Entry and Profits in an Aging Economy: The Role of Consumer Inertia

Gideon Bornstein*
Princeton University & The Wharton School

February, 2020

Abstract

Over the past thirty years, the share of young firms in the US has declined while the share of profits in GDP has increased. This paper explores the role of *consumer inertia*—persistence in households' consumption choices—as a driver of these twin phenomena. The hypothesis is that more consumer inertia makes it more difficult for entrants to establish a customer base, incentivizes large incumbents to raise markups, and shifts production toward large-high-markup firms. Using micro-data on consumer behavior, I estimate the degree of consumer inertia across different age groups and find that young households are significantly less inertial. I develop and estimate a model of firm dynamics with consumer inertia. Through the lens of the model, the aging-induced rise in consumer inertia accounts for about 30% of the twin phenomena. Reduced-form evidence exploiting variation across states and across product categories are consistent with the model's prediction.

*This paper is based on the first chapter of my PhD dissertation. I am indebted to Marty Eichenbaum, Guido Lorenzoni, Sergio Rebelo, Luigi Bocola, and Per Krusell for continuous guidance and support. I thank Sara Moreira for her advice and for sharing with me some of her data codes. I am grateful to David Berger, Larry Christiano, Dean Corbae, Gaby Cugat, Sasha Indarte, Vishal Kamat, Chiara Maggi, Marti Mestieri, Abdoulaye Ndiaye, Jordan Norris, Giorgio Primiceri, Matt Rognlie, Max Tabord-Meehan, Fabrice Tourre, and Arlene Wong for numerous discussions and comments. Researcher own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. All errors are my own. Email: bornstein@princeton.edu.

1 Introduction

Over the past three decades, the US corporate sector has become older and more profitable. The share of young businesses in the economy, defined as businesses which are 5 years old or younger, has declined from 50% in the late 1980s to 30% today. The share of workers employed by young businesses fell from 20% to 10%.[1] Strikingly, over the same period, the share of profits in GDP rose substantially.[2] It is challenging for simple models of firm dynamics to account for the twin phenomena of declining entry and rising profits. Other things equal, higher profits should stimulate firm entry.

In this paper, I study the role of consumer inertia as a driver of these twin phenomena. By consumer inertia, I mean the tendency of consumers to choose the same products over time for reasons other than the fundamental attributes of those products. i.e., not switching a product produced by one firm to one produced by a different firm. The prevalence of consumer inertia has been extensively documented in the industrial organization and marketing literatures.[3] I argue that consumer inertia has increased in the past three decades due to the aging of the US population. I show that the aging-induced rise in consumer inertia accounts for a substantial proportion of the twin phenomena.

The hypothesis is that a rise in consumer inertia leads large incumbents to raise their markups and profits while discouraging entry. I develop my argument in four steps. First, I present a static model that illustrates the logic of the hypothesis. Second, I use detailed micro data to show that younger households display considerably less consumer inertia than older ones. According to my estimates, young households are about 50% more likely to switch their consumption products. Building on this result, I argue that population aging in the US has led to a rise in aggregate consumer inertia. Third, I provide empirical evidence on a causal negative link between consumer inertia and firm formation. I find that product categories with a relatively more inertial consumer base display lower entry rates. In addition, I find that US states which experienced a larger increase in consumer inertia have also undergone a larger decline in the share of young firms. Finally, I develop a quantitative model of entry, exit, and firm dynamics with consumer inertia. I calibrate the model using my estimates of consumer inertia together with data on firm dynamics. The model predicts that the aging-induced rise in consumer inertia accounts for 30%–50% of the twin phenomena between the late 1980s and early 2000s.

Consumer inertia can reflect a variety of forces. The marketing literature tends to stress psy-

[1] See Decker, Haltiwanger, Jarmin and Miranda (2014a), Decker et al. (2014b), Hathaway and Litan (2014), and Pugsley and Şahin (2015).

[2] See Barkai (2016), Gutierrez (2017), and Barkai and Benzell (2018).

[3] Some examples include Luco (2017) and Illanes (2016), which document consumer inertia in the market for pension plans, Handel (2013) and Nosal (2012), which document it in the market for health insurance plans, Anderson, Kellogg, Langer and Sallee (2015), which documents it in the automobile market, Hortaçsu, Madanizadeh and Puller (2015), which documents it in the residential electricity markets, as well as Dubé, Hitsch and Rossi (2010) and Bronnenberg, Dubé and Gentzkow (2012), which document it for consumer packaged goods.

chological and emotional factors, as well as inattention. The industrial organization literature tends to emphasize learning costs, contractual obligations, transaction costs, incomplete information, and search frictions. In my model, I capture consumer inertia in a parsimonious way. Consumers are confronted with a variety of product types (e.g., phones). Each product type consists of a variety of products (e.g., iPhone, Galaxy, Nexus, etc.). In every period, each household consumes a single product from each product type. Households have idiosyncratic preferences over the different products. So when choosing which product to consume, a household takes into account its preferences and the price of the product. Within every product type, the household is locked into its previous choice with some exogenous probability. This simplification allows me to derive an analytic expression for the demand function faced by each firm.

I first study a static economy with two types of firms, incumbents and entrants. There is a measure one of incumbents and an endogenous measure of entrants. The only difference between incumbents and entrants is that incumbents have an initial customer base, i.e., the measure of households who purchased their product in the previous period. Absent consumer inertia, incumbents and entrants behave identically. In the presence of consumer inertia, a fraction of each incumbent's customers will continue to purchase their product regardless of price. Of course, the quantity that they purchase will depend on the price. It follows that the price elasticity of locked-in customers is lower than the price elasticity of unattached customers. Therefore, in equilibrium, incumbents choose a higher price and have higher markups than entrants.

I use the static model to analyze the effects of more consumer inertia. The key result is that more consumer inertia leads to higher aggregate profits, higher incumbent markups, and less entry. So the model qualitatively accounts for the twin phenomena. An increase in consumer inertia reduces the potential customer base of entrants. As a result, the measure of entrants in equilibrium falls. At the same time, a larger share of incumbents' customers are locked-in. Incumbents take advantage of the increasing measure of locked-in customers by raising markups and increasing their profits. So, in addition to the increase in aggregate profits, average markups in the economy go up as well. This result is consistent with a growing body of evidence that the average markup has been increasing over time (e.g., De Loecker and Eeckhout, 2017; Hall, 2018).

In this paper, I focus on changes in consumer inertia arising from changing demographics. To study the level of consumer inertia and how it depends on household demographics, I use Nielsen's Consumer Panel dataset. The dataset includes longitudinal panel information on the purchases of approximately 160,000 households in the US between 2004 and 2015. The panel dimension of the dataset is crucial for identifying the degree of consumer inertia. This dataset allows me to follow the consumption behavior of a household over time, and study how often the household decides to switch products.

A household can choose to consume the same product over time for two reasons: either because of inertia, or because it has a persistent taste for the fundamental attributes of that product. The identification challenge is to distinguish between these two reasons. This empirical challenge fits into a broader class of problems aimed at distinguishing between structural and spurious state-dependence (Heckman, 1981). This distinction matters for the behavior of firms in the economy. If consumers have persistent preferences toward the fundamental attributes of a product, an entrant can introduce a similar product and attract these consumers. On the other hand, if consumers display structural consumer inertia, there is less scope for entry.

To solve the identification problem, I study the behavior of households who moved to a different state. Bronnenberg, Dubé and Gentzkow (2012) documents that households who move between states eventually converge to the typical consumption basket of the state they moved to. I follow their identification strategy, and measure the level of consumer inertia by the speed with which movers converge to the consumption basket of their new state.

In my empirical analysis I group households into four groups, according to the age of the household head: 20–34, 35–49, 50–64, and older than 64. I find that younger households display significantly less consumer inertia than older households. According to my estimates, it takes about 5.5 years for a young household (20–34) to close 50% of the gap between its initial consumption basket and the typical basket of the state to which they moved. In contrast, it takes households in other age groups between 7.5 to 8.5 years to close 50% of the gap. These results are qualitatively consistent with the findings of Bronnenberg et al. (2012).

Having documented that younger households are less inertial, I then study the effects of consumer inertia on firm formation. I do that in two ways. First, I study how entry rates vary across product categories. I construct a measure of consumer inertia for each product category based on the age composition of its customers. I find that there is a significant negative relation between the level of consumer inertia and the firm entry rate into a product category. Other things equal, product categories with a lower level of consumer inertia, i.e., with a relatively younger customer base, have higher entry rates.

Second, I perform a similar analysis at the state level, using annual data on the age composition of each state. I find that an increase in consumer inertia at the state level leads to an economically significant decline in firm formation. My regression specifications account for potential omitted variable bias and use an instrumental variable to identify a causal relationship.

I think of my micro estimates as an upper bound on the impact of consumer inertia on firm formation for the US as a whole. Other things equal, firms will choose to start their business in states with less consumer inertia, because it is easier to penetrate the market in such states. However, this inter-state margin sums to zero, by definition, for the US as a whole. A naive extrapolation of my estimates implies that the rise in aggregate consumer inertia accounts for 70%–80% of the decline in firm formation measures between the late 1980s and early 2000s.

The limitations of the micro estimates prompt me to study the aggregate effects of the rise in consumer inertia using a general equilibrium structural model. I construct a model of entry, exit, and firm dynamics in the presence of consumer inertia. Consumer inertia gives rise to life-cycle dynamics in a firm's customer base and markups. A new firm has no initial customer base, and builds it over time. The firm's optimal markup reflects both a harvesting motive and an investment motive.[4] The harvesting motive refers to firms' incentive to take advantage of their customer base by increasing markups. The investment motive refers to firms' incentive to increase their customer base by lowering markups. The firm's optimal markup reflects the balance between the harvesting and investment motives.

The balance between the harvesting and investment motives varies across the life-cycle of the firm. When a firm joins the economy, it has no initial customer base. As a result, the harvesting motive is muted and the investment motive induces the firm to set a relatively low markup. As a firm builds its customer base over its life-cycle, the harvesting motive becomes more dominant and the investment motive becomes weaker. The investment motive weakens because lowering markups in order to attract new customers reduces the profits coming from locked-in customers. So the optimal markup of a firm is increasing in the customer base it carries from the previous period.

To model firm entry, I assume that a firm must pay a fixed cost in order to begin operation. The measure of entrants in equilibrium is determined by a zero-profit condition. In each period, firms are subject to persistent idiosyncratic productivity shocks. To model firm exit, I assume that firms need to pay a stochastic fixed operating cost in every period. It may be optimal for them to exit, rather than pay this cost. In equilibrium, firms differ both with respect to their customer base and their productivity level.

I calibrate the model's structural parameters to match key features of the US economy in the late 1980s. The level of consumer inertia is calibrated using my micro estimates of consumer inertia, along with data on the age composition of the US population. The other structural parameters are calibrated to match the age distribution of firms, the relative size of firms of different ages, and the share of total profits in GDP.

To assess the quantitative impact of the rise in consumer inertia, I consider an unexpected deterministic shock that moves the level of consumer inertia by an amount implied by the change in the age composition of the US population between the late 1980s and 2050. I then study the transition dynamics of the economy from the initial stationary equilibrium to the new one.

An increase in the level of consumer inertia strengthens both the harvesting motive and the investment motive. As a result, firms with a relatively small customer base, particularly entrants, choose to reduce their markups. On the other hand, firms with a relatively large customer base, for which the harvesting motive dominates the investment motive, choose to raise their markups. So more consumer inertia increases markup dispersion between small and large firms in the economy. Firms become less profitable at the beginning of their life-cycle, and more profitable later in their life-cycle. The changes in profitability across the life-cycle of a firm affect the present value of entrants. Under the calibrated structural parameters I consider, holding fixed the measure of operating firms, a rise in consumer inertia reduces the present value of entrants. As a result, the

[4]These terms go back to Klemperer (1987b), which studies a two period model with switching costs.

measure of entrants in equilibrium falls. In the transition dynamics following a rise in consumer inertia, the rate of entrants in the economy declines.

My key results are as follows. First, the model captures qualitatively the twin phenomena. Second, the model accounts for 40%–50% of the decline in the share of young firms and 30% of the rise in the aggregate profits share between the late 1980s and early 2000s.

Literature. This paper is related to several strands of literature. First, it is related to a recent literature which studies the causes for the decline in business dynamism. One explanation for this decline is changes in the regulatory environment. Davis and Haltiwanger (2014) provide suggestive evidence for this channel. Another possible explanation is skill-biased technical change. Jiang and Sohail (2017), Kozeniauskas (2017), and Salgado (2017) study this channel, and show that the decline in entrepreneurship is more pronounced for highly educated individuals. This channel on its own, however, can only account for the difference in entrepreneurial trends between skilled and unskilled individuals. It cannot account for the overall decline in entrepreneurship rates. Finally, several authors have pointed to demographic changes as drivers of the decline in firm formation. Karahan et al. (2016) argues that a decline in the growth rate of the labor force growth has lead to a decline in the firm entry rate. Liang et al. (2014) and Engbom (2017) study the role of aging workforce as a driver of the decline in firm formation. My paper also focuses on population aging as a driver of the decline in firm formation. But instead of focusing on supply-side channels, I study how population aging affects firm formation through its effect on the demand-side of the economy. In contrast to these papers, I study the implications for profits and find that a rise in consumer inertia is consistent with an increase in incumbents market power and a rise in aggregate profits. Another difference between my paper and the papers in this literature is that the channel I explore should be more prominent for firms whose products are sold to consumers. Decker et al. (2016) documents that the decline in business dynamism in the past three decades has been most prominent for the retail trade, services, and manufacturing sectors.

Second, my paper is related to an emerging literature that studies the reasons for the rise in corporate profits.[5] Autor, Dorn, Katz, Patterson and Van Reenen (2017) argue that the rise of profits is an efficient outcome, reflecting the increasing importance of 'superstar' firms. Grullon, Larkin and Michaely (2017), on the other hand, argue that the rise in profits is an inefficient outcome, resulting from a rise in concentration. Gutierrez and Philippon (2017) provide evidence to support the latter explanation. My model also suggests that the rise in profits is associated with a rise in concentration. The reason is that more consumer inertia leads to a decline in entry rates and incentivizes incumbents to raise their markups.

Another related paper is Neiman and Vavra (2018), which documents that household spending

[5]The magnitude of the rise in profits during the past three decades is the subject of an ongoing debate. See Karabarbounis and Neiman (2018) and the discussion of Rognlie (2018) for the challenges in computing the share of economic profits in GDP. Barkai and Benzell (2018) argue that the rise in profits is robust even when addressing these challenges.

has become more concentrated between 2004 to 2015. Moreover, households with more concentrated product spending also pay more for the products that they purchase. Consistent with these findings, my model implies a rise in spending concentration and a rise in markups. It is also consistent with their finding that spending concentration is distributed across incumbents, and not focused around a few 'superstar' firms. In contrast to their paper, I identify population aging as a driver for the rise in consumer inertia and examine its aggregate implications using a structural model.

My paper is also related to a vast theoretical literature in industrial organization that studies how switching costs affect the pricing decision of firms. For a review of this literature, see Farrell and Klemperer (2007). Several papers in the literature have studied how switching costs affect firm entry and profits (e.g., Klemperer, 1987a; Farrell and Shapiro, 1988; Gabszewicz et al., 1992). These papers restrict attention to duopoly markets and are partial equilibrium in nature. My analysis embodies the channels they stress but focuses on a general equilibrium analysis with a large number of atomistic firms.

Different forms of consumer inertia have also been studied in a macroeconomic context. Early examples include Winter and Phelps (1970) and Rotemberg and Woodford (1991). More recently, Ravn, Schmitt-Grohé and Uribe (2006) studies a dynamic stochastic general equilibrium model where the demand function faced by each firm features inertia, the *deep habits* model. Nakamura and Steinsson (2011) uses a version of this model to micro-found a firm's decision to choose a 'sticky price'. Gilchrist, Schoenle, Sim and Zakrajšek (2017) considers a version of the deep habits model with financial frictions, and show that it can account for inflation dynamics during the financial crisis. Gourio and Rudanko (2014) studies a firm dynamics model where the matching between customers and firms is subject to search frictions. These frictions give rise to consumer inertia. I develop a different approach designed in such a way that the parameter which governs the level of consumer inertia can be mapped to the micro data. The main advantage of my framework is that it allows me to calibrate the model using my micro-based estimates of consumer inertia.

Layout. The paper proceeds as follows. Section 2 presents a static model of consumer inertia. In Section 3, I estimate the degree of consumer inertia for different age groups. In the end of that section, I present the empirical evidence linking consumer inertia and firm formation. Section 4 starts by presenting the structural quantitative model. I then describe its estimation and quantify the contribution of the rise in consumer inertia to the change in profitability and firm formation measures. Section 5 concludes.

2 A Static Model of Consumer Inertia

I start by analyzing a static economy in which households display consumer inertia. There are two types of firms, incumbents and entrants. Incumbents start the period with an existing market share, while entrants have no market share upon entry. Firms cannot price discriminate, i.e., they must charge the same price to all their customers.

The static model serves two purposes. First, it allows me to introduce the way I model consumer inertia in a simple framework. I take a similar modeling approach when estimating consumer inertia using the micro data, and when incorporating consumer inertia in a quantitative model of entry, exit, and firm dynamics. Second, it provides intuition for how consumer inertia alters the behavior of firms.

Consumer inertia can reflect a variety of forces. Examples for such forces include psychological and emotional costs, learning costs, contractual obligations, transaction costs, incomplete information, or search frictions. I model consumer inertia in a parsimonious way. There is a continuum of products, and each household buys only a single product. There is a constant probability a household cannot freely choose what product to buy and is locked into the product it purchased in the previous period.[6]

The main result of the static model is that more consumer inertia leads to higher aggregate profits, higher incumbent markups, and less entry. The intuition is as follows. Higher consumer inertia leads to a fall in the measure of entrants' potential customers. As a result, the measure of entrants in equilibrium goes down. In addition, the fraction of locked-in customers of incumbents goes up. The elasticity of demand of locked-in customers is lower than the demand elasticity of customers who can re-optimize their choice. So the optimal markup set by incumbents is increasing in the level of consumer inertia.

2.1 Demand

The economy is populated by a measure one of ex-ante identical households. Each household consumes a single product out of a variety of products indexed by $j \in (0, J)$, where J is the measure of products in the economy. The utility of household i from consuming product j is given by

$$U_i = u\left(\exp\left(\frac{1}{\sigma - 1}\epsilon_j\right) c_j\right) \ , \tag{1}$$

where $u(\cdot)$ is a strictly increasing function, c_j is the quantity consumed of product j, and ϵ_{j_i} is the idiosyncratic taste of household i for product j.[7] Taste shocks follow a Gumbel distribution and are drawn independently across households and products. As I show below, the parameter σ governs the price elasticity of demand in the demand function faced by the firm.

Households' consumption decision features consumer inertia. There is a probability θ that a household can choose freely which product to buy. With probability $1 - \theta$, the household is locked

[6] In appendix section B.1, I discuss in details how this modeling assumption relates to other structural models of consumer inertia in the literature.

[7] To dispense with notation, I omitted the i superscript in all variables in equation (1) as well as in all subsequent equations. Note that due to different taste shocks, individual households do differ in their consumption behavior ex-post.

into the product it purchased in the previous period. i.e., it must consume the same product it has bought in the previous period. I assume that the previous product consumed by households is uniformly distributed across incumbents. The re-optimization draw is independent across households.

The household's problem can be written as follows,

$$U_i = \max_{\{j,c_j\}} u\left(\exp\left(\frac{1}{\sigma-1}\epsilon_j\right)c_j\right) ,$$
$$\text{s.t.} \qquad p_j c_j \leq w ,$$
$$j = j_i^0 , \quad \text{if } \xi_i = 0 ,$$

where the first constraint is the budget constraint and the second constraint is the consumer inertia constraint. The household's sole income is its salary. It supplies a unit of labor and receives a wage w. The household needs to choose which product to purchase (j), and how much of it to acquire (c_j). ξ^i is an indicator variable that takes the value 0 if the household cannot re-optimize its product choice and is locked into its previous purchased product. This product is denoted by j_i^0. The probability that a household can re-optimize its choice, $\xi_i = 1$, is equal to θ. I refer to households who can re-optimize their product choice as *unattached*.

Before turning to the supply side of the economy, it is useful to derive the demand function faced by the firm. Each firm in the economy produces a single product. To achieve an analytic expression for the demand function faced by firms, I need to derive the optimal consumption decision of households. I start by showing what is the product chosen by an unattached household, given the distribution of prices and taste shocks.

Lemma 1. If the household is unattached, it chooses the product that maximizes $-(\sigma-1)\ln p_j + \epsilon_j$ for $j \in (0, J)$. That is,

$$j_i = \arg\max_{j\in(0,J)} -(\sigma-1)\ln p_j + \epsilon_j \; . \tag{2}$$

The proof of Lemma 1, as well as of all other proofs in the paper, is presented in Appendix A. The intuition of the result above is straightforward. When a household compares between two different products, it values a lower price, as it allows it to buy a larger quantity of the good, and higher taste shock, as it lets it derive higher utility from each good consumed. As taste shocks are assumed to be drawn from a Gumbel distribution, Lemma 1 shows that the consumption choice of *unattached households* follows a multinomial logit.

Using Lemma 1 and the distributional assumption of the taste shocks, I can aggregate the optimal consumption decision of individual households to arrive at the customer-base function faced by the firm. The following proposition presents this result, characterizing the customer base of a firm as a function of its relative price and initial customer base. I denote the customer base of a firm, i.e., the measure of households who purchase a product from the firm, as B, and the initial customer base of a firm by B_0.

Proposition 1. The customer base of a firm with price p is given by

$$B = (1-\theta) B_0 + \frac{\theta}{J} \left(\frac{p}{P} \right)^{1-\sigma} , \tag{3}$$

where B_0 is the initial customer base of the firm, and P is the price index, given by

$$P = \left[\frac{1}{J} \int_0^J p_j^{1-\sigma} dj \right]^{\frac{1}{1-\sigma}} . \tag{4}$$

The first term in equation (3) consists of the locked-in customers, those who were customers of the firm in the previous period and cannot re-optimize. For incumbents, the measure of locked-in customers is equal to $1-\theta$, as their initial customer base is equal to one. Entrants do not have any locked-in customers.

The second term of equation (3) consists of the measure of households who actively choose to buy the firm's good. This term is increasing in the measure of unattached households, θ, decreasing in the measure of operating firms, J, and decreasing in the firm's relative price. Furthermore, it is isoelastic in the firm's relative price. The aggregation of individual multinomial logit consumption decision into an isoelastic demand function faced by the firm goes back to Anderson et al. (1987). Differently from their result, consumer inertia implies that the demand function faced by firms also includes a relatively inelastic term coming from the firm's locked-in customers.

Each household spends w on the product it consumes. As a result, the total revenues of a firm is given by wB. From the firm's perspective, the demand for its good is given by $y = \frac{wB}{p}$. Using the law of motion for a firm's customer base, we have that the demand function faced by the firm is given by

$$y = \left[(1-\theta) B_0 + \frac{\theta}{J} \left(\frac{p}{P} \right)^{1-\sigma} \right] \frac{w}{p} . \tag{5}$$

2.2 Supply

Each firm in the economy sells a single product, and the measure of firms is equal to J. A measure one of these firms are incumbents and an endogenous measure E are entrants, so that $1 + E = J$. The only difference between incumbents and entrants is their initial customer base. An entrant joins the economy with no initial customer base, while incumbents' initial customer base is equal to one. As mentioned above, a fraction $1-\theta$ of an incumbent's initial customer base is locked-in.

Firms use a linear technology in labor to produce their goods. The firm's problem is given by,

$$\begin{aligned} V(B_0) = \max_{\{p,y,l\}} \quad & py - wl , \\ \text{s.t.} \quad & y = (1-\theta) B_0 w p^{-1} + \frac{\theta}{J} P^{\sigma-1} w p^{-\sigma} , \\ & y = l , \end{aligned}$$

where $V(B)$ is the operating profits of a firm with initial customer base B_0. p is the price the firm sets, y is the quantity sold, l is the labor used for production, and P is the price index defined in equation (4). The first constraint is the demand function faced by the firm, and the second constraint is the firm's technology.

In addition to production costs, entrants need to hire f_e units of labor to pay the fixed cost of entry. So the profits of an entrant, net of the fixed entry costs, are given by $V(0) - f_e w$. In equilibrium, a free-entry condition holds so that $V(0) = f_e w$, if the measure of entrants is positive.

Solving the firm's problem, I obtain an equation that implicitly defines the firm's optimal markup.

Proposition 2. The markup charged by the firm, $\mu \equiv p/w$, satisfies the following equation,

$$\mu = \frac{\sigma}{\sigma-1} + B_0 \frac{1}{\sigma-1} \frac{1-\theta}{\theta} J \left(\frac{P}{w}\right)^{1-\sigma} \mu^{\sigma-1} . \tag{6}$$

I assume that $\sigma < 2$, a sufficient condition which ensures that equation (6) uniquely defines the firm's markup as a function of P and J.[8]

Proposition 2 reveals how the firm's markup is set. The demand function faced by the firm consists of two terms: a relatively inelastic one coming from its locked-in customers, and a more elastic one coming from customers who actively choose the firm's product. In its choice of markup, the firm is balancing between these two opposing forces. If the firm only has locked-in customers, it has an incentive to raise its markup as high as possible. While if the firm only has unattached customers, it has an incentive to lower its markup to $\frac{\sigma}{\sigma-1}$. The latter is the markup set by entrants, as those do not have any locked-in customers.

Incumbents markup is increasing in the level of consumer inertia, $1-\theta$, as higher consumer inertia implies a higher share of locked-in customers and a lower share of unattached customers. Incumbents markup is also increasing in the measure of operating firms. The higher is the measure of operating firms, the harder it is for an incumbent to attract unattached customers. As a result, it chooses to increase its markup and harvest the benefits of having a locked-in customer base. Finally, incumbents markup is decreasing in the price index P. Similar to the measure of operating firms, a lower price index makes it harder for incumbents to attract unattached customers and incentivizes them to increase their markup.

2.3 Equilibrium and Comparative Statics

Without loss of generality, I choose the wage as the numeraire and set it to 1. A firm's price is therefore equal to its markup. I denote the price charged by entrants as μ_E and that of incumbents by μ_I.

Definition 1 (Equilibrium)**.** An equilibrium is a set of prices $\{\mu_E, \mu_I, P\}$, consumption policy functions, and a measure of operating firms J such that: (i) consumption policy functions solve the

[8] To see that equation (6) uniquely pins down the level of markup, note that the LHS is linear and increasing in μ and that the RHS is concave and increasing in μ. When $\mu = 0$ the RHS is greater than the LHS, so there exists a unique value of μ which solves the equation.

household's problem, (ii) prices solve the firms' problems, (iii) free entry condition holds, and (iv) the price index P satisfies equation (4).

As the households' policy functions take an explicit form derived in section 2.1, finding an equilibrium boils down to finding a set of four endogenous variables, $\{\mu_E, \mu_I, P, J\}$, such that: (i) P satisfies equation (4), (ii) μ_E solves equation (6) with $B_0 = 0$, (iii) μ_I solves that equation with $B_0 = 1$, and (iv) free entry condition holds.[9]

Proposition 3. There exists a unique equilibrium. In addition, there is a cutoff $\bar{f}_e$ such that if and only if $f_e < \bar{f}_e$ then the measure of entrants is strictly positive.

Having established that an equilibrium exists and is unique, I can study how it depends on the different structural parameters. In particular, I study how the measure of entrants, the markup charged by entrants and incumbents, and the share of profits in the economy, depend on the level of consumer inertia, $1 - \theta$, and on the fixed entry costs, f_e.[10]

Proposition 4. [Comparative statics] Consider two economies, A and B. Suppose the two economies differ by only one structural parameter. The following comparative statics hold:

Parameter	Entrants' markup	Incumbents' markup	Profits share	Entry rate
$\theta_A < \theta_B$	$\mu_E^A = \mu_E^B$	$\mu_I^A > \mu_I^B$	$\Pi_A > \Pi_B$	$E_A < E_B$
$f_e^A > f_e^B$	$\mu_E^A = \mu_E^B$	$\mu_I^A < \mu_I^B$	$\Pi_A > \Pi_B$	$E_A < E_B$

Proposition 4 presents the main result of this section. An increase in the level of consumer inertia (a lower θ) leads to an increase in incumbents markup, an increase in profitability, and a decline in the entry rate. Incumbents choose to increase their markups as the share of locked-in customers relative to unattached customers increases. This results in higher profits for incumbents. Entrants charge the same markup but their customer base shrinks. As a result, if the measure of firms hadn't changed they would make lower profits. The measure of entrants in equilibrium goes down so that the free-entry condition holds.

The proposition above also sheds light on the difference between an increase in entry costs and an increase in consumer inertia. Under both cases, the share of profits in the economy goes up and the entry rate goes down. However, when entry costs increase, incumbents markup goes down. The increase in entry costs leads to a lower measure of entrants. As a result, it is easier for incumbents to attract new customers. Incumbents set a lower markup, as a larger share of their customer base are unattached customers. Since the environment in which incumbents operate is less competitive, despite setting a lower markup, incumbents profits go up.

[9] If the set of parameters that solves these four conditions yields $J < 1$, then the measure of entrants in equilibrium is equal to zero. In that case, we have that $J = 1$, $P = \mu_I$, and μ_I is set so that it solves equation (6) with $B_0 = 1$.

[10] I denote the share of profits in equilibrium by Π. In equilibrium, the share of profits in the economy is equal to the profits earned by incumbents. That is, $\Pi = V(1)$. Entrants profits are equal to the entry cost paid in order to operate, as implied by the free-entry condition.

3 Empirical Analysis

In this section, I estimate the degree of consumer inertia for households of different ages using Nielsen's Consumer Panel dataset. To identify the level of consumer inertia, I study the consumption behavior of households who move between states. This identification strategy was originally proposed by Bronnenberg et al. (2012). I find that young households, defined as households in which the head of the family is younger than 35 years old, display significantly less consumer inertia in comparison to older households. I then exploit two types of variations in the data to study the causal relation between consumer inertia and firm formation. First, I show that product categories in which consumers are more inertial have lower firm entry rates. Second, I show that US states who experienced a larger increase in consumer inertia had larger declines in different firm formation indicators.

3.1 Data

The main dataset I use for the analysis is Nielsen Homescan Consumer Panel Data provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The dataset includes longitudinal panel information on the purchases of approximately 160 thousand households in the US between 2004 and 2015. Households receive a barcode scanner from Nielsen and are directed to scan the barcodes of all consumer packaged goods they purchase. The dataset covers both food and non-food items across all retail outlets in the US.[11]

For each shopping trip of a household, I observe the date of the shopping trip, the products purchased at the Universal Product Code (UPC) level, and the stores in which these goods were purchased. An example of a product at the UPC level is a 16 fl oz plastic bottle of Coca-Cola. Nielsen organizes products at the UPC level into product modules, product categories, and departments. Product modules are the most granular, then product categories, and finally departments. For example, a 16 oz can of "Stapleton's California Whole Prunes packed in Pear Juice" is classified into the product module "canned fruit - prunes", the product category "canned fruit", and into the department "dry grocery".

For each household, I observe a vector of demographic and geographic characteristics including the number of family members and their ages, the household's income, education, and race, as well as its zip code, county and state. The panel composition is designed to be projectable to the total United States population. The repeated reporting of a household's state of residence allows me to identify households that have moved between states during the time sample. Overall, my analysis uses the information of 2,500 households who have moved between states exactly one time during the sample.

I supplement the Consumer Panel dataset with the GS1 US dataset. GS1 US is a not-for-profit information standards organization which administers the UPC barcodes of products in the US.

[11] For a validation study of the dataset, see Einav, Leibtag and Nevo (2010).

Using the GS1 dataset, I link between products at the UPC level from the Nielsen Consumer Panel dataset and the firms producing them.

The reduced-form panel analysis of US states at the end of this section relies on three additional data sources. The first dataset I use for the reduced-form analysis is the Business Dynamics Statistics. It provides me with annual data on different firm formation indicators at the state level from 1980 to 2014. The two firm formation indicators I use in the benchmark analysis are the share of firms aged 5 years or younger, and the share of workers employed by these firms. The two other datasets used for the reduced-form analysis include information on state-level demographic characteristics. I use the National Intercensal data to obtain the age composition of each state between 1970 and 2014. Finally, I use the March CPS dataset to obtain the growth rate of the labor force and the share of workers older than 34 years old in each state between 1980 to 2014.

3.2 Identification and Estimation

A household may choose to purchase the same product repeatedly over time for two reasons: either because of consumer inertia, or because it has persistent taste for the fundamental attributes of that product. Both reasons are a form of state-dependence in the household's consumption decision. Following Heckman (1981), I refer to consumer inertia as structural state-dependence and to persistence in tastes as spurious state-dependence. The identification challenge is to distinguish between these two forms of state-dependence.

The distinction between consumer inertia and persistent unobserved heterogeneity in households' tastes is not only a microeconomic curiosity, but has important macroeconomic implications. While high probability of repeated product choice can arise due to both structural or spurious state dependence, only a change in structural state dependence has any macroeconomic impact. To illustrate that point, suppose household tastes toward the fundamental attributes of a product are serially correlated. A stronger serial correlation results in a higher number of repeated purchases. But household tastes toward the fundamental attributes of a product do not deter entry. An entrant can introduce a product with similar attributes, and attract new customers. If, on the other hand, households display structural inertia in their consumption choice, there is less scope for entry. An entrant can attract a customer that likes a certain product attribute, but cannot attract a customer that is locked into its previous product choice.

I distinguish between consumer inertia and spurious state dependence using the strategy originally proposed by Bronnenberg et al. (2012), henceforth BDG. I study the consumption behavior of households who move between states during the sample. The main identification assumption is that the decision of the household to move between states is orthogonal to its consumption preferences. The consumption decision of a household depends, among other things, on the relative price of a product, the advertisement the household is exposed to, the product placement within a store, as well as consumption norms among household's social circles. When a household moves between states, some of these factors change. Indeed, Bronnenberg et al. (2007) document that there is large and persistent geographic variations in the market shares of national brands. To study the consumption behavior of households, I focus on the top two firms in each product module. In particular, I study the share of purchases of the top firm relative to the purchases of the top two firms combined (henceforth, the *relative purchase share*). By top firms, I mean the two firms with the most purchases in a product module across all households in the sample. The advantage of this approach is that these products are available across all states, so it is unlikely that following the move, a household switches its consumption choice due to the availability of one of these two products. BDG show that the relative purchase share of households who move between states eventually converges to the representative relative purchase share of the state to which they moved. Recall that the structural modeling assumption on consumer inertia is that a household gets to re-optimize its choice with an exogenous probability. So the level of consumer inertia a household displays is given by the speed of convergence of its relative purchase share to that of the state to which it moved. I estimate the structural parameters which govern the speed of convergence by following households' consumption choices after their move.

The estimation procedure follows four steps: (i) computing the relative purchase share for each household, in every year, across every product module, (ii) estimating the representative relative purchase in every state using data on non-movers, (iii) computing the distance between the relative purchase share of movers following their move and the representative relative purchase in the state to which they moved, and (iv) estimating the degree of consumer inertia via non-linear least squares. I describe each of these steps in details below.

First, I denote the relative purchase share of household i at time t in product module j by y_{ijt}. It is given by

$$y_{ijt} = \frac{\sum_{s_{ijt} \in \mathcal{S}_{ijt}} \mathbb{1}_{1 \in r(s_{ijt})}}{\sum_{s_{ijt} \in \mathcal{S}_{ijt}} \left(\mathbb{1}_{1 \in r(s_{ijt})} + \mathbb{1}_{2 \in r(s_{ijt})} \right)} \ , \tag{7}$$

where $\mathcal{S}_{ijt}$ is the set of all shopping trips of household i at time t in which it bought a product from product module j. s_{ijt} denote each individual shopping trip of that household. $r(s)$ is the set of firm rankings from which a product was purchased in shopping trip s. For example, if household i purchased two products from product module j in shopping trip s_{ijt} and these products were of the top firm and 7th firm in that product module, then $r(s_{ijt}) = \{1, 7\}$. A firm's ranking, as mentioned above, is based on the total purchases of all households in the sample.

Next, I estimate the representative relative purchase share in every state. To do so, I run the following regression, while including only households that did not move between states during the sample:

$$y_{ijt} = \mu_{jl(i)} + \Gamma_j X_{it} + \epsilon_{ijt} \ , \tag{8}$$

where $l(i)$ is the state in which household i resides. $\mu_{jl(i)}$ is the representative relative purchase share in product module j at state $l(i)$. X_{it} is a set of household characteristics that includes the age of the household, the number of members in the household, the household's income, the education of the head of the family, as well as its race. Γ_j is a set of product-module-specific coefficients.

I then turn to households that move between states. I restrict attention to households who moved between states only once during the sample. Denote the state to which a household moved by $l^D(i)$, where D stands for destination. Denote the distance between a household's relative purchase share and that of the state to which it moved by δ_{ijt}. This distance is given by

$$\delta_{ijt} = y_{ijt} - \hat{\mu}_{jl^D(i)} - \hat{\Gamma}_j X_{it} \ , \tag{9}$$

where $\hat{\mu}$ and $\hat{\Gamma}_j$ are the regression coefficients from equation (8).

Finally, I estimate the level of consumer inertia for households in different age groups. The structural assumption on consumer inertia is that the distance between a household's relative purchase share and that of the state to which it moved shrinks at rate $(1-\theta_{a(i,t)})$. $a(i,t)$ denotes the age group of household i at time t, and $\theta_{a(i,t)}$ is the probability household i re-optimizes its choice in each product module at time t. That is, the structural assumption is that the following moment condition holds, $\mathbb{E}\left[\delta_{ijt+1} - \left(1-\theta_{a(i,t)}\right)\delta_{ijt}\right] = 0$.

I consider four age categories. Young, maturing, mature, and old households, which correspond to households whose head is between the ages of 20–34, 35–49, 50–64, and older than 64, respectively. I estimate the level of consumer inertia for each age group via non-linear weighted least squares. I minimize the following objective,

$$\sum_i \sum_j \sum_{t>t_i^m} w_{ij} \left[\delta_{ijt} - \left(\prod_{\tau=t_i^m+1}^{t} \left(1-\theta_{a(i,\tau)}\right)\right)\bar{\delta}_{ij0}\right]^2 \ , \tag{10}$$

where t_i^m is the year in which the household moved, and w_{ij} is the weight put on the distance of household i in product module j from the representative purchase share in her destination state. In the estimation, I set w_{ij} to the squared difference of the representative relative purchase share between the origin and destination states of household i in product module j. The logic is that if the measurement error is of similar magnitude for all households, then households who move between states with a higher difference in the representative relative purchase shares provide more information.

I choose not to include information from the year of the move, t_i^m, in equation (10) for two reasons. First, a mover's relative purchase share during the year of the move is a combination of its purchases in its origin state and its purchases in the destination state. Second, a household's move can act as a nudge for the household to reconsider some of its consumption choices. Since my goal is to study consumer inertia for the general population, this type of immediate adjustment should not be included in the estimates of consumer inertia. So the first observation I use for each household takes place one year following the move.

The structural assumption underlying equation (10) is consistent with the way I treat consumer inertia in the static model. Suppose each household can re-optimize its product choice in a fraction

θ of product modules, and that this is the only form of consumer inertia. Then, we expect the distance between the household and the state to which it moved to shrink at rate θ. I take a similar modeling approach in the quantitative model in the next section. This allows me to use the estimated coefficients in the calibration of the quantitative model.

While the identification procedure closely follows Bronnenberg et al. (2012), it is different than theirs along three dimensions. First, I use information on households that moves during the sample. BDG, instead, surveyed households and obtained information on whether and how long in the past these households have moved between states. The Nielsen Consumer Panel dataset today covers a longer time span relative to the time BDG performed their analysis, and the number of households in the sample have more than tripled. So using information on households that move during the sample provides a substantial amount of information, which allows me to exploit the panel dimension of the dataset.[12] Second, I use a different structural model to estimate the level of consumer inertia. In the structural part of their paper, BDG use a stylized model in which a household chooses between two brands. My structural assumption is guided by the consumption decision households display in both the static model presented above and the quantitative model I use in the next section. Third, rather than grouping products into brands, I group them into the firms which produce them. I do so primarily because I study how the population of firms in the economy, and not brands, have changed.[13] Despite these differences, the results I find are very much in line with BDG findings. Namely, there is a large degree of consumer inertia across households, and young households display the lowest amount of consumer inertia.

3.3 Descriptive Statistics and Supportive Evidence for the Identifying Assumptions

The analysis covers 934 different product modules.[14] Table 1 presents the summary statistics of the relative purchase share across states and product modules. The average relative purchase share equals 66%. That is, on average, the total number of purchases from the top firm in a product module is two times larger than the number of purchases from the second best-selling firm.

Table 1 also shows that there is large variation in the representative relative purchase shares across different states. The average standard deviation of the representative relative purchase share in a product module across states is 20%. This variation is used to estimate the level of consumer inertia. If all states had the same relative purchase share, then consumers wouldn't have an incentive to switch their consumption basket following their move. The average absolute difference between the representative relative purchase share in movers' state of origin and their

[12] For comparison, Section II.C. in BDG uses the panel information on households who moved during their sample. Overall, they observe 226 movers while I use information on more than 2,500 movers.

[13] I redid the analysis at the brand level and the results I find are very similar. This result is not surprising. It is often the case that the purchases of the top brand of a firm in a product module overshadow the purchases of other products it produces in that product module.

[14] I exclude product modules in which I observe less than 2,000 purchases, or if the relative purchase share of the top firm is greater than 90%. The estimation results are robust to including these modules.

Table 1: Summary Statistics on Relative Purchase Shares

Number of product modules	934
Relative purchase share of the top firm	
Average across product modules	0.66
Std across product modules	0.11
Std of relative purchase share across states	
Average across all product modules	0.2
Absolute difference between origin and destination states	
Average across all movers and product modules	0.1

Notes: The table presents descriptive statics on the representative relative purchase share of the top firm in each product module across states. The relative purchase share is defined as the ratio between the number of purchases of the top firm to the total number of purchases of the top two firms. These representative shares are estimated using the information of non-mover households via equation (8).

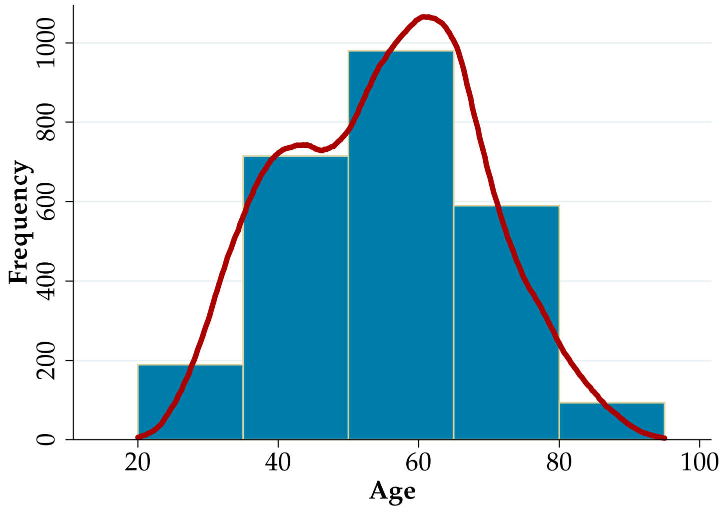
destination state is 10%. So there is substantial variation to identify the level of consumer inertia.

I next turn to describing the characteristics of movers. Figure 1 presents the age distribution of movers at the time of their move. I restrict the analysis to households who moved states exactly one time during the sample. Overall, I use the information on the purchases of 2,567 movers. The majority of movers in the sample are between the ages of 35–65. 189 of the movers are younger than 35 years old and 684 are older than 64 years old. Note that despite having only 189 movers younger than 35 years old, since each one is potentially purchasing products across 1,000 different product modules, I do have tens of thousands of observations on the purchases of young households. The lower number of young households does lead to relatively larger standard errors on the coefficient estimates, as is apparent in the next section.

In Appendix C, I provide two more tables with descriptive statistics on movers. Table 9 presents the migration patterns of movers across the 9 US Census Divisions. Out of the 2,500 movers, only 670 stayed within their Census Division. That is, 75% of movers moved not only between states but also across Census Divisions. Table 10 presents descriptive statistics on the observables across movers and non-movers. I find that, within the defined age-groups, movers and non-movers do not differ substantially across education levels, income, age, and the number of household members. While movers and non-movers do not differ across observables, they may differ across unobservables. In particular, households who move between states may be more open to new experiences relative to households who never move between states. This difference can potentially affect not only the geographic residence choice of households, but also their consumption decisions. In the next section, I discuss how selection into the movers population may affect my estimates.

Let me now provide supportive evidence for two key identifying assumptions. The first identi-

Figure 1: Age Distribution of Movers



Notes: The figure presents the age distribution of movers at the time of the move. Each bar represents the number of movers in the age category used in the analysis. The width of each bar is 15 years, so the first bar consists of movers who are between the ages of 20 and 34. The red line is the kernel density of the age distribution.

fying assumption is on movers' relative purchase share prior to the move. I assume that movers' relative purchase share prior to the move is similar to other households in their origin state. To assess this assumption, I use the following regression,

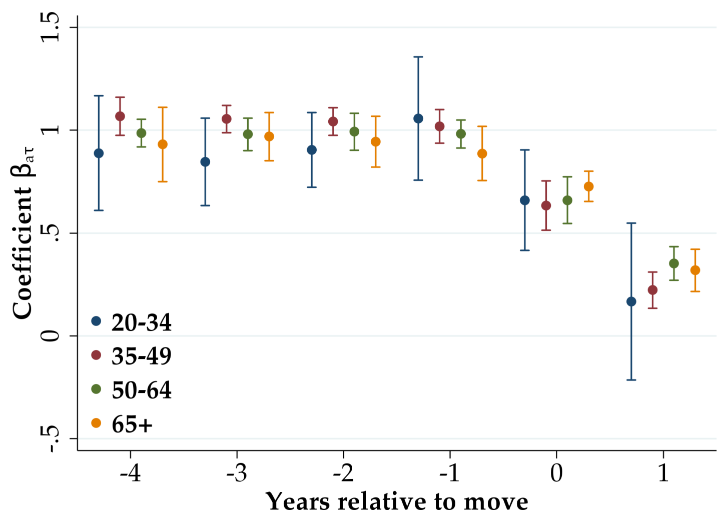
$$\delta_{ijt_i^m+\tau} = \beta_{a\tau}\Delta\mu_{ij} + \epsilon_{ij\tau} \ , \tag{11}$$

where $\delta_{ijt_i^m+\tau}$ is the distance of the relative purchase share of household i, in product module j, τ years after the year of its move (t_i^m), from the representative relative purchase share in the destination state. If $\tau < 0$, this share represents this distance prior to its move. $\Delta\mu_{ij}$ is the difference in the relative purchase share between the origin and destination state, which is equal to $\hat{\mu}_{jl^O(i)} - \hat{\mu}_{jl^D(i)}$. $l^O(i)$ and $l^D(i)$ are the origin and destination states of household i, respectively. The estimates for μ are obtained from running regression (8). $\beta_{a\tau}$ are the coefficients multiplying $\Delta\mu_{ij}$, which can vary with a household's age group a and with the years relative to its move τ.

If movers' relative purchase share is identical to households in their origin state prior to the move, we expect $\beta_{a\tau} = 1$ for all a and $\tau < 0$. I estimate the values of $\beta_{a\tau}$ by running a weighted least squares regression of equation (9). The weights on each observation is given by $(\Delta\mu_{ij})^2$, so that households who move between states with a larger distance in the representative relative purchase share are weighted more heavily. These weights are identical to the weights I use for the main estimation in the next section. I drop observations which have more than one value missing during the 6 year window. Overall, the coefficients are estimated using about 230 thousand observations.[15]

Figure 2 presents the estimation results. The markers represent the point estimates of $\beta_{a\tau}$ and the whiskers represent 95% confidence intervals. Standard errors are clustered by state of origin. Consistently with the identifying assumption, I find that $\beta_{a\tau}$ is not significantly different than one for all $\tau < 0$. Only when $\tau \geq 0$, $\beta_{a\tau}$ is significantly lower than 1.

Figure 2: Movers Consumption Prior to Move



Notes: The figure presents the coefficient estimates from regression (11). A value of 1 corresponds to the typical relative purchase share in the origin state, while a value of 0 corresponds to the relative purchase share in the destination state. The whiskers represent 95% confidence intervals, clustered at the state of origin level. The figure conveys the information that movers' relative purchase share is not significantly different than other households in the origin state prior to their move.

The second identifying assumption is that the relative purchase share in each state and in every product module follows a martingale. In other words, the expected future relative purchase share is equal to the current relative purchase share. Bronnenberg et al. (2012) provide evidence to support this identifying assumption. They compare the relative purchase shares of top brands in every state across 27 models in the period of 1948–1968 with the period 2006–2008. Since my dataset doesn't go as far back, I do a similar exercise but compare the average relative purchase share of the top firm between 2004–2009 and 2010–2015. To do so, I estimate the typical relative purchase share in each state and in every product module using equation (8) for the two periods, separately.

Figure 3 presents a scatter of the typical relative purchase shares in the two periods, 2004–2009 and 2010–2015. The data does not reject the assumption that the expected future relative purchase

[15] The results are robust to restricting the sample to a balanced panel, but the number of observations drops to 90 thousand.

share is equal to the current relative purchase share. The point estimate of the regression equals 0.99, with a standard error of 0.01.

Figure 3: Representative Relative Purchase Shares Over Time

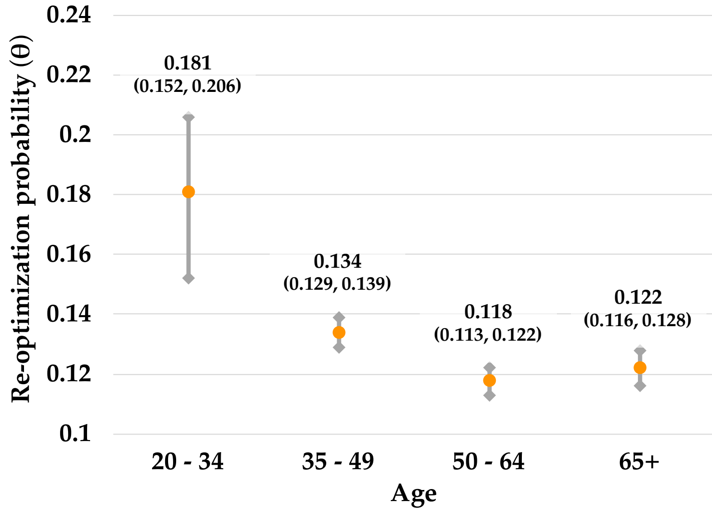
Notes: The figure presents a scatter plot of the relative purchase shares, in each state, in every product module, over two time periods, 2004–2009 and 2010–2015. The solid red line is the linear fit prediction. This figure shows that the identifying assumption that the expected future relative purchase share equals the current relative share is not rejected by the data.

3.4 Results

In the benchmark analysis, I group households into four age groups. For convenience, I refer to a household as young, maturing, mature, and old, if the head of the household is between 20–34, 35–49, 50–64, and older than 64, respectively. The benchmark estimation results are presented in Figure 4. The orange markers in the figure represent the point estimates for the re-optimization probability (θ) across the four age groups considered. The grey whiskers indicate 95% confidence intervals. Inference was done using a bootstrap procedure, clustered at the product-module level. The first thing that stands out is that young households display the highest probability of re-optimization. In other words, young households are less inertial in their consumption choices than older households.

The estimate for the re-optimization probability of a young household is 18.1%. That is, on average, a young household considers switching about one fifth of the products in its consumption basket in every year. The estimates for maturing, mature, and old households respectively, are 13.4%, 11.8%, and 12.2%. So the relationship between consumer inertia and age does not seem to be monotonic. Old households are less inertial than maturing households, though this difference

Figure 4: Consumer Inertia by Different Age Groups



Notes: The figure presents the benchmark estimation results, based on 753,354 observations. The orange markers indicate the point estimates for each age group. The grey whiskers indicate 95% confidence intervals. Inference was done using a bootstrap procedure, clustered at the product module level. The estimate can be interpreted as the exogenous re-optimization probability. So a higher estimate represents a lower level of consumer inertia.

is not significant. One possible explanation for this result is that old household have more time to shop around and consider alternative products.

Consistent with the findings of Bronnenberg et al. (2012), the estimates imply that the convergence of young movers is faster than for older movers. Consider the distance between a mover's relative purchase share one year after the move and the representative relative purchase share of the state to which she moved. A young household closes 50% of this gap in 5.5 years. A maturing, mature, and old households take 7.5, 8.5, and 8.2 years, respectively, to close this gap.

While the results in the previous section show that movers and non-movers are similar across observables, there may still be selection on unobservables into the movers population. Movers may be more open to new experiences relative to non-movers. This may affect not only their location decisions but also their consumption choices. The selection between movers and non-movers may be more pronounced for older households. That is, the similarity between a young mover and a young non-mover may be stronger than for older households. Selection into the movers population may lead to a downward bias in the estimated level of consumer inertia of the general population. If the selection is stronger for older households, then not only the level of consumer inertia would suffer from a downward bias but so would the difference between young and older households. These two downward biases, if exist, attenuate the results I find in the following sections.

Table 2 presents robustness results for the estimated probability of re-optimization. The bench-

Table 2: Consumer Inertia Estimates

Age group	(1)	(2)	(3)	(4)
20–34	0.181	0.189	0.183	0.193
	[0.152, 0.206]	[0.140, 0.263]	[0.155, 0.209]	[0.145, 0.266]
35–49	0.134	0.153		
	[0.129, 0.139]	[0.122, 0.159]		
50–64	0.118	0.119		
	[0.113, 0.122]	[0.107, 0.133]		
65 and over	0.122	0.127		
	[0.116, 0.128]	[0.112, 0.141]		
35 and over			0.123	0.128
			[0.121, 0.126]	[0.119, 0.135]
Module×HH f.e.	✓	✗	✓	✗

Notes: The table presents the estimates of the re-optimization probability under different specifications. The number of observations is 753,354. 95% CI in square brackets. Inference was done using a bootstrap procedure, clustered at the product module level. A higher estimate corresponds to a lower level of consumer inertia.

mark specification includes four age groups, and allows for household-module fixed effects. The latter implies that I allow for the initial adjustment of households following the move to vary by households and product modules. The table presents two types of robustness checks. The first column of the table presents the benchmark estimates. The second column presents the estimates when I don't allow for household-module fixed effects. The estimates do not differ substantially, yet the standard errors expand. Young households are still the most inertial age group. Despite overlapping confidence intervals between the young and maturing age groups, the estimate for young households is significantly larger than the estimate for maturing households at the 5% level.

The last two columns of Table 2 present the estimates when grouping all households who are older than 34 years together. The third column includes household-module fixed effects, and the fourth column presents the results without these controls. In both cases, young households are significantly less inertial than older households. The point estimates indicate that young households are 50% more likely to consider switching their consumption products.

3.5 Consumer Inertia and Firm Formation

I now use the micro-estimates of consumer inertia by age to study the effect of consumer inertia on firm formation. I exploit two types of variations in the data. The first is variation in the level of consumer inertia across customers of different product categories. The second is temporal variation in the level of consumer inertia in different US states. Empirical evidence based on these two types of variations indicate that more consumer inertia leads to a decline in firm formation.

3.5.1 Cross-Product-Category Analysis

This section explores the relationship between consumer inertia and firm entry rate across different product categories. I study how the average consumer inertia among customers in a market affects firm entry rate in that market. To perform the analysis, I construct an indicator of the average level of consumer inertia and compute the entry rate in each product category.

Nielsen's Consumer Panel dataset classifies products into three levels of aggregation: departments, categories, and modules. Modules are the most granular level of aggregation, and departments are the least granular one. There is a trade-off in choosing the level of aggregation for the analysis. The more granular the level of analysis is, the more variation I can exploit. But since new firms often sell products across several product modules, the relevant consumer inertia for their entry decision may be at a less granular level. For example, consider a new carbonated beverage producer which sells its products in the "carbonated beverage" category. Such firm can sell products across both "soft drinks - carbonated" and "soft drinks - low calorie" product modules. As a result, its entry decision would not depend only on consumer inertia among customers of dietary carbonated drinks but on consumer inertia among customers of carbonated beverages in general. For this reason, the analysis is done at the product category level.

The indicator of the average level of consumer inertia in a product category is computed using my micro-estimates of consumer inertia by age together with the age composition of customers. I compute the share of purchases by young, maturing, mature, and old households in every product category across the entire sample (2004–2015).[16] Let the share of purchases by age group a in product category j be noted by s_{aj}. The consumer inertia index is given by

$$\overline{(1-\theta)}_j = \sum_a s_{aj} \left(1 - \hat{\theta}_a\right) \ , \tag{12}$$

where $\overline{(1-\theta)}_j$ is the index of consumer inertia in product category j, and $\hat{\theta}_a$ is the benchmark estimate of the re-optimization probability in age category a.

To compute the annual firm entry rate, I divide the number of firms who sell their product in a product category for the first time by the number of firms who sell products in that category during that year. The firm entry rate used in the analysis is the average annual firm entry rate between 2006–2015.

I regress the firm entry rate on the consumer inertia index, controlling for department fixed effects. I consider four different specifications. First, I present both non-weighted results as well as weighted by the number of firms in the product category. Naturally, entry rates are more volatile in categories with a lower number of firms. So I weight observations to achieve a precise estimate by correcting for hetroskedasticity. Second, I consider the sample of all product categories and one

[16] I use Nielsen weights to construct the share of purchases by each age group.

Table 3: Consumer Inertia and Firm Entry Rate Across Product Categories

	(1)	(2)	(3)	(4)
Consumer inertia index	-3.62***	-2.86**	-4.68*	-2.94*
	[-6.8, -1.4]	[-5.3, -0.6]	[-10.5, 0.3]	[-7.1, 0.3]
Department f.e.	✓	✓	✓	✓
Excluding outliers	✓	✓	✗	✗
Weighted	✗	✓	✗	✓
Obs.	89	89	97	97

Notes: The table presents the cross-product-category regressions. The dependent variable is the entry rate at the product-category level. 95% confidence intervals in brackets. Inference was done using a bootstrap procedure. If weighted, then weights are equal to the number of operating firms in a product category. *** (**) {*} - 99% (95%) {90%} confidence interval does not include zero.

where I exclude the top and bottom 5% of product categories according to their entry rates. The point estimates do not change much when excluding outliers, but confidence intervals are smaller. Table 3 presents the regression results.

The benchmark specification is column (2), in which product categories are weighted by the number of firms operating in them and outliers are excluded. A one percentage point increase in consumer inertia leads to a decline of 2.86 percentage points in the firm entry rate. To get a sense of magnitudes, the standard deviation of the consumer inertia index across product categories is 0.2 percentage points and the standard deviation of entry rates is 3.7 percentage points. So while the relation between the two is negative and significant, consumer inertia accounts for a small fraction of the variation in entry rates across product categories.[17] Product categories may differ along many dimensions, such as fixed operating costs, capital intensity, or regulatory requirements. It is not surprising, then, that consumer inertia cannot explain a large fraction of the variation in entry rates across product categories. The estimates from the other regressions specifications in Table 3 also imply a significant and negative relation between consumer inertia and the firm entry rate in a product category.

3.5.2 Cross-State Analysis

The second empirical analysis I perform uses temporal variation in the level of consumer inertia across US states. I study how the average consumer inertia among the population of a state affects firm formation in that state. Using the age composition of each state in every year, I construct an annual state-specific consumer inertia index based on my micro-estimates of consumer inertia.

As in the previous section, the consumer inertia index I construct is given by the average

[17] The R^2 of specification (2) in Table 3 is 0.36.

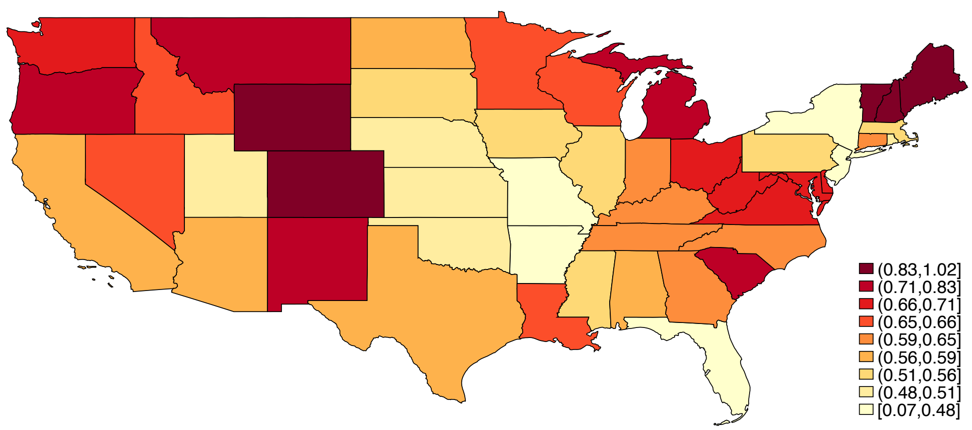
consumer inertia among the adult population (20 years and older) of a state. That is,

$$\overline{(1-\theta)}_{lt} = \sum_a s_{alt}\left(1-\hat{\theta}_a\right) \; , \tag{13}$$

where $\overline{(1-\theta)}_{lt}$ is the consumer inertia index in state l at time t, s_{alt} is the share of population in the age group a out of the adult population in state l at time t. Finally, $\hat{\theta}_a$ is the estimated re-optimization probability of age group a from Section 3.4.

To identify the relation between consumer inertia and firm formation, I exploit variation in the change of consumer inertia across states over time. Different states have experienced different demographic shifts over the 1980–2014 time period. As a result, the level of consumer inertia in some states have increased more during this time period relative to other states. Figure 5 presents the change in the consumer inertia indicator across states between 1980–2010. The map reveals that all states have experienced a rise in consumer inertia. But there is substantial variation in these changes. Florida, which had a small share of young households already in 1980, have experienced a small rise in consumer inertia. Vermont and New Hampshire, on the other hand, have experienced a large increase in consumer inertia.

Figure 5: State-Level Change in Consumer Inertia (1980–2010)



Notes: This figure presents the change in state-level consumer inertia index between 1980 to 2010. The change is in percentage points. This figure shows that all states have experienced an increase in consumer inertia during this time period and that there is substantial variation across states in how consumer inertia increased over time.

For the dependent variable, I use two different firm formation indicators. The first is the share of young firms (5 years or younger) in a state. The second is the share of workers employed by young firms in a state. I choose these indicators as they are less affected by business cycle fluctuations. Appendix D includes results also for the share of entrants and the share of workers employed by entrants. The results are robust to these alternative measures of firm formation.

I use the following regression specification.

$$y_{alt} = \alpha_l + \delta_t + \beta\overline{(1-\theta)}_{alt} + \Gamma X_{alt} + \epsilon_{alt}\ , \tag{14}$$

where y_{alt} is one of the firm formation measures, α_l is a state-level fixed-effect, δ_t is a time fixed-effect, and X_{alt} is a vector of controls. The coefficient of interest is β.

There are two potential biases when estimating regression (14). The first is an omitted-variable bias. Since the constructed index of consumer inertia is based on the age composition in each state, I need to control for other demographic channels which can affect firm formation. Otherwise, I may mistakenly attribute the effects of these other channels to consumer inertia. To address this concern, I include other demographic channels that have been raised by the literature as drivers of the decline in firm formation. The two controls I include are the share of old workers in the workforce and the growth rate of the labor force. Liang et al. (2014) and Engbom (2017) argue that the decline in firm formation is driven by an aging workforce. Karahan et al. (2016) argues that a decline in the growth rate of the labor force can explain a substantial proportion of the decline in firm formation.

Before turning to the second potential bias, it is worth pointing out the differences in the age composition of a state which allow me to separately identify the effects of consumer inertia from the other demographic channels I control for. Consumer inertia depends primarily on the share of young households in the *adult population*. The share of older workers, on the other hand, heavily depends on the share of young households in the *working age population*. So the share of old households (65 and older) in the economy allows me to separately identify the demand channel of consumer inertia from the supply channel of aging workforce. The labor force growth rate depends on the difference between the share of population which moves from the mature to old population (retirees) and the share of young population which joins the workforce. If the two are equal, so that the labor force growth rate is zero, it does not matter how large each of the two components is. For consumer inertia, on the other hand, an increase in the share of young households leads to a decline in consumer inertia regardless of the fraction of mature households that turn old. This is because consumer inertia among the mature and old age groups is fairly similar.

The second potential bias I address is an endogeneity bias. Young households may move to booming states, where the share of young firms is higher. Since young households are less inertial, this can lead to an upward bias. To mitigate this concern, I instrument a state's consumer-inertia index with the 10-year lagged age composition of that state. The approach of using the lagged age composition to control for this potential endogeneity bias is common in the literature and has been used, for example, by Shimer (2001), Karahan et al. (2016), and Engbom (2017). The exclusion restriction is that a household does not move to a state because she thinks that it will boom 10 years in the future.[18] The instrument I use is the predicted level of consumer inertia

[18] Note that if a young household moves to a state because she believes it would boom 10 years into the future, my estimates would suffer from a downward bias. This is because such household would be in the *maturing households*

10-years ahead which is defined as follows,

$$\left[\overline{(1-\theta)}_{lt}^{p}\right]_{t-10} = \sum_{a} \left[s_{alt}^{p}\right]_{t-10} \left(1-\hat{\theta}_{a}\right) \,, \tag{15}$$

where $\left[\overline{(1-\theta)}_{lt}^{p}\right]_{t-10}$ is the predicted value of consumer inertia 10-years ahead. $\left[s_{alt}^{p}\right]_{t-10}$ corresponds to the predicted share of households that would be in age group a, 10 years into the future. For example, the predicted share of young households 10-years ahead is defined as the share of households aged 10–24 out of the population aged 10–79.[19] This instrument has substantial explanatory power. The R^2 from the first-stage regression is equal to 0.91 and the F-stat is equal to 84.

Table 4 presents the regression results. The top panel of the table considers the share of young firms as the dependent variable, and the bottom panel considers the share of workers employed by young firms. The first three columns include the OLS results and the last three columns include the IV results. The OLS results point at a significant negative relationship between the level of consumer inertia and the two measures of firm formation. The OLS specifications indicate that a one percentage point increase in consumer inertia is associated with 17 percentage points decrease in the share of young firms and 8 percentage points decrease in the share of workers employed by young firms.

The IV results lead to a similar conclusion. An increase in the level of consumer inertia leads to a decline in firm formation. In terms of the share of young firms, a rise of one percentage point in consumer inertia leads to a decline of 12–13 percentage points in the share of young firms. To get a sense of magnitudes, the standard deviation of consumer inertia, controlling for state and time fixed-effects, is 7 basis points. The corresponding standard deviation of the share of young firms is 190 basis points. So a one standard deviation increase in the level of consumer inertia leads to 0.9 percentage points decrease in the share of young firms. This corresponds to a decline in the share of young firms of 0.5 standard deviations.

Turning to the bottom panel of Table 4, a one percentage point increase in consumer inertia leads to a decline of 9.3–9.7 percentage points decline in the share of workers employed by young firms. The standard deviation of the share of workers employed by young firms, controlling for state and time fixed-effects, is 130 basis points. A one standard deviation rise in the level of consumer inertia leads to a decline of 65 basis points in the employment share of young firms. Similar to the effect on the share of young firms, this corresponds to a decline in the employment share of young firms of 0.5 standard deviations.

Turning to the growth rate of the labor force, a decline of 1 percentage point in this growth rate is associated with a decline of 6 basis points in the share of young firms. In terms of standard

group 10 years following its move, and consumer inertia of maturing households is larger than the average level of consumer inertia.

[19]Similarly, the predicted share of maturing, mature, and old households correspond to the share of households aged 25–39, 40–54, and 55–79, respectively, out of the population aged 10–79.

Table 4: Consumer Inertia and Firm Formation – State Analysis

	Dependent variable: share of young firms					
	(1)	(2)	(3)	(4)	(5)	(6)
Consumer inertia index	-16.9^{***}	-16.7^{***}	-16.6^{***}	-13.0^{**}	-12.9^{**}	-12.3^{**}
	[-33.7, -9.9]	[-33.3, -9.7]	[-33.1, -9.6]	[-24.8, -1.0]	[-24.8, -1.3]	[-23.8, -0.1]
Labor-force growth rate		0.060^{***}	0.059^{***}		0.064^{***}	0.060^{***}
		[0.04, 0.08]	[0.04, 0.08]		[0.04, 0.09]	[0.04, 0.08]
Share of older workers			-0.01			-0.07
			[-0.09, 0.05]			[-0.23, 0.03]
State and time f.e.	✓	✓	✓	✓	✓	✓
Instrumented	✗	✗	✗	✓	✓	✓
	Dependent variable: share of workers employed by young firms					
	(1)	(2)	(3)	(4)	(5)	(6)
Consumer inertia index	-7.85^{***}	-7.78^{***}	-7.96^{***}	-9.36^{**}	-9.35^{**}	-9.7^{*}
	[-15.7, -2.2]	[-15.6, -2.2]	[-16.7, -2.1]	[-18.5, -0.4]	[-18.4, -0.4]	[-19.1, 0.3]
Labor-force growth rate		0.027^{***}	0.027^{***}		0.025^{***}	0.027^{***}
		[0.01, 0.04]	[0.01, 0.04]		[0.01, 0.04]	[0.01, 0.04]
Share of older workers			0.01			0.04
			[-0.05, 0.07]			[-0.10, 0.14]
State and time f.e.	✓	✓	✓	✓	✓	✓
Instrumented	✗	✗	✗	✓	✓	✓

Notes: This table presents the regression results of (14). The sample period is 1982–2014. In the top panel, the dependent variable is the share of firms that are 5 years or younger. The dependent variable in the bottom panel is the share workers employed by these firms. 95% confidence intervals in brackets. Inference was done using a bootstrap procedure, clustered at the state level. *** (**) - 99% (95%) confidence interval does not include zero.

deviations, a one standard deviation decline in the labor force growth rate is associated with a decline of 0.1 standard deviations in the share of young firms. A similar relationship holds for the employment share of young firms. A one percentage point decline in the labor force growth rate is associated with about 3 basis points decline in the employment share of young firms. A one standard deviation decline in the labor force growth rate is associated with a 0.06 decline in the standard deviation of the employment share of young firms.

The relation between the share of workers older than 35 years old and firm formation is not significant. The point estimates suggest that a larger share of older workers is associated with a negligible decline in the share of young firms and a negligible increase in the employment share of young firms.

The estimated effect of consumer inertia on firm formation is combined of two margins: an extensive margin and an inter-state intensive margin. Consider a decline in the level of consumer inertia in a state. The extensive margin corresponds to the opening of firms who would have otherwise not joined the economy. The inter-state intensive margin corresponds to the opening of firms who would have otherwise joined the economy but locate in a different state. Other things equal, firms will choose to start their business in states with less consumer inertia. They do so because it is easier to penetrate the market in such states. The inter-state margin is particularly relevant for tradable good firms, who sell across many states during their life-cycle.

For the US as a whole, the inter-state intensive margin is a zero-sum game. A decline in aggregate consumer inertia leads to an increase in the number of young firms in the US through the creation of new firms, not through the location decision of firms who would have opened regardless of the level of consumer inertia. Using my micro-estimates of consumer inertia together with the age composition of the US population, I construct an aggregate index of consumer inertia. Between the late 1980s and early 2000s, this index declined by 40 basis points. A naive extrapolation of the point estimates from Table 4 indicate that the decline in aggregate consumer inertia accounts for 70% of the declining share of young firms, and 80% of the declining share of workers employed by these firms between the late 1980s and early 2000s. I cautiously interpret these magnitudes as an upper bound of the aggregate effect of consumer inertia.

Finally, let me address one difference between the data source I use to estimate consumer inertia by age and the dependent variables I use here. While I estimate the level of consumer inertia using information on consumer behavior in the sector of consumer packaged goods (CPG), the firm formation measures I use in the regressions cover all sectors of the economy. The main goal of this section is to establish a qualitative relation between consumer inertia and firm formation. To achieve this goal, the key assumption is that relative consumer inertia by age in the CPG sector is correlated with the relative consumer inertia by age across other sectors.[20] In other words, if consumption choices made by younger households display less inertia than those of older households in the CPG sector, then such relationship is assumed to hold also for other sectors of the economy. I find this assumption reasonable as many of the forces behind consumer inertia are not exclusive to the CPG sector. For example, households can develop habits to the goods and services they consume, or households can display inattentive behavior in their consumption choices.

4 Quantitative Model

In this section, I develop a quantitative general equilibrium model of entry, exit, and firm dynamics that features consumer inertia. The model differs from the static model along three dimensions. Most importantly, as opposed to the static model, the quantitative model is dynamic. This introduces an investment motive in the firm's choice of markup. In particular, the firm has an incentive to lower its markup and increase the measure of its locked-in customers in the following period. Second, households' final consumption good is combined of a continuum of product types. In

[20] Note that the average level of consumer inertia does not affect the results because it is captured by the fixed-effects.

each product type, a household consumes a single product out of a variety of options. So even if a household is locked into a product in a specific product type, it can shift its expenditure toward other product types. As a result, the elasticity of demand coming from locked-in customers is higher than one, its value in the static model, but lower than the demand elasticity coming from unattached customers. Finally, firms are subject to persistent idiosyncratic productivity shocks and stochastic fixed operating costs. The latter gives rise to endogenous exit. The stationary equilibrium of the economy features a time-invariant distribution of firms over the two dimensions of heterogeneity: the size of the customer base and the productivity level.

I abstract from explicitly modeling households of different ages. Instead, I study an economy populated by ex-ante identical households who display the same level of consumer inertia. I then study the transition dynamics of the economy in response to an unanticipated and deterministic rise in the level of consumer inertia. The goal of the quantitative model is to study how a rise in consumer inertia affects the economy, regardless of the driver of this rise.

I calibrate the stationary equilibrium of the model to match key features of the US economy in the late 1980s, such as the age distribution of firms and the aggregate share of profits. The level of consumer inertia is calibrated using my micro-estimates of consumer inertia together with the age composition of the US at the time. To study the impact of an aging-induced rise in consumer inertia, I analyze the effect of an unanticipated and deterministic shock to the level of consumer inertia that moves it from its initial level in the late 1980s according to observed and predicted demographic shifts of the US population.

The model predicts that the rise in consumer inertia accounts for about one half of the decline in the share of young firms and one third of the rise in aggregate profits between the late 1980s and early 2000s.

4.1 Environment

There is a continuum of product types. Each product type consists of a continuous measure of operating firms. In every period, households consume only one product from each product type. Consumer inertia is modeled as introduced in section 2. In each product type, there is an exogenous probability $1-\theta$ that a household is locked into its previous consumption product and cannot re-optimize.

The supply side of the economy builds upon Hopenhayn (1992). The economy is populated by a continuum of firms who differ along two dimensions: their productivity level and the size of their customer base.[21] New firms can enter the economy by paying a fixed entry cost. Entrants start with no initial customer base, and build it along their life-cycle. In addition, firms are subject to persistent idiosyncratic productivity shocks. Finally, incumbents need to pay a stochastic fixed operating cost in every period. This gives rise to endogenous exit.

[21] Firms also differ by the product type in which they operate. This distinction, however, is not important, as I study a symmetric equilibrium in which all product types are identical.

I start by describing the demand side of the economy, and derive the equilibrium demand function for individual firms. I then turn to the supply side of the economy, and lay out the firm's problem.

4.1.1 Demand

There is a continuum of measure one of ex-ante identical households in the economy, indexed by $i \in (0,1)$. The lifetime utility of household i is given by

$$U_i = \sum_{t=0}^{\infty} \beta^t \ln C_{it} \ ,$$

where C_{it} is the aggregate consumption good household i consumes at time t. To ease notation, I omit both the household index i and the time index t from all equations below. The aggregate consumption good is composed of a variety of product types using a Dixit-Stiglitz aggregator,

$$C = \left\{ \int_0^1 \left[\exp\left(\frac{1}{\sigma - 1} \epsilon_{j_m} \right) c_{j_m} \right]^{\frac{\eta-1}{\eta}} dm \right\}^{\frac{\eta}{\eta-1}} \ , \tag{16}$$

where $j_m \in (0, J_m)$ denotes the product the household is consuming from product type m, and c_{j_m} is the quantity consumed of that product. J_m is the measure of products in product type m. ϵ_{j_m} is an idiosyncratic taste shock for consuming product j_m. Taste shocks are independent across products, product types, and households. I describe the distribution of taste shocks after introducing the household's problem below. The budget constraint of the household is given by

$$\int_0^1 p_{j_m} c_{j_m} dm = w + \Pi \ , \tag{17}$$

where p_{j_m} is the price of the chosen product from product type m. I assume household inelastically supply a unit of labor so that their salary is equal to the wage, w. Differently from the static model, I assume that aggregate profits, Π, are rebated to households as a lump sum transfer.

Households are subject to consumer inertia. In the dynamic context I study, consumer inertia alters the household's problem in two ways. First, there is a probability $1-\theta$ that the household cannot re-optimize its consumption choice in product type m and is locked into the product it purchased in the previous period. In case that product is no longer available, the household is forced to re-optimize. The re-optimization draw is iid across households and product types. Second, households do not internalize that they are subject to consumer inertia. So when a household chooses which product to buy, it only considers its current relative price and not its expected relative price in future periods. I discuss the latter assumption in details after introducing the household's problem.

The household's problem is given by

$$\max_{\{\{j_{mt},c_{j_{mt}}\}_{m=0}^{1}\}_{t=0}^{\infty}} \quad \sum_{t=0}^{\infty} \beta^t \ln \left[\int_0^1 \left(e^{\frac{1}{\sigma-1}\epsilon_{j_{mt}}} c_{j_{mt}} \right)^{\frac{\eta-1}{\eta}} dm \right]^{\frac{\eta}{\eta-1}} , \tag{18}$$

$$\text{s.t.} \quad \int_0^1 p_{j_{mt}} c_{j_{mt}} dm = w_t + \Pi_t \ , \quad \text{for all } t \ , \tag{19}$$

$$j_{mt} = \quad j^*_{mt-1}, \quad \text{if} \quad \xi_{mt} = 0 \quad \text{and} \quad \left(j^*_{mt-1} \text{ is available at time } t \right) \ . \tag{20}$$

The household is maximizing its discounted lifetime utility (18), subject to two constraints. The first constraint (19) is the budget constraint of the household, and the second constraint (20) is the consumer inertia constraint. ξ_{mt} is an indicator variable that takes the value 1 in case the household can re-optimize its choice in product type m at time t. ξ_{mt} takes the value 1 with probability θ, drawn independently across product types and households. j^*_{mt-1} is the product the household is locked into, in case it cannot re-optimize. In equilibrium, this product is the one consumed by the household in the previous period. The household, however, does not internalize that it can be locked into the product. To formalize that, I use the $*$ superscript, implying that the household does not directly choose j^*_{mt-1}. j^*_{mt} is defined as follows,

$$j^*_{mt} = \begin{cases} j^*_{mt-1}, & \text{if} \quad (\xi_{mt} = 0) \quad \text{and} \quad \left(j^*_{mt-1} \text{ is available at time } t \right) \\ \arg\max_{j\in(0,J_{mt})} \left[-(\sigma-1)\ln p_{jt} + \epsilon_{jt} \right], & \text{otherwise} \end{cases} \tag{21}$$

I assume that if the household can re-optimize its consumption choice, then its taste shocks follow a Gumbel distribution with location parameter $-\ln J_m$ and unit scale. The assumption on the location parameter implies that there is no love-of-variety in the model. If, instead, the household is locked into its previous consumption choice, I assume that the taste shock satisfies $e^{\frac{\eta-1}{\sigma-1}\epsilon} = \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right)$, where $\Gamma(\cdot)$ is the Gamma function. As Lemma 4 shows below, this assumption implies that if all firms charge the same price, the household spends the same amount on a product which it is locked into and on a product which it chooses freely.

Let me discuss the role of the assumption that households do not internalize that their consumption choices are inertial. Under the possibility of lock-in, a highly sophisticated household needs to take into account not only the current price of a firm but also its expected future prices. So, given the information available to the household, it needs to forecast each firm's future prices. In the presence of endogenous exit, the household also needs to take into account the current and future exit probability of different firms. By choosing a firm with a higher exit probability, a household can effectively decrease its lock-in probability. I assume that households do not internalize the possibility of lock-in. In other words, households are behaving in a myopic fashion when choosing which product to consume.

I choose to use this modeling assumption for three reasons. First, I believe this assumption provides a better approximation to the data than the assumption that households are fully sophisticated. Second, because I model consumer inertia as an exogenous re-optimization probability, households in the model would prefer to buy from a firm that is more likely to exit the economy. If the firm exits the economy, the household is able to re-optimize its product choice with certainty. Suppose, however, that the deep reason for lock-in is some sort of switching cost the household experiences, whether emotional, cognitive or pecuniary. In that case, a household may prefer to buy from a firm that is less likely to exit. If the firm exits the economy, it would have to pay the switching cost. Assuming households only take the current relative price into account allows me to abstract from the fundamental forces underlying consumer inertia.

The third reason for assuming households do not internalize the possibility of lock-in relates to the estimation strategy. A sophisticated consumer would put more weight on a firm's relative price than a myopic consumer. This is because, in equilibrium, relative prices are serially correlated – a firm with a lower relative price this period is likely to have a lower relative relative price also in following periods. To a first order, the difference between the consumption choice of these two households will appear as a difference in their demand elasticity. In particular, the consumption behavior of the sophisticated household would display higher demand elasticity. Since the elasticity of demand is estimated, the degree of sophistication households display can be captured by the resulting demand elasticity. [22]

It is useful to derive the optimal consumption behavior of households before turning to the firm's problem. As in the static model, I am able to derive an analytic formula for the firm's demand function. Some results are similar to the ones derived in section 2. However, the presence of imperfectly substitutable product types adds an aspect to the consumption behavior of households. Households do not only need to choose which product to purchase in every product type, but also how much of their expenditure to allocate toward different products.

First, I show that the household consumption choice can be represented as a multinomial logit discrete choice problem. The following lemma also implicitly implies that in equilibrium $j_{mt} = j^*_{mt}$, for all m and t. In the following lemmas, whenever appropriate, I omit both the time and product type subscripts to ease notation.

Lemma 2. If the household is not locked into a product, it chooses the firm which maximizes

$$j_m \equiv \arg \max_{j \in (0, J_m)} -(\sigma - 1) \ln p_j + \epsilon_j \ .$$

[22]In Appendix B, I consider a dynamic version of the static model where firms' exit probability is exogenous. I analyze the consumption behavior of two consumers. The first is a sophisticated consumer who chooses which product to consume taking into account the possibility of lock-in, and whose primitive elasticity of substitution is σ_1. I further assume that the sophisticated consumer uses a ln-linear approximation to predict the firm's price in future periods given its current price. The second consumer considered is a simple one who does not internalize the lock-in possibility and has an elasticity of substitution of σ_2. I derive an analytic expression for σ_2 as a function of σ_1 and other structural parameters, such that if it holds, the consumption behavior of the two consumers is identical. Furthermore, in the estimated version of the quantitative model, I find that the pricing function of firms is well approximated by a ln-linear law of motion.

Using Lemma 2 and the distributional assumption on the taste shocks, I can derive the equation describing the evolution of a firm's customer base over time.

Proposition 5. The customer base of a firm with price p is given by

$$B' = (1-\theta)B + \frac{F}{J}\left(\frac{p}{P_m}\right)^{1-\sigma} , \tag{22}$$

where B is the customer base it starts the period with, F is the measure of unattached households, and P_m is the price index of the product type, given by

$$P_m = \left[\frac{1}{J}\int_j p_j^{1-\sigma}\right]^{\frac{1}{1-\sigma}} . \tag{23}$$

Both the Proposition above and the preceding Lemma are identical to the ones derived in the static model. The only difference is that the measure of unattached households, F, is endogenous and is greater than θ in equilibrium due to firm exit. Differently from the static model, the presence of imperfectly substitutable product types makes the elasticity of demand of locked-in customers greater than 1. The following Lemma presents the demand of a household conditional on its taste shock toward the chosen product.

Lemma 3. The demand of a household for the chosen product in product type m is given by

$$c^i_{j_m} = \exp\left(\frac{\eta-1}{\sigma-1}\epsilon^i_{j_m}\right)\left(\frac{p_{j_m}}{P}\right)^{-\eta} C ,$$

where p_{j_m} is the price of the chosen product, $\epsilon^i_{j_m}$ is the idiosyncratic taste shock of the household for this product, and P is the aggregate price index given by

$$P = \left[\int_0^1 \exp\left(\frac{\eta-1}{\sigma-1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta}\, dm\right]^{\frac{1}{1-\eta}} .$$

Lemma 3 shows that the demand elasticity for the chosen product is equal to η. In addition, the amount of expenditure allocated to purchase the product depends on the relative taste toward the product. The distribution of tastes toward all other chosen products appears in the aggregate price index P.

As I study a symmetric equilibrium of the economy, I turn my focus to the case where the joint distribution of prices and tastes across products is identical across all product types. By assumption, the taste of a household toward a product it is locked into satisfies $e^{\frac{\eta-1}{\sigma-1}\epsilon} = \Gamma\left(1-\frac{\eta-1}{\sigma-1}\right)$. The following lemma derives the average adjusted taste of a household toward a product it chooses freely, conditional on the price of that product and the distribution of prices of other products in the same product type.

Lemma 4. The average adjusted taste shock of a household that freely chooses to consume a product with price p is given by,

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = \bar{p}\right] = \left(\frac{p}{P_m}\right)^{\eta-1}\Gamma\left(1-\frac{\eta-1}{\sigma-1}\right) ,$$

where $\Gamma(\cdot)$ is the Gamma function.

Lemma 4 shows how the average taste shock among consumers depend on the firm's price. A firm with a low price attracts consumers with a relatively low taste shock, while a firm that sets a high price only attracts households with relatively higher tastes toward their good. Note that, as mentioned earlier, if all prices are equal then the expected taste toward products chosen freely is equal to the taste of the household toward products it is locked into. Using Lemma 4 and the symmetry of product types, I derive the aggregate price index of households as a function of the price distribution across firms. This result is presented in the following Lemma.

Lemma 5. The aggregate price index P is given by,

$$P = \Gamma\left(1 - \frac{\eta - 1}{\sigma - 1}\right)^{\frac{1}{1-\eta}} \left[F P_m^{1-\eta} + (1-F) P_B^{1-\eta}\right]^{\frac{1}{1-\eta}} \ ,$$

where P_m is the product type price index defined above, and P_B is the initial-customer-base-weighted price index given by

$$P_B = \left[\frac{1}{1-F} \int_0^J (1-\theta) B_j p_j^{1-\eta} dj\right]^{\frac{1}{1-\eta}} \ .$$

Note that even though different households are locked into different products across different product types, the continuum of products and product types together with the symmetry of product types implies that all households face the same aggregate price index. The explicit formula for the aggregate price index allows me to aggregate household-level demand to obtain the demand function faced by the firm.

Proposition 6. The demand function faced by each firm is given by

$$y_j = \left[(1-\theta) B p_j^{-\eta} + \frac{F}{M} P_m^{\sigma-\eta} p_j^{-\sigma}\right] \frac{1}{F P_m^{1-\eta} + (1-F) P_B^{1-\eta}} (w + \Pi) \ . \tag{24}$$

The proof of Proposition 6, which is included in the Appendix, reveals that a firm needs to take into account three key trade-offs when choosing its price level. First, its choice of price affects the size of its customer base as indicated in Lemma 5. Second, its price relative to the product-type price index affects, through selection, the average taste shock of its customers toward its product. Third, its price relative to the aggregate price index affects the share of expenditure households allocate to the purchase of its product. With the firm's demand function derived, I can turn to the supply side of the economy.

4.1.2 Supply

Each firm produces a single product in a specific product type. Firms use a linear technology, with labor as the sole input of production. The production of firm j is given by,

$$y_j = a_j l_j \ ,$$

where a_j is the idiosyncratic productivity of firm j and l_j is the labor used for production. Productivity follows an AR(1) process in logs, so that

$$\ln a_j = \rho_a \ln a_j^L + \sigma_a \zeta_j \ ,$$

where ρ_a is the persistence parameter, a_j^L is the productivity level in the previous period, and ζ_j is an iid productivity shock drawn from a standard normal distribution. I denote the conditional productivity distribution by $H(a'|a)$. In the remainder of this section, I omit the j subscript for the individual firm to save on notation.

In each period the firm operates, it needs to pay a fixed operating cost x_o. The operating cost is stochastic and follows an iid log-normal distribution with mean μ_o and standard deviation σ_o. I denote this distribution by $G_o(\cdot)$. After observing the fixed operating cost, the firm can choose to exit the economy instead of paying it.

The timing is as follows. At the beginning of each period, a firm observes its fixed operating cost and productivity level for the period. It then decides whether to pay the cost and operate, or whether to exit the economy. Finally, if the firm decides to operate, it chooses production, labor, and the price it charges. Firms cannot price discriminate, and must charge the same price from all their customers.

A firm has two state variables, its current productivity level (a) and its customer base it carries from the previous period (B). It discounts future profits according to the endogenous interest rate, r. The Bellman equation for the firm's present value, after paying the fixed cost for the current period, is given by

$$\begin{aligned} V(B, a) = \max_{\{p,y,B'\}} \ & py - \frac{w}{a}y + \frac{1}{1+r}\mathbb{E}\left[\max\left\{V\left(B', a'\right) - x_o' w', 0\right\}\right] \\ \text{s.t.} \quad & B' = (1-\theta)B + \frac{F}{J}\left(\frac{p}{P_m}\right)^{1-\sigma} , \\ & y = \left[(1-\theta)Bp^{-\eta} + \frac{F}{J}{P_m}^{\sigma-\eta}p^{-\sigma}\right]\frac{1}{FP_m^{1-\eta} + (1-F)P_B^{1-\eta}}(w + \Pi) \ , \end{aligned}$$

where F is the measure of unattached consumers and J is the measure of operating firms in the product type. P_m and P_B are the price indicies defined in the previous subsection.

Entrants. Potential entrants can join the economy for a fixed entry cost f_e. Upon paying this cost, a potential entrant draws a productivity level from the distribution $H_e(a)$ and observes its stochastic operating cost for the period. It then decides whether to enter the economy and begin operating or not. Entrants join the economy with no initial customer base ($B = 0$). A free entry condition implies that

$$f_e w = \int\int \max\left\{V(0, a) - x_o w, 0\right\} dG_o(x_o) dH_e(a) \ ,$$

if the measure of potential entrants is positive.[23] In the estimation of the model, I assume that entrants' initial prodcutivity is drawn from a log-normal distribution with mean 0 and variance σ_a^2. So an entrant productivity distribution is the same as an incumbent with a previous productivity equal to 1.

To close the model, I assume that a mutual fund holds the shares of all firms in the economy. The mutual fund discounts payoffs with the discount factor of households in the economy. So the interest rate satisfies the following equation,

$$\frac{1}{1+r} = \beta \left(\frac{w+\Pi}{w'+\Pi'}\right)\left(\frac{P'}{P}\right) ,$$

where I have used $C = \frac{w+\Pi}{P}$ and the assumption of $\ln$ utility. The profits (or losses) of the mutual fund, denoted by Π, are rebated in a lump-sum fashion to households.[24] The aggregate profits are defined as follows,

$$\Pi = \int_0^1 \int_0^J \pi_{jm} dj dm - J_e f_e w , \tag{25}$$

where J_e is the measure of potential entrants which pay the entry costs. π_{jm} denotes the operating profits of firm jm which are given by $\pi_{jm} = \left(p_{jm} - \frac{w}{a_{jm}}\right) y_{jm} - x_o^{jm} w$, where x_o^{jm} is the fixed operating cost firm jm incurs.

Solving the firm's problem, I derive an implicit characterization of the firm's markup, defined by $\mu \equiv \frac{p}{w/a}$, as a function of its state variables.

Proposition 7. The markup of a firm with customer base B and productivity a is implicitly defined by the following equation,

$$\mu = \frac{\sigma}{\sigma-1} + \underbrace{\alpha\left(\frac{\eta}{\eta-1} - \frac{\sigma}{\sigma-1}\right)}_{\text{harvesting motive}} \underbrace{- (1-\alpha)\sum_{\tau=1}^{\infty} (\beta(1-\theta))^{\tau}\, \mathbb{E}\left[\mathbb{1}\left(\overline{T} > \tau\right)\gamma_{+\tau}\right]}_{\text{investing motive}} , \tag{26}$$

where $\overline{T}$ is the stopping time indicating that the firm exits the economy in $\overline{T}$ periods. $\mathbb{1}\left(\overline{T} > \tau\right)$ is an indicator that takes the value 1 if the firm still operates τ periods ahead. The variables α and γ_τ are given by

$$\alpha = \frac{(\eta-1)(1-\theta)B}{(\eta-1)(1-\theta)B + (\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{\eta-\sigma}}, \qquad \gamma_{+\tau} = \frac{w_{+\tau}}{a_{+\tau}}\frac{a}{w}\left(\frac{p_{+\tau}}{P_{+\tau}}\right)^{-\eta}\left(\frac{p}{P}\right)^{\eta}(\mu_{+\tau}-1) ,$$

where a variable with subscript $+\tau$ corresponds to the variable τ periods ahead.

[23] The entry cost can be higher than the value of being a potential entrant in equilibrium in case the measure of entrants is equal to zero. In the stationary equilibrium, however, the measure of entrants is positive regardless of the parameter specification.

[24] This assumption is isomorphic to one where households can buy and hold directly claims to the profits of firms.

The Proposition above sheds light on the trade-off a firm faces when choosing its markup. The first term in equation (26), $\frac{\sigma}{\sigma-1}$, is the markup a firm would charge if the economy did not feature consumer inertia. The second term corresponds to the *harvesting motive*. The firm has an incentive to exploit its locked-in customer base and raise its markup. α is between 0 to 1 and is increasing in the share of locked-in costumers of the firm. When the firm does not have any locked-in customers $\alpha = 0$, and when the firm only sells to locked-in customers $\alpha = 1$. The higher the share of locked-in customers of a firm, the stronger the harvesting motive is. If the firm could only sell to locked-in customers, it would charge a markup of $\frac{\eta}{\eta-1} > \frac{\sigma}{\sigma-1}$. Note that even locked-in customers have the option to shift away their expenditure to other product types, which limits the ability of firms to exploit them.

Finally, the third term in equation (26) is the *investing motive*. Firms have an incentive to lower markups in order to attract customers. The customers the firm carries with it allow it to set higher markups and increase profits in following periods. In every period, the firm losses a fraction θ of its locked-in customer base, the customers who get to re-optimize their choice. So the benefits from building a customer base are discounted at rate $\beta(1-\theta)$. The higher the share of locked-in customers of a firm, the weaker is the investing motive. This is because firms cannot price discriminate, and pushing down their markups to attract new customers lowers their profit gains from locked-in customers. So the investing motive is strongest for entrants, who have no initial customer base.

4.2 Stationary Markov-Perfect Equilibrium

I restrict attention to a symmetric equilibrium where all product types are identical. The measure of operating firms in each product type is equal to J, and the price index in each product type is denoted by the scalar P_m.

When describing the environment in the previous section, I referred to firms by their index j. For the equilibrium definition, it is useful to formally define the joint distribution of firms over the two dimensions of heterogeneity: the customer base and productivity levels.

I denote the joint distribution of incumbents across customer base and productivity levels at the beginning of the period, prior to making their exit decision, by $\Lambda(B,a)$. Four components characterize the law of motion for this joint distribution: (i) the exit decision of firms, (ii) the pricing decision of firms, (iii) the measure of entrants, and (iv) the exogenous law of motion for productivity. The law of motion for the joint distribution is defined as follows. For all Borel sets $\mathcal{B} \times \mathcal{A} \subset \mathbb{R}^+ \times \mathbb{R}^+$,

$$\Lambda'(\mathcal{B} \times \mathcal{A}) = \int_{x_o} \int_{\mathscr{B}(B,a,x_o)} \int_{a' \in \mathcal{A}} dH(a'|a) d\Lambda(B,a) dG_o(x_o) + \mathbb{1}(0 \in \mathcal{B}) J'_e \int_{a' \in \mathcal{A}} dH_e(a'), \quad (27)$$

and

$$\mathscr{B}(B,a,x_o) = \left\{(B,a) \quad \text{s.t.} \quad V(B,a) \geq x_o \quad \text{and} \quad B'(B,a) \in \mathcal{B}\right\},$$

where $B'(B,a)$ is the chosen customer base of a firm with initial customer base B and productivity a.

In the stationary equilibrium, the distribution $\Lambda(B,a)$ is constant over time. The definition of the stationary Markov-perfect equilibrium is as follows.

Definition 2 (Equilibrium)**.** A stationary Markov-perfect equilibrium is a set of prices $\{P_m, P_b, P, w\}$, aggregate allocations $\{\Pi, C\}$, policy functions, and a distribution of firms over customer base and productivity levels, $\Lambda(B,a)$, such that:

1. Consumption policy functions solve the household's problem.
2. Pricing policy functions solve the firm's problem.
3. Free entry condition holds.
4. Price indicies equations hold.
5. Markets clear.
6. The distribution of firms is stationary.

4.3 Calibration

The quantitative model contains nine structural parameters. The discount factor β. The re-optimization probability θ. The parameters that govern the elasticity of substitutions between product within the same product type and across product types, σ and η. The two parameters that shape the distribution of fixed operating cost, μ_o and σ_o. The two parameters that shape the productivity distribution, ρ_a and σ_a. And the entry cost f_e.

There is one degree of freedom in the selection of the structural parameters. Scaling f_e, μ_o, and σ_o by a constant, only alters, in equilibrium, the average size of firms and the total measure of firms. It does not change the pricing decision of firms, the exit probability of firms, the entry rate in the economy, or any other endogenous variable. So I normalize the measure of firms in the stationary equilibrium to 1, and calibrate f_e so that the free entry condition holds. This leaves me with eight structural parameters.

One period is assumed to be one year. To be consistent with most of the literature, I calibrate the discount factor β to 0.95. The seven remaining parameters are calibrated so that the calibrated economy matches features of the US economy between 1987–1991. I choose this 5 year period as some of the moments I use for the calibration are only available beginning in 1987.

One advantage of the quantitative model is that I can directly calibrate the degree of consumer inertia $(1-\theta)$ using the micro data. I use my estimates of consumer inertia together with the age composition of the US between 1987–1991 to calibrate the re-optimization probability θ. In particular, I set

$$\theta = \sum_{ag} S_{ag}\hat{\theta}_{ag}$$

where ag is one of four age groups (20–34, 35–49, 50–64, 65+). S_{ag} is the share of US population above 20 years old that is in the age group ag. $\hat{\theta}_{ag}$ are the benchmark estimates of the re-optimization probabilities from Section 3.4.

I calibrate the remaining parameters to match three features of the US business sector in 1987–1991: (i) the age distribution of firms in the economy, (ii) the average size of firms of different ages, and (iii) the aggregate profitability share in the economy. For the age distribution and average size of firms according to age, I use the Business Dynamics Statistics dataset. I categorize firms into four age groups: 0–1, 2–5, 6–10, 11+.[25] I match the average share in each age category, resulting in three moments. For each age category, I compute the log difference of the average size of a firm relative to the youngest age group. This results in another three moments. These six moments are presented in Figure 6. Finally, I use the benchmark estimates of the aggregate profits share from Barkai and Benzell (2018) for this time period. The average hp-filtered aggregate profits share for that time period is 5.6%. The weighting matrix I use is diagonal, with the diagonal elements equal to the inverse of the squared targeted moments. In other words, I minimize the sum of the squared percentage deviations from each moment. I discuss identification in the next section, after presenting the model fit.

Since the level of consumer inertia is estimated for the consumer packaged goods sector, as a robustness check I estimate the parameter θ via GMM together with the other six parameters. The results are presented in Appendix B.3. Strikingly, the estimated level of consumer inertia is not substantially different from the benchmark level of consumer inertia, which is calibrated using the micro data. The estimated level of consumer inertia is 0.129, compared to the benchmark level of 0.146.

4.4 Results

Table 5 presents the calibrated values of the different structural parameters. The re-optimization probability is calibrated to 14.6%. This parameter is calibrated using my micro-estimates of consumer inertia and the age composition in the US between 1987 and 1991. The average share of young households (20–34) in the population in that time period is 36%.

The six remaining parameters are calibrated via a GMM procedure. Figure 6 presents the model fit of the firm age distribution and relative size of different age groups. In addition to these moments, the model yields an aggregate profits share of 5.2%, compared to the targeted 4.8%. The left panel presents the age distribution of firms in the model and in the data. The model does a decent job at matching the firm age distribution. The three targeted moments are the share of

[25] The division between 6–10 and 11+ can only be done beginning in the year 1987. For this reason, I use the period 1987–1991 to calibrate the model.

Table 5: Calibrated Parameters

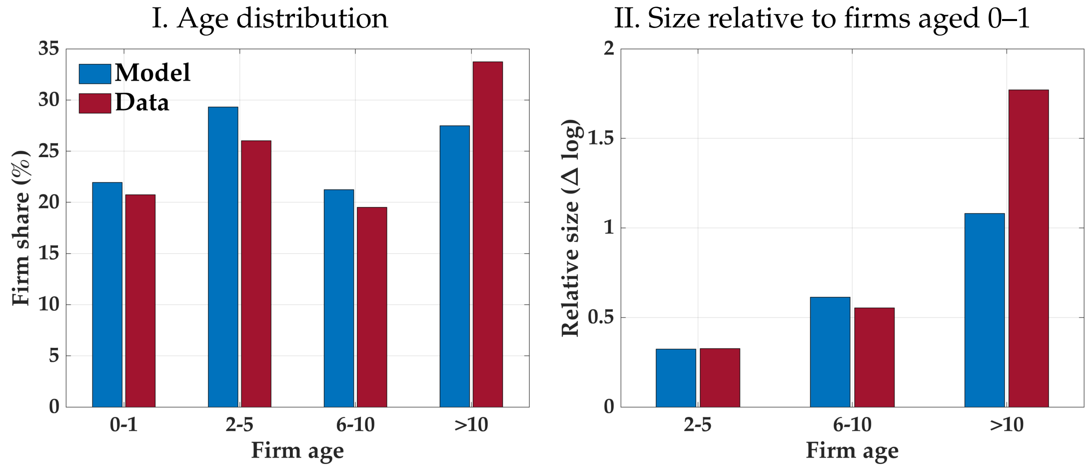
Parameter	Description	Value	Calibration method
β	Discount factor	0.95	Pre-set
θ	Re-optimization probability	0.146	Micro-estimates
μ_o	Average ln fixed operating costs	-3.94	GMM
σ_o	Std of ln fixed operating costs	3.33	GMM
ρ_a	Idiosyncratic productivity persistence	0.98	GMM
σ_a	Std of productivity shocks	0.06	GMM
η	Elasticity of sub. between product types	2.66	GMM
σ	Elasticity of sub. within a product type	3.63	GMM

Notes: This table presents the value of the eight calibrated structural parameters. The bottom six parameters were set to match firm dynamics moments and the aggregate profits share in the US between 1987–1991.

firms which are 0–1, 2–5, and 6–10 years old. The right panel of the figure presents the size of firms of different ages relative to firms aged 0–1. The model matches well the relative size of firms aged 2–5 and 6–10, but does not manage to match the relative size of firms older than 10 years old (old firms). In particular, old firms in the data are 6 times larger than firms aged 0–1 (log difference of 1.8). In the model, old firms are only about 3 times as big (log difference of 1.1). One possible explanation for the inability of the model to match the relative size of old firms is that, in the model, each firm produces a single product. In the data, on the other hand, firms expand the variety of products they sell over their life-cycle.

Identification. While the six structural parameters are jointly estimated via GMM, it is useful to discuss which moments help pin down each parameter. I start by discussing the two fixed cost parameter μ_o and σ_o. The parameter μ_o governs the average exit rate. A higher average fixed operating costs, other things equal, increases the exit probability of all firms. Since firms build their customer base over their life-cycle, the present value of older firms is higher. As a result, older firms can absorb higher operating cost shocks and survive with higher probability. The parameter σ_o governs the shape of the survival probability as a function of age. A low volatility of operating costs imply that the survival probability of old firms is much higher than of young firms. Taking the limit of $\sigma_o \to \infty$, the survival probability becomes independent of age. In the stationary equilibrium, the age distribution of firms is determined by the survival probability of firms of different ages. So these moments help determine the values of μ_o and σ_o, -3.9 and 3.3, respectively. These parameters imply that young firms fixed operating costs are equal to between 12%–13% of their sales. For firms older than 10 years old, the average fixed costs amount to about 10% of sales.

Figure 6: Model Fit



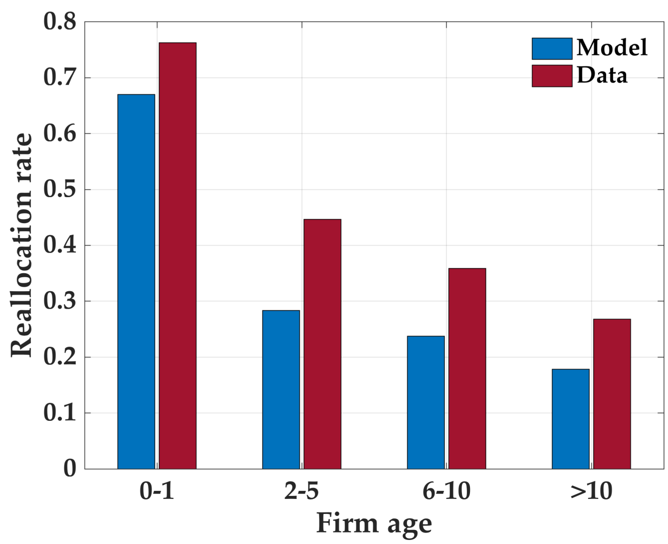
Notes: This figure presents the model fit. The left panel presents the age distribution in the calibrated model and in the data. The right panel presents the relative size in the two. The relative size is the log difference between the average size of a firm to the average size of a firm aged 0–1. In addition to these moments, the last targeted moment is the profitability share in the economy (5.2% and 4.8% in the model and data, respectively).

Next I discuss the productivity parameters ρ_a and σ_a. The persistence parameter ρ_a governs the average productivity by age. The survival probability is increasing in the level of productivity and customer base. Since the customer base increases along a firm's life-cycle, the productivity threshold for exiting the economy decreases with age. As a result, if $\rho_a = 0$, then the average productivity of an operating firm is decreasing with age. Since only the more productive firms survive, higher productivity persistence can flip this result and imply that the average productivity is increasing with age. The relative size of firms by age, helps pin down the persistence parameter ρ_a. The persistence parameter is also important for determining the shape of the survival probability by age. Higher persistence leads to higher difference between the survival probability of old and young firms. The standard deviation of productivity shocks is also important for determining the average productivity of firms by age. Trivially, without productivity shocks the productivity of all firms is the same. But also very high productivity shocks lead to a small difference in the average productivity by age. A combination of high persistence ($\rho_a = 0.98$) and relatively small volatility ($\sigma_a = 0.06$) imply that the average productivity is increasing with age, helping the model match the increasing relative size of firms with age.

The two remaining structural parameters are the ones governing the elasticities of substitution σ and η. These parameters are critical in determining the profit share in the economy. The lower σ and η are, the higher are profits. But these parameters are also important for matching the relative size of firms by age. The elasticity of substitution relevant for young firms is σ, which governs the elasticity of demand for unattached customers. For older firms, who have a higher share of locked-in customers, the more relevant elasticity of substitution is η, which governs the demand elasticity of these customers. These parameters are also key for determining the relative size of firms by age. A high σ implies that young firms set a low markup and produce a large quantity. This dampens the growth in size of young firms. The difference between σ and η determines how markups increase with age. The larger is this difference, the more markups are increasing with age. The stronger is the increase of markups with age, the lower is the size difference between old and young firms. The calibrated values of σ and η are 3.63 and 2.66, respectively. These elasticities of substitution are well within the range of elasticities estimated in Foster, Haltiwanger and Syverson (2008).

Figure 7 presents the reallocation rates by age group in the model and in the data. The reallocation rate is defined as the ratio between job creation and destruction to the total level of employment. The model captures the decreasing reallocation rate with age, and the disproportionally large reallocation rate of the 0–1 age group. However, reallocation rates are between 10–20 percentage points lower in the model relative to the data.

Figure 7: Non-Targeted Moments – Reallocation Rates

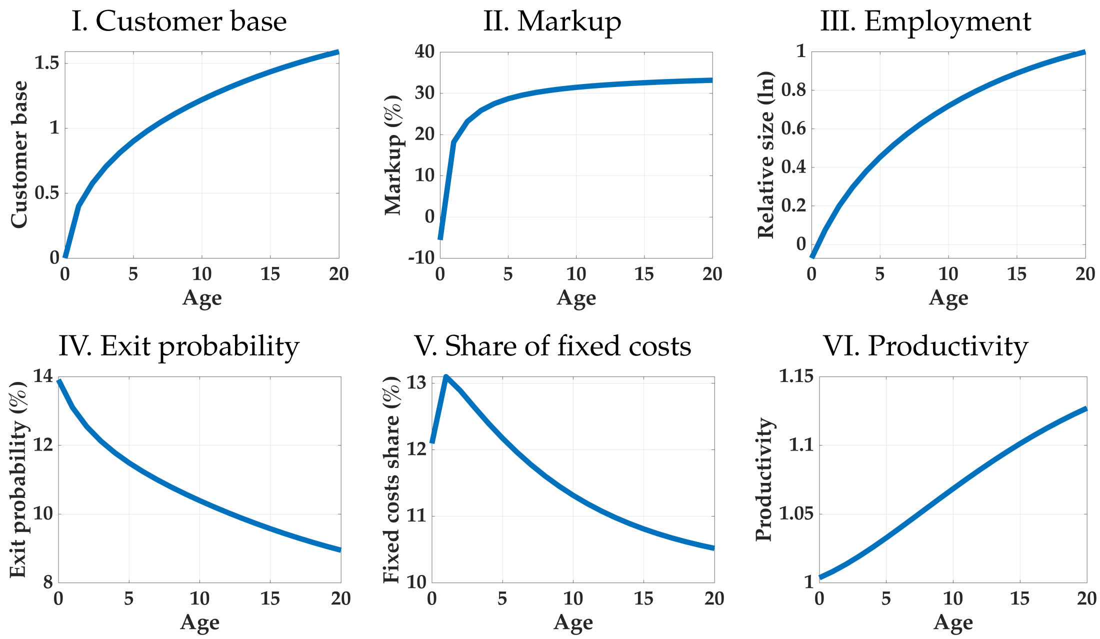


Notes: This figure presents the reallocation rates of firms of different ages in the model and in the data. These moments were not targeted by the estimation procedure. The reallocation rate is the ratio between the sum of job creation and destruction over the sum of employees.

4.5 The Calibrated Stationary Economy

I now describe the properties of the calibrated stationary economy. Figure 8 presents the life-cycle of an average firm in the economy over the first 20 years of its life. A firm enters the economy with no initial customer base. It sets a relatively low markup in order to build its customer base. As it builds its customer base, the harvesting motive turns stronger while the investing motive turns weaker. As a result, the average markup a firm sets is increasing with its age. Despite higher markups, the average size of a firm is growing with age as it sells its product to a larger number of customers.

Figure 8: The Life-Cycle of a Firm



Notes: This figure presents the properties of the average firm across the first 20 years of its life. A firm joins the economy with no customer base and builds it over time. As its customer base grows, the harvesting motive becomes more dominant and the firm raises its markup. Employment is relative to the average firm aged 0–1. Share of fixed costs is with respect to sales.

As a firm builds its customer base, its present value increases. It can absorb larger fixed operating costs, and so its exit probability is decreasing with age. However, as a fraction of sales, younger firms are willing to pay more than older firms. This can be seen in Panel V of Figure 8. Young firms are willing to pay a large fraction of their sales as they expect their profits to rise in the future. Old firms do not expect their profits to increase further, so they are willing to pay a lower fraction of their sales.[26]

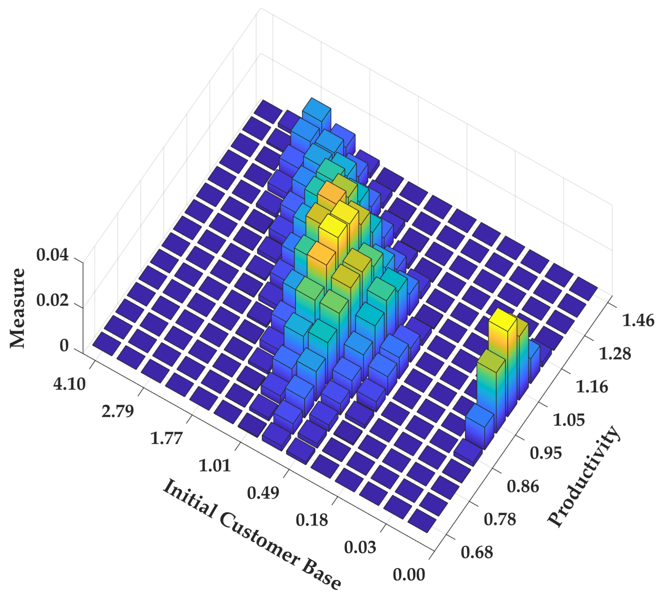
Finally, as discussed in the previous section, the high persistence of idiosyncratic productivity implies that the average productivity is increasing with the firm's age.

Figure 9 presents the stationary distribution of firms along the two heterogeneity dimensions, the initial customer base and productivity levels. For better visualization, I collapse the state

[26]The decreasing relationship between the ratio of fixed costs to sales and age is true for all ages except for entrants. Entrants are not willing to pay as much as 1-year old firms, because the high exit probability they face is stronger than the expected rise in their profits.

space into bins and present the measure of firms in each bin.[27] The firms with initial customer base equal to 0 are the entrants. Depending on their productivity level, entrants set their markup and decide how many customers to attract. More productive firms decide in equilibrium to attract more customers, so we see a positive correlation between the productivity level of a firm and its initial customer base.

Figure 9: The Stationary Distribution of Firms



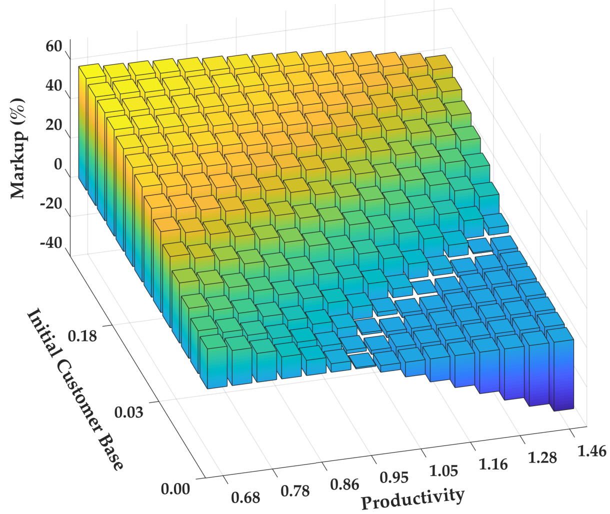
Notes: This figure presents the stationary distribution of firms along the two heterogeneity dimensions:eize the initial customer base and productivity levels. I collapse the state space into bins, and present the measure of firm in each bin. On average, more productive firms also have a higher number of customers.

Figure 10 presents the optimal markup of a firm as a function of its two state variables. Recall that the markup satisfies equation (26). In setting its markup, the firm trades off the harvesting motive and the investing motive. Entrants have no initial customer base, so the harvesting motive is shut down for them. As a firm builds its customer base, a larger fraction of its customer base are locked-in, and the harvesting motive becomes more dominant. As a result, we see that markups are increasing with a firm's customer base.

From equation (26), we see that the harvesting motive is independent of the productivity level of the firm. This is not the case for the investing motive. A more productive firm has a lower exit probability, due to the persistence in productivity. Since the investing motive is increasing in the survival probability, it is also increasing in the productivity level of the firm. Consequently, the optimal markup of a firm is decreasing in a firm's productivity level.

[27] The model is solved on a grid containing 1,001 customer base points and 15 productivity levels.

Figure 10: Markup Policy Function



Notes: This figure presents the equilibrium markup as a function of the firm's customer base and productivity levels. The markup is increasing with the firm's customer base, and decreasing with the firm's productivity.

4.6 The Aggregate Implication of The Rise in Consumer Inertia

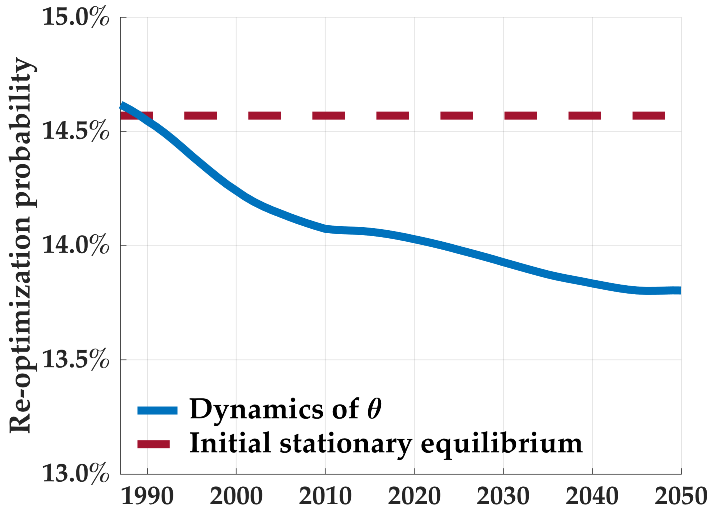
I now turn to study the aggregate implications of the aging-induced rise in consumer inertia. I consider an unexpected deterministic shock that moves the level of consumer inertia from its initial steady state value according to observed and predicted demographic shifts in the US. I use the observed and predicted age composition of the US from the World Bank together with my estimates of the re-optimization probability from Section 3.4 to construct the annual re-optimization probability. The resulting series of re-optimization probability is presented in Figure 11. The aggregate re-optimization probability drops from its initial level of 14.6% to 13.8% in 2050.

I assume that the re-optimization probability remains fixed at its 2050 level and assume the model converges to the new stationary equilibrium in 2100. I study the transition dynamics from the initial stationary distribution to the new one. In particular, I analyze how the aging-induced rise in consumer inertia affects the profits share and different firm formation indicators. The numerical algorithm to solve the transition dynamics is presented in the Appendix E.

Table 6 presents the main results of the quantitative model. I present the change in different firm formation indicators as well as in the profits share between the initial stationary equilibrium, which corresponds to the US economy in 1987–1991, and the implied average levels in the 2000–2004 period. I contrast the results obtained from the model with the observed changes in the data.

The first row in the table presents the change of the profits share in the model and in the data. Profits in the model increased from a level of 5.2% in the stationary equilibrium to 6.3% in the early 2000s, compared to an increase from 4.8% to 9.5% in the data. As for the declining share of

Figure 11: Aggregate Change in Consumer Inertia



Notes: This figure presents the change in consumer inertia, which is the exogenous input to the model when studying transition dynamics. It is constructed using the micro-estimates of consumer inertia and the observed and predicted age composition of the adult population in the US.

Table 6: The Aggregate Implications of the Rise in Consumer Inertia

	Change between late 80s and early 00s		
	Model	Data	Model accounts for:
Aggregate profits share	0.19	0.68	28%
Share of young firms	-0.08	-0.15	50%
Emp. share of young firms	-0.11	-0.26	41%
Entry rate	-0.07	-0.14	52%
Emp. share of entrants	-0.11	-0.21	51%

Notes: This table contrasts the aggregate implications of the rise in consumer inertia in the model with the observed changes in the data. In the model, the change is defined to be the log change between the initial stationary equilibrium and the average value in the 2000–2004 period. In the data, it is the log change between the average in 1987–1991 (the period to which the stationary equilibrium is calibrated) and the average in 2000–2004.

young firms, the model accounts for 50% of the observed decline in the share of young firms, and 40% of the decline in the share of workers employed by young firms. The model accounts for a similar share of the decline in entry rates. The model predicts entry rate declined by 11 log points between the late 1980s and early 2000s, compared to 21 log points in the data. When weighting by the employment share, entry rate declined by 7 log points compared to 14 log points in the data.

Decomposition of operating profits. Normalizing the wage level to 1, operating profits in the economy can be written as follows,

$$\int_j \pi_j dj = (\bar{\mu} - 1) L^D - \bar{x}_o^s (1 + \Pi) \ , \tag{28}$$

where $\bar{\mu}$ is the variable-cost-weighted markup in the economy and $\bar{x}_o^s$ is the sales-weighted share of fixed costs in sales. L^D is the measure of production workers who are not employed to pay for indirect costs such as fixed operating or entry costs.[28] The change in the operating profits in the economy is therefore composed of two components: an increase in average markups, and a decline in the share of fixed operating costs. The model allows me to quantify the contribution of each of these two components to the rise in operating profits between the late 1980s and early 2000s. Both components contributed to this rise: 60% of this rise is due to an increase in average markups and 40% due to a decrease in the share of fixed operating costs.

Decomposition of the decline in the share of young firms. Following the work of Pugsley and Şahin (2015), the decline in the share of young firms can be decomposed into two components: a change in the measure of entrants who join the economy in every period, and a change in the survival probability of operating firms. Holding constant the survival probability of firms at its initial stationary equilibrium, the decline in the entry rate accounts for 80% of the implied decline in the share of young firms. This result is in line with the empirical finding of Pugsley and Şahin (2015) that the majority of the decline in the share of young firms is driven by the decline in the entry rate, and not through a change in the survival probability of operating firms.

Welfare. The growing body of evidence on rising markups has stimulated research on their welfare implications. Prominent examples include Baqaee and Farhi (2017) and Edmond, Midrigan and Xu (2018).[29] My model is not suitable to give a full picture on the welfare costs of the rise in consumer inertia for three main reasons. First, my model features ex-ante identical households who receive an equal share of profits. So higher profits and the associated decline in the labor share does not imply a distributional cost. Second, labor is supplied inelastically. So a higher average markup in the economy does not lead to an inefficiency in the aggregate level of labor. Finally, by construction, my model assumes there is no love-of-variety.[30] If households do enjoy a larger set of options, then the decline in entry and the associated decline in the measure of operating firms entails an additional welfare cost.

Notwithstanding, my model encompasses two channels through which a rise in consumer inertia affects welfare. First, a decline in the share of fixed operating and entry costs is welfare improving as it frees up labor to produce consumption goods. Second, the rise in consumer inertia

[28] The variable-cost-weighted markup is given by $\bar{\mu} = \int_j \frac{l_j^D}{L^D} \mu_j dj$, where $l_j^D = y_j / a_j$.

[29] Baqaee and Farhi (2017) studies production misallocation in a general non-parametric framework and finds that eliminating markups would raise TFP by about 20%. Edmond et al. (2018) decompose the welfare costs of markup in a model of heterogeneous firms and variable markups.

[30] The assumption that delivers no love of variety is that the location parameter of the idiosyncratic taste shocks depends negatively on the measure of operating firms.

alters the distribution of markups across firms. Both because of the shift in the population of firms toward older firms, and because the markup policy function changes. In particular, firms with a small customer base decrease their markups and firms with a relatively large customer base increase their markups. In the model, the welfare gain due to the decline in fixed costs outweighs the misallocation costs of an increase in markup dispersion.

Consistent with the standard theory, my model predicts that firms who set a relatively higher markup are under-producing. In other words, reallocation of production resources toward firms with higher markups is welfare-improving. However, this prediction can change depending on the deep cause for consumer inertia. Consider a firm who has a large customer base and suffers a persistent negative productivity shock. In response to such shock, the firm sets a high markup, exploiting its existing locked-in customer base. What improves welfare by more: reallocation of labor toward this low productivity firm, or closing down the firm and shifting its labor to other operating firms? The answer depends on the welfare costs of switching products.

One interpretation of the model is that the welfare costs of switching are zero with probability θ, and infinity otherwise. In that case, closing down the unproductive firm would be welfare reducing as it entails large welfare costs to the measure of locked-in customers, who would have to switch products. A social planner would prefer to reallocate labor toward this low-productivity firm. Taking an alternative extreme example, suppose that consumer inertia is the result of (irrational) inattention. The welfare costs of switching are zero, but there is a probability $1-\theta$ a household is locked-in. In such a case, the social planner may prefer that the firm exits the economy, freeing up its locked-in customers. These two examples illustrate why one needs to take a stand on the deep cause of consumer inertia to perform a normative analysis. Formally studying the welfare costs of markups in the presence of consumer inertia is left for future research.

5 Conclusion

In this paper, I study the role of a rise in consumer inertia as a driver of the twin phenomena of declining share of young firms and rising profits during the past three decades. Using micro data on consumer behavior, I find that young households are significantly less inertial. So the aging of the baby-boom generation has led to a rise in the aggregate level of consumer inertia. Empirical evidence using variation across product categories and temporal variation across states indicates that consumer inertia has a negative effect on firm formation.

I develop a model of entry, exit, and firm dynamics in the presence of consumer inertia. I calibrate the model so that the stationary equilibrium corresponds to the US economy in the late 1980s. I then consider an unexpected and deterministic shock to the level of consumer inertia, which moves it according to observed and predicted demographic shifts in the US between the late 1980s and 2050. The model implies that the rise in consumer inertia accounts for about one half of the declining share of young firms and one third of the rise in aggregate profits between the late 1980s and early 2000s.

While the model shows how consumer inertia can raise aggregate profits and reduce the share of young firms, several questions remain open for future research. First, what are the welfare costs of rising market power and markup dispersion resulting from an increase in consumer inertia? To answer this question, one needs to take a stand on the fundamental forces underlying consumer inertia. Traditional theory suggests that firms who set a relatively high markup are under-producing, from a social welfare perspective. Suppose, however, that consumer inertia is the result of contractual pecuniary switching costs. Then, a social planner may prefer a low-productive firm with a relatively high markup to cease operation, instead of reallocating production resources toward that firm.

Another question that remains open regards the implication of the declining share of young firms for the growth rate of the economy. Garcia-Macia, Hsieh and Klenow (2016) argues that young firms are not an important driver of economic growth based on the their small share in aggregate employment. But if the average size of young firms is limited by consumer inertia, then the employment share of these firms cannot be used as a sufficient statistic to infer their contribution over time to economic growth. Understanding the welfare and growth implications of the rise in consumer inertia is left for future research.

A Mathematical Appendix

A.1 Stylized Model

Proof of Lemma 1

Lemma 1. If the household is unattached, it chooses the product that maximizes $-(\sigma-1)\ln p_j+\epsilon_j$ for $j\in(0,J)$. That is,

$$j_i = \arg\max_{j\in(0,J)} -(\sigma-1)\ln p_j + \epsilon_j \ . \tag{2}$$

Proof. The household maximizes $u\left(\exp\left(\frac{1}{\sigma-1}\epsilon_j\right)c_j\right)$. Since $u(\cdot)$ is an increasing function, the household's choice, conditional on not being locked-in with its previous product, solves the following problem

$$\max_{\{j,c_j\}} \exp\left(\frac{1}{\sigma-1}\epsilon_j\right)c_j \ ,$$
$$\text{s.t.} \quad c_j p_j = w \ .$$

Substituting the constraint the maximization problem can be rewritten as follows,

$$\max_j \exp\left(\frac{1}{\sigma-1}\epsilon_j\right)\frac{w}{p_j} = w\max_j \exp\left(\frac{1}{\sigma-1}\epsilon_j\right)\frac{1}{p_j} \ .$$

So the optimal product chosen is independent of the level of the wage. Taking logs and multiplying by $\sigma-1>0$, I have that

$$j_i = \arg\max_j -(\sigma-1)\ln p_j + \epsilon_j \ .$$

□

Proof of Proposition 1

Proposition 1. The customer base of a firm with price p is given by

$$B = (1-\theta)B_0 + \frac{\theta}{J}\left(\frac{p}{P}\right)^{1-\sigma} \ , \tag{3}$$

where B_0 is the initial customer base of the firm, and P is the price index, given by

$$P = \left[\frac{1}{J}\int_0^J p_j^{1-\sigma}dj\right]^{\frac{1}{1-\sigma}} \ . \tag{4}$$

Proof. The first term of equation (3) is the share of existing customers who cannot re-optimize their consumption choice. It is simply a share $1-\theta$ of the initial customer base of the firm, B_0.

The second term represents the measure of new customers who actively choose the product. Overall, there is a measure θ of customers who can re-optimize their product choice. I start by considering the probability of a single consumer buying the product given the price distribution and a finite number of firms, J. I then take the limit as the number of firms and number of consumers goes to infinity.

Suppose there are J firms operating. The conditional probability of product j being chosen given that its taste shock is equal to $\bar{\epsilon}$ is

$$\begin{aligned} Pr(j_i = j|\epsilon_j = \bar{\epsilon}) &= \prod_{j' \neq j} Pr\left[-(\sigma-1)\ln p_j + \bar{\epsilon} > -(\sigma-1)\ln p_{j'} + \epsilon_{j'}\right] \\ &= \prod_{j' \neq j} Pr\left[\epsilon_{j'} < (\sigma-1)(\ln p_{j'} - \ln p_j) + \bar{\epsilon}\right] \; . \end{aligned}$$

Using the CDF of the Gumbel distribution and the independence of taste draws I obtain

$$Pr(j_i = j|\epsilon_j = \bar{\epsilon}) = \prod_{j' \neq j} e^{-e^{-\left[(\sigma-1)(\ln p_{j'} - \ln p_j) + \bar{\epsilon}\right]}} \; .$$

Using this equation, I can find the unconditional probability product j is chosen:

$$Pr(j_i = j) = \int_{-\infty}^{\infty} \left(\prod_{j' \neq j} e^{-e^{-\epsilon} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} \right) e^{-e^{-\epsilon}} e^{-\epsilon} d\epsilon$$

If $j' = j$ the term that would appear in the product is $e^{-e^{-\epsilon}}$, so I can multiply the product by $e^{e^{-\epsilon}}$ and let it include all terms:

$$\begin{aligned} Pr(j_i = j) &= \int_{-\infty}^{\infty} \left(\prod_{j'} e^{-e^{-\epsilon} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} \right) e^{e^{-\epsilon}} e^{-e^{-\epsilon}} e^{-\epsilon} d\epsilon \\ &= \int_{-\infty}^{\infty} \left(\prod_{j'} e^{-e^{-\epsilon} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} \right) e^{-\epsilon} d\epsilon \; . \end{aligned}$$

The product can be simplified as follows

$$\prod_{j'} e^{-e^{-\epsilon} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} = e^{-e^{-\epsilon} \sum_{j'} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} = e^{-Q e^{-\epsilon}} \; ,$$

where $Q = \sum_{j'} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}$. Substituting into the integral I have

$$Pr(j_i = j) = \int_{-\infty}^{\infty} e^{-Q e^{-\epsilon}} e^{-\epsilon} d\epsilon \; .$$

I use a change of variables: let $x \equiv e^{-\epsilon}$. The inverse transformation is $\epsilon = -\ln x$, and I have $\frac{d\epsilon}{dx} = -\frac{1}{x}$. So that

$$Pr(j_i = j) = \int_{0}^{\infty} e^{-xQ} dx \; .$$

Now I can integrate to obtain

$$Pr(j_i = j) = \Big|_0^{\infty} -\frac{1}{Q} e^{-xQ} = \frac{1}{Q} \; .$$

That is, we have that

$$Pr(j_i = j) = \frac{1}{\sum_{j'} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} \ .$$

Multiplying both the denominator and the numerator by $e^{-(\sigma-1)\ln p_j}$ the expression above simplifies to

$$Pr(j_i = j) = \frac{e^{-(\sigma-1)\ln p_j}}{\sum_{j'} e^{-(\sigma-1)\ln p_{j'}}} = \frac{p_j^{1-\sigma}}{\sum_{j'} p_{j'}^{1-\sigma}} \ .$$

Taking $J \to \infty$ we have that the density of product j being chosen by the household is

$$f(j_i = j) = \frac{p_j^{1-\sigma}}{\int_0^J p_{j'}^{1-\sigma} dj'} \ .$$

The result above is well known. Defining the price index P, the density of product j being chosen is

$$f(j_i = j) = \frac{1}{J}\left(\frac{p_j}{P}\right)^{1-\sigma} \ ,$$

where

$$P = \left(\frac{1}{J}\int_0^{J_m} p_{j'}^{1-\sigma} dj'\right)^{\frac{1}{1-\sigma}} \ .$$

Since the measure of households who can re-optimize their consumption choice is equal to θ, the measure of households who choose a product with price p is given by

$$\frac{\theta}{J}\left(\frac{p}{P_m}\right)^{1-\sigma} \ .$$

This is the second term of equation (3), thus completing the proof. □

Proof of Lemma 2

Proposition 2. The markup charged by the firm, $\mu \equiv p/w$, satisfies the following equation,

$$\mu = \frac{\sigma}{\sigma-1} + B_0 \frac{1}{\sigma-1}\frac{1-\theta}{\theta} J \left(\frac{P}{w}\right)^{1-\sigma} \mu^{\sigma-1} \ . \tag{6}$$

Proof. The firm's problem, substituting labor using the linear production technology, is given by

$$\begin{aligned} V(B_0) = \max_{\{p,y\}} \quad & py - wy \ , \\ \text{s.t.} \quad & y = (1-\theta) B_0 w p^{-1} + \frac{\theta}{J} P^{\sigma-1} w p^{-\sigma} \ . \end{aligned}$$

The first order conditions are given by

$$\begin{aligned} [y] \qquad & p - w = \lambda \ , \\ [p] \qquad & \frac{py}{\lambda} = (1-\theta) B_0 w p^{-1} + \sigma \frac{\theta}{J} P^{\sigma-1} w p^{-\sigma} \ . \end{aligned}$$

Dividing the second equation by y and using the demand function we have

$$\frac{p}{\lambda}=\frac{(1-\theta)\,B_0wp^{-1}+\sigma\frac{\theta}{J}P^{\sigma-1}wp^{-\sigma}}{(1-\theta)\,B_0wp^{-1}+\frac{\theta}{J}P^{\sigma-1}wp^{-\sigma}}\ .$$

Substituting λ from the first order condition w.r.t y, and simplifying the fraction on the RHS we have

$$\frac{p}{p-w}=\frac{(1-\theta)\,B_0p^{-1}+\sigma\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}{(1-\theta)\,B_0p^{-1}+\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}\ .$$

Raising to the power of -1, the expression becomes

$$1-\frac{w}{p}=\frac{(1-\theta)\,B_0p^{-1}+\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}{(1-\theta)\,B_0p^{-1}+\sigma\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}\ ,$$

so that

$$\frac{w}{p}=1-\frac{(1-\theta)\,B_0p^{-1}+\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}{(1-\theta)\,B_0p^{-1}+\sigma\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}\ .$$

The expression in the RHS can be simplified so that

$$\frac{w}{p}=\frac{(\sigma-1)\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}{(1-\theta)\,B_0p^{-1}+\sigma\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}\ .$$

Raising again to the power of -1 and defining the markup $\mu=\frac{p}{w}$, we have

$$\mu=\frac{(1-\theta)\,B_0p^{-1}+\sigma\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}{(\sigma-1)\frac{\theta}{J}P^{\sigma-1}p^{-\sigma}}\ .$$

Finally the expression on the RHS simplifies to

$$\mu=\frac{\sigma}{\sigma-1}+B_0\frac{1}{\sigma-1}\frac{1-\theta}{\theta}J\left(\frac{P}{w}\right)^{1-\sigma}\mu^{\sigma-1}\ .$$

□

Proof of Proposition 3

Proposition 3. There exists a unique equilibrium. In addition, there is a cutoff $\bar{f}_e$ such that if and only if $f_e<\bar{f}_e$ then the measure of entrants is strictly positive.

Proof. I start by assuming that the entry condition holds with equality. There are four endogenous variables, $\{\mu_E,\mu_I,J,P\}$, that need to satisfy the following four equilibrium equations,

$$[\mu_E]\qquad \mu_E=\frac{\sigma}{\sigma-1}\ ,$$

$$[\mu_I]\qquad \mu_I=\frac{\sigma}{\sigma-1}+\frac{1}{\sigma-1}\frac{1-\theta}{\theta}JP^{1-\sigma}\mu_I^{\sigma-1}\ ,$$

$$[J]\qquad f_e=\frac{1}{\sigma-1}\frac{\theta}{J}P^{\sigma-1}\mu_E^{-\sigma}\ ,$$

$$[P]\qquad P^{1-\sigma}=\frac{1}{J}\int_0^J\mu_j^{1-\sigma}dj\ .$$

The last equation can be simplified to

$$JP^{1-\sigma} = \left[\mu_I^{1-\sigma} + (J-1)\mu_E^{1-\sigma}\right] \ .$$

Using the free entry condition, [J], together with the expression for μ_E we have that

$$JP^{1-\sigma} = \frac{\theta}{\sigma f_e}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}$$

Plugging into equation [μ_I] I obtain

$$\mu_I = \frac{\sigma}{\sigma-1} + (1-\theta)\frac{1}{(\sigma-1)\sigma}\frac{1}{f_e}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\mu_I^{\sigma-1} \ .$$

Multiplying by $\mu_I^{1-\sigma}$ we have

$$\mu_I^{2-\sigma} = \mu_E\mu_I^{1-\sigma} + (1-\theta)\frac{1}{(\sigma-1)\sigma}\frac{1}{f_e}\mu_E^{1-\sigma} \ . \tag{29}$$

This is an equation that uniquely determines μ_I. The LHS is increasing in μ_I and the RHS is decreasing in it. So there exists a unique equilibrium for μ_I. Note also that $\mu_I > \mu_E = \frac{\sigma}{\sigma-1}$, as when $\mu_I = \mu_E$ we have that the RHS is larger than the LHS.

The equation that determines J is given by

$$\left[\mu_I^{1-\sigma} + (J-1)\mu_E^{1-\sigma}\right] = \theta\frac{1}{(\sigma-1)f_e}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \ .$$

We can rearrange it as follows,

$$(J-1)\mu_E^{1-\sigma} = \theta\frac{1}{(\sigma-1)f_e}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} - \mu_I^{1-\sigma} \ .$$

Note that if f_e increases, then μ_I goes down from the implicit equation that determines μ_I. So the LHS increases and RHS decreases. As a result, the equilibrium level of J must decrease. Let $\bar{f}_e$ be the value that makes the resulting $J = 1$. If $f_e < \bar{f}_e$, the measure of entrants is positive as $J > 1$ and we obtained the equilibrium allocations. P is obtained from equation (4).

If $f_e \geq \bar{f}_e$, then there are no entrants in equilibrium. In that case, $P = \mu_I$ and $J = 1$. The equation that pins down μ_I is

$$\mu_I = \frac{\sigma}{\sigma-1} + \frac{1}{\sigma-1}\frac{1-\theta}{\theta} \ .$$

Thus, also in the case where $f_e \geq \bar{f}_e$ the equilibrium is unique. □

Proof of Proposition 4

Proposition 4. [Comparative statics] Consider two economies, A and B. Suppose the two economies differ by only one structural parameter. The following comparative statics hold:

Parameter	Entrants' markup	Incumbents' markup	Profits share	Entry rate
$\theta_A < \theta_B$	$\mu_E^A = \mu_E^B$	$\mu_I^A > \mu_I^B$	$\Pi_A > \Pi_B$	$E_A < E_B$
$f_e^A > f_e^B$	$\mu_E^A = \mu_E^B$	$\mu_I^A < \mu_I^B$	$\Pi_A > \Pi_B$	$E_A < E_B$

Proof. The equation that determines incumbents markup is (29):

$$\mu_I^{2-\sigma} = \mu_E \mu_I^{1-\sigma} + (1-\theta)\frac{1}{(\sigma-1)\sigma}\frac{1}{f_e}\mu_E^{1-\sigma} \, .$$

To see what happens to the measure of entrants we look at J which satisfies,

$$\mu_I^{1-\sigma} + (J-1)\mu_E^{1-\sigma} = \theta\frac{1}{\sigma-1}\frac{1}{f_e}\mu_E^{1-\sigma} \, . \tag{30}$$

Now we can turn to the two different cases:

1. If θ increases, then μ_I goes down as the RHS of (30) goes down. Rearranging equation (30) I obtain,

$$(J-1)\mu_E^{1-\sigma} = \theta\frac{1}{\sigma-1}\frac{1}{f_e}\mu_E^{1-\sigma} - \mu_I^{1-\sigma} \, . \tag{31}$$

Rearranging equation (29) I get

$$\theta\frac{1}{(\sigma-1)}\frac{1}{f_e}\mu_E^{1-\sigma} - \mu_I^{1-\sigma} = \frac{1}{(\sigma-1)}\frac{1}{f_e} - (\sigma-1)\left(\mu_I^{2-\sigma} - \mu_I^{1-\sigma}\right) \, .$$

Since μ_I decreases, $2-\sigma > 0$, and $1-\sigma < 0$, the RHS is increasing. This implies from equation (31) that J must go up as well.

Finally, the share of profits in the economy is equal to the profits made by incumbents as entrants make zero profits net of entry costs in equilibrium. We know that in equilibrium $JP^{1-\sigma} = \theta\frac{1}{\sigma-1}\frac{1}{f_e}\mu_E^{1-\sigma}$. So $JP^{1-\sigma}$ increased. This makes incumbents' profits go down for any given price. Similarly, a larger θ makes profits decrease for any given price. So it must be that incumbents' profits in the new equilibrium are lower than in the economy with a lower θ. Thus, the profit share in the economy is lower.

2. Consider an increase in f_e. As the RHS of equation (30) goes down, incumbents' markup in equilibrium is lower. Entrants' markup is unchanged as it only depends on σ. We see that the RHS of equation (31) goes down as f_e increases and $\mu_I^{1-\sigma}$ increases as well. As a result, the measure of firms in equilibrium, J is lower. Recall that $JP^{1-\sigma} = \frac{\theta}{\sigma f_e}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}$. So $JP^{1-\sigma}$ goes down as f_e goes up. Looking at the incumbent's problem, we see that the quantity sold for any given price is now higher. So it must be that incumbents' profits go up. That is, the share of profits in the economy goes up.

□

A.2 Quantitative Model

Proof of Lemma 2

Lemma 2. If the household is not locked into a product, it chooses the firm which maximizes

$$j_m \equiv \arg \max_{j \in (0, J_m)} -(\sigma - 1) \ln p_j + \epsilon_j \ .$$

Proof. The household maximizes the taste-adjusted consumption, $\exp\left(\frac{1}{\sigma-1}\epsilon_{j_m}\right) c_{j_m}$. For any given level of expenditure on product type m, E_m, the taste-adjusted consumption level obtained by choosing product j is given by

$$e^{\frac{1}{\sigma-1}\epsilon_j} \frac{E_m}{p_j} \ .$$

By taking logs and multiplying by $(\sigma - 1)$ we obtain that the product which maximizes the taste-adjusted consumption level is the one which maximizes

$$-(\sigma - 1) \ln p_j + \epsilon_j \ .$$

□

Proof of Proposition 5

Proposition 5. The customer base of a firm with price p is given by

$$B' = (1 - \theta) B + \frac{F}{J} \left(\frac{p}{P_m} \right)^{1-\sigma} \ , \tag{22}$$

where B is the customer base it starts the period with, F is the measure of unattached households, and P_m is the price index of the product type, given by

$$P_m = \left[\frac{1}{J} \int_j p_j^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \ . \tag{23}$$

Proof. The first term of equation (22) is the share of existing customers who cannot re-optimize their consumption choice. It is simply a share $1 - \theta$ of the existing customer base of the firm, B.

The second term represents the measure of new customers of the firm. Overall there is an endogenous measure F of consumers who can re-optimize their choice in product type m. Note that this measure will be greater than θ in equilibrium due to firm exit. The remainder of this proof proceeds in a similar fashion to the proof of Proposition 1. I want to show that conditional on re-optimizing, the density of households who choose a firm with price p is given by

$$\frac{1}{J} \left(\frac{p}{P_m} \right)^{1-\sigma} \ .$$

Obtaining this, I can multiply this density by the measure of re-optimizing households to obtain the second term of equation (22).

Suppose there are J firms in a product module. Below I show the well-known result about the probability of a household choosing firm j given prices. Using Lemma 2 we know that the household chooses the firm that maximizes $-(\sigma-1)\ln p_j + \epsilon_j$. The conditional probability of firm j being chosen by household i given that the household's taste shock towards the firm is equal to $\bar{\epsilon}$ is

$$\begin{aligned} Pr(j_m = j|\epsilon_i = \bar{\epsilon}) &= \prod_{j' \neq j} Pr\left[-(\sigma-1)\ln p_j + \bar{\epsilon} > -(\sigma-1)\ln p_{j'} + \epsilon_{j'}\right] \\ &= \prod_{j' \neq j} Pr\left[\epsilon_{j'} < (\sigma-1)(\ln p_{j'} - \ln p_j) + \bar{\epsilon}\right] \ . \end{aligned}$$

Using the CDF of the Gumbel distribution with location parameter $-\ln J$ and unit scale as well as the independence of taste draws I obtain[31]

$$Pr(j_m = j|\epsilon_i = \bar{\epsilon}) = \prod_{j' \neq j} e^{-e^{-\left[(\sigma-1)(\ln p_{j'} - \ln p_j) + \bar{\epsilon} + \ln J\right]}} \ .$$

If $j' = j$ the term that would appear in the product is $e^{-e^{-\epsilon - \ln J}}$, so

$$\left(\prod_{j' \neq j} e^{-e^{-\epsilon - \ln J} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}}\right) = \left(\prod_{j'} e^{-e^{-\epsilon - \ln J} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}}\right) e^{e^{-\epsilon - \ln J}} \ .$$

Note that we can rearrange the last term above so that we have,

$$Pr(j_m = j|\epsilon_i = \bar{\epsilon}) = \left(e^{-e^{-\epsilon} \frac{1}{J} \sum_{j'} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}}\right) e^{e^{-\epsilon - \ln J}} = e^{-Qe^{-\epsilon}} e^{e^{-\epsilon - \ln J}} \ , \tag{32}$$

where I define $Q \equiv \frac{1}{J} \sum_{j'} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}$.

Using the equation above, I can find the unconditional probability product j is chosen using the Gumbel distribution of taste shocks:

$$Pr(j_m = j) = \int_{-\infty}^{\infty} e^{-Qe^{-\epsilon}} e^{e^{-\epsilon - \ln J}} e^{-e^{-\epsilon - \ln J}} e^{-\epsilon - \ln J} d\epsilon \ ,$$

which simplifies to

$$Pr(j_m = j) = \frac{1}{J} \int_{-\infty}^{\infty} e^{-Qe^{-\epsilon}} e^{-\epsilon} d\epsilon \ .$$

I use a change of variables: let $x \equiv e^{-\epsilon}$. The inverse transformation is $\epsilon = -\ln x$, and I have $\frac{d\epsilon}{dx} = -\frac{1}{x}$. So that

$$Pr(j_m = j) = \frac{1}{J} \int_{0}^{\infty} e^{-xQ} dx \ .$$

Now I can easily integrate to obtain

$$Pr(j_m = j) = \frac{1}{J} \Big|_0^{\infty} - \frac{1}{Q} e^{-xQ} = \frac{1}{J} \frac{1}{Q} \ . \tag{33}$$

[31] The location parameter is actually irrelevant for this result as all taste shocks have the same location parameter.

That is, we have that

$$Pr(j_m = j) = \frac{1}{\sum_{j'} e^{-(\sigma-1)(\ln p_{j'} - \ln p_j)}} \; .$$

Multiplying both the denominator and the numerator by $e^{-(\sigma-1)\ln p_j}$ we have

$$Pr(j_m = j) = \frac{e^{-(\sigma-1)\ln p_j}}{\sum_{j'} e^{-(\sigma-1)\ln p_{j'}}} = \frac{p_j^{1-\sigma}}{\sum_{j'} p_{j'}^{1-\sigma}} \; .$$

Denote $P_m \equiv = \left[\frac{1}{J}\sum_{j'} p_{j'}^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$. So we have that the probability of a single consumer to choose product j is

$$Pr(j_m = j) = \frac{1}{J}\left(\frac{p_j}{P_m}\right)^{1-\sigma} \; .$$

I now take both the number of consumers and the number of products to infinity. With a slight abuse of notation, let J be the ratio of the measure of consumers to the measure of products. The density of consumers who choose product j is given by

$$f(j_m = j) = \frac{1}{J}\left(\frac{p_j}{P_m}\right)^{1-\sigma} \; .$$

So if the measure of firms is J and the measure of consumers is 1, then the measure of consumers buying from firm j is $f(j_m = j)$ and we have that $\int_0^J f(j_m = j)dj = 1$. □

Proof of Lemma 3

Lemma 3. The demand of a household for the chosen product in product type m is given by

$$c^i_{j_m} = \exp\left(\frac{\eta-1}{\sigma-1}\epsilon^i_{j_m}\right)\left(\frac{p_{j_m}}{P}\right)^{-\eta} C \; ,$$

where p_{j_m} is the price of the chosen product, $\epsilon^i_{j_m}$ is the idiosyncratic taste shock of the household for this product, and P is the aggregate price index given by

$$P = \left[\int_0^1 \exp\left(\frac{\eta-1}{\sigma-1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta}\, dm\right]^{\frac{1}{1-\eta}} \; .$$

Proof. To show this, we start by considering the cost minimization problem of the household. Here, I take as given the joint distribution of prices and taste shocks of chosen products. When deriving the demand faced by the firm, I shall derive this joint distribution explicitly. The cost minimization problem of the household is

$$\min_{\{c_m\}_{m\in(0,1)}} \int_0^1 p_{j_m} c_m dm$$
$$\text{s.t.} \quad \left[\int_0^1 \exp\left(\frac{\eta-1}{\eta(\sigma-1)}\epsilon_{j_m}\right)(c_m)^{\frac{\eta-1}{\eta}}\, dm\right]^{\frac{\eta}{\eta-1}} \geq \bar{C} \; .$$

Taking first order conditions we have

$$p_{j_m} = \lambda \bar{C}^{\frac{1}{\eta}} \exp\left(\frac{\eta-1}{\eta(\sigma-1)}\epsilon_{j_m}\right)(c_m)^{-\frac{1}{\eta}} \ .$$

To find the expression for the multiplier λ we first rearrange the equation above as follows. We raise it to the power of $(1-\eta)$ and rearrange to obtain

$$\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta} = \lambda^{1-\eta}\bar{C}^{-\frac{\eta-1}{\eta}} \exp\left(\frac{\eta-1}{\eta(\sigma-1)}\epsilon_{j_m}\right)(c_m)^{\frac{\eta-1}{\eta}} \ .$$

Integrating over all product modules, and simplifying using the formula for aggregate consumption we have

$$\int_0^1 \exp\left(\frac{\eta-1}{\sigma-1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta}\, dm = \lambda^{1-\eta} \ ,$$

so that

$$\lambda = \left[\int_0^1 \exp\left(\frac{\eta-1}{\sigma-1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta}\, dm\right]^{\frac{1}{1-\eta}} \ .$$

Notice that since λ is the multiplier of the cost minimization problem, it is also the natural price index. We can observe the problem is homothetic by noticing no consumption level enters the equation above. So I define

$$P \equiv \left[\int_0^1 \exp\left(\frac{\eta-1}{\sigma-1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta}\, dm\right]^{\frac{1}{1-\eta}} \ .$$

Plugging into the first order condition I obtain the household's demand for the chosen firm from product module m,

$$c_m = \exp\left(\frac{\eta-1}{\sigma-1}\epsilon^i_{j_m}\right)\left(\frac{p_{j_m}}{P}\right)^{-\eta}\bar{C} \ .$$

□

Proof of Lemma 4

Lemma 4. The average adjusted taste shock of a household that freely chooses to consume a product with price p is given by,

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = \bar{p}\right] = \left(\frac{p}{P_m}\right)^{\eta-1}\Gamma\left(1-\frac{\eta-1}{\sigma-1}\right) \ ,$$

where $\Gamma(\cdot)$ is the Gamma function.

Proof. Let $f_\epsilon\left(\epsilon_j | \, j_m = j, p_j = p\right)$ denote the conditional probability density function of the taste shock ϵ_j given that the price of product j is p and that it was freely chosen by the household. So that,

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = \int_0^\infty \exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right) f_\epsilon\left(\epsilon_j | \, j_m = j, p_j = p\right) d\epsilon_j$$

This density function can be rewritten as follows,

$$\begin{aligned} f_\epsilon\left(\bar{\epsilon}_j \middle| j_m = j, p_j = p\right) &= \frac{f_{\epsilon,\text{ chosen}}\left(\bar{\epsilon}_j, j_m = j \middle| p_j = p\right)}{Pr\left(j_m = j \middle| p_j = p\right)} \\ &= \frac{Pr\left(j_m = j | \epsilon_j = \bar{\epsilon}_j, p_j = p\right)}{Pr\left(j_m = j \middle| p_j = p\right)} e^{-(\bar{\epsilon}_j + \ln J)} e^{-e^{-(\bar{\epsilon}_j + \ln J)}} , \end{aligned}$$

where the last equality uses that the draw of a taste shock is the Gumbel distribution with location parameter $-\ln J$, which is independent of the firm's price. From the proof of Lemma 2 I use equations (32) and (33) to have that,

$$\begin{aligned} Pr\left(j_m = j | \epsilon_j = \bar{\epsilon}_j, p_j = p\right) &= e^{-Qe^{-\bar{\epsilon}_j}} e^{e^{-\bar{\epsilon}_j - \ln J}} , \\ Pr\left(j_m = j \middle| p_j = p\right) &= \frac{1}{JQ} , \end{aligned}$$

where $Q \equiv \frac{1}{J}\int_0^J e^{-(\sigma-1)(\ln p_j - \ln p)} dj$. Combining these together with the PDF of the Gumbel distribution I obtain

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = JQ\int_{-\infty}^{\infty} e^{\frac{\eta-1}{\sigma-1}\epsilon} e^{-Qe^{-\epsilon}} e^{e^{-\epsilon-\ln J}} e^{-e^{-(\epsilon+\ln J)}} e^{-(\epsilon+\ln J)} d\epsilon ,$$

Canceling out terms I have

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = Q\int_{-\infty}^{\infty} e^{\frac{\eta-1}{\sigma-1}\epsilon} e^{-Qe^{-\epsilon}} e^{-\epsilon} d\epsilon ,$$

Using a change of variables $x = Qe^{-\epsilon}$, so that $\epsilon = \ln Q - \ln x$ and $\frac{d\epsilon}{dx} = -\frac{1}{x}$. The integral can then be written as

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = Q\int_0^{\infty} \left(\frac{x}{Q}\right)^{-\frac{\eta-1}{\sigma-1}} e^{-x} Q^{-1} dx ,$$

This can be simplified to

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = Q^{\frac{\eta-1}{\sigma-1}} \int_0^{\infty} x^{-\frac{\eta-1}{\sigma-1}} e^{-x} dx ,$$

This is an Euler integral of the second kind. We have that

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = Q^{\frac{\eta-1}{\sigma-1}} \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right) ,$$

where I've used the definition of the Gamma function $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$.

Using the equality $Q = \left(\frac{p}{P_m}\right)^{\sigma-1}$ we have

$$\mathbb{E}\left[\exp\left(\frac{\eta-1}{\sigma-1}\epsilon_j\right)\middle| j_m = j, p_j = p\right] = \left(\frac{p}{P_m}\right)^{\eta-1} \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right) .$$

□

Proof of Lemma 5

Lemma 5. The aggregate price index P is given by,

$$P = \Gamma\left(1 - \frac{\eta - 1}{\sigma - 1}\right)^{\frac{1}{1-\eta}} \left[F P_m^{1-\eta} + (1-F) P_B^{1-\eta}\right]^{\frac{1}{1-\eta}} \ ,$$

where P_m is the product type price index defined above, and P_B is the initial-customer-base-weighted price index given by

$$P_B = \left[\frac{1}{1-F}\int_0^J (1-\theta) B_j p_j^{1-\eta} dj\right]^{\frac{1}{1-\eta}} \ .$$

Proof. Recall that the definition of the aggregate price index from Lemma 3 is

$$P = \left[\int_0^1 \exp\left(\frac{\eta - 1}{\sigma - 1}\epsilon_{j_m}\right)(p_{j_m})^{1-\eta}\, dm\right]^{\frac{1}{1-\eta}} \ .$$

Since all product modules are symmetric, the aggregate price index is given by

$$P = \left\{\mathbb{E}\left[\exp\left(\frac{\eta - 1}{\sigma - 1}\epsilon_{j_m}\right) p_{j_m}^{1-\eta}\right]\right\}^{\frac{1}{1-\eta}} \ .$$

The expectation term can be split into the fraction of products the household is locked-in with, and the ones it chooses freely. As before, I let F denote the measure of households that are not locked-in with a product in a specific product module. As all product modules are assumed to be symmetric, I have that F is also the measure of product modules each household is locked-in with. So that,

$$P = \left\{(1-F)\mathbb{E}\left[\exp\left(\frac{\eta - 1}{\sigma - 1}\epsilon_{j_m}\right) p_{j_m}^{1-\eta}\middle| \text{lock-in}\right] + F\mathbb{E}\left[\exp\left(\frac{\eta - 1}{\sigma - 1}\epsilon_{j_m}\right) p_{j_m}^{1-\eta}\middle| \text{unattached}\right]\right\}^{\frac{1}{1-\eta}} \ . \tag{34}$$

Using the Law of Iterated Expectations together with Lemma 4 I have that

$$\mathbb{E}\left[\exp\left(\frac{\eta - 1}{\sigma - 1}\epsilon_{j_m}\right) p_{j_m}^{1-\eta}\middle| \text{unattached}\right] = \Gamma\left(1 - \frac{\eta - 1}{\sigma - 1}\right) P_m^{1-\eta} \ .$$

Using the assumption on the taste shock towards locked-in products I have that

$$\mathbb{E}\left[\exp\left(\frac{\eta - 1}{\sigma - 1}\epsilon_{j_m}\right) p_{j_m}^{1-\eta}\middle| \text{lock-in}\right] = \Gamma\left(1 - \frac{\eta - 1}{\sigma - 1}\right)\mathbb{E}\left[p_{j_m}^{1-\eta}\middle| \text{lock-in}\right] \ .$$

I define the last term in the equation to be equal to $P_B^{1-\eta}$. That term is given by,

$$P_B \equiv \mathbb{E}\left[p_{j_m}^{1-\eta}\middle| \text{lock-in}\right]^{\frac{1}{1-\eta}} = \left[\frac{\int_0^J (1-\theta) B_j^L p_j dj}{1-F}\right]^{\frac{1}{1-\eta}} \ .$$

Recall that the definition of F implies that $1 - F = \int_0^J (1-\theta) B_j dj$. The subscript B stands for the the fact the P_B is a power mean of product prices weighted by the initial customer-base of each product, B. Plugging in equation (34) I obtain the desired result

$$P = \Gamma\left(1 - \frac{\eta - 1}{\sigma - 1}\right)^{\frac{1}{1-\eta}} \left[(1-F) P_B^{1-\eta} + F P_m^{1-\eta}\right]^{\frac{1}{1-\eta}} .$$

$\square$

Proof of Proposition 6

Proposition 6. The demand function faced by each firm is given by

$$y_j = \left[(1-\theta) B p_j^{-\eta} + \frac{F}{M} P_m^{\sigma-\eta} p_j^{-\sigma}\right] \frac{1}{F P_m^{1-\eta} + (1-F) P_B^{1-\eta}} (w + \Pi) . \tag{24}$$

Proof. Using Lemma 3 we can write the demand of firm j as follows,

$$y_j = \int_{i: j_m^i = j} \exp\left(\frac{\eta-1}{\sigma-1} \epsilon_j^i\right) \left(\frac{p_j}{P}\right)^{-\eta} C di ,$$

where we integrate over the set of customers who chose firm j out of all firms. Since there is no price discrimination we have

$$y_j = \left[\int_{i: j_m^i = j} \exp\left(\frac{\eta-1}{\sigma-1} \epsilon_j^i\right) di\right] \left(\frac{p_j}{P}\right)^{-\eta} C .$$

We can split customers who chose firm j to ones who actively chose the it and ones who inertially were locked-in with it. So we have

$$y_j = \left[\int_{i: j_m^i = j, \text{lock-in}} \exp\left(\frac{\eta-1}{\sigma-1} \epsilon_j^i\right) di + \int_{i: j_m^i = j, \text{unattached}} \exp\left(\frac{\eta-1}{\sigma-1} \epsilon_j^i\right) di\right] \left(\frac{p_j}{P}\right)^{-\eta} C .$$

Rewriting this in terms of expectations we have

$$y_j = \left\{(1-\theta) B \mathbb{E}\left[e^{\frac{\eta-1}{\sigma-1}\epsilon_j^i} \middle| j_m^i = j, \text{lock-in}\right] + (B' - (1-\theta)B) \mathbb{E}\left[e^{\frac{\eta-1}{\sigma-1}\epsilon_j^i} \middle| j_m^i = j, \text{unattached}\right]\right\} \left(\tfrac{p_j}{P}\right)^{-\eta} C ,$$

where we have used that the measure of inertial customers is equal to $(1-\theta)B$ and the measure of unattached customers equals $B' - (1-\theta)B$.

For inertial households the taste shock is assumed to satisfy $e^{\frac{\eta-1}{\sigma-1}\epsilon_j^i} = \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right)$. Using Lemma 4 we have that for active households $e^{\frac{\eta-1}{\sigma-1}\epsilon_j^i} = \left(\frac{p}{P_m}\right)^{\eta-1} \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right)$. So we can write the demand for firm j as follows,

$$y_j = \left[(1-\theta) B \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right) + (B' - (1-\theta)B) \left(\frac{p}{P_m}\right)^{\eta-1} \Gamma\left(1 - \frac{\eta-1}{\sigma-1}\right)\right] \left(\frac{p_j}{P}\right)^{-\eta} C .$$

Using equation (22) we can rewrite the equation above as follows,

$$y_j = \left[(1-\theta)B + \frac{F}{M}\left(\frac{p_j}{P_m}\right)^{\eta-\sigma}\right]\Gamma\left(1-\frac{\eta-1}{\sigma-1}\right)\left(\frac{p_j}{P}\right)^{-\eta}C\,.$$

Finally, using the expression for the aggregate price level P and that $C = w/P$, the Γ function cancels out and we have:

$$y_j = \left[(1-\theta)Bp_j^{-\eta} + \frac{F}{M}{P_m}^{\sigma-\eta}p_j^{-\sigma}\right]\frac{1}{\theta P_m^{1-\eta} + (1-\theta)P_B^{1-\eta}}w\,.$$

□

Proof of Lemma 7

Proposition 7. The markup of a firm with customer base B and productivity a is implicitly defined by the following equation,

$$\mu = \frac{\sigma}{\sigma-1} + \underbrace{\alpha\left(\frac{\eta}{\eta-1} - \frac{\sigma}{\sigma-1}\right)}_{\text{harvesting motive}}\underbrace{-(1-\alpha)\sum_{\tau=1}^{\infty}(\beta(1-\theta))^{\tau}\,\mathbb{E}\left[\mathbb{1}\left(\overline{T}>\tau\right)\gamma_{+\tau}\right]}_{\text{investing motive}}, \tag{26}$$

where $\overline{T}$ is the stopping time indicating that the firm exits the economy in $\overline{T}$ periods. $\mathbb{1}\left(\overline{T}>\tau\right)$ is an indicator that takes the value 1 if the firm still operates τ periods ahead. The variables α and γ_τ are given by

$$\alpha = \frac{(\eta-1)(1-\theta)B}{(\eta-1)(1-\theta)B + (\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{\eta-\sigma}}, \qquad \gamma_{+\tau} = \frac{w_{+\tau}}{a_{+\tau}}\frac{a}{w}\left(\frac{p_{+\tau}}{P_{+\tau}}\right)^{-\eta}\left(\frac{p}{P}\right)^{\eta}(\mu_{+\tau}-1)\ ,$$

where a variable with subscript $+\tau$ corresponds to the variable τ periods ahead.

Proof. The firm's problem is given by

$$\begin{aligned} V(B,a) = \max_{\{p,y,B'\}} py - \frac{w}{e^a}y + \frac{1}{1+r}\mathbb{E}\left[\max\left\{V\left(B',a'\right) - x_o w, 0\right\}\right] \\ \text{s.t.}\quad B' = (1-\theta)B + \frac{F}{J}\left(\frac{p}{P_m}\right)^{1-\sigma}\ , \\ y = \left[(1-\theta)Bp^{-\eta} + \frac{F}{J}{P_m}^{\sigma-\eta}p^{-\sigma}\right]\frac{1}{FP_m^{1-\eta} + (1-F)P_B^{1-\eta}}(w+\Pi)\ . \end{aligned}$$

Denote the Lagrange multiplier on the first constraint by λ_B and on the second by λ_y. Taking first order and envelope conditions, I have

$$[y]: \qquad \lambda_y = p - \frac{w}{e^a} \;, \tag{35}$$

$$[p]: \qquad py = \frac{\eta(1-\theta)Bp^{-\eta} + \sigma\frac{F}{J}{P_m}^{\sigma-\eta}p^{-\sigma}}{FP_m^{1-\eta} + (1-F)P_B^{1-\eta}}\left(w+\Pi\right)\lambda_y + (\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{1-\sigma}\lambda_B \;, \tag{36}$$

$$[B']: \qquad \lambda_B = \frac{1}{1+r}\mathbb{E}\left[V_B(B',a')\mathbb{1}\left(V(B',a') > x_o'w'\right)\right] \;, \tag{37}$$

$$[B]: \qquad V_B(B,a) = (1-\theta)\lambda_B + \frac{(1-\theta)p^{-\eta}}{FP_m^{1-\eta} + (1-F)P_B^{1-\eta}}\left(w+\Pi\right)\lambda_y \tag{38}$$

Iterating equations (37) and (38) and using $\frac{1}{1+r} = \beta\frac{w+\Pi}{w'+\Pi'}\frac{P'}{P}$, I get

$$\lambda_B = \frac{w+\Pi}{\left(FP_m^{1-\eta}+(1-F)P_B^{1-\eta}\right)^{\frac{1}{1-\eta}}}\sum_{\tau=1}^{\infty}\beta^\tau(1-\theta)^\tau\mathbb{E}\left[\mathbb{1}\left(T\geq\tau\right)\left(\frac{p_{+\tau}}{\tilde{P}_{+\tau}}\right)^{-\eta}\left(p_{+\tau} - e^{-a_{+\tau}}w_{+\tau}\right)\right] \;, \tag{39}$$

where $\mathbb{1}(T > \tau)$ is an indicator that takes the value 1 if the firm is still operating τ periods in the future. The $+\tau$ subscript corresponds to variables τ periods ahead. $\tilde{P} \equiv \left(FP_m^{1-\eta} + (1-F)P_B^{1-\eta}\right)^{\frac{1}{1-\eta}}$. Combining equations (35) and (36) together with the demand function and rearranging, I obtain

$$\mu = \frac{\eta(1-\theta)B \quad + \quad \sigma\frac{F}{J}\left(\frac{p}{P_m}\right)^{\eta-\sigma}}{(\eta-1)(1-\theta)B+(\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{\eta-\sigma}} - \frac{(\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{1-\sigma}p^\eta}{(\eta-1)(1-\theta)B+(\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{\eta-\sigma}}\frac{e^a}{w}\frac{\theta P_m^{1-\eta}+(1-\theta)P_B^{1-\eta}}{(w+\Pi)}\lambda_B \;, \tag{40}$$

where $\mu = \frac{p}{w/e^a}$ is the firm's markup. Using the expression for λ_B from equations (37) and (39), I obtain

$$\mu = \frac{\sigma}{\sigma-1} + \alpha\left(\frac{\eta}{\eta-1} - \frac{\sigma}{\sigma-1}\right) - (1-\alpha)\sum_{\tau=1}^{\infty}\beta^\tau(1-\theta)^\tau\mathbb{E}\left[\mathbb{1}\left(\overline{T}>\tau\right)\gamma{+\tau}\right] \;, \tag{41}$$

where

$$\alpha = \frac{(\eta-1)(1-\theta)B}{(\eta-1)(1-\theta)B + (\sigma-1)\frac{F}{J}\left(\frac{p}{P_m}\right)^{\eta-\sigma}} \;,$$

and

$$\gamma_{+\tau} = \frac{w_{+\tau}}{w}\frac{e^a}{e^{a_{+\tau}}}\left(\frac{p_{+\tau}}{\tilde{P}_{+\tau}}\right)^{-\eta}\left(\frac{p}{\tilde{P}}\right)^{\eta}\left(\mu_{+\tau}-1\right) .$$

□

B Other Results

B.1 Comparison to Other Structural Models of Consumer Inertia

In this section, I highlight the advantages of the structural model of consumer inertia proposed in this paper relative to other prominent structural models of consumer inertia. I focus on three examples: a model of switching costs, a model with additive deep habits, and a model with multiplicative deep habits.

The common way of modeling consumer inertia in the industrial organization and marketing literatures is by introducing switching costs. Such switching costs stand for learning costs, transaction costs, psychological and emotional costs, incomplete information, or artificial contractual costs imposed by firms. While switching costs can be introduced in a variety of ways, the approach the literature often takes is assuming a multinomial logit discrete choice problem of the following form. The indirect utility of household i from choosing product j when its price is p_j is given by,

$$U_{ij} = -(\sigma - 1) \ln p_j + (1 - \theta)\mathbb{1}\left[j_i^0 = j\right] + \epsilon_j^i \ , \tag{42}$$

where j_i^0 is the previous consumption choice of household i. $\mathbb{1}\left[j_i^0 = j\right]$ takes the value one if the product is the same as the one consumed in the previous period, and zero otherwise. I choose the notation above so that the interpretation of parameters is similar to the stylized model. The term $(1 - \theta) > 0$ governs the size of switching costs in the market. Note that, for simplicity, I assumed that products only differ by their price, and not along other characteristics.

The multinomial logit switching costs specification is not suitable to study the behavior of firms in general equilibrium for two reasons. First, as I show below, if firms are atomistic, this specification implies that the switching costs, regardless of their size, do not alter the firm's problem. In particular, the firm's problem does not depend on its initial customer base. Second, this specification implies that the price elasticity of demand of previous customers is the same as the demand elasticity of new customers. I discuss the implications of assuming an equal elasticity of demand for new and past customers in the final part of this section, when presenting the multiplicative deep habits model.

Suppose there is a finite number of firms operating in the market, and denote their number by J. Let the previous customers of firm j relative to the average customer base of a firm be denoted by B_j^0. Without loss of generality, I assume that the measure of consumers in the market is equal to one so that the average customer base of a firm is $\frac{1}{J}$. So the customer base of firm j is given by $\frac{1}{J}B_j^0$. Using the multinomial logit assumption that the distribution of the ϵ term follows a standard Gumbel distribution, I obtain the following law of motion for the customer base of the firm:

$$B_j = \sum_{j'} B_{j'}^0 \frac{e^{\mathbb{1}_{j=j'}(1-\theta)} p_j^{1-\sigma}}{e^{1-\theta} p_{j'}^{1-\sigma} + \sum_{z \neq j'} p_z^{1-\sigma}} \ . \tag{43}$$

We can rewrite the expression above as follows,

$$B_j = \left(\frac{p_j}{P_j}\right)^{1-\sigma} + \frac{B_0^j}{J}\left(\frac{p_j}{\tilde{P}_j}\right)^{1-\sigma} , \tag{44}$$

where

$$P_j = \left[\sum_{j'\neq j} \frac{B_{j'}^0}{J} \frac{1}{\frac{1}{J}\sum_{z\neq j'} p_z^{1-\sigma} + \frac{1}{J}e^{1-\theta}p_{j'}}\right]^{-\frac{1}{1-\sigma}} ,$$

and

$$\tilde{P}_j = \left[\frac{1}{\frac{1}{J}\sum_{z\neq j} p_z^{1-\sigma} + \frac{1}{J}e^{1-\theta}p_j}\right]^{-\frac{1}{1-\sigma}} .$$

Taking the number of firms to infinity, $J \to \infty$, I have that

$$B_j = \left(\frac{p_j}{P}\right)^{1-\sigma} , \tag{45}$$

where

$$P = \left[\frac{1}{J}\sum_{j'} p_{j'}^{1-\sigma} dj'\right]^{\frac{1}{1-\sigma}} .$$

Therefore, when the number of firms goes to infinity and firms are atomistic, the switching costs do not affect the firm's problem. This is because, relative to the measure of customers who look for a product, the measure of customers who previously bought from the firm are negligible. For this reason, the multinomial logit discrete choice problem with switching costs is not suitable to study the behavior of firms in equilibrium if firms are atomistic.

The second structural model of consumer inertia I consider is the additive deep habits model, which was introduced by Ravn et al. (2006). The utility function of the household in that model is given by

$$U = \left[\int_i \left(c_i - (1-\theta)c_i^0\right)^{\frac{\sigma-1}{\sigma}} di\right]^{\frac{\sigma}{\sigma-1}} , \tag{46}$$

where c_i is the consumption of product i, and c_i^0 is the consumption of that product in the previous period. The parameter $(1-\theta) \in (0,1)$ governs the degree of consumer inertia. The resulting demand function of the firm is

$$y_j = (1-\theta)y_j^0 + \left(\frac{p_j}{P}\right)^{-\sigma} \tilde{C} , \tag{47}$$

where y_j^0 is the production of the firm in the previous period, P is the aggregate price index, and $\tilde{C}$ is an aggregate quantity that governs the level of demand. Compare the demand function (47) to the demand function in the static mode:

$$y_j = (1-\theta)B_j w p_j^{-1} + \left(\frac{p_j}{P}\right)^{-\sigma} \tilde{C} . \tag{48}$$

We see that the two models deliver a similar demand function. Indeed, the implications of my model are very similar to the additive deep habits model. Similar to my quantitative model, the additive deep habits model includes both the harvesting and investment motives. Firms have incentive to push up prices and exploit their customer base, as well as an incentive to reduce markups in order to attract customers and increase the demand for their products in future periods.

The key difference is that the demand elasticity coming from previous production is equal to 0 in the deep habits model, and is equal to 1 in my static model.[32] The fully inelastic term in equation (47) implies that the firm has an incentive to set the markup to infinity and make infinite profits. This is in contrast to my model, where the optimal markup of the firm is finite. Schmitt-Grohé and Uribe (2007) show numerically that if households can drop varieties out of their consumption basket, then firms have no incentive to deviate from the finite price that solves the first order condition of the problem.

Since the markup that maximizes the unconstrained problem of the firm in the deep habits model is equal to infinity, one cannot use a value function iteration to solve for the optimal behavior of firms. As a result, the quantitative firm dynamics literature has adopted variants of the multiplicative deep habits model, instead. Some examples include Foster et al. (2016), Moreira (2016), Gilchrist et al. (2017), and Sedláček and Sterk (2017).

In its simplest form, the multiplicative deep habits model, which was also introduced in Ravn et al. (2006), assumes the following utility function,

$$U = \left[\int_i \left(\frac{1}{\left(c_i^0\right)^{1-\theta}} c_i \right)^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}} .$$

The resulting demand function is given by

$$y_j = \left(y_j^0\right)^{1-\theta} \left(\frac{p_j}{P}\right)^{-\sigma} \tilde{C} \,, \tag{49}$$

so the firm's past production, y_j^0, acts as a demand shifter. A key implication of the multiplicative deep habits model is that the elasticity of demand faced by the firm is independent of the initial customer base of the firm, or of its past production. This implies that the harvesting motive is muted by assumption. That is, the firm has no direct incentive to increase markups due to a larger customer base. In a static version of the multiplicative deep habits model, firms would set the same markup independent of the degree of consumer inertia, $1-\theta$. In a dynamic version of the model, the investment motive is present. Firms have an incentive to reduce their markups in order to increase production, and, as a result, increase the demand for their products in future periods.

The lower is the previous production of the firm, the stronger is the investment motive. This is because lowering markups reduces profits by more, if current demand is higher. Therefore, as

[32] In my quantitative model, the demand elasticity coming from locked-in customers is equal to $\eta \in (1, \sigma)$.

in my model, the markup set by firms is increasing in their past production. In my model, as well as in the additive deep habits model, this is both because the harvesting motive is increasing in past production and because the investment motive is decreasing in past production. In the multiplicative deep habits model, this is only because of the investment motive. While this distinction may not be essential when the degree of consumer inertia is constant over time, it is crucial when considering a change in consumer inertia. In the multiplicative deep habits model, an increase in consumer inertia leads to a decline in the markups set by all firms, as it strengthens the investment motive. The decline in markups is simply a result of the muted harvesting motive in the multiplicative deep habits model.

B.2 The Sophisticated Consumer and the Simple Consumer: An Equivalence Result

In this subsection, I compare the consumption behavior of a sophisticated consumer and a simple consumer in a dynamic version of the static model. The simple consumer does not internalize it can get locked into the product it chooses, while the sophisticated consumer does. The sophisticated consumer uses a log-linear approximation for the firm's pricing decision.

I find that the consumption behavior of the sophisticated consumer is identical to that of a simple consumer with a different elasticity of substitution. I derive the explicit formula for the elasticity of substitution of the simple consumer that makes the consumption decisions of the two consumers isomorphic. If the persistent coefficient in the firm's law of motion is non-negative, then the elasticity of substitution of the simple consumer is higher than that of the sophisticated consumer.

I assume the household's utility is given by

$$\sum_{t=0}^{\infty} \beta^t \ln\left(e^{\frac{1}{\sigma-1}\epsilon_{jt}} c_{jt}\right) \ ,$$

where c_{jt} is the consumption of the household at time t, and ϵ_{jt} is the idiosyncratic taste of the household towards the good consumed. As in the static model, I assume the household spends its income w on the chosen product. I assume that the wage is time-invariant so that the consumption of the chosen product is given by

$$c_{jt} = \frac{w}{p_{jt}} \ .$$

The consumption decision of the simple consumer is the same as in Section 2. The chosen product of the simple household at time 0 is given by,

$$j = \arg\max_{j'} -(\sigma-1)\ln p_{j'0} + \epsilon_{j'0} \ . \tag{50}$$

The sophisticated household, on the other hand, needs to take into account not only the current price of the good but also the expected relative price of the good in future periods. I assume that firms exit the economy with exogenous probability δ, and that the probability of re-optimization is θ. So, with probability $(1-\delta)(1-\theta)$ the household is locked into the product it purchased in the preceding period, and it can re-optimize its product choice otherwise. Since this probability is the same across all products, when choosing what product to consume, the household only needs to consider its utility given that it is locked into the product in the future. In particular, the sophisticated household is maximizing

$$j = \arg\max_{j'} \sum_{t=0}^{\infty} \beta^t (1-\delta)^t (1-\theta)^t \ln\left(e^{\frac{1}{\sigma-1}\epsilon_{j't}} \frac{w}{p_{j't}}\right) .$$

Rewriting this problem, I obtain

$$j = \arg\max_{j'} \sum_{t=0}^{\infty} \beta^t (1-\delta)^t (1-\theta)^t \left[\epsilon_{j't} + (\sigma-1)\ln(w) - (\sigma-1)\ln p_{j't}\right] .$$

And since w is independent of the chosen product, the product chosen by the sophisticated household solves

$$j = \arg\max_{j'} \sum_{t=0}^{\infty} \beta^t (1-\delta)^t (1-\theta)^t \left[\epsilon_{j't} - (\sigma-1)\ln p_{j't}\right] .$$

I assume that the taste shock for a product to which the household is locked into is constant and independent of past tastes toward that product. This assumption implies that the chosen product satisfies

$$j = \arg\max_{j'} -(\sigma-1)\sum_{t=0}^{\infty} \beta^t (1-\delta)^t (1-\theta)^t \ln p_{j't} + \epsilon_{j'0} .$$

Suppose the sophisticated household uses a log-linear rule to forecast a firm's price, so that

$$\ln p^e_{jt+1} = \rho \ln p_{jt} + (1-\rho)\ln \bar{p} ,$$

where ρ is the persistence on the relative price of the firm, and $\bar{p}$ is the price to which a firm eventually converges as it ages. Substituting into the expression that pins down the chose product, I obtain

$$j = \arg\max_{j'} -(\sigma-1)\sum_{t=0}^{\infty} \beta^t (1-\delta)^t (1-\theta)^t \left[\rho^t \ln p_{j'0} + (1-\rho^r)\ln \bar{p}\right] + \epsilon_{j'0} .$$

Removing the constant and simplifying the sum, I get the following expression which pins down the chosen product:

$$j = \arg\max_{j'} -\frac{\sigma-1}{1-\beta\rho(1-\delta)(1-\theta)} \ln p_{j'0} + \epsilon_{j'0} . \tag{51}$$

We see that the only difference in the consumption decision of the sophisticated consumer (51) relative to the simple consumer (50) is in the coefficient multiplying the price. Therefore, if the parameter σ of the simple consumer satisfies

$$\sigma_{\text{simple}} = 1 + \frac{\sigma_{\text{sophisticated}} - 1}{1-\beta\rho(1-\delta)(1-\theta)} ,$$

then the consumption choice of the simple consumer is identical to the consumption choice of the sophisticated consumer. Consequently, if we estimate the parameter σ to match data moments, the two models are isomorphic. The only difference is in the interpretation of the estimated coefficient.

B.3 Quantitative Model Results with an Estimated Level of Consumer Inertia

In this section, I present the estimation results where the degree of consumer inertia is estimated, instead of being calibrated using my micro estimates. The second to last column of Table 7 presents the estimation results, and the last column presents the benchmark results where the re-optimization probability is calibrated using the micro estimates of consumer inertia. The re-optimization is strikingly only slightly smaller than the calibrated one. Recall that the micro estimates are based on data from of consumer packaged goods. These are goods households actively choose, as opposed to, for example, their auto-renewed insurance policy. It is not surprising, then, that the resulting re-optimization probability, estimated using data on the population of businesses in the US, is lower. The estimate implies that for the US economy as whole, the re-estimation probability is about 10% lower than for the consumer packaged goods sector. The other estimated parameters are also fairly similar to the benchmark specification. Given that the estimated re-optimization probability is not substantially different than the benchmark estimate, the similarity of the other parameters is not surprising.

Table 7: Estimation Results with the Level of Consumer Inertia Estimated

Parameter	Description	θ estimated	Benchm.
θ	Re-optimization probability	0.129	0.146
μ_o	Average ln fixed operating costs	-4.16	-3.94
σ_o	Std of ln fixed operating costs	3.26	3.33
ρ_a	Idiosyncratic productivity persistence	0.99	0.98
σ_a	Std of productivity shocks	0.058	0.06
η	Elasticity of sub. between product types	2.77	2.66
σ	Elasticity of sub. within a product type	3.98	3.63

Notes: This table presents value of the structural parameters estimated via GMM. The first column out of the two presents the estimation results where the re-optimization probability is estimated together with the other six structural parameters via GMM. The second column presents the benchmark estimation results, where the re-optimization probability is calibrated using the micro estimates of consumer inertia.

Table 8 presents the model fit. The table shows that estimating the re-optimization probability slightly improves the model fit.

Table 8: Model Fit

Moment	Data	θ estimated	Benchmark
Share of firms aged 0–1	20.7%	20.2%	21.9%
Share of firms aged 2–5	26.0%	27.8%	29.3%
Share of firms aged 6–10	19.5%	21.1%	21.2%
Relative size of firms aged 2–5	32.6	31.3	32.3
Relative size of firms aged 6–10	55.3	64.6	61.4
Relative size of firms aged >10	177.1	115.1	108.1
Share of profits in GDP	4.8%	5.1%	5.2%

Notes: This table presents model fit for the benchmark estimation (last column) as well as the estimation where the re-optimization probability is estimated via GMM (second to last column). Allowing the re-optimization probability to be estimated provides only a slight improvement in the fit of the model.

C Additional Tables and Figures

Table 9: Moving Patterns

	Division of destination									Total
Division of origin	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	
1	33	18	7	4	60	6	5	8	3	**114**
2	21	50	19	7	166	8	16	20	11	**318**
3	5	12	78	29	139	40	33	40	17	**393**
4	4	9	25	46	34	8	27	35	14	**202**
5	16	49	54	23	244	54	48	42	23	**553**
6	1	3	19	8	61	23	16	6	9	**146**
7	2	8	21	34	49	15	59	26	17	**231**
8	5	9	26	15	40	5	47	62	46	**255**
9	9	15	20	17	46	10	34	114	61	**326**
Total	**96**	**173**	**269**	**183**	**839**	**169**	**285**	**353**	**201**	**2,568**

Notes: This table presents the movers migration patterns. The row indicates the division of origin, while the column indicates the destination division. The 9 divisions in order are: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific. Out of the 2,568 movers, only 656 stayed within their division.

Table 10: Difference between Movers and Non-Movers

		Young	Maturing	Mature	Old
Education	Movers	[SC, GC, GC]	[SC, GC, GC]	[SC, SC, GC]	[HS, SC, GC]
	Non-movers	[SC, SC, GC]	[HS, SC, GC]	[HS, SC, GC]	[HS, SC, GC]
Income	Movers	[27.5, 47.5, 65]	[37.5, 65, 85]	[37.5, 55, 85]	[27.5, 37.5, 55]
	Non-movers	[32.5, 47.5, 85]	[37.5, 55, 85]	[32.5, 55, 85]	[22.5, 37.5, 55]
Household members	Movers	[1,2,4]	[2,3,4]	[2,2,2]	[1,2,2]
	Non-movers	[2,3,4]	[2,3,4]	[2,2,3]	[1,2,2]
Age	Movers	[28, 31, 32]	[39, 42, 46]	[53, 57, 60]	[67, 70, 75]
	Non-movers	[29, 31, 33]	[39, 43, 46]	[53, 56, 60]	[67, 71, 76]

Notes: This table presents descriptive statistics of movers and non-movers. Each cell contains the 25 quantile, the median, and the 75 quantile, respectively. For education, HS, SC, and GC, correspond, respectively, to high school graduate, some college education, and college graduate. For income, the numbers represent thousands of dollars. Since the data contains only income bins, the numbers correspond to the average value of each income bin. The data is computed using the first observation for each household. The table shows that movers do not differ substantially from non-movers.

D Additional Regressions Results

Table 11: Consumer Inertia and Entry Rates – State Analysis

	Dependent variable: firm entry rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Consumer inertia index	-5.06^{***}	-4.92^{***}	-4.98^{***}	-4.3^{**}	-4.3^{**}	-4.2^{**}
	[-10.2, -2.5]	[-9.8, -2.5]	[-9.6, -2.6]	[-8.6, -0.2]	[-8.5, -0.3]	[-8.3, -0.1]
Labor-force growth rate		0.05^{***}	0.05^{***}		0.05^{***}	0.05^{***}
		[0.03, 0.07]	[0.03, 0.07]		[0.04, 0.07]	[0.03, 0.07]
Share of older workers			0.004			-0.01
			[-0.02, 0.03]			[-0.06, 0.02]
State and time f.e.	✓	✓	✓	✓	✓	✓
Instrumented	✗	✗	✗	✓	✓	✓
	Dependent variable: share of workers employed by entrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Consumer inertia index	-2.0^{***}	-2.0^{***}	-2.1^{***}	-2.2^{*}	-2.2^{*}	-2.3
	[-4.1, -0.6]	[-4.1, -0.6]	[-4.4, -0.6]	[-4.5, 0.2]	[-4.4, 0.1]	[-4.7, 0.3]
Labor-force growth rate		0.013^{***}	0.013^{***}		0.013^{***}	0.013^{***}
		[0.01, 0.02]	[0.01, 0.02]		[0.01, 0.02]	[0.01, 0.02]
Share of older workers			0.006			0.009
			[-0.01, 0.02]			[-0.03, 0.04]
State and time f.e.	✓	✓	✓	✓	✓	✓
Instrumented	✗	✗	✗	✓	✓	✓

Notes: This table presents the regression results of (14). The sample period is 1980–2014. In the top panel, the dependent variable is the share entrants out of all firms. The dependent variable in the bottom panel is the share workers employed by entrants. 95% confidence intervals in brackets. Inference was done using a bootstrap procedure, clustered at the state level. *** (**) - 99% (95%) confidence interval does not include zero.

E Numerical Appendix

E.1 Algorithm for Stationary Equilibrium

I use a value function iteration procedure to compute the stationary equilibrium of the economy. The algorithm consists of the following two steps. I describe the steps in more details below.

0) Start with a guess for the present value of the firm on the grid points, V, the measure of unattached consumers, F, the two price indices, P_m and P_B, and the measure of profits in the economy, Π.[33]

1) Solve the optimal pricing decision of firms given a guess for the value function in the following period, and aggregate endogenous variables. Obtain an updated guess for V, together with policy functions of firms.

2) Compute the ergodic distribution of firms across the two dimensions of heterogeneity. Obtain an updated guess for the endogenous aggregate variables F, P_m, P_b, and Π. Check distance between value function and aggregate variables from previous guess. If the difference is not sufficiently small, repeat from step (1).

Grid. I use a two-dimensional grid to represent the state variables of the firm: their customer base and their productivity. I construct the productivity using the Tauchen and Hussey (1991) approach. The grid points are denoted by a_i, where $i = 1, ..., N_a$. I set N_a to 15. I denote the resulting Markov transition matrix by Π_a, which is a 15×15 matrix whose columns sum to one.

The customer base grid is constructed in a similar way to the capital grid in Maliar et al. (2010), so that the grid is denser for lower values of customer base. In particular, the customer base grid points are given by

$$B_j = \left(\frac{j}{N_B}\right)^{\kappa} B_{max}, \quad \text{for } j = 1, \ldots, N_B \ ,$$

where N_B is the number of customer grid points, B_{max} is the largest measure of customer base considered, and κ governs the degree of the polynomial grid. I set N_B to 1,000, B_{max} to 10, and κ to 2.5. The high maximal level of customer base ensures that for different parameter specifications, which I consider when estimating via GMM, the ergodic distribution of firms contains no firms at the upper bound of customer base. The polynomial grid ensures that despite having a high upper bound for customer base, most of the grid points are in the region where the majority of firms in equilibrium lie.

Step 1 - Solving the Firm's Problem. The present value of a firm, gross of paying the fixed operating costs, is defined over the grid points. I denote it by V, an $N_a \times N_B$ matrix. The value

[33] For the initial stationary equilibrium, I normalize the measure of firms to 1 and calibrate f_e to match the present value of entrants in equilibrium. For the final stationary equilibrium, where f_e is taken as given,

function is given by

$$V(B,a) = \max_{\{p,y,B'\}} py - \frac{w}{a}y + \frac{1}{1+r}\mathbb{E}\left[\max\left\{V\left(B',a'\right) - x'_o w, 0\right\}\right]$$

$$\text{s.t.} \quad B' = (1-\theta)B + \frac{F}{J}\left(\frac{p}{P_m}\right)^{1-\sigma} \quad ,$$

$$y = \left[(1-\theta)Bp^{-\eta} + \frac{F}{J}{P_m}^{\sigma-\eta}p^{-\sigma}\right]\frac{1}{FP_m^{1-\eta} + (1-F)P_B^{1-\eta}}(w+\Pi) \, ,$$

where $\frac{1}{1+r} = \beta$ in the stationary equilibrium, and w is normalized to 1 without loss of generality. We can substitute the constraints to arrive at the following firm's problem

$$V(B,a) = \max_{B'} \pi(B,a,B') + \frac{1}{1+r}\mathbb{E}\max\left\{V\left(B',a'\right) - x'_o, 0\right\} \, ,$$

where $\pi(B,a,B')$ is the periodic profits of a firm with initial customer base B, productivity a, which chooses to sell to a measure B' of customers. The current guess for the value function allows me to construct the expected discounted future profits of the firm, net of operating costs. I then maximize the firm's present profits on the grid of customer base levels. The resulting maximization procedure yields the present value for each point on the grid, as well as the policy choices over the grid $B'(B,a)$, $p(B,a)$, and $y(B,a)$. With the present value at hand, I derive the survival probability of each firm which I denote by $s(B,a)$. $s(B,a)$ corresponds to the probability that its present value of profits net of paying the fixed operating costs is greater than zero.

Step 2 - Computing the Ergodic Distribution. With the policy functions at hand, I construct the transition matrix. There are $N_a N_B$ states, and I denote the ergodic distribution by a vector Λ of that size. $\Lambda(B,a)$ denotes the measure of firms that have an initial customer base B, productivity a. 4.2, where the distribution is defined prior to the exit decision.

The transition matrix, denoted by Π_T, is of size $N_a N_B \times N_a N_B$. The law of motion for the distribution is given by

$$\Lambda(B',a') = \sum_B \sum_a \Lambda(B,a)s(B,a)\mathbb{1}\left(B'(B,a) = B'\right)\Pi_a(a',a) + \mathbb{1}(B'=0)\Pi_e(a) \, ,$$

where $\Pi_a(a',a)$ is the exogenous probability of drawing the productivity a' given past productivity a. $Pi_e(a)$ is the probability of an entrant drawing productivity a. Note that by writing the law of motion in this way, I assume that the measure of entrants is equal to one. I start from a guess for Λ and iterate until convergence. Once I obtain the resulting Λ, I divide it by a constant that ensures that the measure of operating firms is equal to 1.

After obtaining Λ, I can compute the implied endogenous aggregate variables $\{P_m, P_B, \Pi, F\}$. If the maximal difference between the implied aggregate difference and their initial guess is sufficiently small, and the maximal difference between the implied present value of a firm on each grid points is sufficiently close to the initial guess, I have found the stationary equilibrium. Otherwise, I update the guess. For the new guess, I use a convex combination between the initial guess and the implied one.

E.2 Algorithm for Transition Dynamics

The initial stationary equilibrium is calibrated to match features of the US economy in the late 1980s. I denote the probability of re-optimization in the initial stationary equilibrium by θ_0. I then consider an unexpected and deterministic shock to that probability that according to observed and predicted demographic shifts in the US population. In particular, I consider a vector of θ of length 51, corresponding to the period between the early 1990s and 2050. I assume that after 2050 the age composition does not change so that the re-optimization probability is kept constant at its 2050 level. I further assume that the model converges to the new stationary equilibrium by 2100. I denote the period 2100 by T.

Similar to the previous section, I start by computing the terminal stationary equilibrium. The only difference is that when computing the terminal equilibrium, the measure of firms J is not normalized to 1 but is instead endogenous. It is set so that the expected zero-profits condition in equilibrium holds. So instead of iterating over 4 aggregate endogenous variables, I iterate over 5 aggregate endogenous variables.

Solving the transition dynamics of the economy consists of two main steps:

0) Start with an initial guess for the aggregate endogenous variables along the transition path, $\{P_{mt}, P_{Bt}, J_t, F_t, \Pi_t\}_{t=1}^{T}$.

1) Solve backwards the present value of firms, and obtain the policy rules in each period. The terminal present value of firms is that of the terminal stationary equilibrium.

2) Iterate forward to compute the distribution of firms in every period. Using the distribution of firms and the policy rules, compute the implied aggregate endogenous variables. If the difference between the guess and implied variables is not close enough, update the guess and repeat from step (1). The initial distribution of firms is the ergodic distribution of firms in the initial stationary distribution.

References

ANDERSON, S. P., DEPALMA, A. and THISSE, J.-F. (1987). The ces is a discrete choice model. *Economics Letters*, **24** 139.

ANDERSON, S. T., KELLOGG, R., LANGER, A. and SALLEE, J. M. (2015). The intergenerational transmission of automobile brand preferences. *The Journal of Industrial Economics*, **63** 763–793.

AUTOR, D., DORN, D., KATZ, L. F., PATTERSON, C. and VAN REENEN, J. (2017). *The fall of the labor share and the rise of superstar firms*. National Bureau of Economic Research.

BAQAEE, D. R. and FARHI, E. (2017). Productivity and misallocation in general equilibrium. Tech. rep., National Bureau of Economic Research.

BARKAI, S. (2016). Declining labor and capital shares. *Stigler Center for the Study of the Economy and the State New Working Paper Series*, **2**.

BARKAI, S. and BENZELL, S. G. (2018). 70 years of us corporate profits. *Stigler Center for the Study of the Economy and the State New Working Paper Series*.

BRONNENBERG, B. J., DHAR, S. K. and DUBÉ, J.-P. (2007). Consumer packaged goods in the united states: National brands, local branding. *Journal of Marketing Research*, **44** 4–13.

BRONNENBERG, B. J., DUBÉ, J.-P. H. and GENTZKOW, M. (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, **102** 2472–2508.

DAVIS, S. J. and HALTIWANGER, J. (2014). Labor market fluidity and economic performance. Tech. rep., National Bureau of Economic Research.

DE LOECKER, J. and EECKHOUT, J. (2017). The rise of market power and the macroeconomic implications. Tech. rep., National Bureau of Economic Research.

DECKER, R., HALTIWANGER, J., JARMIN, R. and MIRANDA, J. (2014a). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, **28** 3–24.

DECKER, R., HALTIWANGER, J., JARMIN, R. and MIRANDA, J. (2014b). The secular decline in business dynamism in the us. *Unpublished draft, University of Maryland*.

DECKER, R. A., HALTIWANGER, J., JARMIN, R. S. and MIRANDA, J. (2016). Where has all the skewness gone? the decline in high-growth (young) firms in the us. *European Economic Review*, **86** 4–23.

DUBÉ, J.-P., HITSCH, G. J. and ROSSI, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, **41** 417–445.

EDMOND, C., MIDRIGAN, V. and XU, D. Y. (2018). How costly are markups? Tech. rep., National Bureau of Economic Research.

EINAV, L., LEIBTAG, E. and NEVO, A. (2010). Recording discrepancies in nielsen homescan data: Are they present and do they matter? *QME*, **8** 207–239.

ENGBOM, N. (2017). Firm and worker dynamics in an aging labor market. Tech. rep., Technical Report, Working Paper.

FARRELL, J. and KLEMPERER, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of industrial organization*, **3** 1967–2072.

FARRELL, J. and SHAPIRO, C. (1988). Dynamic competition with switching costs. *The RAND Journal of Economics* 123–137.

FOSTER, L., HALTIWANGER, J. and SYVERSON, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, **98** 394–425.

FOSTER, L., HALTIWANGER, J. and SYVERSON, C. (2016). The slow growth of new plants: Learning about demand? *Economica*, **83** 91–129.

GABSZEWICZ, J., PEPALL, L. and THISSE, J.-F. (1992). Sequential entry with brand loyalty caused by consumer learning-by-using. *The Journal of Industrial Economics* 397–416.

GARCIA-MACIA, D., HSIEH, C.-T. and KLENOW, P. J. (2016). How destructive is innovation? Tech. rep., National Bureau of Economic Research.

GILCHRIST, S., SCHOENLE, R., SIM, J. and ZAKRAJŠEK, E. (2017). Inflation dynamics during the financial crisis. *American Economic Review*, **107** 785–823.

GOURIO, F. and RUDANKO, L. (2014). Customer capital. *Review of Economic Studies*, **81** 1102–1136.

GRULLON, G., LARKIN, Y. and MICHAELY, R. (2017). Are us industries becoming more concentrated?

GUTIERREZ, G. (2017). Investigating global labor and profit shares. *Unpublished Working Paper, SSRN*.

GUTIERREZ, G. and PHILIPPON, T. (2017). Declining competition and investment in the us. Tech. rep., National Bureau of Economic Research.

HALL, R. E. (2018). New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy. Tech. rep., National Bureau of Economic Research.

HANDEL, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, **103** 2643–82.

HATHAWAY, I. and LITAN, R. (2014). The other aging of america: the increasing dominance of older firms. *Brookings Institution*.

HECKMAN, J. J. (1981). Heterogeneity and state dependence. In *Studies in labor markets*. University of Chicago Press, 91–140.

HOPENHAYN, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society* 1127–1150.

HORTAÇSU, A., MADANIZADEH, S. A. and PULLER, S. L. (2015). Power to choose? an analysis of consumer inertia in the residential electricity market. Tech. rep., National Bureau of Economic Research.

ILLANES, G. (2016). Switching costs in pension plan choice. *Unpublished manuscript*.

JIANG, H. and SOHAIL, F. (2017). Skill biased entrepreneurial decline. Tech. rep., Mimeo, Washington University in St. Louis.

KARABARBOUNIS, L. and NEIMAN, B. (2018). Accounting for factorless income. In *NBER Macroeconomics Annual 2018, volume 33*. University of Chicago Press.

KARAHAN, F., PUGSLEY, B. and SAHIN, A. (2016). Demographic origins of the startup deficit. In *Technical Report*. New York Fed, mimeo.

KLEMPERER, P. (1987a). Entry deterrence in markets with consumer switching costs. *The Economic Journal*, **97** 99–117.

KLEMPERER, P. (1987b). Markets with consumer switching costs. *The quarterly journal of economics*, **102** 375–394.

KOZENIAUSKAS, N. (2017). What's driving the decline in entrepreneurship. *manuscript, New York University*.

LIANG, J., WANG, H. and LAZEAR, E. P. (2014). Demographics and entrepreneurship. Tech. rep., National Bureau of Economic Research.

LUCO, F. (2017). Switching costs and competition in retirement investment.

MALIAR, L., MALIAR, S. and VALLI, F. (2010). Solving the incomplete markets model with aggregate uncertainty using the krusell–smith algorithm. *Journal of Economic Dynamics and Control*, **34** 42–49.

MOREIRA, S. (2016). Firm dynamics, persistent effects of entry conditions, and business cycles.

NAKAMURA, E. and STEINSSON, J. (2011). Price setting in forward-looking customer markets. *Journal of Monetary Economics*, **58** 220–233.

NEIMAN, B. and VAVRA, J. (2018). The rise in household spending concentration.

NOSAL, K. (2012). Estimating switching costs for medicare advantage plans. *Unpublished manuscript, University of Mannheim*.

PUGSLEY, B. W. and ŞAHIN, A. (2015). Grown-up business cycles. *The Review of Financial Studies*.

RAVN, M., SCHMITT-GROHÉ, S. and URIBE, M. (2006). Deep habits. *The Review of Economic Studies*, **73** 195–218.

ROGNLIE, M. (2018). Comment on "accounting for factorless income"(karabarbounis and neiman, nber macro annual 2018).

ROTEMBERG, J. J. and WOODFORD, M. (1991). Markups and the business cycle. *NBER macroeconomics annual*, **6** 63–129.

SALGADO, S. (2017). Technical change and entrepreneurship.

SCHMITT-GROHÉ, S. and URIBE, M. (2007). Why the monopolist does not deviate from the symmetric equilibrium in a model with good-specific subsistence points. *Note, Duke University*.

SEDLÁČEK, P. and STERK, V. (2017). The growth potential of startups over the business cycle. *American Economic Review*, **107** 3182–3210.

SHIMER, R. (2001). The impact of young workers on the aggregate labor market. *The Quarterly Journal of Economics*, **116** 969–1007.

TAUCHEN, G. and HUSSEY, R. (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica: Journal of the Econometric Society* 371–396.

WINTER, S. and PHELPS, E. (1970). Optimal price policy under atomistic competition. *Microeconomic Foundations of Employment and Inflation Theory* 309–337.