

The Extensive Margin of Aggregate Consumption Demand

Claudio Michelacci
EIEF and CEPR

Luigi Paciello
EIEF and CEPR

Andrea Pozzi*
EIEF and CEPR

May 1, 2018

Preliminary and Incomplete

Abstract

We document that around one half of the cyclical variation in aggregate non-durable consumption expenditures by US households comes from changes in the products entering their consumption basket. Most of this variation is due to changes in the rate at which households add new products to their basket, while removals from the basket are relatively acyclical. These patterns hold true within narrowly defined sectors of products or quality categories and are only partly driven by changes in the price of products or their availability in the market. We rationalize this evidence by incorporating a conventional random utility model of discrete choice of products into a standard household dynamic optimization problem. Household preferences over products in her consideration set randomly vary over time and because of this a larger set reduces the welfare relevant household price index. The household can save in financial assets and decides how much to spend in experimenting for new products to be added to her consideration set. In response to income shocks the household increases savings and experiments more, which allows to smooth consumption by persistently reducing her future price index. The calibrated model predicts that experimentation expenditures fluctuate by around 15 percent from peak to bottom in the business cycle. This experimentation channel has novel implications for consumption smoothing, the measurement of household level inflation, and the role of aggregate demand stabilization policies.

*We acknowledge the financial support of the European Research Council under ERC Starting Grant 676846. We thank Rossella Mossucca and Federica Di Giacomo for excellent research assistance. We thank Jeff Campbell and Guido Menzio for useful discussions. E-mail: gigipaciello@gmail.com.

1 Introduction

There has been much renewed interest in the theoretical determinants of aggregate consumption expenditures, see for example Lorenzoni (2009), Philippon and Midrigan (2011), Kaplan and Violante (2014), Bai, Ríos-Rull, and Storesletten (2012) and Kaplan and Menzio (2016). In this paper we emphasize that a household can increase her consumption expenditures by spending more in products that were already in her consumption basket in previous periods (intensive margin), by increasing the number of products in the basket (extensive margin) or both. We propose a simple methodology to decompose changes in aggregate consumption expenditures along the intensive and extensive margin. We further decompose the extensive margin into a component due to product *additions*, which arises because households add new products to their consumption basket, and another one due to product *removals*, which arises because households stop buying products previously in their consumption basket. The overall response of the extensive margin corresponds to *net* additions, equal to the difference between product additions and product removals.

We rely on the Kilts-Nielsen Consumer Panel (KNCP) over the 2000's—that tracks consumption in nondurables for a representative panel of US households—to document that about half of the cyclical dynamics of aggregate consumption expenditures is due to variation in net additions, with most of this variation due to the procyclical properties of the rate at which households add new varieties to their basket, while removals from the basket are relatively acyclical. We also find that the new varieties added to the consumption basket of the household in a period tend to remain persistently into the basket of the household also in future periods, independently of the future dynamics of expenditures. These patterns hold true at different frequencies (quarterly or annual), within narrowly defined sectors of products or quality categories and they are only partly driven by changes in the price of products or their availability in the market.

We interpret this evidence by emphasizing that households do not like all varieties available in the market and to get to know whether they like one of them they should first buy and consume it. We model this demand for experimentation, by incorporating a conventional random utility model of discrete choice of products into a standard household dynamic optimization problem. Building on McFadden (1974a, 1974b, 1978, and 1981) we assume that household preferences over the varieties in her consideration set randomly vary over time according to a generalized extreme value distribution and the household chooses

to consume the variety she likes the most in the period.¹ There is a continuum of sectors and in each sector the consideration set of the household contains a discrete number of varieties. Due to the preference shocks, a larger set reduces the welfare relevant household price index, which provides a micro-foundation for love of variety. The household can save in financial assets and decides how much to spend in experimenting for new varieties to be added to her consideration sets. Consumption can be adjusted along the extensive margin (the number of varieties purchased), as well as along the intensive margin (how much to spend in the varieties purchased in the previous period). In response to income shocks the household increases savings and experiments more, which allows to smooth consumption by persistently reducing her future price index.

Due to the discrete choice framework, the consumption basket of the household in a period is a subset of the consideration set of the household. Expenditures in additions to the consumption basket of the household are the sum of (i) expenditures for experimentation, (ii) consumption expenditures in varieties truly newly added to the consideration set of the households, in brief *true additions*, and (iii) expenditures in varieties already part of the consideration set of the household in previous periods, that we refer to as *false additions* to the consideration set. We rely on the panel structure of the KNCP to separately identify the relative importance of these three components and calibrate the model targeting detailed scanner data. As in Kaplan and Violante (2014) we allow for some costs to adjust household holdings of financial assets. The adjustment costs are calibrated to match the response of household consumption expenditures to the 2008 economic stimulus payments by the US federal government to households, which as in Parker, Souleles, Johnson, and McClelland (2013) and Broda and Parker (2014), we found implied a marginal propensity to consume of fifty percent in response to a payment equivalent to around ten percent of quarterly consumption expenditures. This marginal propensity to consume is the sum of a propensity to consume along the extensive margin and one along the intensive margin. In response to the stimulus payment, the overall marginal propensity to consume is explained by more than two thirds by the increase in expenditures in products newly added to the basket. The contribution of removals to the response of the extensive margin is small, while the intensive margin accounts for the remaining one third of the overall marginal propensity

¹Luce (1959) used an axiomatic approach to characterize the equilibrium probability of a given choice under a discrete set of alternatives. McFadden (1974b) provides the first micro-foundation for the equilibrium probabilities in Luce (1959) by considering the problem of individuals who maximizes their utility under additive random utility shocks distributed according to a type-I extreme-value distribution. The formulation of the random utility model of discrete choice with random utility shocks characterized by Generalized Extreme Value distributions is due to McFadden (1978, 1981).

to consume.

To identify how experimentation expenditures fluctuates over the cycle, we assume that household income fluctuates due to changes in the demand for labour and capital, which for simplicity we assume are driven by technology shocks—as in standard (real) business cycle models which typically reproduce well the cyclical properties of consumption expenditures. We back up the shocks of the model to match the time series profile of aggregate consumption expenditures in the data. The calibrated model matches well the cyclical properties of the intensive margin as well the time series profile of gross and net additions. The model predicts that households’ expenditures for experimentation are strongly pro-cyclical and substantially more volatile than aggregate expenditures: experimentation is 10% below normal times at the bottom of the recession, while the corresponding fall in total consumption expenditure is just 3%. The fall in experimentation expenditures implies that the number of varieties in the consideration set of households falls sluggishly and it reaches its trough one year later than the trough in consumption expenditures. During the recovery, households experiment more and gradually rebuild their consideration set. The procyclical properties of experimentation expenditures makes household level inflation less procyclical, which implies that aggregate inflation statistics tends generally to overestimate the welfare relevant measure of inflation for households.

Some references to the literature Several authors have emphasized that the consumption basket of households changes over the business cycle, but the distinction between the extensive and the intensive margin in aggregate consumption expenditures has been typically overlooked. Jaimovich, Rebelo, and Wong (2017) show that, over the cycle, households substitute across products of different quality while Argente and Lee (2017) emphasize that in recessions wealthier households have greater ability to substitute towards lower quality products because they normally consume high quality products. In these papers substitution occurs along the intensive margin. Other papers have emphasized that households can pay less for the products they buy by searching more intensively for lower prices: Aguiar and Hurst (2007) document that older households pay lower prices for the same products because they allocate more time to shopping activities; Coibion, Gorodnichenko, and Hong (2015) and Campos and Reggio (2017) provide evidence that unemployed households obtain lower prices thanks to greater search intensity while Kaplan and Menzio (2016) study the implications of the mechanism for business cycle analysis. No one of these papers analyze love of variety and household decision to experiment to expand her consideration set of products, which provides an alternative mechanism to reduce household level prices

persistently. Neiman and Vavra (2018) document some important changes in consumption expenditures along the extensive margin: a fall in the number varieties purchased by households and an increased segmentation in the type of products purchased by households with different income levels. Jaravel (2018) shows that this segmentation could be the result of endogenous technological progress which is directed towards households with greater income and thereby greater spending capacity. Here we emphasize that even within detailed quality categories of products and income groups of households, there is heterogeneity in the varieties purchased by households, which we interpret as reflecting (endogenous) differences in consideration sets. Baker, Johnson, and Kueng (2017) document how sales taxes changes the shopping behavior of households, but they do not look at whether temporary shocks have persistent effect on the consumption basket of households and thereby on their utility, which is the key novel insight of our analysis.

There is a large literature both on the measurement of the marginal propensity to consume of households to income shocks (Agarwal, Marwell, and McGranahan 2017, Blundell, Pistaferri, and Preston 2008 Broda and Parker 2014, Campbell and Hercowitz 2009, Johnson, Parker, Souleles, and McClelland 2006, and Parker et al. 2013) as well on its theoretical determinants, see Kaplan and Violante (2010, 2014), Attanasio and Pavoni (2011), Heathcote, Violante, and Storesletten (2014), and Kueng (2018). In our model the overall marginal propensity to consume is the sum of a propensity along the intensive margin and another one along the extensive margin, and we find that most of the response of the extensive margin comes from gross additions. The idea that consumption expenditures reflect an experimentation motive provides a different (complementary) interpretation for the excessive sensitivity of consumption expenditures to transitory income shocks documented by the literature, as a component of spending is motivated by a better attempt to achieve consumption smoothing. Because of this, in our model, expenditures in non durable goods and services do not reflect consumption flows.

Since the seminal work by Dixit and Stiglitz (1977) and Krugman (1979, 1980), many models have focused on love of variety as an important determinant of household welfare, see Broda and Weinstein (2006, 2010) for direct measurement of its relevance for welfare. But to love a variety, households should first get to know whether they like it which might require some costly experimentation that endogenously varies over the cycle. There is also a large literature on constructing accurate welfare measure of inflation see for example Aghion, Bergeaud, Boppart, Klenow, and Li (2017), Bils and Klenow (2001), Jones and Klenow (2016), Kaplan and Schulhofer-Wohl (2017), and Redding and Weinstein (2016) for

some important recent contributions. Our theory emphasizes that the consideration sets of households reflect past investment in experimentation and that increases in the cardinality of the set reduces inflation at the household level with consequences for welfare.

Since the pioneering contribution of Blanchard and Diamond (1990) and Davis and Haltiwanger (1990), several papers have shown the usefulness of looking at flows to characterize the evolution of stock variables. But the flow approach has typically neglected the distinction between intensive and extensive margin, which is the focus of this paper. The application to household consumption expenditures is also novel.

Section 2 show how the intensive and the extensive margin contribute to the fluctuations of consumption expenditures. Section 3 presents the model. Section 5 discusses calibration. Section 6 measures the cyclical properties of experimentation expenditures. Section 7 concludes. Appendix contains details on theoretical derivations, data and model computation.

2 The intensive and extensive margin of consumption expenditures

We discuss how we decompose the cyclical fluctuations of consumption expenditures into a component due to the intensive margin and another one due to the extensive margin. The extensive margin is further decomposed into a component due to *additions* of products to the consumption basket of households and another one due to *removals* of products from the basket. We then discuss the data used before turning to the empirical evidence.

2.1 Methodology

We build on the methodology initially developed by Davis and Haltiwanger (1992) to study the reallocation of workers across plants and by Bartelsman and Dhrymes (1998) and Campbell (1998) to analyze the evolution of aggregate productivity. We denote by e_{hvt} the expenditures by household $h \in \mathcal{H}$ for *variety* $v \in \mathcal{V}$ at time t , divided by the aggregate number of households present in the economy at time t . Here $\mathcal{H}$ denotes the set of all households present in the economy at *some* time t , while $\mathcal{V}$ denotes the set of all varieties available in the market at some time t . At time t the total expenditures of household h (divided by the aggregate number of households in the economy at time t) are obtained by

summing the expenditures of the household across all varieties to obtain

$$e_{ht} \equiv \sum_{\nu \in \mathcal{V}} e_{\nu ht}, \quad (1)$$

which allows to express aggregate expenditures per household as equal to

$$E_t = \sum_{h \in \mathcal{H}} e_{ht}.$$

The growth rate of expenditures per household can then be expressed as follows:

$$\Delta E_t \equiv \frac{E_t - E_{t-1}}{E_{t-1}} = \sum_{h \in \mathcal{H}} \frac{e_{ht} - e_{ht-1}}{e_{ht-1}} \times \frac{e_{ht-1}}{E_{t-1}}. \quad (2)$$

The change in the expenditures of household h , e_{ht} in (1), comes partly from changes in the expenditures of products that were also purchased in the previous period, i.e. the *intensive margin*, and partly from the expenditures in *net additions* of products to the consumption basket of the household, i.e. the *extensive margin*. Net additions can be further decomposed into the difference between the today expenditures of the household in (gross) *additions* of products to the consumption basket and the previous period expenditures in *removals* of products from the today basket of the household. This leads to the decomposition

$$\frac{e_{ht} - e_{ht-1}}{e_{ht-1}} = i_{ht} + a_{ht} - r_{ht} \quad (3)$$

where

$$i_{ht} = \sum_{\nu \in \mathcal{V}} \frac{e_{\nu ht} - e_{\nu ht-1}}{e_{ht-1}} \times \mathbb{I}(e_{\nu ht-1} > 0) \times \mathbb{I}(e_{\nu ht} > 0) \quad (4)$$

$$a_{ht} = \sum_{\nu \in \mathcal{V}} \frac{e_{\nu ht}}{e_{ht-1}} \times \mathbb{I}(e_{\nu ht-1} = 0) \times \mathbb{I}(e_{\nu ht} > 0) \quad (5)$$

$$r_{ht} = \sum_{\nu \in \mathcal{V}} \frac{e_{\nu ht-1}}{e_{ht-1}} \times \mathbb{I}(e_{\nu ht-1} > 0) \times \mathbb{I}(e_{\nu ht} = 0) \quad (6)$$

In the expressions above $\mathbb{I}(\cdot)$ denotes the indicator function. i_{ht} in (4) measures the percentage change in household h expenditures due to variations in the expenditures of products that the household has bought in the previous period and continues to buy also in the current one. a_{ht} in (5) measures the percentage increase in household h expenditures due to additions of products to the consumption basket of the household. Finally r_{ht} in (6) measures the percentage fall in household h expenditures due to the removals of products from the consumption basket. After combining (2) with (3), we obtain that the percentage growth in aggregate expenditures per household can be expressed as equal to

$$\Delta E_t = I_t + N_t \quad (7)$$

where I_t and N_t denote the changes in aggregate expenditures due to the intensive margin and to net additions, respectively. The contribution of the intensive margin is obtained by summing across households, weighted by their expenditures, the terms i_{ht} in (4) so that

$$I_t = \sum_{h \in \mathcal{H}} i_{ht} \times \frac{e_{ht-1}}{E_{t-1}} \quad (8)$$

while net additions are defined as equal to

$$N_t = A_t - R_t. \quad (9)$$

where

$$A_t = \sum_{h \in \mathcal{H}} a_{ht} \times \frac{e_{ht-1}}{E_{t-1}} \quad (10)$$

and

$$R_t = \sum_{h \in \mathcal{H}} r_{ht} \times \frac{e_{ht-1}}{E_{t-1}} \quad (11)$$

denote the changes in aggregate expenditures due to the additions of products to the consumption basket of households, and the removals of products from the basket of households, respectively. Formally, they are obtained by summing across households (again weighted by their expenditures) the terms a_{ht} and r_{ht} as given in (5) and (6), respectively.

2.2 Data

We apply the methodology using data from the Kilts-Nielsen Consumer Panel (KNCP). The KNCP tracks the expenditures on a wide range of non durable consumption products purchased in a variety of stores by a representative sample of US households.² The sample consists of around 60,000 households. Data have a rotating panel structure, with the median household staying in the sample three years. The data start in 2004, but we focus on the period 2007-2014 because the sample size as well the set of products covered by the KNCP expanded substantially in 2007 and we exclude 2015 because part of the end-of-the-year expenditures in 2015 are reported in the data release for 2016, which is not yet available to us. Each household in the KNCP uses a scanning device provided by Nielsen

²The product categories covered by the KNCP are dry grocery, frozen food, dairy, deli, packaged meat, fresh-food, non-food grocery, alcohol, general merchandise, and health and beauty aids, which represent around 13% of total consumption expenditures (including durables and non-durables) as calculated by the Consumer Expenditure Survey (CEX), see the Appendix for further details. The set of stores covered in the KNCP includes traditional grocery shops, drugstores, supermarkets, superstores and club stores.

to register the prices and quantities of all products she purchased in any possible shopping trip. Products are identified by their Universal Product Code (UPC).³ As in Einav, Leibtag, and Nevo (2010) and Broda and Parker (2014) we retain in the sample all households who report at least one shopping trip per month in each month of a year. Results are unchanged when selecting households who report nonzero monthly expenditures for at least 10 months in a year. We exclude from the analysis the expenditures in the category “general merchandise,”—which are durable goods reported only spottily by households in the KNCP,—and in products with no UPC (such as fresh products and bakery goods).⁴ In aggregating across households we use the projection weights provided by Nielsen. These weights could change over time. Since we consider households across two consecutive periods, we use as a household weight in period t the average between the Nielsen weight at $t - 1$ and at t . In the data, we identify a variety as a brand product-module pair. The brand of a product is a variable created by the University of Chicago. It refers to a collection of UPCs that are sold under the same characterizing name or logo, independently of the size and other cosmetic features of the package. Examples of brands in the soft drink category are “Diet Coke” and “Cherry Coke”. The product module is instead a categorization created by Nielsen to partition the set of UPCs into 735 homogeneous product groups. Examples of product modules are carbonated beverages, laundry supplies, diapers and frozen pizza. We treat all white labels in the same product-module as the same variety. In an average year, there are around 70,000 different varieties purchased by households in the sample. An average household buy around 350 of them in a year. We checked that results are robust to using UPCs to identify varieties. As a baseline period, we take the year—which naturally control for seasonal changes in the composition of the consumption basket—, but we also present results at the quarterly frequency.

2.3 Evidence

Panel (a) of Figure 1 plots the growth rate of consumption expenditures ΔE_t (blue solid line) together with the contribution of the intensive margin I_t (black dotted line) and net additions N_t (red dashed line). The three components comove positively, with a correlation above ninety per cent, see Table 1. Expenditure growth rates start falling in 2009 and fall

³In practice, the UPC price could be either entered manually by the household or imputed by Nielsen using the Kilts-Nielsen Retail Scanner Data (KNRS). The KNRS contains information on prices and sales of all UPCs sold in a representative sample of around 40,000 stores. Only if the household buys products in a store not covered by the KNRS, she is asked to report prices, which should incorporate any discount (through coupons, fidelity cards or other) enjoyed by the household.

⁴In the KNCP, only a subsample of households report expenditures on these products.

further down in 2010, to recover in 2011. Net additions accounts for a significant share of the fall in expenditure growth during the recession, and more generally for its fluctuations over time. Panel (b) further decomposes net additions N_t into gross additions A_t (red dashed line) and removals R_t (black dotted line). The fall in net additions in the recession is largely driven by the fall in the rate at which households add new products to their basket, while removals are relatively acyclical. Panels (c) and (d) of Figure 1 are analogous to panels (a) and (b) but consumption flows are calculated at the quarterly rather than at the yearly frequency. To remove seasonality, the quarterly series are reported as a 4-quarters moving average. Additions and removals are now larger in level, but the cyclical properties of E_t , I_t , N_t , A_t and R_t remain similar in that N_t explains a large share of the fluctuations of E_t and A_t is more volatile than R_t . Relatively to the data at the yearly frequency, the contribution of net additions to fluctuations in expenditures is now somewhat larger while additions and removals tend to comove slightly more strongly.

Table 1 provides some descriptive statistics. At the yearly frequency, the intensive margin I_t and net additions N_t have roughly the same volatility; gross additions A_t are more volatile than removals R_t and additions are more strongly correlated with expenditure growth rates than removals. The row labeled “ β -decomposition” in Table 1 reports the OLS estimated coefficient β_X from regressing the variable X_t in column, $X_t = I_t, N_t, A_t, R_t$, against expenditure growth ΔE_t :

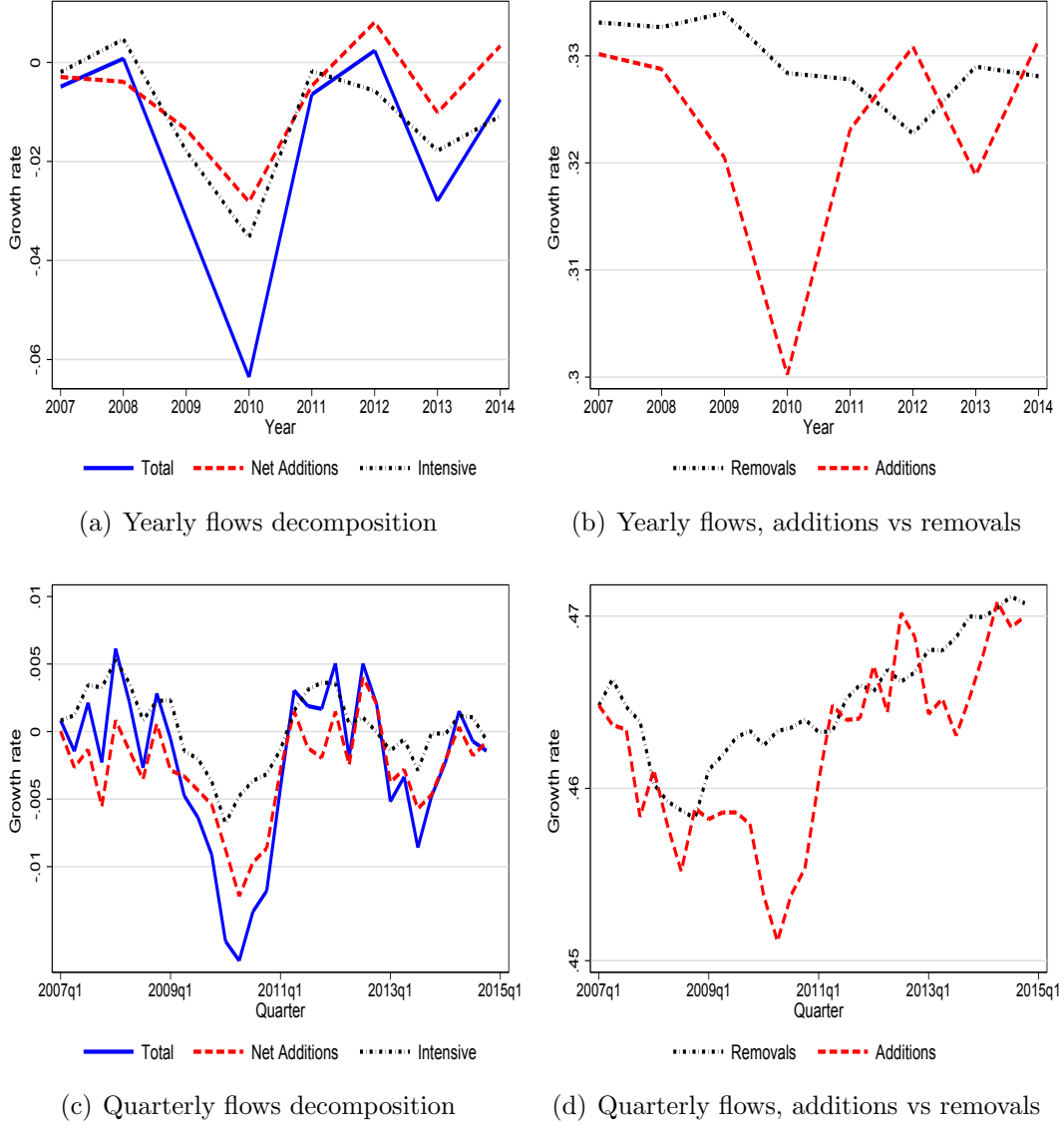
$$X_t = \text{cte} + \beta_X \Delta E_t + \text{error}.$$

From the properties of OLS, it follows that the coefficients for the intensive margin and net additions sum to one, $\beta_I + \beta_N = 1$ and the coefficient for net additions equals the difference between the coefficient for additions and the one for removals, $\beta_N = \beta_A - \beta_R$. Because of this, β_X measures the contribution of variable X_t to the cyclical fluctuations of ΔE_t . Using this metric at the yearly frequency, we conclude that net additions account for 54% of the fluctuations of expenditure growth, and that gross additions account for 85% of the contribution of net additions. At the quarterly frequency, the contribution of net additions to the volatility of expenditure growth goes up to 61%, with gross additions accounting for almost ninety per cent of the contribution of net additions.

2.4 The role of supply, sectors, and quality substitution

Additions and removals can be decomposed into *within* and *between* group components. Groups could be defined in terms of sectors and/or quality categories, or could classify

Figure 1: Aggregate expenditures and consumption expenditure flows



Note: Panel (a) and (c) plot the growth rate of aggregate expenditures (blue solid line) together with the contribution of the intensive margin I (black dotted line) and net additions N (red dashed line) at the yearly and quarterly frequency, respectively. Panel (b) and (d) plot gross additions A_t (red dashed line) and removals R_t (black dotted line) at the yearly and quarterly frequency, respectively. A variety is defined as a brand product-module pair. The quarterly series are calculated as a 4-quarters moving average to remove seasonality.

products depending on whether they are newly introduced into the market at time t , they are withdrawn from sale at t or they have been available both at $t - 1$ and t . These are possible partitions of the set of relevant varieties at time t , $\mathcal{V}_t$. Formally, let $\mathcal{V}_t^g$, $g = 1, 2, \dots, \mathcal{G}$

Table 1: Descriptive statistics of expenditure flows

	ΔE_t	I_t	N_t	A_t	R_t
a) Yearly frequency					
Standard deviation (%)	2.3	1.2	1.3	1.1	0.44
Correlation with ΔE_t	1.00	0.93	0.94	0.97	-0.41
β -Decomposition, β_X	1.00	0.46	0.54	0.46	-0.08
b) Quarterly frequency					
Standard deviation (%)	0.60	0.26	0.39	0.50	0.36
Correlation with ΔE_t	1.00	0.87	0.94	0.65	0.12
β -Decomposition, β_X	1.00	0.39	0.61	0.54	-0.07

Notes: A variety is identified by a brand product-module pair. The row labeled “ β -Decomposition” reports the OLS estimated coefficient β_X from regressing the variable in the corresponding column, $X_t = I_t, N_t, A_t, R_t$, against the percentage change in expenditures, $X = \beta_0 + \beta_X \Delta E + \epsilon$. The properties of OLS imply that $\beta_I + \beta_N = 1$ & $\beta_N = \beta_A - \beta_R$.

denote a partition of $\mathcal{V}_t$: $\forall t, \bigcup_{g=1}^{\mathcal{G}} \mathcal{V}_t^g = \mathcal{V}_t$ and $\mathcal{V}_t^g \cap \mathcal{V}_t^{g'} = \emptyset, \forall g \neq g'$. Also let

$$e_{ht}^g = \sum_{\nu \in \mathcal{V}_t^g} e_{\nu ht}$$

denote the time- t expenditures of household h in varieties that belong to group g . Given the partition $\mathcal{G}$, the *within-group* component of additions and removals is equal to the additions and removals within any group $g = 1, 2, \dots, \mathcal{G}$ where the household has spent a (strictly) positive amount both at $t - 1$ and at t . The time- t *between-group* component of additions A_t^b is equal to the sum of all additions in groups where households had zero expenditures at $t - 1$, while the between-group component of removals R_t^b is equal to the sum of all removals in groups where households have zero expenditures at t , so that

$$A_t^b = \sum_{g=1}^{\mathcal{G}} A_t^g \tag{12}$$

$$R_t^b = \sum_{g=1}^{\mathcal{G}} R_t^g \tag{13}$$

where

$$A_t^g = \frac{\sum_{h \in \mathcal{H}} e_{ht}^g \times \mathbb{I}(e_{ht-1}^g = 0) \times \mathbb{I}(e_{ht}^g > 0)}{E_{t-1}}$$

and

$$R_t^g = \frac{\sum_{h \in \mathcal{H}} e_{ht-1}^g \times \mathbb{I}(e_{ht-1}^g > 0) \times \mathbb{I}(e_{ht}^g = 0)}{E_{t-1}}.$$

measures the contribution of group g to the overall between-group component of additions and removals, respectively. Then, the within-group component of additions A_t^w and removals R_t^w are obtained as a residual as follows:

$$A_t^w \equiv A_t - A_t^b \quad (14)$$

$$R_t^w \equiv R_t - R_t^b. \quad (15)$$

We perform this between/within group decomposition for different partitions of the set of existing varieties at time t . Table 2 reports β -decompositions analogous to those reported in Table 1: each entry is the estimated OLS coefficient of a regression where the dependent variable is the within or the between component of the variable in column and the independent variable is expenditure growth ΔE_t . The first row of the table also reports the total contribution of variable X_t (the sum of the between and within component), which are equal to the values in Table 1. For expositional simplicity we focus the discussion on the yearly decomposition, but also report results at the quarterly frequency.

Demand versus supply factors We start partitioning varieties at time t depending on whether they are newly introduced at t , they are withdrawn at t , or they have been available both at $t-1$ and at t . The question is whether the cyclical properties of additions and removals are driven by the net entry of new varieties in the market, which is procyclical (Broda and Weinstein 2010), or whether they arise also *within* the set of varieties continuously available in the market. We identify a variety as new at time t if some households in the KNCP report some positive expenditures on that variety at t , and no households in the KNCP and no firms in the KNRS data—which report sales per variety for a representative sample of US stores—have ever reported positive expenditures on that variety before time t . A withdrawn variety is identified analogously.⁵ Panel A) of Table 1 shows that, at the yearly frequency, β_N^w divided by β_N is equal to 95%, which indicates that net additions within the group of continuing varieties account for the great majority of the contribution

⁵We also experimented with an alternative definition where market availability is defined at the Scantrak market level, which is a partition into 54 geographical areas of the US, with respect to which the KNCP is fully representative.

Table 2: Within/Between component of expenditure flows, β -decompositions

Frequency: Variable X_t :	Yearly			Quarterly		
	N_t	A_t	R_t	N_t	A_t	R_t
Total contribution of X , $\beta_X^w + \beta_X^b$	0.54	0.46	-0.08	0.61	0.54	-0.07
a) Available varieties						
Within, β_X^w	0.51	0.38	-0.13	0.54	0.47	-0.07
Between, β_X^b	0.03	0.08	0.05	0.07	0.07	0.00
b) Changes in Sectoral composition						
Within, β_X^w	0.29	0.32	0.03	0.21	0.29	0.08
Between, β_X^b	0.25	0.14	-0.11	0.40	0.25	-0.15
c) Quality substitution						
Within, β_X^w	0.23	0.27	0.04	0.16	0.24	0.08
Between quality within product group, β_X^b	0.23	0.17	-0.06	0.25	0.25	0.00
Between product group, β_X^b	0.08	0.02	-0.06	0.20	0.15	-0.05

Notes: Each entry is the estimated OLS coefficient β_X^s from regressing the between component $s = b$ or the within component $s = w$ of the variable $X_t = N_t, A_t, R_t$ in column against expenditure growth: $X_t^s = \beta_X^s \Delta E_t + \epsilon$. The first row reports the total contribution of variable X_t (the sum of the between and within component) to the fluctuations of ΔE_t , which as in Table 1. A variety is identified by a brand-product module pair. The sectors in panel (b) correspond to the 735 product modules defined by Nielsen. In panel (c), there are 950 groups defined in terms of 95 Nielsen defined product-groups (e.g. cereals, snacks,...) each divided in 10 quality bins corresponding to the deciles of the distribution of (average) product prices within the product-group.

of net additions to the cyclical variation of ΔE_t . Gross additions on continuing and newly introduced varieties are both procyclical, but the former comoves much more strongly with expenditure growth, $\beta_A^w = 0.38$ while $\beta_A^b = 0.08$. The cyclical properties of removals on continuing and withdrawn varieties are instead slightly different: removals on continuing varieties are counter-cyclical with $\beta_R^w = -0.13$, removals due to withdrawn varieties are just mildly pro-cyclical with $\beta_R^b = 0.05$. Overall, this evidence is consistent with the idea that firms product innovation is pro-cyclical, but changes in the net supply of varieties cannot fully account for the previous findings, suggesting that household demand factors play an important role.

Sectoral changes in the consumption basket Income fluctuations could modify the sectoral preferences of households and thereby change the consumption basket of households along the extensive margin. We now investigate the importance of this mechanism and find that it plays a limited role in explaining the previous results. We partition varieties according to 735 detailed sectoral categories, corresponding to the product modules defined by Nielsen (e.g. cold cereals, hot cereals,...). Panel B) of Table 2 documents that, even at this very detailed sectoral breakdown, a significant share of the cyclical contribution of net additions to expenditure growth is due to the within-sector component. When focusing on gross additions, we see that, both at the yearly and quarterly frequency, the within-sector component plays a dominant role. For example, at the yearly frequency, it accounts for around 70 percent of the overall cyclical contribution of gross additions to expenditure growth. The between-sector component of removals is counter-cyclical and contributes more than the within-sector component to the cyclical fluctuations of expenditure growth—at the yearly frequency $\beta_R^b = -0.11$ the $\beta_R^w = 0.03$ —, but the overall contribution of removals remains limited.

Quality substitution Some authors have emphasized that, over the cycle, households substitute across products of different quality (Jaimovich, Rebelo, and Wong 2017, Argente and Lee 2017), but quality substitution can occur along the intensive or the extensive margin. To evaluate how much quality substitution matters for the cyclical patterns of net and gross additions, we first construct a measure of quality of variety $\nu \in \mathcal{V}$ by running the following regression:

$$\ln(p_{\nu t}) = d_{\nu} + \tau_t + e_{\nu t}, \quad (16)$$

where $p_{\nu t}$ is the price of the variety ν at time t , while d_{ν} and τ_t are dummies for variety and time, respectively. Nielsen defines 95 product-groups (e.g. cereals, snacks,...), and within each of them the price of varieties is expressed in the same unit. We take d_{ν} as a measure of the relative quality of the variety ν within the product group. We rank d_{ν} within each product group, we calculate the deciles of the associated quality distribution, and then assign each variety to the corresponding quality bin. As a result we obtain 950 groups of varieties defined in terms of 95 product modules each divided in 10 quality bins. By applying the same logic as in Section 2.4, we then decompose additions and removals depending on whether they occur within the same quality bin of a given product group, between quality bins of the same product group or between product groups. The β -decomposition associated with this partition of the set of varieties is reported in panel

(c) of Table 2. Quality substitution plays a role in explaining the cyclical properties of net and gross additions, but it does not fully account for them: at the yearly frequency, the contribution to expenditure growth of the within-quality component of net additions is as important as the contribution of the between-quality component. For gross additions, the contribution to expenditure growth of the within-quality component is fifty percent larger than the contribution of the between-quality component. Changes in the sectoral composition of the basket also matter for the contribution of net and gross additions to expenditure growth, but their quantitative importance is generally small relative to the contribution of the within-quality component. The contribution of removals to expenditure growth remains limited both within and between quality or sectoral groups.

2.5 Robustness and extensions

We consider the robustness of our results when identifying varieties using UPCs rather than brand product-module pairs and when measuring expenditures at constant rather than at current prices. We also investigate how results change when focusing on products with different durability, and when considering more restrictive definitions of additions and removals, intended to better identify persistent changes in the consumption basket of the household. Table 5 in the Appendix report descriptive statistics of the underlying series. Table 3 reports β -decompositions analogous to those reported in Table 1 and 2.

Varieties as UPCs Our baseline definition of a variety is that of brand-product module pair. We our descriptive analysis with the methodology described in Section 2.1 counting each Universal Product Code (UPC) as a different variety. In this alternative definition the number of varieties grows because multiple UPCs -for instances referring to different package sizes- were grouped into the same brand in our baseline exercise. For instance, the 33 cl. can and the 12-cans pack of Diet Coke share the same brand code while they have different UPCs. The second raw of Table 3 reports the β -decomposition of product additions and removals calculated under this new definition of varieties. Not surprisingly, net addition is much more volatile and cyclical than the intensive margin as there is more turnover of varieties at the UPC level. There is a higher comovement of net addition with aggregate expenditure which is mostly driven by the higher cyclicity of removal. Hence, comparing these results with the ones of Table 1 we conclude that there is substantial cyclical removal within brand, consistent with the view that in a recession households substitute towards bulk size UPCs within the same brand to benefit from quantity discounts.

Table 3: Further robustness, β -decompositions

Variable X_t :	N_t	A_t	R_t
Varieties as UPCs	0.74	0.51	-0.23
Constant prices	0.56	0.47	-0.09
Durability			
-Less than 12 months	0.45	0.36	-0.09
-More than 12 months	0.52	0.34	-0.18
Robust additions & removals	0.40	0.30	-0.10
Persistent vs temporary additions & removals			
-Persistent	0.18	0.13	-0.05
-Temporary	0.30	0.29	-0.01

Notes: All data are at the yearly frequency. Each entry is the estimated OLS coefficient β_X from regressing the variable $X_t = N_t, A_t, R_t$ in column against expenditure growth: $X_t = \beta_X \Delta E_t + \epsilon$. A variety is identified by a brand product-module pair, with the exception of the first two rows where a variety is a UPC. The row “Constant prices” calculates additions and removals at constant over time prices. The row “Durability” computes additions and removals separately for varieties that are storable for less or more than 12 months, using the index calculated in Alessandria, Kaboski, and Midrigan (2010). In the row “Robust additions & removals”, an addition is defined as robust if the variety added at t was purchased neither at $t - 1$ nor at $t - 2$, while the removal is defined as robust if the variety removed at t was purchased both at $t - 1$ and at $t - 2$. In the row “Persistent vs temporary additions & removals”, an addition is defined as persistent if the variety added at t is also purchased at $t + 1$, while it is temporary if the variety it is not purchased at $t + 1$. Analogously, a removal is defined as persistent, if the variety removed at t , it is not purchased at $t + 1$, while it is temporary if the variety is purchased again at $t + 1$.

Constant prices Our object of interest in the baseline specification is the dollar expenditure of household h in variety v , $e_{\nu ht}$. As product additions are evaluated at period t prices whereas product removals are evaluated at period $t - 1$ prices procyclical price inflation could in principle cause procyclical net product addition. To test against this hypothesis, we consider an alternative specification where we measure household expenditure at constant prices over time. Let $\bar{e}_{\nu ht} = u_{\nu ht} \bar{p}_v$ denote expenditure by household h on variety v in period t evaluated at the constant price $\bar{p}_v$, where $u_{\nu ht}$ are the units purchased by the household (quantities are expressed in appropriate unit of measure for each variety: ml, ounces..). For this exercise, a variety is identified by a Universal Product Code as homogenized by KNCP (i.e. UPC-ver). We compute a baseline price for each variety, p_v , as the

average of the posted prices over time and across retailers (additional details are provided in the appendix),

$$p_v = \sum_{i \in I^v} \sum_m \sum_{j \in m} \sum_{t \in t_j} \frac{1}{t_j} \omega_i \omega_m \omega_{ji} \hat{p}_{ijmt},$$

where $\hat{p}_{ijmt}$ is the unit price of UPC i part of the brand v posted by store j in market m in quarter t . We first take an average over the observed unit prices in quarter t for store j . Then, we take a weighted average over the stores active in a given market, so to obtain an average unit price for each market. Next, we average over markets to obtain a unit-UPC price at the market level. Finally, we take an average across all UPCs within a same brand to obtain a single price for each variety.

We then use the constant price measure of expenditure to compute the decomposition of aggregate expenditure growth per household in product additions and removals with the methodology described in Section 2.1. In the first row of Table 3 we report the β -decomposition of product additions and removals. The decomposition computed using the constant prices so obtained is remarkably similar to our baseline at current prices: net additions are strongly procyclical due to the strong comovement of gross addition with aggregate expenditure, whereas product removals fluctuate much less with the cycle. The same is true if we experiment with an alternative way of computing constant prices (not reported), which allows expenditure to change if households allocate their consumption to UPCs that have different unit price within the same brand. Hence, $\bar{e}_{\nu ht} = \sum_{i \in v} u_{iht} \bar{p}_i$ where $\bar{p}_i$ is the constant unit price of UPC i part of brand v , given by $\bar{p}_i = 1/T \sum_{t=1}^T p_{it}$ with

$$p_{it} = \sum_m \omega_m \sum_{j \in m} \omega_{ji} \sum_{\tau \in t_j} \frac{1}{t_j} \hat{p}_{ij\tau}$$

The advantage of this definition is that it takes into account that even UPCs within the same brand may be fairly heterogeneous and, therefore, widely differ in unit price.

Durability Goods differ in their degree of durability, so that some goods can be stockpiled and accumulated by households so that they could be purchased at low frequencies. We want to investigate if the consumption expenditure flows vary across goods with different durability. To do so, we report in 3 the β -decomposition of product additions and removals calculated for products that have a durability of (weakly) less than 12 months and (strictly) more than 12 months respectively.⁶ In both cases the addition rate is substantially pro-

⁶We also thank Guido Menzio and Leena Rudanko for sharing the durability indexes constructed in Alessandria, Kaboski, and Midrigan (2010).

cyclical, and comparable in magnitude across the two groups. The removal rate is also pro-cyclical, but more so for higher durability products.

Robust additions and removals In our baseline measurement of product additions to and removals from the consumption basket, we measured an addition at year t as a variety purchased at t but not at $t - 1$, and removals as varieties purchased in year $t - 1$ but not at t . In this section we consider a more conservative definition of additions and removals, which we label *robust*, by defining an addition in year t as a variety purchased at t but not both at $t - 1$ and $t - 2$; similarly, robust removals are varieties purchased in both $t - 2$ and $t - 1$ but not in year t . We notice that in order to implement this exercise we need to restrict the set of households over which we compute consumption expenditure flows, in particular we consider in each year t the subset of households who have been in the sample for at least other two years in the past. The third row of Table 3 reports the β -decomposition of product additions and removals calculated under the robust definition. Not surprisingly the contribution of net additions to expenditure growth falls relatively to our baseline case as the measurement of addition is more restrictive, but robust additions remain strongly procyclical.

Persistent versus temporary additions and removals In the last two rows of Table 3 we report results based on the set of households who report positive expenditure not only in year t but also in years $t - 1$ and $t + 1$. We define as *persistent* addition, the expenditure in year t for a household-product match which is characterized by zero expenditure at $t - 1$ and positive in year $t + 1$. If instead the expenditure on a match which was new at time t becomes zero at $t + 1$, we call this addition *temporary*. Similarly, persistent removal at t implies positive expenditure in the product by the household at $t - 1$, and zero expenditure in both years t and $t + 1$. If the expenditure in year $t + 1$ is instead positive, we label this temporary removal. About 2/3 of the variation in addition is temporary, whereas about 2/3 of the variation in removal is persistent. Similarly, about 2/3 of the comovement of addition with aggregate expenditure is due to temporary addition.

3 McFadden meets Ramsey

We introduce a model where households add and remove varieties from the consumption basket from one period to another, both because of random preference shocks that change the preferred variety purchased by a household for a given consideration set and because

of experimentation for new varieties that may enlarge the consideration set. In particular we consider a household that in a given period chooses to consume the (single) variety that maximizes her utility within a finite set of alternatives. On the supply side, there is a mass $s > 0$ of different sectors, and in each of them there is a measure one of varieties. Varieties are perfectly substitutable within sectors, and differentiated across sectors with a constant elasticity of substitution $\sigma > 1$. Think of s as a large number, which for convenience is assumed to be set on the positive real line, as typically done in the literature that builds on the Dixit and Stiglitz model. As a result there are $[0, s] \times [0, 1]$ varieties supplied in the economy. We assume here that the household is aware of all s sectors. We will relax this assumption later. In the spirit of the literature on discrete choice models (McFadden (1974a)), we assume that a household can only choose among a finite set of alternatives within each sector $j \in [0, s]$. Even if the supply of varieties is infinitely large in each sector, the household is only aware of a finite number of varieties $n_{jt} \in \mathcal{N}$ that she likes in sector j (Masatlioglu, Nakajima, and Ozbay (2012), Manzini and Mariotti (2014)), with $\Omega_{jt} \subseteq [0, 1]$ denoting her *consideration set* in sector j .

While varieties are perfectly substitutable within sectors, households choose to consume only one of them because of random idiosyncratic preference shocks, that scale the utility from consumption of a particular variety. In particular, the effective consumption utility obtained from consuming $q_{\nu j}$ units of variety ν in sector j in period t is given by $z_{\nu jt} q_{\nu j}$. We assume that in a fraction ζ of sectors idiosyncratic preference shocks $z_{\nu jt}$ are drawn at the beginning of each period t for each variety $i \in \Omega_{jt}$ from a Frechét distribution with shape parameter $\alpha > 0$. Preference shocks are identically and independently distributed across time, varieties and sectors in these sectors. In the remaining $1 - \zeta$ fraction of sectors, preferences are unchanged from the previous period with probability $1 - \psi$,

$$z_{\nu jt} = \begin{cases} z_{\nu jt-1} & \text{with probability } 1 - \psi \\ \varepsilon_{\nu jt} & \text{otherwise} \end{cases},$$

where $\varepsilon_{\nu jt}$ is also drawn from a Frechét distribution with shape parameter $\alpha > 0$, and is identically and independently distributed across time and varieties.

Hence, conditional on the realization of $z_{\nu jt}$ for each i and j , and given a total consumption expenditure v_t , the household solves the following problem in period t :

$$c_t = \left[\int_0^s \max_{\{q_{\nu j} \geq 0 \ \forall i \in \Omega_{jt}\}} \left(\int_{i \in \Omega_{jt}} z_{\nu jt} q_{\nu j} di \right)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}} \quad (17)$$

subject to

$$\int_0^s \int_{i \in \Omega_{jt}} q_{\nu j} di dj = v_t. \quad (18)$$

Proposition 1 *Let $f_t(n)$ denote the fraction of sectors j where the household's consideration set is composed of $n_j = n$ varieties, for all $n \geq 0$. Let the preference shock $z_{\nu jt}$ be i.i.d. distributed according to a Frechét distribution with shape parameter $\alpha > \sigma - 1$. Let v_t denote the total expenditure expressed in units of varieties. The total expenditure in a sectors $j \in [0, s]$ with $n_j = n$ varieties in the consideration set is given by*

$$v_{nt} \equiv \frac{n^{\frac{\sigma-1}{\alpha}}}{\sum_{m=0}^{\infty} m^{\frac{\sigma-1}{\alpha}} f_t(m)} v_t. \quad (19)$$

The optimal consumption basket $\{c_{\nu jt}\}$, $\forall i \in \Omega_{jt}$ and $\forall j \in [0, 1]$ is such that

$$c_t = \frac{v_t}{p_t} \quad \text{with} \quad p_t = \left[s \Gamma \left(1 - \frac{\sigma-1}{\alpha} \right) \sum_{n=0}^{\infty} n^{\frac{\sigma-1}{\alpha}} f_t(n) \right]^{-\frac{1}{\sigma-1}}. \quad (20)$$

Proof. See the appendix. The optimal price index p_t is characterized by a constant elasticity of $-1/(\sigma - 1)$ with respect both to the number of sectors in the economy, s , and to the weighted average of the size of the consideration set, n , scaled by the ratio $(\sigma - 1)/\alpha$. The number of sectors matter for the price index due to the *love for variety* built into the preferences of the household with respect to differentiated products. The size of the consideration set in a sector matters for the price index even if varieties are ex-ante identical, because of a larger consideration set increases the possibility of the household to find a better match to his preference shocks. Hence, a larger consideration set reduces the risk of a bad variety draw for the household. We notice that the effect of the size of the consideration set on household utility decreases in α and increase in σ . A higher α is associated to a smaller chance of getting a bad preference draw from the Frechét distribution, whereas a higher σ allows for more substitution across sectors in case of a bad preference draw.

3.1 The experimentation technology

So far we have treated the consideration set of a household in each sector as exogenous, here we introduce the technology that allows the household to change the number of varieties in her consideration set, and obtain the law of motion for the cross-sectoral distribution $f_t(n)$. While a household is aware of only n_j varieties she likes in sector j , she is aware of the existence of other varieties in that sector but does not know her preferences with respect

to these varieties before purchasing them for the first time. Varieties are of two types: those that always give zero utility, i.e. with $z_{\nu jt} = 0$ for all t , and those that give random strictly positive utility with $z_{\nu jt}$ distributed according to Frechét with shape parameter α . As a household doesn't know her preferences toward a variety before consuming it for the first time, purchasing any of them is relatively more risky with respect to those already in the consideration set (which by definition contains only varieties that give strictly positive utility). In particular, to learn whether the household likes the variety, the household has to purchase (and consume) at least one unit of the variety. For simplicity, we assume that varieties consumed in the experimentation phase provide no utility, implying that experimentation is costly both because it is risky and because it provides no immediate utility. Hence the household will not purchase more than one unit of the variety for the first time. If the household spends x units on a mass x of different varieties the probability that the household discovers one she likes is λx^η , the probability that the household likes none of them is $1 - \lambda x^\eta$, while the probability that she likes more than one is of order smaller than x . The parameter $\eta \in (0, 1]$ captures the possible decreasing returns in the intensity of experimentation. It follows that, upon spending x on experimentation, the probability of discovering $n \geq 0$ new varieties she likes is distributed according to a Poisson with mean λx^η . For simplicity we assume that experimentation is fully random in that the household when buying a new variety does not know the sector of the variety. The sector associated to a variety gets known only after having experimented the variety. The idea is that how much a variety results substitutable to other varieties depends on the preferences of the individual household. Finally, to ensure stationarity, we introduce exogenous attrition in the consideration set so that the household no longer likes a variety, which then drops from her consideration set, with probability δ . Product removals from the consideration set are i.i.d. over time and across products.

The next proposition shows that if the initial (at time zero) number of varieties the household likes in a sector is distributed as a Poisson distribution, then the stationary distribution is also a Poisson.

Proposition 2 *If the distribution $f_{t-1}(n)$ characterizing the fraction of sectors with a consideration set of size n at time $t - 1$ is a Poisson distribution with mean μ_{t-1} , then the same distribution at time t , $f_t(n)$, is a Poisson distribution with mean μ_t given by*

$$\mu_t = (1 - \delta) \mu_{t-1} + \frac{\lambda x_t^\eta}{s}. \quad (21)$$

Proof. See the appendix. The cross-sectoral distribution of the size of the consideration

set is Poisson with mean μ_t that falls over time because of the exogenous attrition in the set of varieties that the household likes, δ , and increases as a function of the endogenous experimentation by the household, x_t , which adds varieties to the consideration set of a randomly chosen sector at a rate $\lambda/s x_t^\eta$. Hence the dynamics of μ_t captures the dynamics in the average size of the consideration set of a household across sectors.

4 McFadden and Ramsey meet Prescott

In this section we embed our model of experimentation and consumption into a canonical real business cycle model and study the interaction of consumption, saving and experimentation decisions by a household. Each variety ν in any sector j is produced with a Cobb-Douglas production function,

$$y_{\nu jt} = z_t k_{\nu jt}^\omega \ell_{\nu jt}^{1-\omega}$$

where z_t is an aggregate productivity shock evolving according to

$$\log(z_t) = \rho_z \log(z_{t-1}) + \varepsilon_t, \quad (22)$$

$\ell_{\nu jt}$ and $k_{\nu jt}$ are labor and capita input. There is perfect competition in the product market hence all firms set the same price $q_{\nu jt} = q$ and make zero profits and demand labor and capital so that

$$\omega q z_t \left(\frac{k_{\nu jt}}{\ell_{\nu jt}} \right)^{\omega-1} = \rho_{kt} \quad (23)$$

$$(1 - \omega) q z_t \left(\frac{k_{\nu jt}}{\ell_{\nu jt}} \right)^\omega = w_t. \quad (24)$$

All firms choose the same mix of capital and labor implying that the aggregate production function is given by

$$y_t = z_t k_t^\omega \ell_t^{1-\omega},$$

where, abusing notation, $y_t = \int_0^s \int_{0,1} y_{\nu jt} d\nu dj$, is the total supply of varieties, and $k_t = \int_0^s \int_{0,1} k_{\nu jt} d\nu dj$, and $\ell_t = \int_0^s \int_{0,1} \ell_{\nu jt} d\nu dj$, is the aggregate demand of capital and labor. The numeraire of the economy is the supply of varieties whose price q is normalized to 1.

There is a representative household who lives for an infinite number of periods and maximizes the expected discounted utility flows from consumption, discounting one period ahead at rate $\beta < 1$. The state of the household problem at the beginning of period t is given by the amount of capital from the previous period k_{t-1} , a fraction δ_k of which

has depreciated, paying $\rho_{kt} k_{t-1}$ in period t , and the average number of varieties per sector given by $\mu_{t-1} (1 - \delta)$. The household supplies labor ℓ_t and receives nominal labor income $w_t \ell_t$. She allocates her resources among experimentation x_t , consumption c_t , and savings in physical capital k_t . New varieties are added to the consideration set so that the mean of the Poisson distribution, μ_t , is given by equation (21). Changes in holding of capital require paying quadratic adjustment costs, $0.5 \bar{\chi} (k_t/k_{t-1} - 1)^2 k_{t-1}$ with $\bar{\chi} > 0$. Adjustment costs in capital allow deviations of consumption patterns from the permanent income hypothesis. Transitory shocks that have little effects on permanent income can have disproportionately large effect on current consumption to avoid incurring in large asset transaction cost, in the spirit of Kaplan and Violante (2014). Consumption expenditure $v_t \equiv p_t c_t$ is allocated across the varieties in the consideration set of the household according to the model described above which provides one consumption unit for each $1/p_t$ units of expenditure. Finally, the government finances expenditures g , equally allocated across all varieties, which provide no utility to the household and are financed by a lump-sum tax $\tau = g$.

The household problem The household problem in recursive form is given by

$$V(k_{t-1}, \mu_{t-1}, z_{t-1}) = \max_{\{c_t \geq 0, \ell_t \geq 0, x_t \geq 0, k_t \geq b\}} \frac{c_t^{1-\gamma}}{1-\gamma} - \frac{\ell_t^{1+\phi}}{1+\phi} + \beta \mathbb{E}_t [V_{t+1}(k_t, \mu_t, z_t)]$$

s.t.

$$p_t c_t + x_t + i_t = \rho_{kt} k_{t-1} + w_t \ell_t - \tau, \quad (25)$$

$$i_t = k_t - k_{t-1} (1 - \delta_k) - \frac{\bar{\chi}}{2} \left(\frac{k_t - k_{t-1}}{k_{t-1}} \right)^2 k_{t-1}, \quad (26)$$

and subject to the equation determining the price p_t as a function of μ_t in (20) and the law of motion for μ_t in equation (21), and z_t in equation (22). Parameters are such that $\gamma > 0$ and $\phi \geq 0$. The optimal saving and experimentation decisions by a household are given by:

$$1 + \bar{\chi} \Delta k_t = \beta \mathbb{E}_t \left\{ \frac{p_t}{p_{t+1}} \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \left[1 - \delta_k + \rho_{kt+1} + \bar{\chi} \Delta k_{t+1} \left(1 + \frac{\Delta k_{t+1}}{2} \right) \right] \right\}, \quad (27)$$

$$1 - \varepsilon_{p\mu t} \eta \frac{\lambda x_t^\eta}{s \mu_t} \frac{p_t c_t}{x_t} = \beta \mathbb{E}_t \left[\frac{p_t}{p_{t+1}} \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \right] (1 - \delta), \quad (28)$$

holding with equality if, respectively, $k_t > \underline{b}$ and $x_t > 0$, where $\Delta k_t \equiv (k_t - k_{t-1})/k_{t-1}$ is the growth rate in bond holdings, and

$$\varepsilon_{p\mu t} \equiv \left| \frac{\partial p_t}{\partial \mu_t} \frac{\mu_t}{p_t} \right| = \frac{1}{\sigma - 1} \left(\frac{\sum_{n=0}^{\infty} n^{\frac{\sigma-1}{\alpha}+1} f(n; \mu_t)}{\sum_{n=0}^{\infty} n^{\frac{\sigma-1}{\alpha}} f(n; \mu_t)} - \mu_t \right) \in (0, 1), \quad (29)$$

is the elasticity of the basket price to a change in the stock of varieties, captured by μ_t . Abusing notation, the function $f(n; \mu_t)$ denotes a Poisson distribution with mean μ_t . The right-hand-side of equation (27) measures the cost of investing one more unit in assets at t , whereas left-hand-side measures the benefit at $t + 1$ given by the nominal return ρ_{t+1} plus the savings on the adjustment costs, properly discounted by the inflation rate and ratio of marginal utilities.

Similarly, the left-hand-side of equation (28) captures the cost at time t of one extra unit spent on experimentation. We notice that such cost is reduced by the positive effect of more varieties on the current price level, captured by the second term of the left-hand-side of equation (28), and hence on consumption. We notice that, keeping $\varepsilon_{p\mu t}$ constant, a permanent increase in expenditure e_t that leaves the right hand side unchanged, i.e. $p_t = p_{t+1}$ and $c_t = c_{t+1}$, causes experimenting x_t to increase one for one with consumption expenditure $v_t = p_t c_t$ because of a *scaling* effect. The intuition is that the incentives to experiment scale up with the level of consumption expenditures as the benefits of a better fit of preferences are amplified by larger expenditures. When consumption expenditure changes transitorily there is also a *smoothing* effect on experimenting: when expenditure is higher today relatively to tomorrow the household has incentives to allocate a larger share to experimenting in order to smooth consumption utility. Hence experimenting x_t increases more than one for one with consumption expenditure v_t . All in all, this suggests that experimenting is more elastic to shocks that cause transitory changes in expenditure as opposed to permanent changes.

Finally, the optimal supply of labor by the household satisfies

$$\ell_t^\phi = c_t^{-\gamma} \frac{w_t}{p_t}. \quad (30)$$

Given the law of motion for μ_t , the budget constraint in equation (25) and the definition of p_t in (20), these first order conditions determine the optimal values of $c_t, x_t, a_t, \ell_t, \mu_t$.⁷

⁷Let $u \equiv (\sigma - 1)/\alpha$. We notice that $E(n^{1+u})/E(n^u) - E(n) = 1$ if $u = 1$ given that n is Poisson distributed. Moreover, $E(n^{1+u})/E(n^u)$ is increasing in u implying that $\varepsilon_{p\mu t} < 1$ given $u < 1$.

Steady state Consider a household in steady state, with expenditure $\bar{e} = \bar{y} - \bar{g}$. The steady state return on capital is given by

$$\bar{\rho}_k = \frac{1}{\beta} - 1 + \delta_k.$$

The steady state expenditure in experimentation, $\bar{x}$, and consumption, $\bar{v}$, are given by

$$\bar{x} = \bar{e} \frac{\bar{\varepsilon}_{p\mu}}{\bar{\varepsilon}_{p\mu} + \beta \left(1 + \frac{1/\beta - 1}{\delta}\right)}, \quad \bar{v} = \bar{e} \frac{\beta + \frac{1-\beta}{\delta}}{\bar{\varepsilon}_{p\mu} + 1 + \frac{1-\beta}{\delta}} \quad (31)$$

The steady state value of μ_t is given by

$$\bar{\mu} = \frac{\lambda \bar{x}^\eta}{\delta s}, \quad (32)$$

which determines the value of $\bar{\varepsilon}_{p\mu}$ in equation (29); $\bar{\mu}$ is larger the larger the discovery rate of new varieties relatively to the attrition rate of old ones, and the larger the expenditure in experimentation per sector. Hence, the steady state value of $\bar{x}$ is a solution to a fixed point problem. We notice that there exists a unique solution to such problem as $\lim_{\bar{\mu} \rightarrow 0} \bar{\varepsilon}_{p\mu} = +\infty$ and $\bar{\varepsilon}_{p\mu}$ is strictly decreasing in $\bar{\mu}$, and hence in $\bar{x}$. The next proposition shows that the equilibrium level of experimentation in steady state is strictly decreasing in the level of the interest rate, and that its elasticity with respect to disposable income is smaller than one.

Proposition 3 *In steady state experimentation $\bar{x}$ is decreasing in the equilibrium interest rate $\bar{\rho}$, i.e. $\partial \bar{x} / \partial \bar{\rho} < 0$, and its elasticity with respect to disposable income, $\bar{e} = \bar{y} - \bar{g}$, is smaller than 1, i.e. $\partial \log(\bar{x}) / \partial \log(\bar{e}) \in (0, 1)$.*

Proof. See the appendix. The reason for the negative relationship between experimentation and the interest rate follows from the fact that experimentation is a substitute to saving in bonds. Experimenting increases less than one for one with disposable income because the elasticity of the price index to the stock of varieties, $\bar{\varepsilon}_{p\mu}$, is decreasing in the level of μ and, given the correlation between income and μ , in the level of income.

5 Calibration

The aim of this section is to use available data on the addition of products to the consumption basket to calibrate the parameters of the model. We first discuss the type of statistics we would need to identify the key parameters of the model, then we introduce household level data that allow us to obtain such statistics.

5.1 Observable statistics: flows in consumption expenditures

Additions to the consumption basket of the household at time t are defined as equal to the sum of the expenditures in all varieties purchased at time t which were not purchased at time $t - 1$, and they are equal to

$$A_t = \frac{x_t}{e_{t-1}} + \frac{\lambda x_t^\eta}{s \mu_t} \frac{v_t}{e_{t-1}} + \xi_t^a \quad (33)$$

The first term in the right hand side of (33) measures the expenditures at time t in experimentation for new varieties x_t , which amount to additions to the household's consumption basket independently of whether the household will actually incorporate them into her consideration set. The second term measures the contribution of the expenditures at time t in varieties which are *true additions* to the consideration set of the household at time t : expenditures for the consumption of varieties which have been added to the consideration set at time t . Notice that under our assumptions, the expenditures for experimentation in varieties that the household turns out to incorporate into her consideration set have zero measure. Finally there are additions to the consumption basket of the household which are not true additions to her consideration set. These *false* additions are due to temporary shocks in preferences that make the household chooses to purchase a variety at time t that was already in her consideration set at time $t - 1$ —but that at the time was not purchased because it was preferred by other varieties in the set. The sum of the expenditures at time t in all varieties that were not consumed at time $t - 1$ even if they were already in the consideration set at time $t - 1$ is equal to

$$\xi_t^a = \frac{(1 - \delta)^2 \mu_{t-2}}{\mu_t} \frac{v_t}{e_{t-1}} - \frac{1}{e_{t-1}} \sum_{n=1}^{\infty} b_{nt}^1 f(n; \mu_t), \quad (34)$$

which is equal to the difference between the expenditures at time t in all varieties already part of the consideration set of the household at time $t - 2$ and the total expenditures at time t in varieties already present in the consideration set at $t - 2$ that are purchased both at $t - 1$ and at t . Notice that the expenditures at time t in varieties added to the consideration set at time $t - 1$ are never part of the false additions, since they are purchased both at $t - 1$ (as part of the experimentation expenditure x_{t-1}) and at time t —and thereby they are part of the intensive margin. In (34) ω_{nt}^1 denotes the (expected) expenditures at time t in varieties that (i) were already in the consideration set of the household at time $t - 2$ and (ii) were also consumed at time $t - 1$, conditional on being today in a sector with n varieties in the consideration set. The analytical expression for ω_{nt}^1 is given by (54) in the Appendix.

Removals from the consumption basket are defined as equal to the sum of all expenditures at time $t - 1$ in varieties which are not consumed at time t , which satisfies

$$R_t = \frac{x_{t-1}}{e_{t-1}} + \delta \frac{v_{t-1}}{e_{t-1}} + \xi_t^r. \quad (35)$$

The first term in the right hand side of (35) corresponds to the expenditures in experimentation at time $t - 1$, which lead to removals from the basket of the household at time t because, under our assumptions, the expenditures for experimentation in varieties that the household turns out to incorporate into her consideration set have zero measure. The second term in the right hand side of (35) corresponds to *true* removals from the consideration set of the household which happen with probability δ . Finally the third term in the right hand side of (35) measures removals due to shocks to preferences which make the household opt for a variety different from the one consumed at time $t - 1$ even if this is still in her consideration set today. These are *false* removals from the consideration set of the household. They are equal to the sum of the expenditures in all varieties purchased at time $t - 1$, that are still in the consideration set of the household at time t , but that are not purchased at time t just because they are preferred by other varieties in set. They are given by

$$\xi_t^r = (1 - \delta) \frac{v_{t-1}}{e_{t-1}} - \frac{1}{e_{t-1}} \sum_{n=1}^{\infty} b_{nt}^0 f(n; \mu_{t-1}), \quad (36)$$

which is equal to the difference between the total expenditures at time $t - 1$ in varieties that are also in the consideration set of the household at time t minus the total expenditures at time $t - 1$ in varieties that are purchased both at $t - 1$ and at t . In the expression, $1 - \delta$ is the probability that the variety consumed at time $t - 1$ remains in the consideration set of the household at time t , while ω_{nt}^0 denotes the expenditures at time $t - 1$ in varieties which are also consumed at time t , conditional on being at $t - 1$ in a sector with n varieties in the consideration set, whose analytical expression is given by (55) in the Appendix.

The intensive margin is defined as equal to the sum of the changes in the expenditures for varieties which are purchased both at time $t - 1$ and at time t . After remembering the definition of ω_{nt}^0 in (55) and the one of ω_{nt}^1 in (54), it can be calculated as equal to

$$I_t = \frac{(1 - \delta) \lambda x_{t-1}^\eta}{s \mu_t} \frac{v_t}{e_{t-1}} + \frac{1}{e_{t-1}} \sum_{n=1}^{\infty} [b_{nt}^1 f(n; \mu_t) - b_{nt}^1 f(n; \mu_{t-1})], \quad (37)$$

The first term measures the total expenditures at time t in varieties added to the consideration set at time $t - 1$. Notice that these varieties are purchased for sure at $t - 1$ as part of the experimentation expenditure x_{t-1} . The second term measures the total expenditures

at time t in varieties already present in the consideration set at $t - 2$ that are purchased both at $t - 1$ and at t . The sum of these two terms measures the expenditures at time t in varieties that were also purchased at time $t - 1$. From this we subtract the total expenditures at time $t - 1$ in varieties that are purchased both at $t - 1$ and at t , which is the third term in (37).

By combining (33), (35) and (37), we immediately see that household's total expenditures satisfies the identity

$$\frac{e_t - e_{t-1}}{e_{t-1}} = \frac{x_t - x_{t-1}}{e_{t-1}} + \frac{v_t - v_{t-1}}{e_{t-1}} = N_t + I_t, \quad (38)$$

where N_t denotes the net additions to the consumption basket of the household, equal to

$$N_t = A_t - R_t. \quad (39)$$

5.2 Parameter values calibrated ex-ante

The first set of parameters is chosen to standard values in the literature. The period of the model is a quarter. The time discount factor, β , is chosen such that the equilibrium yearly return on capital is 5%. We consider the household shopping behavior in a steady state where expenditure in steady state is denoted by $\bar{e} = 1$. Hence, $\bar{v} = 1 - \bar{x}$ is the share of consumption expenditure. The intertemporal elasticity of substitution is assumed to be equal to one, implying $\gamma = 1$. We set the parameter governing the elasticity of product substitution across sectors to $\sigma = 3$, consistent with the estimate of the elasticity of the intensive margin of demand for grocery products obtained by Paciello, Pozzi, and Trachter (). The elasticity of output to capital is set to $\omega = 1/3$, implying a labor share of output of $2/3$. The parameter controlling the curvature of the disutility from labor is set to $\phi = 2$ so to match a $1/2$ Frisch elasticity of labor supply to the wage. The depreciation rate of physical capital is assumed to be $\delta_k = 0.25$. Government spending is set to $g = 0.33$ corresponding to twenty percent of total output in steady state as in Hall (2009).

5.3 Parameter values calibrated in steady state

The set of parameters $\{\alpha, \lambda, \delta, s, \zeta, \eta, \psi\}$ determines the rate at which products are added and removed from the consumption basket, as described in Section 5.1. We calibrate these parameters to match observable statistics produced by our model under the assumption that the economy is in non-stochastic steady state, where $\varepsilon_t = 0$ and $z_t = 1$ for all t . Essentially, the calibration of these parameters is independent of the general equilibrium

of the model, so that we can focus on the consumption flows of the household in isolation. We first discuss the calibration of the attrition rate δ . Let $\tilde{A}_t^\tau$ denote the expenditure share in a generic period t on varieties added to the consideration set $\tau > 0$ quarters ago. Such value in steady state is given by

$$\tilde{A}_t^\tau = \delta (1 - \delta)^\tau \bar{v}, \quad (40)$$

where $\delta = \lambda \bar{x}^\eta / (s \bar{\mu})$ is the steady state share of expenditure on varieties added τ periods ago, $(1 - \delta)^\tau$ is the survival probability of each of them, and $\bar{v} = 1 - \bar{x}$ is the steady state share of expenditure on consumption. Hence, for given $\bar{x}$, any $\tilde{A}_t^\tau$ for some τ identifies δ . We choose $\tau = 5$, so to target the expenditure on varieties added one year before, in the same quarter of the year, and address the issue of seasonality. We next describe how we estimate $\tilde{A}_t^\tau$ in the KNCP data described in Section 2. We restrict the sample to those households that have been in the sample in each of the ten years from 2004 to 2013, amounting to 5,782 households. We then identify the additions to the consideration set of a given household in each of the four quarters of 2013, i.e. $t = 2013Q1 - 2013Q4$, as the expenditure in brands that have never been purchased by the household in any of the preceding nine years. Once we have identified the set of products added to the consideration set in a given reference quarter t of 2013, we compute the share of expenditure in those products $\tilde{A}_t^\tau$ at $\tau = 5$. We then take an average across households. We repeat this for each quarter of 2013, and take an average across the reference quarters t of 2013 for the same τ . This procedure gives $\tilde{A}_t^5 = 0.017$, implying $\delta = 0.02$.

We calibrate the value of the shape of the distribution governing preference shocks, α , the share of sectors where preferences change every period, ζ , and, conditional on being in a sector with persistent preferences, the probability that a household changes preferences change from one quarter to another, ψ as well as the steady state number of varieties per sector, $\bar{\mu}$, to fit the addition rate to quarter t that is robust to a rolling window of τ quarters in the past, $\hat{A}_t^\tau$, defined as

$$\hat{A}_t^\tau = \sum_{h \in \mathcal{H}_t} \hat{a}_{ht}^\tau \times \frac{e_{ht-1}}{\sum_{s \in \mathcal{H}_t} e_{st-1}}, \quad (41)$$

$$\hat{a}_{ht}^\tau = \sum_{\nu \in \mathcal{V}_t} \frac{e_{\nu ht}}{e_{ht-1}} \times \mathbb{I}(e_{\nu ht} > 0) \times \prod_{j=1}^{\tau} \mathbb{I}(e_{\nu ht-j} = 0), \quad (42)$$

where $\hat{a}_{ht}^\tau$ is the robust addition rate of household h given by the growth rate in expenditure in quarter t due to the purchase of varieties ν that have never been purchased in the preceding τ quarters; the aggregate robust addition rate $\hat{A}_t^\tau$ is a weighted average across

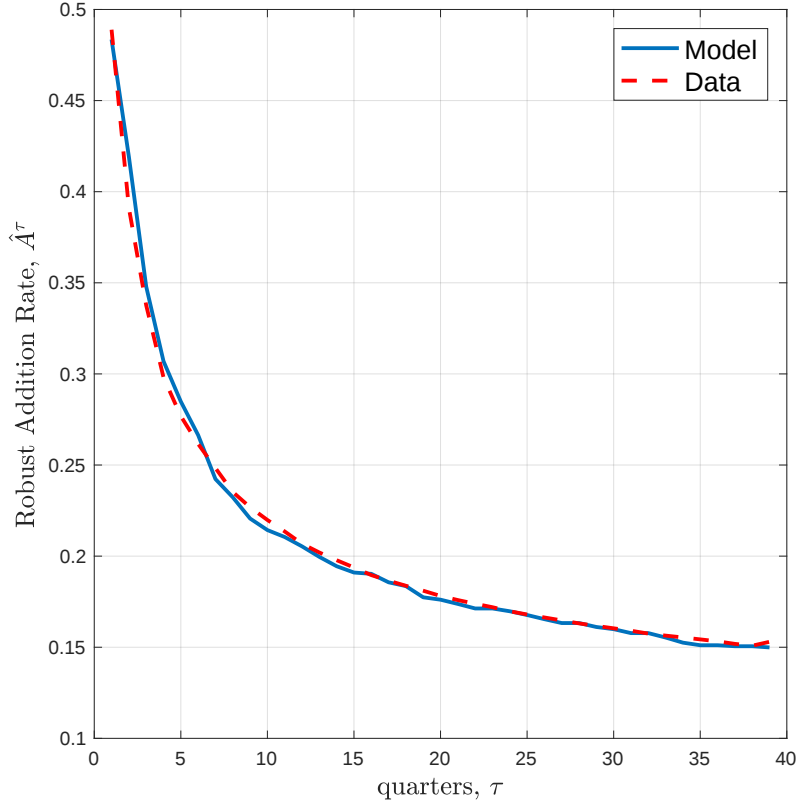
Table 4: Baseline calibration

Model		Data	
Parameter	Value	Moment	Value
$\bar{e}$	1	Normalization	1
g	0.33	Government spending share of output	0.2
β	0.988	Yearly real rate	5%
γ	1	Elasticity of inter-temporal substitution	1
σ	3	Elasticity of demand	3
ϕ	2	Frisch elasticity	0.5
δ_k	0.25	Physical capital yearly depreciation	0.1
ρ_z	0.85	Autocorrelation of GDP per capita	0.85
χ	9.4	Consumption expenditure response to rebate	2%
ω	0.33	Labor share of output	0.66
δ	0.02	Expenditure share of new goods after 1 year, $\tilde{A}^\tau$ with $\tau = 5$	1.7%
ζ	0.62	Fit robust addition rate $\hat{A}^\tau$ for $\tau = 1, 2, \dots, 39$	Figure 2
ψ	0.75	Fit robust addition rate $\hat{A}^\tau$ for $\tau = 1, 2, \dots, 39$	Figure 2
α	5.18	Fit robust addition rate $\hat{A}^\tau$ for $\tau = 1, 2, \dots, 39$	Figure 2
λ/s	0.89	Fit robust addition rate $\hat{A}^\tau$ for $\tau = 1, 2, \dots, 39$	Figure 2
s	88.2	Share of expenditure on each variety $\bar{q}$	1%
η	0.45	Net addition response to rebate	2/3 of Δe

the households. The red dashed line of Figure 2 plots the average robust addition rate $\hat{A}_t^\tau$, across the four quarters of 2013 and across all households that are in the sample in each year until 2004.

We notice that, in the model, the 1-quarter robust addition rate, $\hat{A}_t^1$, is given by equation

Figure 2: Addition rate robust to τ quarters, $\hat{A}^\tau$



Note: The figure plots $\hat{A}_\tau$, the share of expenditure due additions to the consumption basket in any quarter t of varieties that have not been purchased in any of the preceding τ quarters. The solid blue line is the prediction of the model at $\delta = 0.02$, $\bar{x} = 0.12$, $\bar{\mu} = 5$, $\zeta = 0.62$, $\psi = 0.75$ and $\alpha = 5.2$. The red dashed line is the data.

(33), whereas the addition rate robust to an infinite number of quarters, coincides with the growth rate of expenditure due to true additions plus experimentation,

$$\lim_{\tau \rightarrow \infty} \hat{A}_t^\tau = \frac{x_t}{e_{t-1}} + \frac{\lambda x_t^\eta}{s\mu_t} \frac{v_t}{e_{t-1}}.$$

We notice that, in the steady state with $\bar{e} = 1$, $\lim_{\tau \rightarrow \infty} \hat{A}_t^\tau = \bar{x} + \delta(1 - \bar{x})$, so that, given δ , $\lim_{\tau \rightarrow \infty} \hat{A}_t^\tau$ identifies the share of expenditure in experimentation $\bar{x}$. The quarter on quarter addition rate to the basket, $\hat{A}^1$, is informative about the average number of varieties per sector $\bar{\mu}$, whereas ζ and ψ determine the speed at which $\hat{A}^\tau$ converges to its asymptote, which is pin down by $\bar{x}$ and δ . Hence, we choose the values of $\bar{x}$, $\bar{\mu}$, ζ and ψ that best fit $\hat{A}^\tau$ for $\tau = 1, 2, \dots, 39$ reported in the red dashed line of Figure 2. We obtain $\bar{\mu} = 5$, $\zeta = 0.62$,

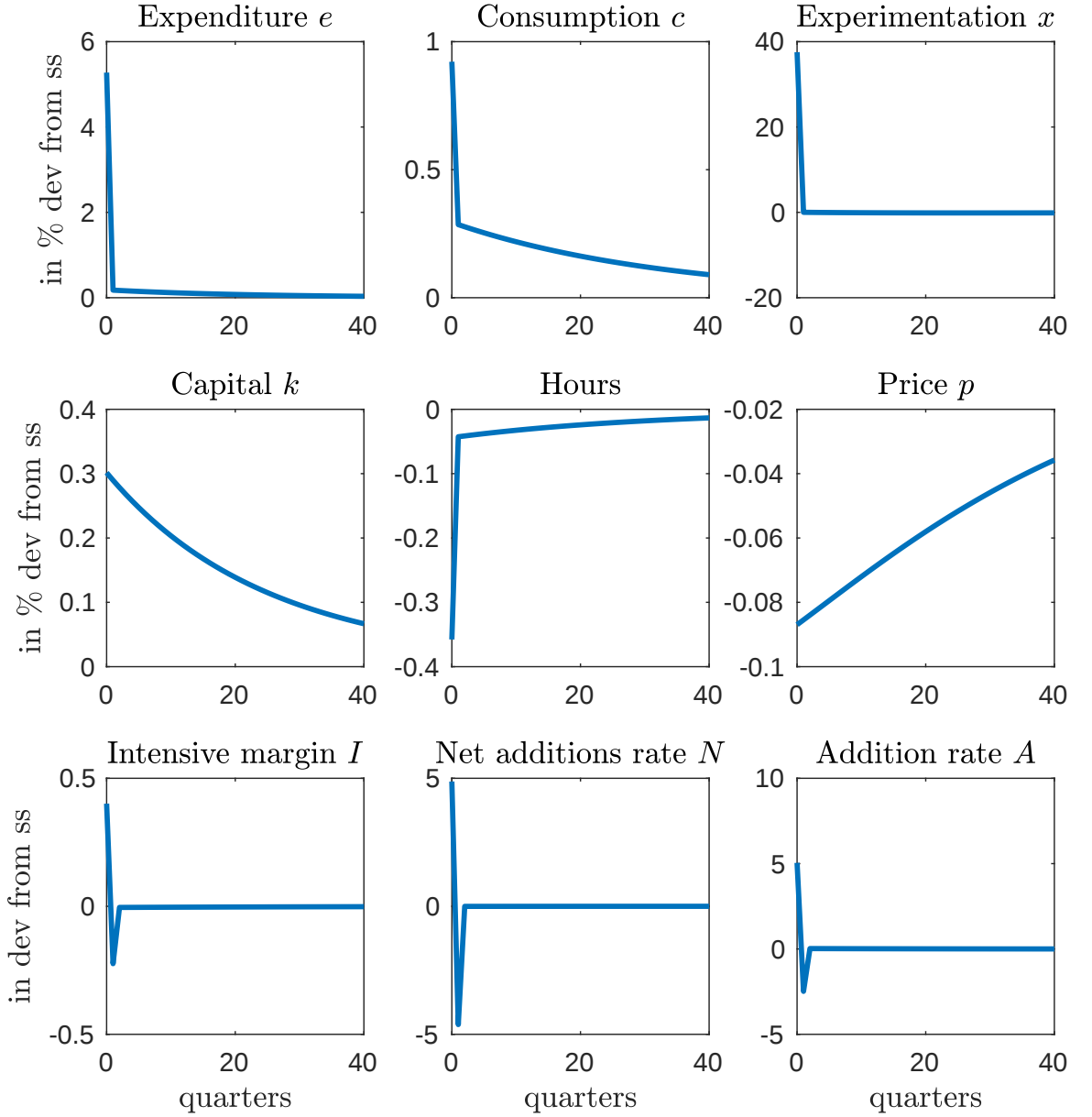
$\psi = 0.75$ and $\bar{x} = 0.12$. Next, given the first order condition for $\bar{x}$, we are able to backup $\alpha = 5.18$ from this combination of parameters and endogenous variables. Next, we calibrate the mass of different products s supplied in the economy. The value of s determines instead the level of the basket price $\bar{p}$ and hence, given the share of expenditure on consumption $1 - \bar{x}$, the actual consumption flow $\bar{c}$. We proceed as follows in the calibration of s . We notice that $\bar{v}$ can be written as $\bar{v} = s [1 - \exp(-\bar{\mu})] \bar{e}_\nu$, where $\bar{e}_\nu$ is the average share of expenditure on a variety conditional on purchasing it, and $s [1 - \exp(-\bar{\mu})]$ is the number of varieties actually purchased. Hence, in the data, we measure $\bar{e}_\nu$ as the share of expenditure on each variety added at a quarter t to the consideration set and repurchased at $t + 1$. Essentially in the measurement of $\tilde{A}^\tau$ for $\tau = 1$ we divide for the number of varieties that are repurchased. We do this for each of the four quarters of 2013 and take an average, obtaining $\bar{e}_\nu = 0.0104$. Then, given $\bar{v} = 1 - \bar{x}$ and the value of $\bar{\mu}$ we backup the mass of sectors $s = 88.2$.

5.4 Parameter values calibrated in response to a fiscal stimulus

Finally, we are left to calibrate the parameters λ , η and χ . We notice that, for given value of η , λ is given by $\lambda = s \delta \bar{\mu} / \bar{x}^\eta$. Hence, we next describe the calibration strategy for η and χ . The value of χ determines the frictions in asset adjustments that household face. A larger χ is associated to a smaller response of asset accumulation to shocks and, as a result, to a larger response of consumption expenditure, as the household is less able to smooth consumption over time. The parameter η governs instead the elasticity of experimenting to shocks. Everything else being equal, a larger η is associated to a larger response of net additions to shocks. We identify a transitory income shock to households in our data with the 2008 Economic Stimulus Act (discussed later), and simulate the response of our economy to such a shock by considering a one time, unexpected, one quarter reduction in the lump-sum tax τ , financed by an equivalent reduction in government expenditure g . In particular, we consider our economy where $\bar{k}$ and $\bar{\mu}$ are at the initial steady state at $t = 0$, where the household expects $z_t = 1$ and $\tau = 0.2$ for all $t \geq 0$. We study the effects on household consumption and experimentation behavior of a one time unforeseen reduction in τ equivalent to ten percent of quarterly expenditure $\bar{e}$, hence $\Delta\tau = -0.1$, which is equivalent in size to the observed policy experiment.

Operationally, we take a first order Taylor expansion of equations (27)-(29), together with equations (25) and (20), at the equilibrium values of the economy at $t = -1$, i.e. $\{\bar{k}, \bar{\mu}, \bar{x}, \bar{v}, \bar{p}\}$, resulting from the calibration in Table 4, and simulate the economy to the

Figure 3: Impulse responses to a fiscal stimulus



Notes: Numbers are reported in percentage units. Impulse responses to a one time decrease in income τ that lasts only for one period; the period is a quarter. The size of the shock is $\Delta\tau = -0.1 \times e_0$.

10% reduction in τ . In doing so, we set $\chi = 1.8$ to match the response in the quarter of impact of expenditures e_t to the fiscal stimulus equal to 47% of the increase in disposable income in the quarter of the rebate, as obtained from Table 5 in Broda and Parker (2014).⁸

⁸The authors estimate the marginal propensity to consume in the quarter of reception of the rebate

Moreover, we choose η to target a response of net additions corresponding to 2/3 of the overall increase in expenditure in the quarter of the shock, as reported below.

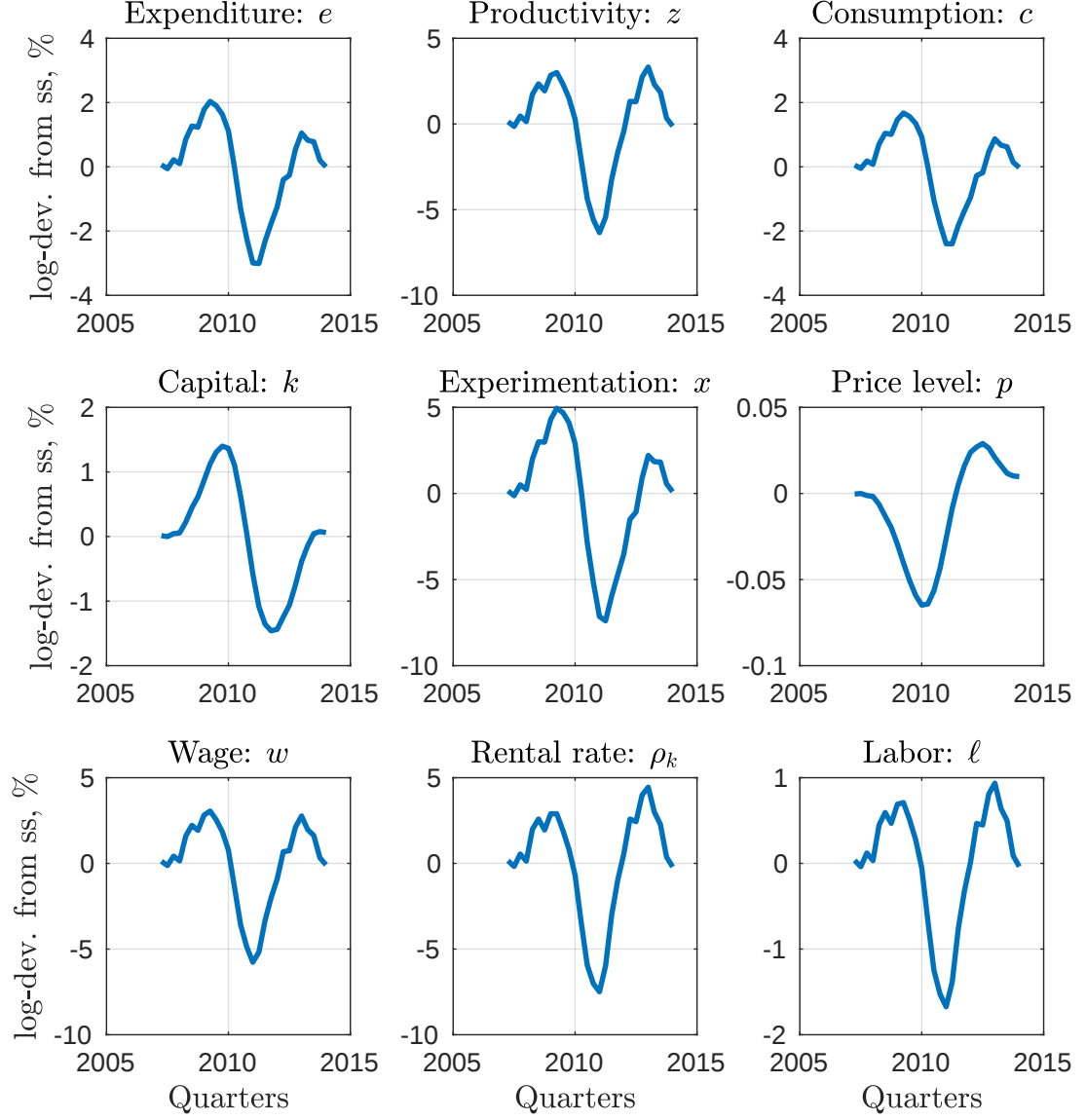
Figure 3 plots the impulse responses of selected variables to the tax reduction. Consumption expenditures, e_t increase on impact by 4.7% as a result of the calibration. Consumption, however, only increases by 0.9% on impact as part of the increase in consumption expenditures goes to finance higher experimentation x_t , which increases by 1/3 on impact of the shock. As a result of higher income the household increases her savings in capital to smooth consumption over time. However, due to the adjustment costs savings only go up by about 0.31% on impact relatively to their initial steady state. As a result of higher experimentation, the average number of varieties per sector in the consideration set of the household, μ , goes up, causing a fall in the price of the consumption basket of 0.07% on impact. We notice that, despite the shock lasts only for one period, its effect on consumption patterns is quite persistent. These results point to a new channel through which the household can smooth the positive effect of an increase in income over time. The motivation for the large response of experimentation to the shock is found in the first order condition for x_t in equation (28). When frictions reduce the ability of the household to smooth consumption through assets, experimentation is a valid alternative.

6 Consumption and experimentation in the cycle

We use our calibrated model to backup the time series for consumption experimentation in the U.S. business cycle, and compare the predicted time series for additions to their empirical counterparts. We proceed as follows. We assume that the economy is in steady state at $t = -1$, where $z_{-1} = 1$, $\mu_{-1} = \bar{\mu}$ and $k_{-1} = \bar{k}$. In our model the business cycle is driven by shocks to productivity ε_t . We next backup the sequence of shocks ε_t for $t \in [0, T]$, to productivity in equation (22) so that the growth rate of expenditure, e_t , implied by the log-linearized model calibrated as in Table 4 matches one-for-one the hp-filtered quarterly growth rate of consumption expenditure in the KNCP data from 2007:Q1 to 2014:Q4, reported in the bottom-left panel of Figure 1. Hence, in our exercise, 2007 : Q1 corresponds to $t = 0$ while 2014 : Q4 corresponds to $t = T$.

Once we have the shock process, ε_t , we can simulate our model and obtain the implied patterns of experimentation, x_t , consumption, c_t , the price level, p_t , as well as other aggregate variables as reported in Figure 4. The data gives us the increase in expenditures $(\partial \log(C_t) / \partial \log(Y_t))$ in the same set of data and households that we have used to calibrate our model.

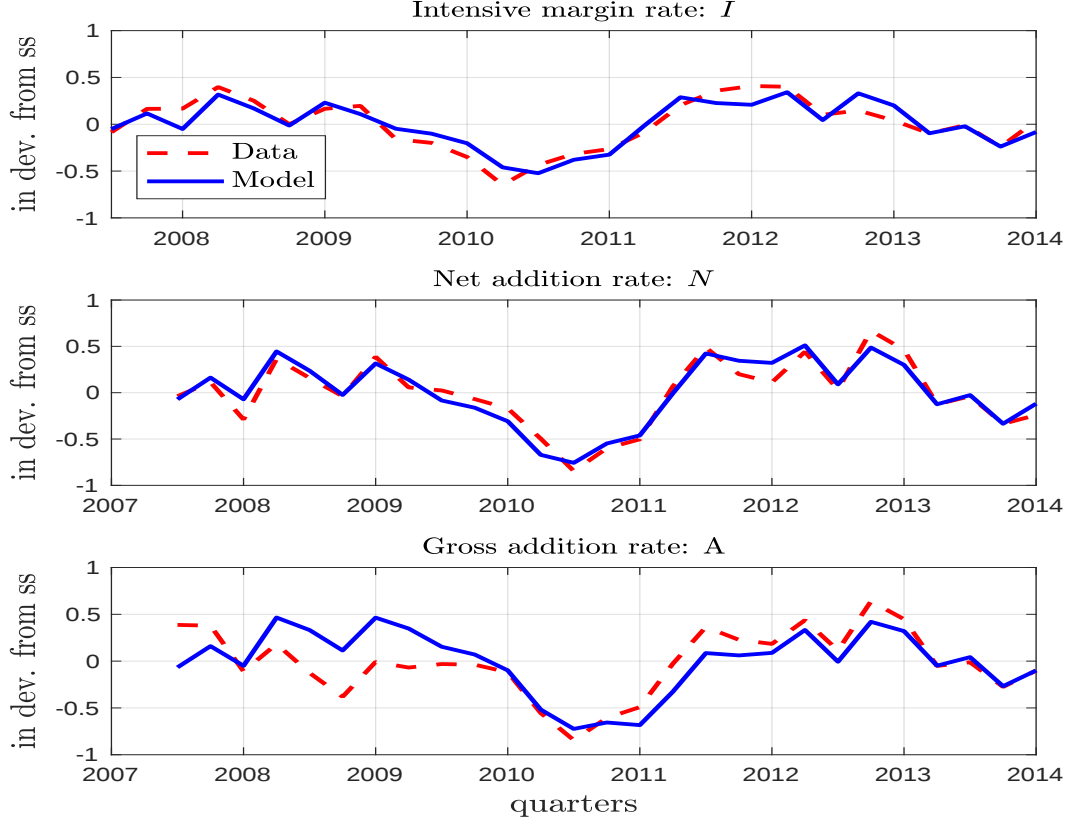
Figure 4: Model predicted business cycle dynamics



Notes: Dynamics of economic variables in the model in log-deviations from the initial steady state. The expenditure series matches the expenditure in the KNCP data from 2007:Q1 to 2014:Q4.

until 2008, where expenditure is 2% up relatively to the initial value, then drops to reach -3% of the initial value in 2010, and recovers to +1% in 2012. Expenditure, e_t , consumption c_t and experimentation x_t essentially comove with very similar patterns. This can be explained by the first order condition for x_t in equation (28), where a persistent shock to productivity as the one we are considering here (recall $\rho_z = 0.85$) reduces the needs for

Figure 5: Consumption expenditure flows: model vs data

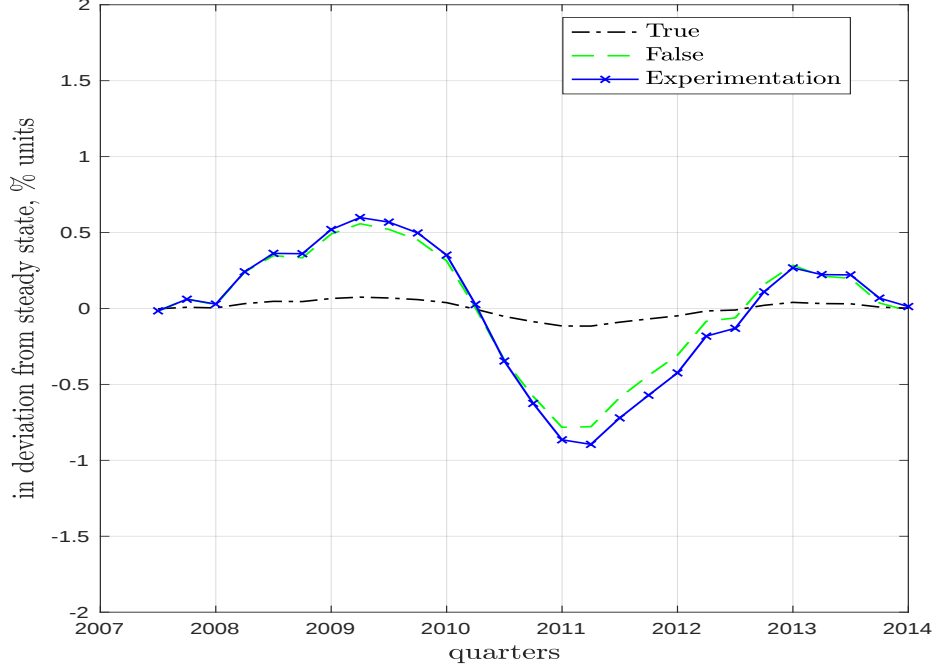


Notes: in % units. The intensive margin, net and gross additions are defined in equations (33)-(37).

savings due to consumption smoothing, and hence experimentation increases one for one with expenditure due to a scaling effect. The behavior of capital, labor and prices is then very much similar to a standard business cycle model. The price of the consumption basket in fact varies quantitatively little over the cycle as the stock of varieties μ_t is sluggish in adjusting to experimentation. It is interesting to notice however, that the pattern of experimentation lags with respect to consumption expenditure and productivity, and is more aligned with the path of physical capital. This is because of the decreasing returns to scale in experimentation $\eta < 1$ causing experimentation to increase slower over the cycle.

In terms of contribution to volatility of total expenditure e_t , we notice that experimentation accounts for 12% of the volatility over this sample. This implies that actual consumption volatility is 12% smaller than volatility of consumption expenditures in our model. This is a clear point of distinction from a standard RBC model, where there would not be any fluctuation in experimentation and consumption expenditure e_t would coincide with actual consumption c_t .

Figure 6: Decomposition of additions

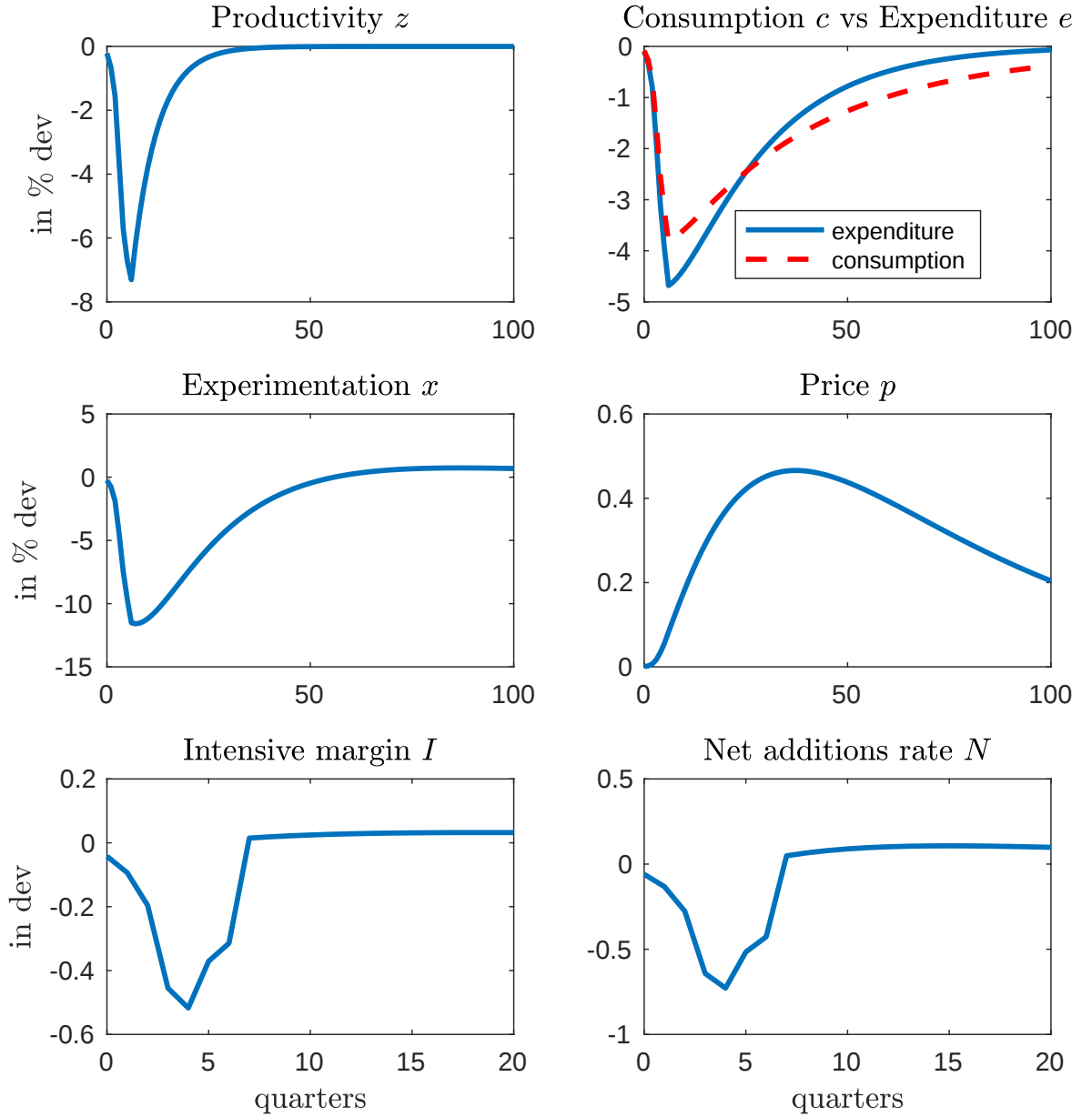


Notes: Model predicted decomposition of total additions as of equation (33) from 2007:Q1 to 2014:Q4.

To validate the predictions of our calibrated model, we compute the model-predicted quarter-on-quarter intensive margin, gross and net additions rate, i.e., I_t , A_t and N_t respectively as defined in equations (33)-(37). We then compare the time series predicted by the model to the hp-filtered empirical counterpart derived in Section 2.3 and plotted in Figure 1. We report our predicted series against the data counterpart in Figure (5). The model does a good job at capturing the cyclical behavior of additions and the intensive margin during the recession. In Figure (6) we use the model to decompose the contribution of additions to the change in demand from steady state in its three components as of equation (33): experimentation, true additions and false additions. We notice that fluctuations in experimentation, false and true additions are pro-cyclical. Experimentation and false additions account for almost all of the fluctuations in overall additions, with a similar contribution.

Finally, we focus on the response of the economy during the recession. We extrapolate the sequence of negative shocks to productivity, ε_t , between 2009:Q2 and 2010:Q4 estimated in the exercise reported in Figure 4, and look at the responses of economic variables in deviation from the 2009:Q1 level assuming that $\varepsilon_t = 0$ for $t > 2010 : Q4$. This exercise allows us to isolate the effect of the recession on the path of economic variables. We report

Figure 7: Response of the economy to shocks from 2009:Q2 to 2010:Q4



Notes: Impulse responses to the sequence of shocks to ε_t estimated between 2009:Q2 and 2010:Q4, assuming $\varepsilon_t = 0$ for $t > 2010 : Q4$. Numbers are reported in percentage units and in deviation from the 2009:Q1 level.

such responses in Figure 7. We emphasize that the response of actual consumption deviates substantially from the response of expenditures. In particular, consumption falls less at the peak of the recession as the household reduces experimentation expenditures. However, the fall in consumption is more persistent than expenditure due to the persistent increase

in the price index resulting from the reduction in experimentation.

7 Conclusions

We have provided novel evidence on the cyclical properties of consumption and its composition. Our results suggest that a large share of the fluctuations in non-durable consumption over the business cycle is due to households adding products that were not purchased in the previous periods to their consumption basket - the *extensive margin* of demand. Interestingly, the rate at which products are removed from the consumption basket displays much less pronounced cyclical properties. Using a model, we identify the part of the dynamics of extensive margin of demand that is due to the household experimentation for new products. We find that experimentation is highly pro-cyclical. Our novel framework highlights a new margin through which households can smooth consumption utility over time, namely experimenting to accumulate more products available for consumption to the household. A larger set of products, *consideration set*, allows the household to better fit his time-varying random preferences, and hence to increase her expected utility. When adjusting assets is costly, experimentation is a valid substitute to smooth consumption utility.

References

- Agarwal, S., N. Marwell, and L. McGranahan (2017). Consumption responses to temporary tax incentives: Evidence from state sales tax holidays. *American Economic Journal: Economic Policy* 9(4), 1–27.
- Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2017). Missing growth from creative destruction. NBER Working Paper 24023.
- Aguiar, M. and E. Hurst (2007). Life-cycle prices and production. *American Economic Review* 97(5), 1533–1559.
- Alessandria, G., J. P. Kaboski, and V. Midrigan (2010). Inventories, lumpy trade, and large devaluations. *American Economic Review* 100(5), 2304–2339.
- Argente, D. and M. Lee (2017). Cost of living inequality during the Great Recession. University of Chicago, mimeo.
- Attanasio, O. and N. Pavoni (2011). Risk sharing in private information models, with asset accumulation: Explaining the excess smoothness of consumption. *Econometrica* 79(4), 1027–1068. Mimeo, University College London.
- Bai, Y., J.-V. Ríos-Rull, and K. Storesletten (2012). Demand shocks that look like productivity shocks. Mimeo, University of Rochester.
- Baker, S. R., S. Johnson, and L. Kueng (2017). Shopping for lower sales tax rates. Mimeo, Northwestern University.

- Bartelsman, E. J. and P. J. Dhrymes (1998). Productivity dynamics: US manufacturing plants, 1972–1986. *Journal of Productivity Analysis* 9(1), 5–34.
- Bils, M. and P. Klenow (2001). Quantifying quality growth. *American Economic Review* 91(4), 1006–1030.
- Blanchard, O. J. and P. Diamond (1990). The cyclical behavior of the gross flows of U.S. workers. *Brookings Papers on Economic Activity* (2), 85–155.
- Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review* 98(5), 1887–1921.
- Broda, C. and J. A. Parker (2014). The Economic Stimulus Payments of 2008 and the aggregate demand for consumption? *Journal of Monetary Economics* (68), S20–S36.
- Broda, C. and D. E. Weinstein (2006). Globalization and the gains from variety. *American Economic Review* 2(121), 541–585.
- Broda, C. and D. E. Weinstein (2010). Product creation and destruction: Evidence and price implications. *American Economic Review* 3(100), 691–723.
- Campbell, J. and Z. Hercowitz (2009). Liquidity constraints of the middle class. *CentER Discussion Paper Series No. 2015-009*.
- Campbell, J. R. (1998). Entry, exit, embodied technology, and business cycles. *Review of economic dynamics* 1(2), 371–408.
- Campos, R. G. and I. Reggio (2017). Do the unemployed pay lower prices? A reassessment of the value of unemployment insurance. Banco de España, mimeo.
- Coibion, O., Y. Gorodnichenko, and G. H. Hong (2015). The cyclicality of sales, regular and effective prices: Business cycle and policy implications. *American Economic Review* 105(3), 993–1029.
- Davis, S. J. and J. Haltiwanger (1990). Gross job creation and destruction: Microeconomic evidence and macroeconomic implications. *NBER Macroeconomics Annual* 5, 123–186.
- Davis, S. J. and J. Haltiwanger (1992). Gross job creation, gross job destruction, and employment reallocation. *Quarterly Journal of Economics*, 819–863.
- Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review* 67(3), 297–308.
- Einav, L., E. Leibtag, and A. Nevo (2010). Recording discrepancies in Nielsen Homescan data: Are they present and do they matter? *Quantitative Marketing and Economics* 8(2), 207–239.
- Hall, R. E. (2009). By how much does gdp rise if the government buys more output? Technical report, National Bureau of Economic Research.
- Heathcote, J., G. L. Violante, and K. Storesletten (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review* 104(7), 2075–2126.
- Jaimovich, N., S. Rebelo, and A. Wong (2017). Trading down and the business cycle. Northwestern university, mimeo.

- Jaravel, X. (2018). The unequal gains from product innovations: Evidence from the US retail sector. *Quarterly Journal of Economics* forthcoming.
- Johnson, D. S., J. A. Parker, N. S. Souleles, and R. McClelland (2006). Household expenditure and the income tax rebates of 2001. *American Economic Review* 96(5), 1589–1610.
- Jones, C. and P. J. Klenow (2016). Beyond GDP? welfare across countries and time. *American Economic Review* 106(9), 2426–2457.
- Kaplan, G. and G. Menzio (2016). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy* 124(3), 771–825.
- Kaplan, G., K. Mitman, and G. Violante (2017). Non-durable consumption and housing net worth in the Great Recession: Evidence from easily accessible data. *Journal of Public Economics*.
- Kaplan, G. and S. Schulhofer-Wohl (2017). Inflation at the household level. *Journal of Monetary Economics* 91(4), 19–38.
- Kaplan, G. and G. Violante (2010). How much consumption insurance beyond self-insurance? *American Economic Journal: Macroeconomics* 2(4), 53–87.
- Kaplan, G. and G. L. Violante (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica* 82(4), 1199–1239.
- Krugman, P. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics* 9(4), 469–479.
- Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review* 70(5), 950–959.
- Kueng, L. (2018). Excess sensitivity of high-income consumers. *Quarterly Journal of Economics* forthcoming.
- Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review* 99(5), 2050–2084.
- Luce, R. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.
- Manzini, P. and M. Mariotti (2014). Stochastic choice and consideration sets. *Econometrica* 82(3), 1153–1176.
- Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012). Revealed attention. *American Economic Review* 102(5), 2183–2205.
- McFadden, D. (1974a). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.
- McFadden, D. (1974b). The measurement of urban travel demand. *Journal of Public Economics* 3, 303–328.
- McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (Eds.), *Spatial Interaction: Theory and Planning Models*, pp. 75–96. North Holland.

- McFadden, D. (1981). Econometric models of probabilistic choice. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data and Econometric Applications*, pp. 198–272. MIT Press.
- Neiman, B. and J. Vavra (2018). The rise in household spending concentration. University of Chicago, mimeo.
- Paciello, L., A. Pozzi, and N. Trachter. Price dynamics with customer markets. *International Economic Review* forthcoming.
- Parker, J. A., N. S. Souleles, D. S. Johnson, and R. McClelland (2013). Consumer spending and the economic stimulus payments of 2008. *American Economic Review* 103(6), 2530–2553.
- Philippon, T. and V. Midrigan (2011). Household leverage and the recession. NBER Working Paper 16965.
- Redding, S. J. and D. E. Weinstein (2016). A unified approach to estimating demand and welfare. NBER Working Paper 22479.

APPENDIX

A Proofs

A.1 Proof of Proposition 1

Let

$$\hat{z}_{jt} \equiv \max_{i \in \Omega_{jt}} \{z_{\nu jt}\} \quad (43)$$

denote the maximum value of the varieties present in the consideration set of sector j at time t , denoted by Ω_{jt} . For simplicity we adopt the convention $\hat{z}_j = 0$ if the consideration set of the household is empty, $\Omega_{jt} = \emptyset$. Without loss of generality, we assume that the household purchases at most one variety per sector: if two or more varieties have the same value of z —which is a zero probability event—the household randomly buys only one of them. The total expenditure of the household in sector j at time t can then be expressed as follows

$$q_{jt} = c_t (\hat{z}_{jt})^{\sigma-1} \iota_t^{-\sigma}, \quad (44)$$

where ι_t is the Lagrange multiplier on the budget constraint (18). Let $v_t = \int_0^s q_{jt} dj$. After integrating (44) over j , we can then solve for ι_t to obtain

$$\iota_t^\sigma = \frac{c_t}{v_t} \int_0^s (\hat{z}_{jt})^{\sigma-1} dj, \quad (45)$$

which can be used to replace ι_t in (44). The resulting expression can be substituted in (17) to yield:

$$c_t = v_t \left[\int_0^s (\hat{z}_{jt})^{\sigma-1} dj \right]^{\frac{1}{\sigma-1}}. \quad (46)$$

Since the price index p_t should satisfy the identity $p_t c_t = v_t$, (46) implies that p_t is equal to

$$p_t = \left(\int_0^s (\hat{z}_{jt})^{\sigma-1} dj \right)^{-\frac{1}{\sigma-1}}. \quad (47)$$

To calculate (47), we use the law of iterated expectations and partition the sectors according to their number of varieties in the consideration set of the household at time t , n_{jt} , which allows us to write

$$p_t = \left\{ \sum_{n=0}^{\infty} E[(\hat{z}_{jt})^{\sigma-1} \mid n_{jt} = n] f(n; \mu_t) \right\}^{-\frac{1}{\sigma-1}}. \quad (48)$$

To evaluate (48), notice that, given Ω_{jt} , the preference shocks $z_{\nu jt}$'s in (43) are i.i.d (independently on when they are drawn). Also notice that if in a sector there are n_{jt} varieties in the consideration set and the preference shocks $z_{\nu jt}$'s are drawn (independently) from a Frechét distribution G with shape parameter α and scale parameter equal to one, then the

CDF of $\hat{z}_{jt}^{\sigma-1}$ is given by

$$\Pr \left(\max_{i=1,2,\dots,n_j} (z_{\nu_{jt}})^{\sigma-1} \leq u \mid n_{jt} = n \right) = \prod_{i=1}^{n_j} G \left(u^{\frac{1}{\sigma-1}} \right) = \exp \left(- \left(n_j^{-\frac{\sigma-1}{\alpha}} u \right)^{-\frac{\alpha}{\sigma-1}} \right),$$

which implies that $\hat{z}_{jt}^{\sigma-1}$ is distributed according to a Frechét distribution with shape parameter $\alpha/(\sigma-1)$ and scale parameter equal to $n_j^{\frac{\sigma-1}{\alpha}}$. Hence, if $\alpha > \sigma-1$ we have that

$$E[(\hat{z}_{jt})^{\sigma-1} \mid n_{jt} = n] = n^{\frac{\sigma-1}{\alpha}} \Gamma \left(1 - \frac{\sigma-1}{\alpha} \right), \quad (49)$$

which can be substituted into (48) to yield (??).

After substituting (45) into (44), we obtain that the expected expenditures in a sector with $n_{jt} = n$ varieties in the consideration set are equal to

$$v_{nt} = v_t \frac{E[(\hat{z}_{jt})^{\sigma-1} \mid n_{jt} = n]}{\int_0^s (\hat{z}_{jt})^{\sigma-1} dj} = v_t \frac{n^{\frac{\sigma-1}{\alpha}} \Gamma \left(1 - \frac{\sigma-1}{\alpha} \right)}{\int_0^s (\hat{z}_{jt})^{\sigma-1} dj},$$

where the second equality uses (49). By using (??) and (47) we finally obtain (??).

A.2 Proof of Proposition 2

We prove that the distribution at time t is equal to the distribution of the sum of two independent Poisson random variables, one with mean $\lambda_1 = (1-\delta)\mu_{t-1}$, the other with mean $\lambda_2 = \frac{x_t}{s}$. The proof then follows from the fact that the sum of two independent Poisson random variables is again a Poisson random variable with mean $\lambda_1 + \lambda_2$. The first random variable X_1 corresponds to the number varieties in the consideration set of the household in the sector after the taste shock δ is realized and before experimenting for new varieties at time t . X_1 is Poisson because if Z_1 is Poisson with mean μ_{t-1} and the distribution of X_1 conditional on $Z_1 = k$ is a binomial distribution with number of trial k and success probability $1-\delta$, $X_1 \mid (Z_1 = k) \sim \text{Binomial}(k, 1-\delta)$, then X_1 is a Poisson random variable with mean $(1-\delta)\mu_{t-1}$, i.e. $X_1 \sim \text{Pois}((1-\delta)\mu_{t-1})$. The second random variable X_2 corresponds to the number of new varieties discovered in a sector by spending x_t on experimentation which we show is a Poisson random variable with expected value $\lambda x_t/s$, i.e. $X_2 \sim \text{Pois}(\lambda x_t/s)$. We prove this latter result using two alternative logics that clarify the meaning of undirected experimentation.

1. *Logic one* Assume that the household consists of s shoppers. Each shopper is randomly assigned to a sector, but the identity of the sector is unknown. The shopper spends x_t/s in new varieties. If the shopper finds a variety the household likes, she also discovers the sector she was assigned to. Otherwise the shopper reports to the household that nothing was discovered and the identity of the sector remains unknown. The number of new varieties discovered by the shopper in the sector is Poisson with expected value $\lambda x_t/s$, since the probability of discovering exactly k new varieties in

the sector is equal to

$$\lim_{\phi \rightarrow 0} \binom{\frac{x_t}{s\phi}}{k} (\lambda\phi)^k (1 - \lambda\hat{\delta})^{\frac{x_t}{s\phi} - k} = \frac{\left(\frac{\lambda x_t}{s}\right)^k e^{-\frac{\lambda x_t}{s}}}{k!} \quad (50)$$

where the term in the left hand side corresponds to the probability of observing k successes under a binomial distribution with number of trials $x_t/(s\phi)$ and success probability $\lambda\hat{\delta}$ and the equality uses the Stirling formula, which allows to conclude that $X_2 \sim Pois(\lambda x_t/s)$.

2. *Logic two* Suppose that the varieties in the different sectors are indexed over the unit interval and that the unit interval is partitioned in $1/\phi$ equal intervals of measure ϕ . Also assume that the household consists of a $1/\phi$ shoppers who sample varieties across sectors rather than within a sector. Each shopper is specialized in sampling a specific portion of the unit interval of measure ϕ . Each shopper spends $x_t\phi$ units in buying varieties in that portion in different sectors. The sector of the variety is initially unknown. Once the shopper discovers a new variety the household likes, the shopper also discovers the identity of the sector, otherwise the sector remains unknown. The probability that the variety belongs to a specific sector is equal to ϕ/s . Given that the shopper has discovered exactly m varieties overall, the number of varieties found in a specific sector is given by a binomial distribution with success probability ϕ/s and number of trials equal to m . Moreover for ϕ that goes to zero the total number of new varieties discovered by a shopper is a Poisson random variable with mean λx_t , which follows from a logic analogous to (50) after taking into account that each shopper spends x_t . Then the unconditional probability that a shopper discovers exactly n new varieties in a specific sector is given by the probability distribution of a Poisson random variable with mean $\lambda x_t \hat{\delta}/s$ —equal to the product of the mean of the Poisson random variable and the mean of the Binomial, see XXX. But in the household there $1/\phi$ shoppers who experiments independently. Since the sum of independent Poisson random variables is again a Poisson random variable with mean equal to the sum of the means of the independent random variables, we then conclude that the number of new varieties in a sector is characterized by a Poisson random variable with mean $\lambda x_t/s$, i.e. $X_2 \sim Pois(\lambda x_t/s)$.

A.3 Proof of Proposition 3

$$\frac{\partial \bar{x}}{\partial \rho} = \bar{e} \frac{1 - \frac{1}{\delta}}{[\bar{\varepsilon}_{p\mu}(1 + \rho) + 1 + \frac{\rho}{\delta}]^2 - (1 + \rho) \left(1 + \frac{\rho}{\delta}\right) \frac{\lambda}{s} \frac{\partial \bar{\varepsilon}_{p\mu}}{\partial \mu}} < 0,$$

where $\partial \bar{\varepsilon}_{p\mu}/\partial \mu < 0$ is given by

$$\begin{aligned} (\sigma - 1) \frac{\partial \bar{\varepsilon}_{p\mu}}{\partial \mu} &= \frac{E(n^{u+2}) E(n^u) - [E(n^{u+1})]^2}{[E(n^u)]^2 \mu} - 1 \\ &= \left[\frac{E(n^{u+2})}{E(n^{u+1})} - \frac{E(n^{u+1})}{E(n^u)} \right] \frac{E(n^{u+1})}{E(n^u) E(n)} - 1 \in (-1, 0) \end{aligned}$$

with $E(n^u) \equiv \sum_{n=0}^{\infty} n^u f(n; \bar{\mu})$ and $u = \frac{\sigma-1}{\alpha}$. The proof that $\partial \bar{\varepsilon}_{p\mu} / \partial \mu < 0$ follow from noticing that, given the properties of the Poisson distribution, $\partial \bar{\varepsilon}_{p\mu} / \partial \mu = 0$ when $u = 1$, and the argument of the square bracket is increasing in u , so as the ratio that multiplies the square bracket. Hence at $\kappa = \frac{\sigma-1}{\alpha} < 1$ $\partial \bar{\varepsilon}_{p\mu} / \partial \mu$ must be smaller than at $u = 1$. The proof that $\partial \bar{\varepsilon}_{p\mu} / \partial \mu > -1$ follows from noticing that the first term is strictly positive. Next we show that the elasticity of experimentation to steady state expenditure is smaller than one. This follows immediately from equation (31) and $\partial \bar{\varepsilon}_{p\mu} / \partial \mu < 0$:

$$\frac{\partial \bar{x}}{\partial \bar{e}} \frac{\bar{e}}{\bar{x}} = \frac{1}{1 - \frac{\bar{x}^2}{\bar{e}} \frac{\lambda}{s} \frac{\rho \bar{e} + 1 + \frac{\rho}{s}}{\bar{e}(1+\rho)} \frac{\partial \bar{\varepsilon}_{p\mu}}{\partial \mu}} \in (0, 1).$$

B Extensions of the discrete choice model

B.1 Expected utility under the Gumbel distribution

Let $z_{\nu j}$ be independently and identically distributed according to a Gumbel distribution with shape parameter κ , and zero location parameter, i.e. with CDF $\exp(-e^{-\kappa z})$, in each sector j , then $\hat{z}_j = \max_{i=1, \dots, n_j} \{z_{\nu j}\}$ is distributed according to a Gumbel distribution with shape parameter κ and location parameter $\log(n_j)/\kappa$, i.e. $\exp(-e^{-\kappa(z - \log(n_j)/\kappa)})$. The mean of $\hat{z}_j$ is then given by $(0.5772 + \log(n_j))/\kappa$. We notice that the expectation of $\hat{z}_j^{\sigma-1}$ does not generally have a closed form solution. In the special case $\sigma = 3$, we obtain $E[\hat{z}_j^{\sigma-1}] = (\pi^2/6 + (\log(n_j) + 0.5772)^2)/\kappa^2$, which is characterized by an elasticity that is not constant with respect to n .

B.2 Probability of drawing variety with heterogeneous quality

As all varieties are ex-ante identical within each sector, they all have the same expenditure share which is $1/n_j$ in sector j for a given household; $1/n_j$ is also the probability that a specific variety is purchased in sector j . If we were to allow for fixed factors of heterogeneity, $\theta_{\nu j}$, across varieties within the same sector j , so that the utility flow from consuming one unit of variety i in sector j were to be given by $q_{\nu j} = \theta_{\nu j} + z_{\nu j}$, then the probability of choosing variety i is given by

$$P(i|j) = \frac{\theta_{\nu j}^{\alpha}}{\sum_{i=1}^{n_j} \theta_{\nu j}^{\alpha}}.$$

B.3 Analytical additions and removals

Proof of (33) We now prove that the expression for total additions in (33) holds true by deriving its second and third term.

Derivation of second term We start deriving the expression for the second term of (33) which measures the contribution of true additions. This term is calculated as follows. If the new variety is purchased in a sector with n varieties in the consideration set, the household spends (in expected value) v_{nt} in the variety. Conditional on being in a sector with n varieties, the probability that the household purchases a newly added variety is

equal to

$$\sum_{m=0}^n b\left(m; \frac{\lambda x_t}{s\mu_t}, n\right) \frac{m}{n} = \frac{\lambda x_t}{s\mu_t}, \quad (51)$$

which allows to calculate the overall expenditures in varieties newly added to the consideration set as equal to

$$\sum_{n=1}^{\infty} v_{nt} \frac{\lambda x_t}{s\mu_t} f(n; \mu_t) = \frac{\lambda x_t}{s\mu_t} v_t. \quad (52)$$

To calculate the probability in (51) we used the fact that if X_i , $i = 1, 2$ are independent Poisson random variables with mean λ_i , then the distribution of X_1 given $X_1 + X_2$ is a binomial distribution with success probability $\frac{\lambda_1}{\lambda_1 + \lambda_2}$ and number of trials equal to $X_1 + X_2$ (see XXXX). In our case $X_1 + X_2$ corresponds to the n varieties in the consideration set of the household at time t which is the sum of the $m \geq 0$ new varieties resulting from the successful experimentation of the household at time t (which is a Poisson random variable with mean $\frac{x_t \lambda}{s}$) and of the old varieties inherited from the consideration set of the household at time $t - 1$ (which is a Poisson random variable with mean $(1 - \delta)\mu_{t-1}$). Since each variety in the consideration set has an equal probability to be purchased, the left-hand side of (51) calculates the probability of purchasing a new variety in a sector with n varieties in the consideration set, which, after using the property of the mean of a binomial random variable, leads to the last equality in (51).

Derivation of third term To derive the expression for the third term of (33) as specified in (34), notice that the expenditures at time t in varieties added to the consideration set at time $t - 1$ are never part of the additions at time t , since the variety was necessarily consumed at $t - 1$ as part of the experimentation expenditure x_{t-1} . The first term in (34) corresponds to the expenditures at time t in all varieties already part of the consideration set of the household at time $t - 2$ and is obtained by using a logic analogous to the one used to derive (52). We now derive the expression for the second term in (34), which measures the total expenditures at time t in varieties already present in the consideration set at $t - 2$ that are purchased both at $t - 1$ and at t . Remember that ω_{nt}^1 denotes the (expected) expenditures at time t in varieties that (i) were already in the consideration set of the household at time $t - 2$ and (ii) were also consumed at time $t - 1$, conditional on being today in a sector with n varieties in the consideration set. To derive an expression for ω_{nt}^1 we first prove that the number of varieties k which have exited the consideration set in a sector between time $t - 1$ and t given that $\hat{u}$ varieties in the sector have survived until time t is Poisson with mean $\delta\mu_{t-1}$ —which also implies that k is independent of $\hat{u}$. To prove this, notice that the joint probability that in the sector there were $\hat{u} + k$ varieties in the consideration set at $t - 1$ and that $\hat{u}$ of them have survived at t is equal to

$$\binom{\hat{u} + k}{k} (1 - \delta)^{\hat{u}} \delta^k \frac{(\mu_{t-1})^{\hat{u} + k}}{(\hat{u} + k)!} e^{-\mu_{t-1}}$$

The unconditional probability than $\hat{u}$ varieties in the sector have survived at time t is equal to

$$\frac{[(1 - \delta)\mu_{t-1}]^{\hat{u}}}{\hat{u}!} e^{-(1 - \delta)\mu_{t-1}}$$

Then the probability that k varieties have exited the consideration set between $t - 1$ and t given that $\hat{u}$ varieties have remained in the set between $t - 1$ and t is equal to

$$\frac{\binom{\hat{u}+k}{k} (1-\delta)^{\hat{u}} \delta^k \frac{(\mu_{t-1})^{\hat{u}+k}}{(\hat{u}+k)!} e^{-\mu_{t-1}}}{\frac{[(1-\delta)\mu_{t-1}]^{\hat{u}}}{\hat{u}!} e^{-(1-\delta)\mu_{t-1}}} = \frac{(\delta\mu_{t-1})^k e^{-\delta\mu_{t-1}}}{k!} = f(k; \delta\mu_{t-1}),$$

which is a Poisson distribution with mean $\delta\mu_{t-1}$.

We now use the fact that the number of varieties which have exited the consideration set from $t - 1$ to t , k , and the number of varieties which have survived, $\hat{u}$, are independent (Poisson) random variables. Moreover, notice that the number of varieties which have survived in the consideration set is the sum of two independent (Poisson) random variables: those which survived from the consideration set at $t - 2$ (denoted by u in the expression below) and those newly added to the consideration set at $t - 1$ which have survived until t (denoted by j in the expression below). Then the (expected) expenditures at time t in varieties that (i) were already in the consideration set of the household at time $t - 2$ and (ii) were also consumed at time $t - 1$, conditional on being today in a sector with n varieties in the consideration set can be calculated as follows:

$$\omega_{nt}^1 = \sum_{k=0}^{\infty} \sum_{u=1}^n \sum_{j=0}^{n-u} \left\{ f(k; \delta\mu_{t-1}) m\left(u, j; \frac{(1-\delta)^2\mu_{t-2}}{\mu_t}, \frac{(1-\delta)\frac{\lambda_{x_{t-1}}}{s}}{\mu_t}, n\right) \times \frac{u}{k+u+j} \left[\varphi^{\frac{k+u+j}{n+k}} v_{n+kt} + (1-\varphi) \frac{v_{nt}}{n} \right] \right\}. \quad (53)$$

where

$$m\left(u, j; \frac{(1-\delta)^2\mu_{t-2}}{\mu_t}, \frac{(1-\delta)\frac{\lambda_{x_{t-1}}}{s}}{\mu_t}, n\right) = \frac{n!}{(n-u-j)!u!j!} \left(\frac{\frac{\lambda_{x_t}}{s}}{\mu_t}\right)^{n-u-j} \left[\frac{(1-\delta)^2\mu_{t-2}}{\mu_t}\right]^u \left[\frac{(1-\delta)\frac{\lambda_{x_{t-1}}}{s}}{\mu_t}\right]^j$$

is the multinomial distribution characterizing the probability that given n varieties in the consideration set of the sector at time t , u of them were also in the consideration set at time $t - 2$, j of them were added to the consideration set at time $t - 1$, while the remaining $n - u - j$ were added just at t . Given n , and the properties of independent Poisson random variables we immediately have that $n - u - j$, u and j are multinomial random variables with success probabilities equal to $\frac{\lambda_{x_t}}{\mu_t}$, $\frac{(1-\delta)^2\mu_{t-1}}{\mu_t}$ and $\frac{(1-\delta)\frac{\lambda_{x_{t-1}}}{s}}{\mu_t}$, respectively. Given k , u and j , the probability that the variety consumed at $t - 1$ was in the consideration set at $t - 2$ and has remained in the consideration set at t is equal to $u/(k + u + j)$. Given k , u and j , and that the variety consumed at $t - 1$ has survived, the term in square brackets of (53) calculates the expected expenditures at time t in the variety, by considering separately the case where the preferences towards the varieties in the consideration set at time $t - 1$ have remained unchanged (probability $\varphi \equiv \varphi_0\varphi_1$) from the case where they are redrawn at time t (probability $1 - \varphi$). Notice that the variety consumed at time $t - 1$ was preferred to all the $k + u + j$ varieties present in the consideration set at time $t - 1$. Then, without redrawing of preferences and given the symmetry in preference shocks, the variety consumed at $t - 1$ is also consumed at time t with probability $\frac{k+u+j}{n+k}$ (this is the probability that a variety has maximal value over $n + k$ varieties given that it is the maximal value over $k + u + j$ varieties), and, conditional on consumption, expenditures are equal to v_{n+kt} , since yesterday

the variety was preferred to all the $k + u + j$ varieties present in the consideration set and today it is also preferred to all the $n - u - j$ varieties newly added to the consideration set at time t . This explains the first term in square brackets of (53). The second term considers the case when there is a redrawing of preferences. In this case the variety consumed at $t - 1$ is chosen with probability $1/n$ and, conditional on consumption, expenditures are equal to v_{nt} , since the variety is the one preferred among the n varieties in the today consideration set. We now simplify (53) by canceling out the term $k + u + j$. Then by using the property of the value of the mean of a multinomial random variable, we finally obtain that

$$\omega_{nt}^1 = \sum_{k=0}^{\infty} \left\{ f(k; \delta\mu_{t-1}) \left[\frac{\varphi n}{n+k} \frac{(1-\delta)^2 \mu_{t-2}}{\mu_t} v_{n+kt} + (1-\varphi) \frac{v_{nt}}{n} \vartheta_{n,k,t} \right] \right\} \quad (54)$$

where

$$\vartheta_{n,k,t} = \sum_{u=0}^n \sum_{j=0}^{n-u} m \left(u, j; \frac{(1-\delta)^2 \mu_{t-2}}{\mu_t}, \frac{(1-\delta) \frac{\lambda x_{t-1}}{s}}{\mu_t}, n \right) \frac{u}{k+u+j}$$

Analytical expression for ω_{nt}^0 Remember that ω_{nt}^0 denotes the expenditures at time $t - 1$ in varieties which are also consumed at time t , conditional on being at $t - 1$ in a sector with n varieties in the consideration set. We prove that ω_{nt}^0 is equal to

$$\omega_{nt}^0 = v_{nt-1} (1 - \delta) \sum_{m=0}^{\infty} f\left(m; \frac{\lambda x_t}{s}\right) \left[\varphi \frac{n}{m+n} + (1 - \varphi) \sum_{u=0}^{n-1} b(u; 1 - \delta, n - 1) \frac{1}{m + u + 1} \right] \quad (55)$$

where we denoted by

$$b(u; 1 - \delta, n) = \binom{n}{u} (1 - \delta)^u \delta^{n-u}, \quad u = 0, 1, \dots, n$$

the probability of u successes in the case of a binomial random variable with success probability $1 - \delta$ and number of trials equal to n . To understand (55), start noticing that at time $t - 1$, the expenditures in a sector with n varieties in the considerations set are v_{nt-1} and there is a probability $(1 - \delta)$ that the variety consumed at $t - 1$ remains in the consideration set at time t . This accounts for the first two terms in the right hand side of (55). To understand the two summatories in (55) notice that the index m refers to the number of new varieties which have entered the consideration set of the household at time t , while $u \leq n - 1$ refers to the number of old varieties (different from the one consumed at time $t - 1$) present in the consideration set of the household both at time $t - 1$ and at time t . Notice that m and u are two independent random variables: m is Poisson with mean $\frac{\lambda x_t}{s}$; u is binomial with success probability $1 - \delta$ and number of trials equal to $n - 1$. Then, for given number of new varieties m , the term in square brackets calculates the probability that the variety consumed at time $t - 1$ is still consumed at t by separately considering the case where the preferences towards the varieties in the consideration set at time $t - 1$ remain unchanged (probability $\varphi \equiv \varphi_0 \varphi_1$) from the case where they are redrawn at time t (probability $1 - \varphi$). In the former case, the variety consumed at $t - 1$ —which is the one preferred among the n varieties in the consideration set at $t - 1$ —is still consumed at time t , provided that it is *not* dominated by any of the m varieties newly added to the consideration set, which, due to symmetry in preference shocks, happens with probability $n/(m+n)$ —which is the probability that a variety has maximal value over $m+n$ varieties given that it is the maximal value over n varieties. If preferences are redrawn, the variety consumed at time $t - 1$ is also consumed at time t only if the two following conditions are both verified: it should be preferred to the $m \geq 0$ new varieties added to the consideration set of the household at time t ; and it should remain preferred to the $u \leq n - 1$ other old varieties inherited from the consideration set of the household at time $t - 1$, whose preferences have been redrawn. Due to symmetry in preferences, these two conditions are simultaneously satisfied with probability $\frac{1}{m+u+1}$. By summing over the possible realizations of u , we finally obtain the term in square brackets in (55), while by summing over m we calculate the probability that, conditional on survival, the variety consumed at $t - 1$ is also consumed at time t .

B.3.1 Aggregations

We can compute true additions over an interval of τ periods. In particular, the contribution of expenditure due to true additions plus experimenting over $\tau = 1, 2, 3, \dots$ periods, to total expenditure in those τ periods, i.e. $\tau \bar{e}$, is given by

$$\hat{A}^\tau \equiv \frac{\bar{x}}{\bar{e}} + \delta \frac{\bar{e} - \bar{x}}{\bar{e}} \frac{\sum_{j=0}^{\tau} \sum_{k=0}^j (1-\delta)^{j-k}}{\tau} = \frac{\bar{x}}{\bar{e}} + \left(1 - \frac{\bar{x}}{\bar{e}}\right) \left[1 - \frac{1-\delta - (1-\delta)^{\tau+1}}{\tau \delta}\right]. \quad (56)$$

Similarly, the share of expenditure for τ periods, say from t_0 to $t_0 + \tau - 1$ of varieties added to the consideration set in the preceding τ periods, say from $t_0 - \tau$ to $t_0 - 1$, is given by

$$\tilde{A}^\tau = \left(1 - \frac{\bar{x}}{\bar{e}}\right) \delta \frac{1}{\tau} \sum_{k=0}^{\tau-1} (1-\delta)^{-k} \sum_{j=\tau}^{2\tau-1} (1-\delta)^j = \left(1 - \frac{\bar{x}}{\bar{e}}\right) \frac{1-\delta}{\tau \delta} [1 - (1-\delta)^\tau]^2. \quad (57)$$

C Log-linearization of the partial equilibrium model

Let $\hat{x}_t = dx_t/x_{-1}$ denote the log-deviation of the equilibrium variable (in this case x) from its initial value before the shock. The linearized first order conditions imply:

$$\begin{aligned} \chi(\hat{k}_t - \hat{k}_{t-1}) - \beta \chi(\hat{k}_{t+1} - \hat{k}_t) &= (\gamma - 1)(\hat{p}_{t+1} - \hat{p}_t) - \gamma(\hat{v}_{t+1} - \hat{v}_t) + (1 - \beta(1 - \delta))\hat{\rho}_{kt+1}, \\ \hat{k}_t &= (1 + \bar{\rho}_k - \delta)(\hat{k}_{t-1} + \hat{\rho}_{kt}) - \frac{\bar{v}}{k} \hat{v}_t - \frac{\bar{x}}{k} \hat{x}_t + \frac{\bar{\ell}\bar{w}}{k}(\hat{w}_t + \hat{\ell}_t), \\ \hat{\mu}_t &= (1 - \delta)\hat{\mu}_{t-1} + \delta \eta \hat{x}_t, \\ \frac{1 - \beta(1 - \delta)}{\beta(1 - \delta)} [(1 - \iota_{\epsilon\mu})\hat{\mu}_t - \hat{v}_t + (1 - \eta)\hat{x}_t] &= (\gamma - 1)(\hat{p}_{t+1} - \hat{p}_t) - \gamma(\hat{v}_{t+1} - \hat{v}_t), \\ \hat{w}_t &= \alpha(\hat{k}_t + \hat{\ell}_t) + \hat{z}_t \\ \hat{\rho}_{kt} &= (\alpha - 1)(\hat{k}_t + \hat{\ell}_t) + \hat{z}_t \\ \phi \hat{\ell}_t &= -\gamma \hat{v}_t + \hat{w}_t + (\gamma - 1)\hat{p}_t \end{aligned}$$

where $\hat{p}_t = \epsilon_{p\mu} \hat{\mu}_t$ with $\epsilon_{p\mu} \equiv (dp/d\mu) \times (\mu/p)$, and $\iota_{\epsilon\mu} \equiv (d\epsilon_{p\mu}/d\mu) \times (\mu/\epsilon_{p\mu})$, given by

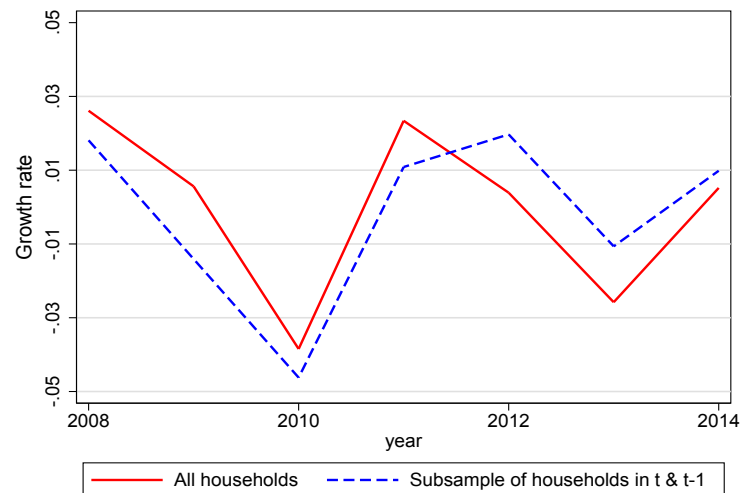
$$\begin{aligned} \epsilon_{p\mu} &= -\frac{1}{\sigma - 1} \left[\frac{h(n^{\frac{\sigma-1}{\alpha}+1})}{h(n^{\frac{\sigma-1}{\alpha}})} - h(n) \right] \\ \iota_{\epsilon\mu} &= -\frac{h(n)}{\epsilon_{p\mu}} \frac{1}{\sigma - 1} \left[\frac{h(n^{\frac{\sigma-1}{\alpha}+2}) h(n^{\frac{\sigma-1}{\alpha}}) - [h(n^{\frac{\sigma-1}{\alpha}+1})]^2}{[h(n^{\frac{\sigma-1}{\alpha}})]^2 h(n)} - 1 \right] \\ h(n^\rho) &\equiv \sum_{n=0}^{\infty} n^\rho f(n; \mu_{-1}) \end{aligned}$$

with $h(n) = \mu$.

D Growth rates in subsample

Figure 9 plots the growth rate in aggregate demand computed using all households in the sample at t (those who report shopping at least once a month) and the growth rate conditional on the subsample of households who buy both in year t and in year $t - 1$. The correlation between the two time series is high, 0.83. The growth rate in aggregate demand

Figure 8: Growth rate of aggregate demand: all households and subsample



Note: The figure plots the growth rate in aggregate demand computed using all households in the sample at t (those who report shopping at least once a month) and the growth rate conditional on the subsample of households who buy both in year t and in year $t - 1$.

when including all households is always greater than the growth rate when looking only at the subset of households that are present in both periods.

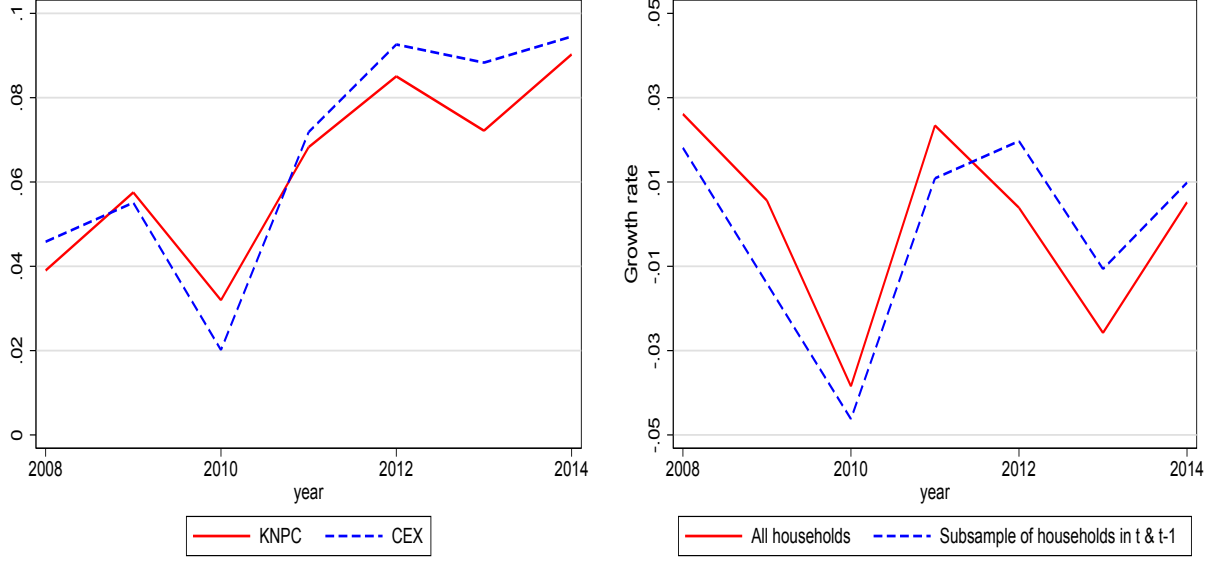
E Data Appendix

E.1 Growth rates in subsample

In Figure ?? we plot the (log of) the US yearly aggregate expenditures per household as computed from aggregating household expenditures in the KNPC (as a red continuous line) and a measure of consumption expenditures per household from the Consumer Expenditure Survey (CEX) maintained by the US Census Bureau (as a blue-dashed line). The CEX series is constructed selecting a series of categories comparable to those tracked in the KNPC data: food at home, alcoholic beverages, tobacco, housekeeping supplies, health and personal care. These categories represent around 13% of total annual expenditure for households in the CEX. Both series are normalized to zero in 2007. Expenditures in the KNPC are aggregated using (using the projection weights provided by Nielsen) using the projection weights provided by Nielsen. Aggregate expenditures per household in CEX and KNPC track each other closely (the correlation between the two series is 0.89). This is in line with the evidence in Kaplan, Mitman, and Violante (2017) who show that the Kilts-Nielsen Retail Scanner Dataset (KNRS), a panel dataset of total sales (quantities and prices) at the UPC (barcode) level for around 40,000 geographically dispersed stores in the US tracks well various definitions of non-durable consumption expenditures in NIPA. Panel (b) of Figure 9 plots the growth rate in expenditures using all households in the sample at t and the growth rate conditional on the subsample of households present in the KNPC both in year t and in year $t - 1$. The correlation between the two time series is high, 0.83. The growth rate in the full sample is generally higher than in the subsample of continuing households, but on average, the two series are strongly correlated.

E.2 Robustness

Figure 9: Consumption expenditures per household: KNPC vs CEX data



(a) Level: KNPC vs CEX

(b) Growth rates: all households and subsample

Notes: Panel (a) plots the yearly aggregate expenditures per household from KNPC (red continuous line) and CEX (blue dashed line); both series are in logs and normalized to zero in 2007. The CEX series is constructed summing expenditures in food at home (mnemonic `cxufoodhomelb0101m`), alcoholic beverages (mnemonic `cxualcbevglb0501m`), tobacco (mnemonic `cxutobaccolb0201m`), drugs (mnemonic `cxudrugslb0501m`), health (mnemonics `cxumedservslb0501m` and `cxumedsuppllb0501m`), housekeeping supplies (mnemonic `cxuhkpgsupplb0201m`), and health and personal care (mnemonic `cxuperscarelb0201m`). Panel (b) plots the growth rate in aggregate demand computed using all households in the sample at t (those who report shopping at least once a month) and the growth rate conditional on the subsample of households who buy both in year t and in year $t - 1$.

F Estimating customer spillovers

Step 1: Varieties creation

We can define the variety creation of household i in product module p in quarter t living in scantrack-market m as equal to the number of varieties purchased in the current quarter which were not purchased in the prior one:

$$c_{ipmt}^n = \sum_{j \in S_p} 1\{e_{ijmt} > 0 \& e_{ijmt-1} = 0\} \quad (58)$$

and analogously for variety destruction

$$d_{ipmt}^n = \sum_{j \in S_p} 1\{e_{ijmt} = 0 \& e_{ijmt-1} > 0\} \quad (59)$$

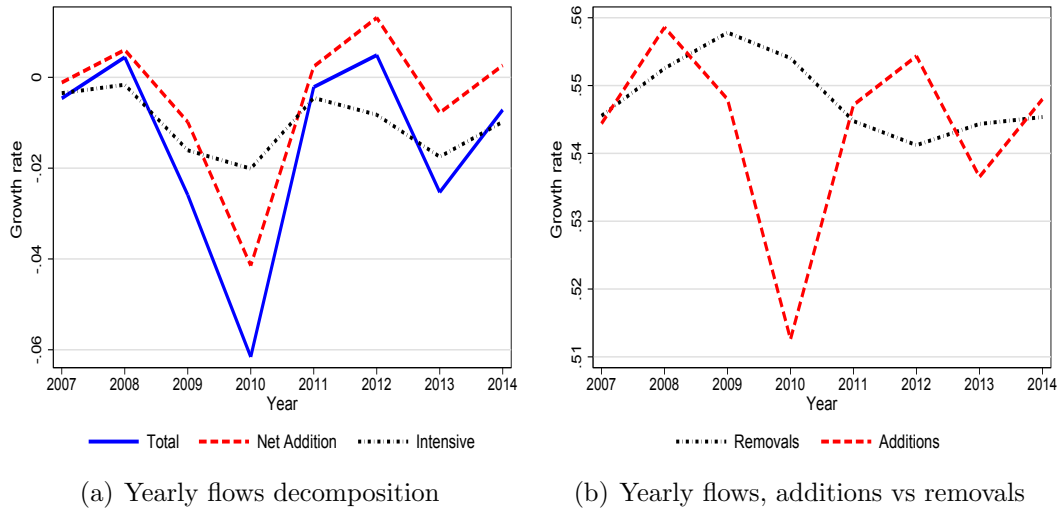
Table 5: Descriptive statistics of yearly consumption flows

	ΔE_t	I_t	N_t	A_t	R_t
A) Constant prices					
Mean	-0.041	-0.026	-0.014	0.320	0.335
Standard deviation	0.034	0.016	0.020	0.016	0.005
Correlation with ΔE_t	1	0.944	0.962	0.990	-0.535
B) UPCs					
Mean	-0.009	-0.009	0.001	0.546	0.545
Standard deviation	0.023	0.006	0.018	0.013	0.008
Correlation with ΔE_t	1	0.836	0.982	0.912	-0.665
C) Storability ≤ 12 months					
Mean	-0.013	-0.011	-0.001	0.278	0.279
Standard deviation	0.023	0.014	0.012	0.009	0.004
Correlation with ΔE_t	1	0.926	0.897	0.947	-0.489
D) Storability > 12 months					
Mean	-0.013	-0.004	-0.009	0.328	0.337
Standard deviation	0.023	0.012	0.014	0.009	0.009
Correlation with ΔE_t	1	0.875	0.894	0.794	-0.449
E) Robust					
Mean	-0.008	-0.145	0.137	0.260	0.123
Standard deviation	0.023	0.015	0.011	0.007	0.004
Correlation with ΔE_t	1	0.939	0.883	0.907	-0.521
F) Persistent					
Mean	0.001	0.137	-0.136	0.126	0.262
Standard deviation	0.022	0.018	0.006	0.004	0.004
Correlation with ΔE_t	1	0.967	0.624	0.711	-0.247
G) Transitory					
Mean	0.001	-0.139	0.140	0.203	0.063
Standard deviation	0.022	0.016	0.009	0.007	0.002
Correlation with ΔE_t	1	0.945	0.798	0.857	-0.257

Notes: A variety is identified by a brand-product module pair.

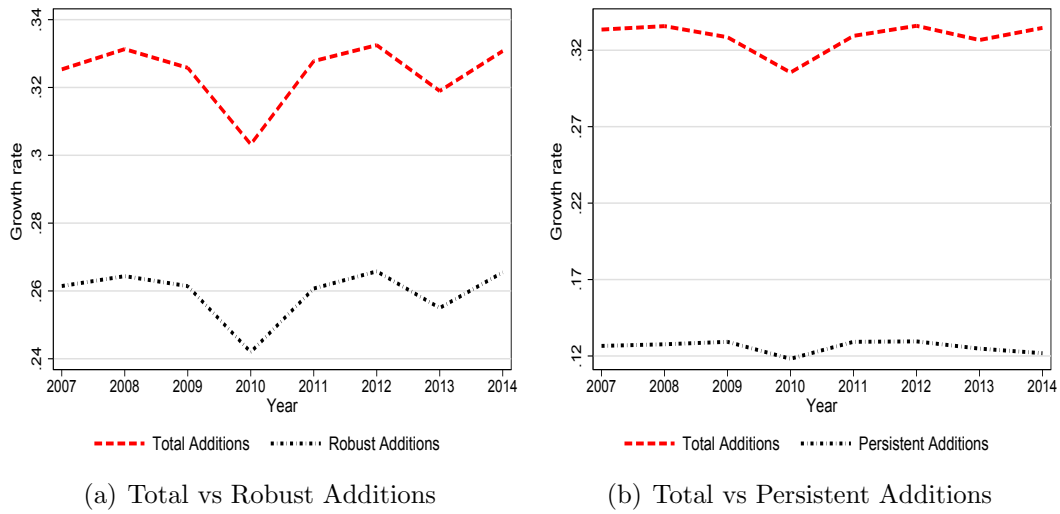
where S_p is the set of all varieties in product module p while e_{ijmt} is the quarter t expenditure (in dollars) by household i living in Scantrack-market m in a variety j . Alternatively to

Figure 10: Aggregate demand decomposition: a product is a UPC



Note: A product is a UPC (upc-vec in NKPC).

Figure 11: Robust and persistent additions



Note: A product is a UPC (upc-vec in NKPC).

measuring varieties creation in units (e.g. number of varieties created), we can do it in value by calculating the expenditure by households in new varieties. This is akin to what we do in the expenditure decomposition exercises. We then define customer creation, customer

destruction, net customer creation and changes in the intensive margin

$$c_{ipmt}^v = \sum_{j \in S_p} e_{ijmt} \cdot 1\{e_{ijmt} > 0 \& e_{ijmt-1} = 0\} \quad (60)$$

$$d_{ipmt}^v = \sum_{j \in S_p} e_{ijmt-1} \cdot 1\{e_{ijmt} = 0 \& e_{ijmt-1} > 0\} \quad (61)$$

$$n_{ipmt}^v = c_{ipmt}^v - d_{ipmt}^v \quad (62)$$

$$i_{ipmt}^v = \sum_{j \in S_p} (e_{ijmt} - e_{ijmt-1}) \cdot 1\{e_{ijmt} > 0 \& e_{ijmt-1} > 0\} \quad (63)$$

Clearly we can also define the change in expenditures

$$\Delta e_{ipmt}^v = n_{ipmt}^v + i_{ipmt}^v \quad (64)$$

where

$$e_{ipmt}^v = \sum_{j \in S_p} e_{ijmt}$$

are the expenditures of household i in all varieties in product module p . Rather than in terms of value we can define the above variables in rates:

$$c_{ipmt}^r = \frac{c_{ipmt}^v}{\sum_{j \in S_p} (e_{ijmt} + e_{ijmt-1})/2} \quad (65)$$

$$d_{ipmt}^r = \frac{d_{ipmt}^v}{\sum_{j \in S_p} (e_{ijmt} + e_{ijmt-1})/2} \quad (66)$$

$$n_{ipmt}^r = \frac{n_{ipmt}^v}{\sum_{j \in S_p} (e_{ijmt} + e_{ijmt-1})/2} \quad (67)$$

$$i_{ipmt}^r = \frac{i_{ipmt}^v}{\sum_{j \in S_p} (e_{ijmt} + e_{ijmt-1})/2} \quad (68)$$

$$e_{ipmt}^r = \frac{\Delta e_{ipmt}^v}{\sum_{j \in S_p} (e_{ijmt} + e_{ijmt-1})/2} \quad (69)$$

All measures can be aggregated at the Scantrack-product module-quarter level, to obtain the aggregate creation, destruction, net creation, intensive margin in the scantrack market in province module. The aggregation can be done for the level and the rates, with weights or without weights. In general we will keep the weights constant (average weights) so that $\bar{w}_{it}$ and all households weights should add-up to one, $\sum_{i \in H_m} \bar{w}_{it} = 1$, where H_m denotes the set of households in the scantrack market m . The variable will then have the interpretation of being per capita. $\bar{w}_{it}$ can be defined using the Nielsen weights (as equal to the average of the weights across the two years to have them constant) or without using any weight (so equal to one over the number of households in the sample in that given quarter in the scantrack market). So we have that the aggregate value in scantrack market

m in product module p are equal to

$$C_{pmt}^v = \sum_{i \in H_m} \bar{w}_{it} c_{ipmt}^v \quad (70)$$

$$D_{pmt}^v = \sum_{i \in H_m} \bar{w}_{it} d_{ipmt}^v \quad (71)$$

$$N_{pmt}^v = \sum_{i \in H_m} \bar{w}_{it} n_{ipmt}^v \quad (72)$$

$$I_{pmt}^v = \sum_{i \in H_m} \bar{w}_{it} i_{ipmt}^v \quad (73)$$

$$E_{pmt}^v - E_{pmt-1}^v = \sum_{i \in H_m} \bar{w}_{it} \Delta e_{ipmt}^v = N_{pmt}^v + I_{pmt}^v \quad (74)$$

and analogously for rates

$$C_{pmt}^r = \frac{\sum_{i \in H_m} \bar{w}_{it} c_{ipmt}^v}{(E_{pmt}^v + E_{pmt-1}^v) / 2} \quad (75)$$

$$D_{pmt}^r = \frac{\sum_{i \in H_m} \bar{w}_{it} d_{ipmt}^v}{(E_{pmt}^v + E_{pmt-1}^v) / 2} \quad (76)$$

$$N_{pmt}^r = \frac{\sum_{i \in H_m} \bar{w}_{it} n_{ipmt}^v}{(E_{pmt}^v + E_{pmt-1}^v) / 2} \quad (77)$$

$$I_{pmt}^r = \frac{\sum_{i \in H_m} \bar{w}_{it} i_{ipmt}^v}{(E_{pmt}^v + E_{pmt-1}^v) / 2} \quad (78)$$

$$\frac{E_{pmt}^v - E_{pmt-1}^v}{(E_{pmt}^v + E_{pmt-1}^v) / 2} = N_{pmt}^r + I_{pmt}^r \quad (79)$$

where

$$E_{pmt}^v = \sum_{i \in H_m} \sum_{j \in S_p} \bar{w}_{it} e_{ijmt} = \sum_{i \in H_m} \bar{w}_{it} e_{ipmt}^v \quad (80)$$

denotes the aggregate expenditures in scantrack market m in product module p at time t whose definition might vary depending on the definition of the weights $\bar{w}_{it}$ that we use. Notice that the data set should contain E_{pmt}^v to aggregate all variables (creation, destruction, net, the intensive margin) across product modules in the same scantrack market.

Step 2: “Stock” of varieties consumed

We define the stock of varieties consumed in product module p in quarter t by household i living in Scantrack market m as the number of varieties consumed in the current quarter:

$$s_{ipmt} = \sum_{j \in S_p} 1\{e_{ijmt} > 0\}$$

where e_{ijmt} is the quarter t expenditure (in dollars) by household i living in Scantrack m in a variety j which belongs to product module p .

We can then define the number of varieties per capita in product module p and scantrack market m as equal to

$$S_{pmt} = \sum_{i \in H_m} \bar{w}_{it} s_{ipmt} \quad (81)$$

We can also define

$$S_{pmt}^v = \frac{\sum_{i \in H_m} \bar{w}_{it} e_{ipmt}^v s_{ipmt}}{E_{pmt}^v} \quad (82)$$

S_{pmt} in (81) and S_{pmt}^v in (82) differ depending on whether they are expenditure weighted: in S_{pmt} we weight the number of varieties consumed by a household just with the Nielsen weights (or the equal households weight equal to one over the number of households in the scantrack market); in S_{pmt}^v we weight households' varieties also using the expenditures of the households. We can also define expenditures per variety as equal to

$$ES_{pmt} = \frac{E_{pmt}^v}{S_{pmt}} \quad (83)$$

which is a measure of the intensive margin. Clearly we have that $E_{pmt}^v = S_{pmt} \times ES_{pmt}$, which allows to decompose aggregate expenditures in an intensive and extensive component.

Step 3: Instrument for the stock of varieties consumed

We will instrument the stock of variety consumed using a strategy reminiscent of the "friends of friends instruments" using the average stock of varieties consumed by migrants to Scantrack-market m from another Scantrack-market m' . Go to <https://www.irs.gov/uac/soi-tax-stats-migration-data> and download the migration pattern county-to-county from 1990 to 2014. Use the county-to-Scantrack mapping to construct a Scantrack-to-Scantrack migration matrix using the county data you downloaded. Each county should be assigned to just one and only one Scantrack. For every Scantrack pair (m, m') we start calculating $M_{my}^{m'}$ equal to the number of migrants from Scantrack m' to Scantrack m in year y as obtained from the IRS dataset. Note that $M_{mt}^{m'} \neq M_{m't}^m$. We could construct three types of instruments: (i) just the number of varieties change over time; (ii) just the migration weights change over time; (iii) both the number of varieties and the migration weights change over time;.

Case 1: Constant migration flows

1. We aggregate the migration flows over the the years 1992-2014 as follows:

$$\omega_m^{m'} = \frac{\sum_y M_{my}^{m'}}{\sum_y \sum_k M_{my}^k}$$

The denominator is equal to the total number of migrants to scantrack market m . Note that $\omega_m^{m'} \neq \omega_{m'}^m$, and that the time unit of reference is the year, y , given that the frequency of the IRS data on migration

2. We calculate

$$\tilde{S}_{pmt} = \sum_l \omega_m^l S_{plt} \quad t = 2007 : 1, \dots, 2014 : 4 \quad (84)$$

and

$$\tilde{S}_{pmt}^v = \sum_l \omega_m^l S_{plt}^v \quad t = 2007 : 1, \dots, 2014 : 4 \quad (85)$$

We also calculate

$$\widetilde{ES}_{pmt} = \sum_l \omega_m^l ES_{plt} \quad t = 2007 : 1, \dots, 2014 : 4$$

which is an instrument for the intensive margin. Notice that the set of instruments could include lags of $\tilde{S}_{pmt}$, $\tilde{S}_{pmt}^v$ or $\widetilde{ES}_{pmt}$. Notice that these instruments vary by quarter within a year.

Case 2: Constant expenditures in product modules

1. For every Scantrack pair (m, m') we calculate the following migration weights:

$$\omega_{my}^{m'} = \frac{M_{my}^{m'}}{\sum_k M_{my}^k} \quad y = 2007, \dots, 2014$$

where $M_{my}^{m'}$ is the number of migrants from Scantrack m' to Scantrack m computed in year y in the IRS dataset.

2. We then calculate

$$\bar{S}_{pmy} = \sum_l \omega_{my}^l \left(\frac{1}{4} \sum_{t \in y} S_{plt} \right) \quad y = 2007, \dots, 2014 \quad (86)$$

and

$$\bar{S}_{pmy}^v = \sum_l \omega_{my}^l \left(\frac{1}{4} \sum_{t \in y} S_{plt}^v \right) \quad y = 2007, \dots, 2014 \quad (87)$$

$t \in y$ denotes the set of quarters which belongs to year y . We also calculate

$$\overline{ES}_{pmy} = \sum_l \omega_m^l \left(\frac{1}{4} \sum_{t \in y} ES_{plt} \right) \quad y = 2007, \dots, 2014$$

which is an instrument for the intensive margin. Notice that the set of instruments could include lags of $\bar{S}_{pmt}$, $\bar{S}_{pmt}^v$ or $\overline{ES}_{pmt}$. Notice that this instrument varies by year not by quarter within a year.

Case 3: Both migration and number of varieties vary

1. For every Scantrack pair (m, m') we calculate again

$$\omega_{my}^{m'} = \frac{M_{my}^{m'}}{\sum_k M_{my}^k} \quad y = 2007, \dots, 2014$$

where $M_{my}^{m'}$ is the number of migrants from Scantrack m' to Scantrack m computed in year y in the IRS dataset. Note that $\omega_m^{m'} \neq \omega_{m'}^m$.

2. Let $y(t)$ denote the calendar year where t belongs to. We calculate:

$$\hat{S}_{pmt} = \sum_l \omega_{my(t)}^l S_{plt} \quad t = 2007 : 1, \dots, 2014 : 4$$

and

$$\hat{S}_{pmt}^v = \sum_l \omega_{my(t)}^l S_{plt}^v \quad t = 2007 : 1, \dots, 2014 : 4 \quad (88)$$

We also calculate

$$\widehat{ES}_{pmt} = \sum_l \omega_{my(t)}^l ES_{plt} \quad t = 2007 : 1, \dots, 2014 : 4$$

Notice that the set of instruments could include lags of $\hat{S}_{pmt}$, $\hat{S}_{pmt}^v$ or $\widehat{ES}_{pmt}$. Notice that these instruments vary by quarter within a year.

Step 4: Constructing the database

In the final database the level of observation will be a Scantrack-product-module-quarter triplet. The information on each row of the database should be the following:

- Scantrack
- Year
- Quarter
- Product Module
- Market/Product-module expenditure, current: E_{pmt}^v
- Creation in value and in rate: C_{pmt}^v , C_{pmt}^r
- Destruction in value and in rate: D_{pmt}^v , D_{pmt}^r
- Intensive margin in value and in rate: I_{pmt}^v , I_{pmt}^r
- Net creation in value and in rate: N_{pmt}^v , N_{pmt}^r
- Stock of varieties consumed: S_{pmt} , S_{pmt}^v

- Expenditure per variety: ES_{pmt}
- Instruments case 1: $\tilde{S}_{pmt}, \tilde{S}_{pmt}^v, \widetilde{ES}_{pmt}^v$
- Instruments case 2: $\bar{S}_{pmt}, \bar{S}_{pmt}^v, \overline{ES}_{pmt}^v$ (notice that these come at yearly frequency. repeat same item per each quarter of the year)
- Instruments case 3: $\hat{S}_{pmt}, \hat{S}_{pmt}^v, \widehat{ES}_{pmt}^v$
- All these variables should appear twice depending on whether we use Nielsen weights $((w_{it} + w_{it})/2)$ or equal weights (equal to one over the number of households in the sample in that year)