a cross-price elasticity), and these approaches are typically different from the theoretically-consistent measure of cannibalization effects in our framework. As discussed in Section 4.4 of the paper, our measure of cannibalization is the partial elasticity of the sales of existing varieties with respect to an increase in the number of varieties (reproduced below):

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

for firm f and product group g at time t. In our setup, cannibalization 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. Our framework emphasizes that cannibalization varies across product groups (with the elasticities of substitution σ_g^U and σ_g^F), across firms (with firm expenditure shares S_{fgt}^F and number of products N_{fgt}^U), and across newly-introduced products (with the expenditure shares of these products $S_{N_{fgt}^U}^U$). Therefore, even using a common measure of cannibalization effects, our framework implies that this measure should vary across sectors, firms, products and time periods.

Much of the existing marketing literature measures cannibalization in terms of the effect of a price reduction (or promotion or coupon) for one variety on the sales of existing varieties. In one of the first studies in this empirical literature, Gupta (1988) estimates a multinomial choice model using data for ground caffeinated coffee in Pittsfield, MA. Three measures of brand promotion are considered: (a) feature-or-display, (b) feature-and-display, (c) promotional price reduction, where "feature" corresponds to an in-store flyer and "display" corresponds to displayed in store. Of the overall elasticity of sales with respect to these three forms of promotion, around 84 percent is accounted for by brand switching, 14 percent or less by purchase time acceleration, and 2 percent by stockpiling. Similar results are found in a series of subsequent studies, including Chiang (1991), Chintagunta (1993), Bucklin, Gupta, and Siddarth (1998), and Bell, Chiang, and Padmanabhan (1999).

Clarifying the interpretation of these results, Van Heerde, Gupta and Wittink (2003) point out that finding 84 percent of the sales elasticity comes from brand switching does not imply that 84 percent of the sales increase comes from existing brands. Using a different decomposition based on sales and the same data as in some of these earlier studies, Van Heerde, Gupta and Wittink (2003) demonstrate a smaller contribution from brand switching of around one third. In subsequent work, Van Heerde, Leeflang and Wittink (2004) decompose sales increases due to price promotions into cross-period (stockpiling), cross-brand (brand switching) and category expansion (consumption) effects. Using data for four product categories (tuna, tissue, shampoo and peanut butter), these effects on average account for around one third of the sales increase. Finally, using data on twelve grocery product categories (seven US, three UK, two Australian), Dawes (2012) investigates the effect of price promotions for one pack size of a brand on the sales of other pack sizes of the same brand. On average, around 22 percent of the increase in sales for a promoted brand pack size comes from reduced sales of other pack sizes of the same brand.

Closer to our measure of cannibalization, Raghavan Srinivasan, Ramakrishnan and Grasman (2005a) undertake a case study of the effect of the introduction of a new product set using data from an anonymous U.S. consumer beverage company. Following the introduction of the new product set 3, sales of the existing product sets 1 and 2 are found to drop by 44 and 24 percent respectively, without any change in prices for these existing product sets. Using the same data, Raghavan Srinivasan, Ramakrishnan and Grasman (2005b) estimate the magnitude of cannibalization using an augmented Autoregressive Integrated Moving Average (ARIMA) model, and find that cannibalization reduces predicted sales by 28-46 percent. Our estimate of the cannibalization rate for a firm in the carbonated beverages sector ($\sigma^F = 5.18$ and $\sigma^U = 7.82$) ranges from 39 percent to 51 percent assuming the firm has a market share of under 20 percent and introduces a product that has a market share equal to the typical product of that firm. The cannibalization rate would be lower otherwise. Therefore, despite the caveats noted above about the different definitions of cannibalization and heterogeneity across firms and sectors, our cannibalization rate is quite similar to the marketing literature's estimates.

S12 Firm Sales Decomposition by Product Group

In this section, we report the results of our decomposition of firm sales variation into the contributions of (a) firm appeal, (b) product scope and (c) an average product appeal-adjusted prices residual for the 10 largest product groups. From the decomposition in the paper, firm sales can be written as:

$$\begin{aligned} \ln E^F_{fgt} &= (\sigma^F - 1)\ln \varphi^F_{fgt} + \ln E^G_{gt} + (\sigma^F - 1)\ln P^G_{gt} + \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\ln N^U_{fgt} \\ &+ \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\ln\left(\frac{1}{N^U_{fgt}}\sum_{u\in\Omega^U_{fgt}}\left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1-\sigma^U}\right), \end{aligned}$$

and hence:

$$\Delta^g \ln E^F_{fgt} = \left\{ \underbrace{(\sigma^F - 1)\Delta^g \ln \varphi^F_{fgt}}_{\text{firm demand}} \underbrace{+ \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\Delta^g \ln N^U_{fgt}}_{\text{firm scope}} \underbrace{+ \left(\frac{\sigma^F - 1}{\sigma^U - 1}\right)\Delta^g \ln\left(\frac{1}{N^U_{fgt}}\sum_{u\in\Omega^U_{fgt}}\left(\frac{P^U_{ut}}{\varphi^U_{ut}}\right)^{1-\sigma^U}\right)}_{\text{average product demand-adjusted prices residual}} \right\}, \tag{71}$$

where Δ^g is the difference operator relative to the geometric mean for product group g. The average product appeal-adjusted prices residual, which we call $\Delta^g \ln \rho_{fgt}$, includes the net contribution of (i) product appeal, (ii) the mark-up, (iii) average marginal cost, (iv) cost dispersion.

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 (71) on log firm sales as follows:

$$\begin{aligned}(\sigma_g^F - 1)\Delta^g \ln \varphi_{fgt}^F &= \alpha_g^\varphi \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^\varphi, \\ \left(\frac{\sigma_g^F - 1}{\sigma_g^U - 1}\right)\Delta^g \ln N_{fgt}^U &= \alpha_g^N \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^N, \\ \Delta^g \ln \rho_{fgt}^F &= \alpha_g^\rho \Delta^g \ln E_{fgt}^F + \varepsilon_{fgt}^\rho, \end{aligned} \tag{72}$$

where we have again differenced relative to the geometric mean of the product group. We allow the coefficients $\{\alpha_g^\varphi, \alpha_g^N, \alpha_g^\rho\}$ 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 $\alpha_g^\varphi + \alpha_g^N + \alpha_g^\rho = 1$.

Table S5 reports the results of this decomposition of variation in log firm sales for the 10 largest product groups. Although we find some differences in the relative importance of each component across product groups, we find that firm appeal and scope account for the majority of firm sales variation across all 10 product groups.

Table S5: Variance Decomposition (Top 10 Product Groups)

	Firm Appeal	Scope	Residual	σ_F	σ_U
Carbonated Beverages	0.408	0.278	0.313	5.18	7.82
	(0.014)	(0.002)	(0.013)	(0.20)	(0.53)
Pet Food	0.458	0.298	0.245	5.51	8.15
	(0.021)	(0.002)	(0.021)	(0.26)	(0.46)
Bread And Baked Goods	0.495	0.280	0.225	4.44	6.62
	(0.004)	(0.001)	(0.004)	(0.07)	(0.33)
Paper Products	0.763	0.244	-0.007	2.92	4.44
	(0.007)	(0.001)	(0.007)	(0.09)	(0.31)
Tobacco And Accessories	0.419	0.214	0.366	4.33	6.98
	(0.017)	(0.002)	(0.016)	(0.22)	(0.77)
Prepared Foods-Frozen	0.755	0.142	0.103	3.93	8.34
	(0.004)	(0.001)	(0.004)	(0.09)	(0.61)
Snacks	0.629	0.219	0.152	4.37	7.32
	(0.006)	(0.001)	(0.005)	(0.09)	(0.60)
Milk	0.772	0.141	0.088	2.41	5.05
	(0.003)	(0.001)	(0.002)	(0.06)	(0.21)
Packaged Meats-Deli	0.737	0.192	0.071	2.67	4.86
	(0.003)	(0.001)	(0.002)	(0.05)	(0.27)
Juice, Drinks - Canned, Bottled	0.705	0.187	0.108	2.95	5.25
	(0.004)	(0.001)	(0.004)	(0.05)	(0.21)

Note: Rows in rank order. 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. Standard errors in parentheses.

S13 List of Product Groups

Table S6: Product Groups

Product Group	Sales Share	No. Firms
Automotive	0.003	624
Baby Food	0.005	90
Baby Needs	0.003	562
Baked Goods-Frozen	0.004	559

Baking Mixes	0.004	470
Baking Supplies	0.005	883
Batteries And Flashlights	0.007	610
Beer	0.015	704
Bread And Baked Goods	0.033	2430
Breakfast Food	0.009	342
Breakfast Foods-Frozen	0.005	371
Butter And Margarine	0.007	370
Candy	0.023	1953
Carbonated Beverages	0.036	773
Cereal	0.022	433
Cheese	0.026	994
Coffee	0.011	706
Condiments, Gravies, And Sauces	0.014	2932
Cookies	0.011	1424
Cosmetics	0.005	719
Cot Cheese, Sour Cream, Toppings	0.006	445
Cough And Cold Remedies	0.008	442
Crackers	0.008	639
Deodorant	0.003	162
Desserts, Gelatins, Syrup	0.005	496
Desserts/Fruits/Toppings-Frozen	0.003	363
Detergents	0.015	291
Diet Aids	0.002	151
Disposable Diapers	0.007	125
Dough Products	0.004	230
Dressings/Salads/Prep Foods-Deli	0.021	2002
Eggs	0.007	563
Electronics, Records, Tapes	0.017	1132
Feminine Hygiene	0.001	181
First Aid	0.004	798
Flour	0.001	327
Fresheners And Deodorizers	0.005	647
Fruit - Canned	0.004	534
Fruit - Dried	0.004	564
Glassware, Tableware	0.003	1251
Grooming Aids	0.002	907
Gum	0.002	325
Hair Care	0.011	661
Hardware, Tools	0.004	1209
Household Cleaners	0.007	833
Household Supplies	0.008	1437
Housewares, Appliances	0.014	809
Ice Cream, Novelties	0.016	714
Insecticds/Pesticds/Rodenticds	0.003	443
Jams, Jellies, Spreads	0.005	952
Juice, Drinks - Canned, Bottled	0.026	1493
Juices, Drinks-Frozen	0.001	150
Kitchen Gadgets	0.004	1991
Laundry Supplies	0.008	864
Light Bulbs, Electric Goods	0.007	763
Liquor	0.012	485
Medications/Remedies/Health Aids	0.024	1759
Milk	0.029	702
Nuts	0.008	829
Oral Hygiene	0.009	496
Packaged Meats-Deli	0.029	1143
Packaged Milk And Modifiers	0.006	451
Paper Products	0.033	855
Pasta	0.004	672
Personal Soap And Bath Additives	0.006	888
Pet Care	0.012	1295

Pet Food	0.035	641
Photographic Supplies	0.004	191
Pickles, Olives, And Relish	0.004	863
Pizza/Snacks/Hors Doeurves-Frzn	0.012	773
Prepared Food-Dry Mixes	0.010	941
Prepared Food-Ready-To-Serve	0.011	1399
Prepared Foods-Frozen	0.031	1774
Pudding, Desserts-Dairy	0.001	168
Salad Dressings, Mayo, Toppings	0.007	712
Sanitary Protection	0.004	118
Seafood - Canned	0.004	521
Shaving Needs	0.004	286
Shortening, Oil	0.005	715
Skin Care Preparations	0.006	924
Snacks	0.031	2513
Snacks, Spreads, Dips-Dairy	0.002	1011
Soft Drinks-Non-Carbonated	0.012	1460
Soup	0.011	679
Spices, Seasoning, Extracts	0.005	1606
Stationery, School Supplies	0.011	2196
Sugar, Sweeteners	0.005	346
Table Syrups, Molasses	0.002	424
Tea	0.007	779
Tobacco And Accessories	0.032	510
Unprep Meat/Poultry/Seafood-Frzn	0.011	997
Vegetables - Canned	0.010	950
Vegetables And Grains - Dried	0.002	512
Vegetables-Frozen	0.010	474
Vitamins	0.019	1204
Wine	0.012	1717
Wrapping Materials And Bags	0.008	469
Yogurt	0.009	313

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.

References

[1] Bell, David R., Jeongwen Chiang and Padmanabhan (1999) "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18(4), 504-26.

[2] Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding and Peter K. Schott (2007) "Firms in International Trade," *Journal of Economic Perspectives,* 21(3), 105-130.

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

[4] Broda, Christian and David E. Weinstein (2006) "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541-86.

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

[6] Bucklin, Randolph E., Sunil Gupta, and S. Siddarth (1998), "Determining Segmentation in Sales Response across Consumer Purchase Behaviors," *Journal of Marketing Research*, 35(2), 189-197.

[7] Chiang, Jeongwen (1991), "A Simultaneous Approach to the Whether, What and How Much to Buy Questions," *Marketing Science*, 10(4), 297-315.

[8] Chintagunta, Pradeep K. (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*, 12(2), 184-208.

[9] Dawes, John G. (2012) "Brand-Pack Size Cannibalization Arising from Temporary Price Promotions," *Journal of Retailing*, 88(3), 343-55.

[10] Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2004) "Dissecting Trade: Firms, Industries, and Export Destinations," *The American Economic Review Papers and Proceedings*, , Vol. 94, No. 2, , pp. 150-154.

[11] Feenstra, Robert C. (1994) "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157-177.

[12] Feenstra, Robert and David E. Weinstein (2015) "Globalization, Markups, an the U.S. Price Level," *Journal of Political Economy*, forthcoming.

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

[14] Gupta, Sunil (1988) "Impact of Sales Promotions on When, What and How Much to Buy," *Journal of Marketing Research*, November, 342-55.

[15] Raghavan Srinivasan, Sundara, Sreeram Ramakrishnan and Scott E. Grasman (2005a) "Identifying the Effects of Cannibalization on the Product Portfolio," *Marketing Intelligence and Planning*, 23(4), 359-71.

[16] Raghavan Srinivasan, Sundara, Sreeram Ramakrishnan and Scott E. Grasman (2005b) "Incorporating Cannibalization Models into Demand Forecasting," *Marketing Intelligence and Planning*, 23(5), 470-85.

[17] Van Heerde, Herald J., Sachin Gupta and Dick R. Wittink (2003), "Is 75% of the Sales Promotion Bump Due to Brand Switching? No, Only 33% is," *Journal of Marketing Research*, 40(4), 481-491.

[18] Van Heerde, Herald J., Peter. S. H. Leeflang and Dick R. Wittink (2004), "Decomposing the Sales Promotion Bump with Store Data," *Marketing Science*, 23(3), 317-334.