Doctoral Thesis

The Individual Anomalies and Aggregate Impacts of Households' Consumption and Savings Decisions

Nick Pretnar

Carnegie Mellon University
Tepper School of Business

April 14, 2020

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Dissertation Committee

Finn Kydland
Carnegie Mellon University
Tepper School of Business
NBER & UCSB

Alan Montgomery
Carnegie Mellon University
Tepper School of Business

Stephen Spear
Carnegie Mellon University
Tepper School of Business

Ali Shourideh
Carnegie Mellon University
Tepper School of Business

Christopher Y. Olivola
Carnegie Mellon University
Tepper School of Business

Espen Henriksen
BI Norwegian Business School
Oslo, Norway

Abstract

An economy, at its core, consists of people doing things. People work and play, eat and drink, create and think, and all the while perform activities both measured and unmeasured that nonetheless provide value to themselves, their families, and society at large. Within this structure, the most fundamental unit of our modern economy is the household, which consists of at least one consumer who spends time consuming some positive amount of net resources. This consumer may also spend time supplying labor or investing in new ideas or various speculative enterprises so that he can afford his desired consumption level. Even if he does not work or invest, he still must consume in order to ensure his sustenance and continued living. Thus, while the consumer comprises the core unit of the modern economy, it is consumption itself which is the economy's most central activity. In the United States, consumption amounts to almost 70% of all output. To better understand the many important factors that drive broad aggregate economic outcomes, it behooves us to study the behavior of individual consumers themselves.

A goal of mine as a researcher is to continue working to expand the frontier of knowledge pertaining to the motivations and outcomes of household behaviors. And while this thesis does not reflect all of my economic research projects currently underway, it does fully reflect the variegated flavors of the questions and puzzles I consider. Here, I focus on two particular frontiers in household economic research. In Chapters 1 and 2, along with co-authors, I consider how the allocation of off-market time directly affects consumption, savings, the labor supply, and aggregate growth. In Chapters 3 and 4, along with co-authors, I apply theories from behavioral economics, like mental accounting and the non-fungibility of various forms of liquidity, to explore the high-frequency spending and savings patterns of individual consumers. While the first two chapters consider how mechanisms behind household decisions drive macroeconomic outcomes, the latter chapters deal with granular, microeconomic behavior. This thesis thus represents the breadth of my many research interests spanning both macro- and microeconomics.

Since Becker (1965), economists focussing on household decisions have grappled with the classic consumption/leisure tradeoff in various applications. However, rich decision structures, like Becker's original theoretical model featuring multiple off-market time utilization decisions, have been only minimally explored. Data limitations may be to blame: in the United States quality time use data has only been available since 2003, while other developed countries lack a comprehensive time use survey of households. Still though, we can gain further insight into many well-established economic puzzles, like the reason for the rise in the U.S. services share I explore in Chapter 1, or the decline in long-run U.S. GDP growth I explore in Chapter 2, among others, by examining these phenomena in model environments where households face rich, previously unexplored time use decisions.

Chapter 1 features joint work with co-author William Bednar. We consider an application of a home production model toward United States structural change where households decide how much time to spend consuming various market purchases. In the context of our problem, structural change is the process by which the share of consumption devoted toward intangible services has risen while the share of consumption devoted toward physical, manufactured goods has fallen. There is contentious debate in the literature as to whether supply-side factors or demand-side factors have driven the rise in the services share of aggregate output. Income effects from non-homothetic consumer preferences have been touted as a primary contributor. We test this implication in a model that accounts for consumers' joint consumption and off-market time allocation decisions. When accounting for time to consume in this manner, we show that homothetic utility functions can still generate non-linear expansion paths as wages increase. In the model, differences in the time intensities of different home production activities affect how consumers adjust their consumption allocation in response to relative price and real wage changes. Our findings suggest that the rise in the U.S. services share since 1948 is primarily due to relative price changes which dominate income effects from wage growth, contrasting with many findings in the literature.

In Chapter 2, I work with co-author Finn Kydland to address how population aging will impact aggregate GDP growth when accounting for the time working-age adults spend caring for their elders. As the population of the United States ages, the number of elderly people who require living assistance is increasing. To understand how this impacts aggregate output, we calibrate an overlapping generations model where growth endogenously depends on the care young agents choose to provide for their parents. Relative to an economy with a constant population distribution, we project that population aging will reduce GDP 17% by 2056 and 39% by 2096. Curing old-age diseases such as Alzheimer's and dementia can lead to 5.4% higher output relative to the baseline, while improving welfare for consumers of all ages.

The mental accounting theories of Richard Thaler have inspired a generation of behavioral scientists to reconsider how consumer behavior departs from predictions of classic economic models of rationality. Until recently, Thaler's theories have been largely isolated to applications in experimental and other controlled settings. Further, there lacks broad consensus as to what types of behavior exactly correspond to mental accounting. In Chapters 3 and 4, along with co-authors Alan Montgomery and Christopher Y. Olivola, I apply Thaler's various theories of mental accounting to transaction-level field data from consumers' bank balance ledgers. Generally, we seek to understand the degree to which mental accounting behavior is actually observed in consumption spending data and also the degree to which including flexible, behavioral features like mental accounting can improve the fit and predictions of structural models of demand.

In Chapter 3 we construct a unifying theory of two stage budgeting and mental accounting in order to reconcile heterogeneity in consumer-level weekly spending and savings patterns. Mental accounting and rational inattention induce behavioral wedges between first stage expenditure budgets and second stage actual expenditure. Specifically, consumers engage in Gabaix (2014) sparse maximization, re-assessing only a subset of their spending budgets every period. Over or under spending affects future budgeting and expenditure decisions. With agent-level weekly expenditure data, we use latent Bayesian inference to structurally estimate the degree to which low-income consumers appear constrained by mental accounting frictions. We find that consumers optimally set only 25-50% of first stage budgets each period. A sparse max model with mental accounting fits the data best, compared to alternative models without one and/or the other. In a counterfactual experiment, relaxing rationality constraints leading to greater budget attentiveness is not necessarily welfare improving if consumers can easily adjust budgets on the fly to mitigate the disutility of over expenditure. We are the first to estimate the structural parameters and latent decisions of a two stage budgeting model with sparse maximization and mental accounting. In doing so, demand shifters in our estimation are endogenous, resulting from behavioral frictions.

In the final chapter I engage in a second application of mental accounting theory to explore empirical evidence that consumers use liquidity from debit cards and credit cards differently. Thaler (1999) describes one of the primary components of mental accounting as the budgeting of specific utility-providing activities which can depend on the resources used to fund those activities. The analysis presented in this chapter focusses on household expenditure of durable and non-durable goods and the liquidity sources used to fund these different expenditures. Specifically, we exploit a linked dataset of credit and debit card users to examine consumer purchasing patterns of durable and non-durable consumption commodities under both methods of payment. Our findings suggest that on average durable purchases are more sensitive to increases in available credit than non-durable purchases, and most consumers are more likely to increase total consumption due to increases in available credit than increases in available checking account balances. We empirically show that the standard neo-classical consumption/savings model, the equilibrium conditions of which implicitly assume that the household's available resources (liquidity and investments) are perfectly fungible, fails to rationalize our data for the median/modal consumer in our sample. However, our results are rich because we also show that the behavioral distribution of consumers includes both households which treat liquidity as fungible and those that do not. Given the heterogeneity we find, future work should test whether these results would matter on aggregate.

Contents

1. Home Production with Time to Consume

joint with William Bednar

1.1. Introduction

When considering how households use market purchases, complementarities exist between the consumption of these purchases and non-work time. In frameworks with only one consumption commodity and elastic labor supply, complementarities between leisure and consumption are explicitly considered. However, this is usually not the case in models where consumers derive utility from multiple consumption commodities. Under this premise, we explore the fundamental question as to why household consumption allocations vary in relative prices and income. Our contention is that demand for different market purchases depends ultimately on how households spend time using their purchases in various home production activities. Indeed, Gary Becker recognized this in his seminal paper on home production, "A Theory of the Allocation of Time" (Becker 1965). In a model where households choose both market purchases and how to allocate time toward their consumption, both the relative productivities and labor intensities of different home production processes determine the responsiveness of the consumption allocation to relative prices and income. In this sense, we provide a micro-foundational explanation rooted in home production for why demand is sensitive to price and income changes. To our knowledge, this paper is the first to show that income effects can still be generated by a homothetic utility function as long as Beckerian time use and consumption complementarities are considered.

Our home production formulation differs from those of Gronau (1977), Graham and Green (1984), Benhabib, Rogerson, and Wright (1991), Rupert, Rogerson, and Wright (1995), Ingram, Kocherlakota, and Savin (1997), and Boerma and Karabarbounis (2019) who all dispense with the idea that market goods *and* time are combined together in some home production function to produce a final commodity. Notable exceptions are McGrattan, Rogerson, and Wright (1993), Gomme, Kydland, and Rupert (2001), and Bridgman, Duernecker, and Herrendorf (2018) which allow for durable capital to be combined with time toward the production of home consumption commodities, but additional non-durable market goods are still absent from their home production formulations. Micro-data analyses, like those in Aguiar and Hurst (2005, 2007) and Aguiar, Hurst, and Karabarbounis (2013), provide evidence that consumer expenditure and off-market time use decisions are interdependent, suggesting home production models should take such complementarities into consideration.

The relationship between off-market time use and market purchases is most evident when considering how consumers use physical, consumption commodities, i.e. goods. However, even when using market services, like for example hiring a home cleaning service, consumers must at least briefly spend time finding the maid and explaining to the hired hand what cleaning needs to be done. While this amount of time is less than the amount of time required to clean the house oneself, the amount of time is *not* zero. Time use complementarities are more readily apparent when thinking about the inputs required in such a process as making a meal. Meanwhile, this model of household behavior challenges the conventional assumption made in modern macroeconomic models to classify as leisure all time spent outside of a formal, income-producing job.[1] The Beckerian approach, rather, is to model all off-market time as spent engaging in some home production activity, one of which could be leisure.

To illustrate the importance of time use and market purchase complementarities, consider someone who purchases a market good, like say a boat, and does not spend any time using it. Can we truly say that the boat is providing value to this consumer? He works overtime nights and weekends to pay for a boat that he never uses. It is well-established common knowledge that boats depreciate in value very quickly, so surely this consumer does not view the purchase as a capital investment. He must make time in his schedule to use the boat and derive utility from it. He thus faces an implicit tradeoff between working more to earn more money to pay for the boat and all the maintenance it requires versus spending time actually using the boat and deriving utility from it. If he knows he does not have the time to both supply labor to pay for the boat and enjoy the boat once he has it, he will not make the decision to purchase the boat.[2] In this way, allowing for time use and consumption complementarities also provides a richer characterization of the tradeoff households face when making labor supply decisions.

Modeling explicit complementarities between time use and market purchases assumes two ideas relating to consumer behavior: 1) consumers enjoy market purchases and derive final utility from their consumption by spending time using them; 2) the return consumers derive from using market purchases in home production processes depends on the ability of the consumer to engage in transformation of market purchases to final consumption. Our model is flexible enough to capture all limiting conditions, where time is valued on its own, absent market purchase complimentarities, and vice-versa.

We apply our micro-founded model of the household toward a general equilibrium quantitative analysis of United States structural change.[3] This is a natural application of our framework given observed changes in relative prices and income are associated with the changing composition of consumption expenditure. Using aggregate U.S. goods and services expenditure data, we estimate our model to understand the degree to which wage growth versus relative price effects are responsible for the rising services share. The results suggest that when both accounting for time use complementarities in home production and households using service flows from consumer durables, relative price effects have almost exclusively driven the rise

[1] An exception here is the analysis in Aguiar, Hurst, and Karabarbounis (2013) which distinguishes between off-market leisure time and off-market non-leisure time.

[2] Or at least he shouldn't.

[3] Others have used Stone-Geary preferences to capture the responsivenss of relative demand changes to changes in relative prices and income (Geary 1950; Stone 1954). For examples, see Kongsamut, Rebelo, and Xie (2001), Buera and Kaboski (2009), Matsuyama (2009), Herrendorf, Rogerson, and Valentinyi (2013), Uy, Yi, and Zhang (2013), and Bridgman, Duernecker, and Herrendorf (2018). Alternatively, Boppart (2014) uses a form of "price independent generalized linearity" (PIGL) preferences described by Muellbauer (1975, 1976) that admit aggregation across households.

in the services share since 1948. This empirical finding lends credence to theories contending that supply-side factors, like changes to relative sectoral productivities, skill-biased technical change, or differentials in sectoral capital deepening, are most responsible for the structural evolution of the post-war U.S. economy.

Accounting for how households combine off-market time with market purchases is especially important when considering the utility they derive from the service flows of consumer durables. To the best of our knowledge, the structural change literature thus far fails to account for how the service flows from consumer durables generate consumption utility. Our quantitative results regarding structural change are robust to the inclusion of different measures of consumer durable asset service flows in the overall goods series. After counterfactually fixing relative prices at those observed in 1948, wage gains do not alone affect changes in the relative consumption basket. Thus, using our home production framework, supply-side factors affecting market prices appear to be the primary drivers of structural change. This lends credence to stories suggesting differential rates of sectoral capital deepening and/or technological advancement have primarily contributed to the rise in the services share of the United States economy. For example, both Ngai and Pissarides (2007) and Kehoe, Ruhl, and Steinberg (2018) show that structural transformation can be driven by differentials in sectoral productivities which lead to capital deepening and force labor to shift to the less productive sector, driving up market prices in that sector (services) relative to the more productive one (goods). The quantitative analyses in Kehoe, Ruhl, and Steinberg (2018) take place in an open economy setting with a global savings glut featuring differences in capital flows between countries as well as non-homothetic preferences. In fact, the results in Kehoe, Ruhl, and Steinberg (2018) suggest that non-homothetic preferences are not particularly important for generating structural change. Acemoglu and Guerrieri (2008) achieve similar sectoral dynamics with sectoral differences in factor shares for production inputs, leading to differential rates of capital accumulation. Buera and Kaboski (2012), on the other hand, show that the relative rate of structural change depends on the productivity advantage of high-skilled workers in predominantly service industries versus their low-skilled counterparts working in manufacturing. Autor and Dorn (2013) tell a story with implications that contradict the premises but not the results in Buera and Kaboski (2012): automation in manufacturing has driven low-skilled workers to low-wage service industry jobs. As consumer preferences have evolved to favor variety over specialization, new services are created, but these jobs are occupied mostly by low-skilled workers.

While our empirical results justify continued exploration of the degree to which supply-side phenomena may be driving structural change, we focus mostly in this paper on the decision process of households engaging in home production, allowing the supply-side of the economy to exhibit a fairly standard sector-specific structure. This paper proceeds as follows. First, in Section 1.2 we outline a model economy where households engage in Beckerian home production. We document the model's implications for home production value-added, income and relative price effects, and household labor supply decisions. Then in Section 1.3 we present both aggregated and cross-sectional empirical facts and estimate our structural model using a Hamiltonian Monte Carlo (HMC) routine to pin down the posterior distribution of structural parameters conditional on aggregate data. With parameters in hand, we then engage in counterfactual simulations on our demand representation to understand how increases in wages and changes to relative prices have contributed to demand-driven structural change. In Section 1.4 we conclude.

1.2. Model Economy

Consider a closed economy where market prices are determined endogenously by general equilibrium market clearing conditions. The economy consists of finitely many infinite-lived firms and finitely many infinite-lived households. Firms make market goods in some stylized production process which utilizes capital and labor leased from households. Households purchase market goods and use them along with their time spent away from work to produce final consumption goods in the home, taking market prices as given. Ultimately, households consume and derive final utility from the non-traded output of these off-market home production activities.

In Section 1.2.1 we construct a stylized model of household decisions which captures fundamental aspects of our theory. In our set up, the activities of firms follow fairly standard conventions in aggregate economics, so we only briefly touch on them in Section 1.2.2, where we also outline the general equilibrium market clearing conditions which must hold. The equilibrium is defined in Section 1.2.3. Finally, in Section 1.2.4 we demonstrate how our model of the household allows for a homothetic utility structure that can surprisingly also accommodate income effects, while additonally showing implications for our set up regarding labor supply decisions.

1.2.1. General Model of the Household

We will develop the model at the household level indexed by h. Time is discrete and indexed by t. Each period a household derives utility from the consumption of I final goods c_{ith} each produced under the home production process f_{ith}, which takes as inputs a J_i dimensional vector of market goods $\boldsymbol{q}_{ith}$ and time n_{ith}.[4] Final consumption is such that

$$c_{ith} = f_{ith}(\boldsymbol{q}_{ith}, n_{ith}) \quad \forall i, t, h \tag{1.1}$$

We dispense with modeling fixed durable assets as inputs. Assume instead that some components of market inputs contain the service flows from durables, which ultimately are what the household uses when engaging in home production. In our quantitative exercises we will construct the data series so that the value of service flows from durable goods held by households is included in the value of market goods as if the household rents durables from some firm "producing" new service flows each period. Price indices will be adjusted to account for this via procedures described in Appendix A.1.

Let $\boldsymbol{q}_{th}$ be a J dimensional vector whose components are each of the commodities on the market place. Let $\boldsymbol{P}_t$ be a J dimensional vector of market prices and $\boldsymbol{x}_{th}$ the associated expenditure vector, the components of which are such that $P_{jt}q_{jth} = x_{jth}$.

[4] All vectors are column vectors and denoted using bold font.

Assumption 1. There is no joint production using market goods. That is, if q_{jth} is a component of $\boldsymbol{q}_{ith}$ then q_{jth} cannot be a component of $\boldsymbol{q}_{i'th}$ for any $i' \neq i$.

Assumption 2. All input resources are used up in the production of final consumption commodities. That is $\forall j \in \{1, \ldots, J\}$, $\exists i \in \{1, \ldots, I\}$ such that q_{jth} is a component of $\boldsymbol{q}_{ith}$.

Assumptions 1 and 2 together imply that $J = \sum_{i=1}^{I} J_i$.[5]

Each period households derive flow utility $u(\boldsymbol{c}_{th})$ from consumption of the I dimensional vector of final goods $\boldsymbol{c}_{th}$.[6]

Assumption 3. $u(\boldsymbol{c}_{th})$ is separable across components of $\boldsymbol{c}_{th}$. That is, for all $k \neq i$ and $k \neq j$, the marginal rate of substitution between final consumption derived from processes i and j is independent of consumption in process k.

The separability imposed by Assumption 3 will allow us to invoke an aggregation theorem to collapse the market commodity space into indices such that the production process for each final good takes one and only one market good as input. This is described in Lemma 1 and its proof, which requires specification of equilibrium conditions. Before moving on to that, let us finish characterizing the household's choices.

Let $\overline{n}$ denote the total time available to the household, and assume that all households face the same time constraints. Households earn wages w_{th} from supplying labor l_{th} on the open market place. Households do not care how their labor is allocated toward production of market goods. They choose n_{ith} which is time spent engaging in home production activity i. Let $\boldsymbol{n}_{th}$ denote the I-dimensional vector describing time spent on home activities. Total time allocated to market and home activities must satisfy:[7]

$$l_{th} + \sum_{i=1}^{I} n_{ith} \leq \overline{n} \tag{1.2}$$

Households choose to invest in capital $k_{t+1,h}$, receiving gross return $R_t k_{th}$ on current investments. Given this information, we can write the household's problem:

$$\max_{\boldsymbol{q}_{th}, \boldsymbol{n}_{th}, k_{t+1,h}} \quad \sum_{t=0}^{\infty} \beta^t u(\boldsymbol{c}_{th}) \tag{1.3}$$

$$\text{subject to} \quad \sum_{j=1}^{J} P_{jt} q_{jth} + w_{th} \sum_{i=1}^{I} n_{ith} \leq w_{th}\overline{n} + R_t k_{th} - k_{t+1,h} \tag{1.4}$$

$$c_{ith} = f_{ith}(\boldsymbol{q}_{ith}, n_{ith}) \qquad \forall i, t, h \tag{1.5}$$

[5]Assumption 1 allows us to avoid the parameter identification issues in home production models with joint production described in a back and forth between Bill Barnett, Robert Pollak, and Michael Wachter in the 1970s (Pollak and Wachter 1975; Barnett 1977). Special thanks to Javier Birchenall for pointing this out.

[6]These "goods" are the outputs of home production activities. In the spirit of Becker (1965), a single component of $\boldsymbol{c}_{th}$ might capture the total enjoyment one feels both from spending time cooking *and* time eating a meal. We also refer to the components of $\boldsymbol{c}_{th}$ as distinct household "activities."

[7]Note that each home production process is associated with multiple market inputs but only one time input. Therefore, we have separate indices for market inputs used in each home production activity, but the time use vector is I dimensional.

We do not need to explicitly specify a parameterization for $f_{ith}(\boldsymbol{q}_{ith}, n_{ith})$. Given the assumptions above, we can invoke an aggregation theorem over commodities to make the analysis more compact.[8] This is described in Lemma 1.

Lemma 1. Assume each household is a utility maximizer. Under Assumptions 1, 2, and 3, and under Theorem 1 of Green (1964) attributed to Leontief (1947), we can restrict our analysis to

$$\widetilde{u}_{th}(q_{1th}, \ldots, q_{ith}, \ldots, q_{Ith}, n_{1th}, \ldots, n_{Ith}) \tag{1.6}$$

where q_{ith} is some index that describes the grouping of market goods $\{q_{i1th}, \ldots, q_{ij_ith}, \ldots, q_{iJ_ith}\}$.

Proof. See Appendix A.2.3. □

The intuition behind Lemma 1 is that Assumptions 1 thru 3 guarantee that the intratemporal marginal rate of substitution for two goods used in the same home production process is independent of other goods not used in that process. Under Lemma 1 we can then form a single composite commodity q_{ith} which describes the market value of the entire vector of commodities $\boldsymbol{q}_{ith}$ used in production process f_{ith}.[9] Thus rather than specifying a functional form for $f_{ith}(\boldsymbol{q}_{ith}, n_{ith})$ from here on our analysis operates on $c_{ith} = \widetilde{f}_{ith}(q_{ith}, n_{ith})$, the commodity-aggregated home production function for final good c_{ith}.[10]

Assumption 4. The aggregated home production function $\widetilde{f}_{ith}(q_{ith}, n_{ith})$ is strictly increasing, quasi-concave, and homogeneous of degree one.

Assumption 4 is used in our equilibrium results to characterize the value added to household market purchases from engaging in home production activities. Composing $u(\boldsymbol{c}_{th})$ with each $\widetilde{f}_{ith}(q_{ith}, n_{ith})$ gives us $\widetilde{u}_{th}$:

$$u(\boldsymbol{c}_{th}) = u\Big(\widetilde{f}_{1th}(q_{1th}, n_{1th}), \ldots, \widetilde{f}_{ith}(q_{ith}, n_{ith}), \ldots, \widetilde{f}_{Ith}(q_{Ith}, n_{Ith})\Big) = \widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th}) \tag{1.7}$$

Unlike $u(\boldsymbol{c}_{th})$, composed utility depends on the possibly time-varying home production process, so we index it with both t and h to account for this. Each home production process is associated with time and one market input, which implies that there are the same number of market goods as final goods, i.e. $J = I$, in the commodity-aggregated problem.

1.2.1.1. Household Equilibrium Conditions

Let μ_{th} denote the period t marginal utility of wealth, i.e. the Lagrange multiplier on the budget constraint. Each period household choices must satisfy the budget constraint plus the following

[8] Thanks to Laurence Ales for pointing this result out.

[9] Scalar q_{ith} is the composite good. Again, bold font is reserved for vectors.

[10] We admit our usage of the word "aggregated" is slightly abusive throughout this paper. To be clear, $\widetilde{f}_{ith}$ is an "aggregated" home production function in the sense that it takes the composite q_{ith} as an input, where q_{ith} is the sum over the quantities of all commodities in its class. We will also use the word "aggregate" to describe total expenditure and consumption in the entire United States economy for specific commodity classes, i.e. "goods" and "services."

intratemporal and intertemporal conditions:

$$\frac{\partial u}{\partial c_{ith}}\frac{\partial \widetilde{f}_{ith}}{\partial q_{ith}} = P_{it}\mu_{th} \qquad \forall i,t,h \tag{1.8}$$

$$\frac{\partial u}{\partial c_{ith}}\frac{\partial \widetilde{f}_{ith}}{\partial n_{ith}} = w_{th}\mu_{th} \qquad \forall i,t,h \tag{1.9}$$

$$\mu_{th} = \beta\, \mathbb{E}_t R_{t+1}\mu_{t+1,h} \qquad \forall i,t,h \tag{1.10}$$

For each final activity c_{ith} we can combine the equilibrium conditions for the marginal utilities of q_{ith} and n_{ith} to arrive at an expression describing the marginal rate of technical substitution between time and market inputs for process i:

$$\frac{\partial \widetilde{f}_{ith}}{\partial q_{ith}}\Big/\frac{\partial \widetilde{f}_{ith}}{\partial n_{ith}} = \frac{P_{it}}{w_{th}} \tag{1.11}$$

Of interest are the tradeoffs faced by consumers when engaging in market purchases. Consider the following expression describing the marginal rate of substitution between different final activities c_{jth} and c_{ith}:

$$\frac{\partial u}{\partial c_{jth}}\Big/\frac{\partial u}{\partial c_{ith}} = \frac{P_{jt}}{\frac{\partial \widetilde{f}_{jth}}{\partial q_{jth}}}\frac{\frac{\partial \widetilde{f}_{ith}}{\partial q_{ith}}}{P_{it}} \tag{1.12}$$

The two terms on the right hand side of (1.12) are the shadow prices, with respect to the internal household marketplace, of consuming c_{ith} and c_{jth}. If the market price P_{it} is in dollar units, then the shadow price of c_{ith} is equal to the dollar-value of market inputs per unit of output from process $\widetilde{f}_{ith}$.

Lemma 2. The shadow price of activities c_{ith} associated with the consumption of q_{ith} is equal exactly to P_{it} if and only if $\frac{\partial \widetilde{f}_{ith}}{\partial q_{ith}} = 1$.

Proof. See Appendix A.2.3. □

Lemma 2 shows that the marginal value of final consumption is only exactly equal to the market price of inputs in the absence of home production frictions inducing diminishing returns. When the marginal product of market inputs is unity the production function is perfectly linear in q_{ith}. Thus, increases in market inputs do not result in diminishing marginal activities. That is, the amount of activity associated with the consumption of q_{ith} constantly increases at the same rate across all levels of market inputs. Let us consider why diminishing marginal returns may make more sense. First, holding time n_{ith} fixed, adding another unit of market inputs q_{ith} will conceivably lead to a smaller increase in final consumption output due to the fact that with more market purchases and the same amount of time, the amount of time consumers have to use each specific input decreases, leading to unusable purchases, i.e. waste. Similarly, holding q_{ith} fixed and increasing n_{ith}, consumers now devote more time toward using each market purchase, wasting time on frivolous tasks using such purchases after the totality of their usefulness has been reached.

It is often common in economic analyses to think of the household as ultimately a consumer rather than a producer. Truly, the household engages in both production and consumption tasks, using its time to manipulate market purchases into a final consumable item, like for example a meal. At each step along a supply chain the value of the new outputs created using inputs is at least the sum of the values of the inputs used as long as producers are making non-negative economic profits. Thus, we should expect that if time and market inputs are used to produce an output in the home, additional value added should ensue as in all other steps along the supply chain. The explicit market costs of home production are simply the cost of market purchases. Proposition 1 states that the value added from home production, above and beyond explicit costs, is the market value of time used in the home production process. Unlike in production at the firm level, the laborer and the ultimate end-user of output are necessarily the same. Thus, the market value of the time the consumer spends engaging in specific home production tasks directly quantifies the additional value his efforts provide him, since he faces the opportunity cost of not working on the market and earning more income which he could use to purchase additional market inputs. The model structure, in fact, imposes an opportunity cost calculation to value-added in contrast to the replacement wage approach used in other papers such as Bridgman, Duernecker, and Herrendorf (2018).

Proposition 1. For each i, the value of the production of final good c_{ith} net of the cost of market expenditure (value added) is equal to $w_{th}n_{ith}$, the market value of time spent on task i.

Proof. See Appendix A.2.3. □

The proof of Proposition 1 requires two applications of Euler's theorem for homogeneous functions and is left for inspection in the attached appendix. Corollary 1 demonstrates why failing to account for time use complementarities in home production leads to extreme results.

Corollary 1. If $c_{ith} = q_{ith}$, so that consumers derive utility directly from market purchases, then home production provides no additional value to the household.

Corollary 1 may seem obvious: of course there can be no value added from a process that never happens. But what this says is that if we fail to account for how households spend their off-market time using market purchases, then we are essentially saying that engaging in home production provides no additional value to the household beyond the value of the market purchases themselves. This is an extreme statement that says a meal cooked and prepared in the household is only as valuable as the sum of all the market commodities used to prepare it. Under such a modeling assumption, the intrinsic skills of the homemaker contribute nothing to the value of the final meal. This result thus demonstrates how splitting the off-market time allocation decision into a vector of decisions over activities while also allowing for time use complementarities with market purchases provides a mechanism for quantifying and capturing the value of engaging in home production.

1.2.1.2. Parameterization of Demand

To perform comparative statics and ultimately estimate the model, we make standard functional form assumptions. We choose a constant elasticity of substitution (CES) form for $u(\boldsymbol{c}_{th})$:

$$u(\boldsymbol{c}_{th}) = \left(\sum_{i=1}^{I} \theta_i c_{ith}^{\rho} \right)^{\frac{1}{\rho}} \tag{1.13}$$

ρ parameterizes the intratemporal elasticity of substitution which is $\frac{1}{1-\rho}$. Let z_{ith} describe the total factor productivity of process $\widetilde{f}_{ith}$. This term captures several things: 1) a household's exogenously evolving ability to spend time using market commodity i in order to produce the final consumption commodity; 2) the intrinsic value to a household of quality gains to q_{ith}; 3) folk knowledge possessed by the household as to how best accomplish the production of c_{ith}. We specify a flexible CES form for the composite-commodity aggregated home production functions $\widetilde{f}_{ith}(q_{ith}, n_{ith})$:[11]

$$\widetilde{f}_{ith}(q_{ith}, n_{ith}) = z_{ith} \left(\omega_i q_{ith}^{\nu_i} + (1-\omega_i) n_{ith}^{\nu_i} \right)^{\frac{1}{\nu_i}} \tag{1.14}$$

ω_i is assumed to be interior to the unit interval while $\nu_i < 1$ for concavity. Further ω_i and ν_i are assumed to be the same for all households, so that final consumption heterogeneity across households results only from variation in productivities and market wages. In Appendix A.2.1 we derive the first-order conditions under this parameterization.

Under CES home production and CES utility, we can use the inframarginal rate of technical substitution between time and market inputs for process i to write n_{ith} as a linear function of q_{ith}:

$$n_{ith}(q_{ith}) = q_{ith} \left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)} \right]^{\frac{1}{\nu_i - 1}} \tag{1.15}$$

After substituting (1.15) into $c_{ith}(q_{ith}, n_{ith}(q_{ith}))$, the corresponding expression for final consumption is

$$c_{ith}(q_{ith}, n_{ith}(q_{ith})) = q_{ith}\, z_{ith} \left(\omega_i + (1-\omega_i) \left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)} \right]^{\frac{\nu_i}{\nu_i - 1}} \right)^{\frac{1}{\nu_i}} \quad \forall i \tag{1.16}$$

After substituting out the inframarginal conditions, the parameterized marginal rate of substitu-

[11] In our comparative static exercises in Section 1.2.4 we will consider a less flexible Cobb-Douglas structure in order to cleanly demonstrate the tradeoffs captured by our home production formulation.

tion for market inputs to processes i and j is:

$$\frac{\theta_i \omega_i}{\theta_j \omega_j}\left(\frac{q_{ith} z_{ith}}{q_{jth} z_{jth}}\right)^{\rho}\left(\omega_i + (1-\omega_i)\left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)}\right]^{\frac{\nu_i}{\nu_i - 1}}\right)^{\frac{\rho-\nu_i}{\nu_i}} \times\left(\omega_j + (1-\omega_j)\left[\frac{w_{th}\omega_j}{P_{jt}(1-\omega_j)}\right]^{\frac{\nu_j}{\nu_j - 1}}\right)^{\frac{\nu_j-\rho}{\nu_j}} = \frac{P_{it}}{P_{jt}} \tag{1.17}$$

This expression is derived in detail in Appendix A.2.1 and is the primary expression we use in our structural estimation in Section 1.3.2 in order to infer the parameter set associated with the underlying data generating process observed through the lens of aggregate time series.

The Euler equation describing consumption dynamics, meanwhile, depends on the levels of home productivities in a non-linear manner:

$$\begin{aligned}&\left(\sum_{i=1}^{I}\theta_i c_{ith}^{\rho}\right)^{\frac{1-\rho}{\rho}}\theta_i c_{ith}^{\rho-1}\left(\omega_i q_{ith}^{\nu_i} + (1-\omega_i) n_{ith}^{\nu_i}\right)^{\frac{1-\nu_i}{\nu_i}} z_{ith}\omega_i q_{ith}^{\nu_i - 1}\\ &= \beta\mathbb{E}_t\left\{R_{t+1}\left(\sum_{i=1}^{I}\theta_i c_{i,t+1,h}^{\rho}\right)^{\frac{1-\rho}{\rho}}\theta_i c_{i,t+1,h}^{\rho-1}\left(\omega_i q_{i,t+1,h}^{\nu_i} + (1-\omega_i) n_{i,t+1,h}^{\nu_i}\right)^{\frac{1-\nu_i}{\nu_i}} z_{i,t+1,h}\omega_i q_{i,t+1,h}^{\nu_i - 1}\right\}\end{aligned} \tag{1.18}$$

Each c_{ith} is a function of z_{ith}, thus the entire vector of home productivity levels $\boldsymbol{z}_{th}$ enters into both sides of the Euler equation. Notice, however, in the intratemporal condition shown in (1.17) only relative productivities matter. We will exploit this feature of our equilibrium conditions in our estimation routine described in Section 1.3.2 to estimate the structural parameters with a relative productivity residual after log-linearizing (1.17). We can then safely ignore the Euler equation in our estimation since the structural errors and all structural parameters enter into both (1.17) and (1.18) without any additional exogenous processes driving consumption dynamics. For this reason, a model estimated using both (1.17) and (1.18) will be stochastically singular. We can thus estimate the structural parameters assuming the Euler equation holds, forming the likelihood around the intratemporal condition.

1.2.2. Firms and Market Clearing

The supply side of the market consists of $I+1$ firms with I of them producing the I aggregated market goods and an additional firm producing an investment good. Firms that produce market goods have idiosyncratic Cobb-Douglas production technology but must pay the same final wages W_t and final capital rental rate r_t to households. W_t describes the aggregate average value of an hour of labor. With finitely many infinite-lived households, this amounts to $W_t = \frac{\sum_h w_{th} l_{th}}{\sum_h l_{th}}$, so that W_t is nominal labor income divided by total hours. r_t is the net rate of capital rental, and $R_t = r_t + 1 - \delta$ is the gross rate. Finally, we assume that the firm which produces the investment good costlessly transforms capital one-to-one to investment using no labor.

Firms are characterized by idiosyncratic Cobb-Douglas production technology which features firm-specific, exogenous time-varying total factor productivity A_{it} and firm-specific, exogenous time-varying labor productivity ζ_{it}. α_i is the output elasticity of capital and Y_{it} describes the totality of market good i produced in period t.[12] Output in sector i thus satisfies

$$Y_{it}(K_{it}, L_{it}) = A_{it} K_{it}^{\alpha_i} (\zeta_{it} L_{it})^{1-\alpha_i} \tag{1.19}$$

Each period, firm $i \in \{1, \ldots, I\}$ chooses inputs of capital K_{it} and labor L_{it} to produce market good Y_{it} by solving a static profit maximization problem:

$$\max_{L_{it}, K_{it}} P_{it} A_{it} K_{it}^{\alpha_i} (\zeta_{it} L_{it})^{1-\alpha_i} - r_t K_{it} - W_t L_{it} \tag{1.20}$$

Profit maximization implies

$$\frac{r_t}{P_{it}} = A_{it} \alpha_i \left(\frac{K_{it}}{\zeta_{it} L_{it}} \right)^{\alpha_i - 1} \quad \forall i \tag{1.21}$$

$$\frac{W_t}{P_{it}} = A_{it} \zeta_{it} (1 - \alpha_i) \left(\frac{K_{it}}{\zeta_{it} L_{it}} \right)^{\alpha_i} \quad \forall i \tag{1.22}$$

By assuming a single capital rental rate and wage rate, in general equilibrium the marginal products of capital and labor must equate across all firms.

As in all competitive market economies, equilibrium prices ensure that markets clear. Market clearing for goods implies

$$\sum_h q_{ith} = Q_{it} = Y_{it}(K_{it}, L_{it}) \tag{1.23}$$

where we use capital letters to denote aggregates. The investment good is produced linearly from investment capital $K_{I+1,t}$:

$$\sum_h s_{th} = K_{I+1,t} \tag{1.24}$$

where s_{th} is period t savings of household h. Further, we normalize prices so that the price of the investment good is unity every period. Capital markets clear when

$$\sum_h k_{th} = \sum_{i=1}^{I+1} K_{it} \tag{1.25}$$

Labor markets clear when

$$\sum_h l_{th} = \sum_{i=1}^{I} L_{it} \tag{1.26}$$

Note that households do not choose how to allocate their capital and labor toward sector-specific production. Rather, the allocations happen as a function of productivity and capital intensity differentials between sectors. In this manner, it is as if there exists an invisible intermediary firm that costlessly allocates capital and labor toward their production-activities in a Pareto efficient manner.

[12] We use capital letters to denote a firm's choice variables.

1.2.3. Equilibrium Definition

Given an initial stock of aggregate capital K_0 and known stochastic processes for productivities $\left\{\{\boldsymbol{z}_{th}\}_h, \boldsymbol{A}_t, \boldsymbol{\zeta}_t\right\}_{t=0}^{\infty}$, a general equilibrium economy under Beckerian home production consists of[13]

1. Sequences of household policies for market purchases, time use, and capital investment for each household

$$\left\{\{\boldsymbol{q}_{th}^{*}, \boldsymbol{n}_{th}^{*}, l_{th}^{*}, k_{t+1,h}^{*}\}_h\right\}_{t=0}^{\infty} \tag{1.27}$$

2. Sequences of firm policies for capital and labor inputs

$$\{K_{it}^{*}, L_{it}^{*}\}_{t=0}^{\infty} \quad \forall i \in \{1, \ldots, I\} \tag{1.28}$$

$$\{K_{I+1,t}^{*}\}_{t=0}^{\infty} \tag{1.29}$$

3. Sequences of aggregate average prices

$$\{\boldsymbol{P}_t^{*}, r_t^{*}, W_t^{*}\}_{t=0}^{\infty} \tag{1.30}$$

such that

i. Each period the household budget constraint holds with equality and choices satisfy (1.8) thru (1.10).

ii. Each period the choices of firms satisfy (1.21) thru (1.22).

iii. Each period market clearing conditions (1.23) thru (1.26) are satisfied.

1.2.4. Comparative Statics For Household Decisions

To illustrate the important theoretical implications of our model formulation, we engage in several comparative statics. In order to characterize and describe equilibrium tradeoffs in as transparent of a manner as possible, suppose for simplicity that home production functions are Cobb-Douglas so that $\nu_i = 0$ for all i. From here on, we will dispense with time t and household h subscripts for simplicity to consider a single household operating in a static environment. Thus we have for home production, $\widetilde{f}_i(q_i, n_i) = z_i q_i^{\omega_i} n_i^{1-\omega_i}$. The household uses its fixed income net of savings $\overline{y}$ to purchase two market commodities q_1 and q_2 at prices P_1 and P_2, which enter into home production functions, along with time n_1 and n_2, to produce final commodities c_1 and c_2. Assume households know market prices and home productivities z_1 and z_2 and take them as given. For simplicity, suppose the utility weights on final goods are such that $\theta_1 = \theta_2$.[14] The exercises can be grouped into two camps: 1) analyzing relative productivity, relative price, and wage effects when the household inelastically supplies a fixed amount of labor $\overline{l}$; 2) analyzing how off-market time use varies when households can adjust their labor supply elastically.

[13] We use superscript stars to denote equilibrium objects.

[14] Assuming otherwise does not change the qualitative nature of our results.

Under Cobb-Douglas home production relative demand for market good j to market good i can be written

$$\left(\frac{q_j}{q_i}\right) = \left[\frac{\theta_i\omega_i[(1-\omega_i)/\omega_i]^{(1-\omega_i)\rho}}{\theta_j\omega_j[(1-\omega_j)/\omega_j]^{(1-\omega_j)\rho}}\right]^{\frac{1}{\rho-1}} P_j^{\frac{1-\rho+\rho\omega_j}{\rho-1}} P_i^{\frac{1-\rho+\rho\omega_i}{1-\rho}} w^{\frac{\rho(\omega_i-\omega_j)}{\rho-1}} \left[\frac{z_i}{z_j}\right]^{\frac{\rho}{\rho-1}} \tag{1.31}$$

Relative market good consumption is thus a power function of prices, wages, and unobserved relative productivities. The same procedure can be applied so as to express the time devoted toward production process j relative to process i as follows:

$$\left(\frac{n_j}{n_i}\right) = \left[\frac{\theta_i(1-\omega_i)[\omega_i/(1-\omega_i)]^{\rho\omega_i}}{\theta_j(1-\omega_j)[\omega_j/(1-\omega_j)]^{\rho\omega_j}}\right]^{\frac{1}{\rho-1}} P_j^{\frac{\rho\omega_j}{\rho-1}} P_i^{\frac{\rho\omega_i}{1-\rho}} w^{\frac{\rho(\omega_i-\omega_j)}{\rho-1}} \left[\frac{z_i}{z_j}\right]^{\frac{\rho}{\rho-1}} \tag{1.32}$$

See Appendix A.2.1 for the detailed derivations of these expressions.

In Section 1.2.4.1 we will set $\omega_1 = \omega_2 = \omega$ so that differentials in relative factor inputs to different home production processes are entirely driven by price and productivity differentials. This strong assumption will allow us to conduct our indifference-curve analysis on clean, closed-form expressions. Note that when $\omega_1 = \omega_2$, relative market consumption will not change as wages change. This can be readily verified by confirming that the coefficient on wages in (1.31), $\frac{\rho(\omega_2-\omega_1)}{\rho-1}$, is 0 when $\omega_1 = \omega_2$, so that relative consumption is independent of wages under this parameterization. In Section 1.2.4.2, we relax the assumption that output elasticities across processes are equal in order to analyze how income changes induce relative market purchase changes. Note that when $\omega_1 \neq \omega_2$, relative prices can remain constant and yet relative demand will change as wages grow. Thus, demand-side structural change can occur despite the fact that the underlying utility function is homogeneous of degree one in $\boldsymbol{q}$. It can be readily verified that since $\boldsymbol{n}$ is a linear function of $\boldsymbol{q}$, we can always scale $\boldsymbol{q}$ by some scalar a, even after composing $\widetilde{u}(\boldsymbol{q},\boldsymbol{n})$ with $\boldsymbol{n}(\boldsymbol{q})$ so that $\widetilde{u}\big(a\,\boldsymbol{q},\boldsymbol{n}(a\,\boldsymbol{q})\big) = a\,\widetilde{u}(\boldsymbol{q},\boldsymbol{n})$, regardless of whether market input intensities ω_i are the same. Finally, in Section 1.2.4.3 we relax the assumption that labor is inelastically supplied and demonstrate that on the intensive margin household labor supply is independent of market prices if home production time use intensities are identical, $\omega_1 = \omega_2$. We then show that when $\omega_1 \neq \omega_2$ both the signs and magnitudes of household labor supply responses to relative price changes depend on the degree of final consumption substitutability ρ and which market commodity q_i is associated with the more time intensive production process.

1.2.4.1. Consumption and Off-Market Time Use Tradeoffs Under Inelastic Labor $\bar{l}$ and Identical Factor Intensities $\omega_1 = \omega_2 = \omega$

Consider the following expressions for the marginal rates of substitution for market inputs and time use when $\nu_i = 0$ and after substituting out the inframarginal rates of technical substitution between market purchases and time:

$$MRS(q_1, q_2) = \left(\frac{z_1}{z_2}\right)^{\rho} \left(\frac{q_1}{q_2}\right)^{\rho-1} \left(\frac{P_1}{P_2}\right)^{\rho(1-\omega)} \tag{1.33}$$

$$MRS(n_1, n_2) = \left(\frac{z_1}{z_2}\right)^{\rho} \left(\frac{n_1}{n_2}\right)^{\rho-1} \left(\frac{P_2}{P_1}\right)^{\rho\omega} \tag{1.34}$$

Proposition 2 summarizes the effects of changes in $\frac{z_1}{z_2}$ on household equilibrium choices.

Proposition 2. Fix $P_1 = P_2 = 1$ and z_2. Consider the following cases separately:

i. If the outputs of home production are substitutes so that $\rho \in (0, 1)$, then an increase (decrease) in z_1 is welfare improving and results in an increase (decrease) in equilibrium $\frac{q_1}{q_2}$ and an increase (decrease) in $\frac{n_1}{n_2}$.

ii. If the outputs of home production are complements so that $\rho \in (-\infty, 0)$, then an increase (decrease) in z_1 is welfare improving and results in a decrease (increase) in equilibrium $\frac{q_1}{q_2}$ and a decrease (increase) in $\frac{n_1}{n_2}$.

Proof. See Appendix A.2.3. □

Analyzing how relative market purchases and home time use respond to relative market price changes may perhaps be more interesting to readers given prices are what we observe, not productivities. Proposition 3 demonstrates that for fixed relative productivities with P_2 as numeraire, the elasticity of substitution for the outputs of home production dictates the sign of co-movements in $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ as a result of changes to P_1.

Proposition 3. Fix $z_1 = z_2 = 1$ and P_2. Consider the following cases separately:

i. If the outputs of home production are substitutes so that $\rho \in (0, 1)$, then an increase (decrease) in $\frac{P_1}{P_2}$ leads to a decrease (increase) in equilibrium $\frac{q_1}{q_2}$ and a decrease (increase) in equilibrium $\frac{n_1}{n_2}$.

ii. If the outputs of home production are complements so that $\rho \in (-\infty, 0)$, then an increase (decrease) in $\frac{P_1}{P_2}$ leads to a decrease (increase) in equilibrium $\frac{q_1}{q_2}$ and an increase (decrease) in $\frac{n_1}{n_2}$.

Proof. See Appendix A.2.3. □

Turning now to relative price effects, when home production outputs are substitutes the ratios move together. When they are complements, increases in P_1 lead to decreases in $\frac{q_1}{q_2}$, but consumers offset the decline in q_1 by shifting time toward the production of c_1, so $\frac{n_1}{n_2}$ increases. How can market consumption and time use ratios move in different directions? Note first that price changes induce shifts in both the budget constraint and the quasi-indifference curve that shows the tradeoff between consumption of q_1 and q_2. Yet, with respect to time use, only the quasi-indifference curves shift. When $\rho \in (0, 1)$, we show in the proof to Proposition 3 that $\frac{\rho\omega}{\rho-1}$, the coefficient on relative prices $\frac{P_1}{P_2}$ in the relative time use equation, is negative. Consumers thus substitute their resources away from the process associated with the market good whose prices are increasing. Yet when $\rho < 0$, $\frac{\rho\omega}{\rho-1} > 0$, and consumers devote less market resources to process c_1 but relatively more time. In this case complementarities within the household, specifically complementarities between the outputs of different home production processes, dominate the traditional substitution effect induced by raising relative prices. In this way, when final activities are complements, consumers can insure themselves against adverse price shocks by substituting time for market purchases toward the activity for which market prices increased. These results demonstrate the complexity of various substitution effects when the time-allocation vector is split up among different tasks in the manner we impose.

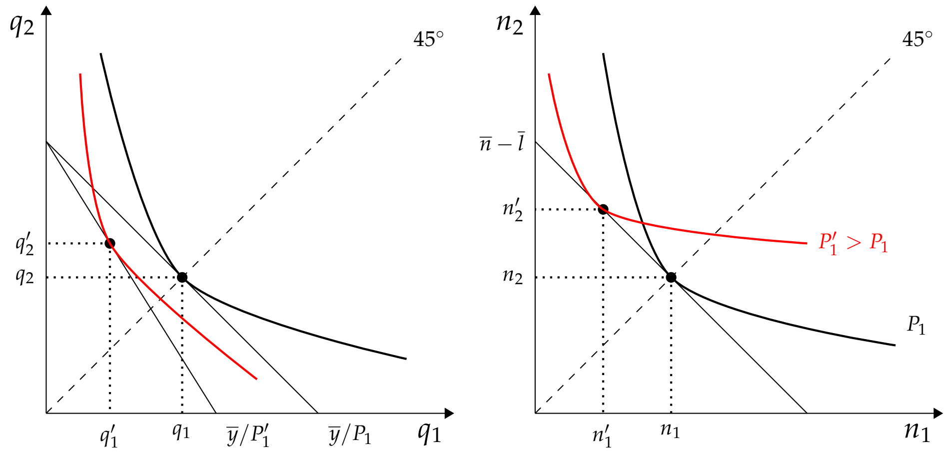


Figure 1.1.: $(0 < \rho < 1)$ — The plot demonstrates how equilibrium outcomes change when $\frac{P_1}{P_2}$ increases relative to baseline unit relative prices, i.e. $P_1 = P_2$. A change in $\frac{P_1}{P_2}$ when final consumption commodities are perfect substitutes induces positive co-movement of $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$. The proof of Proposition 3 details the exact mathematical mechanisms causing this phenomenon.

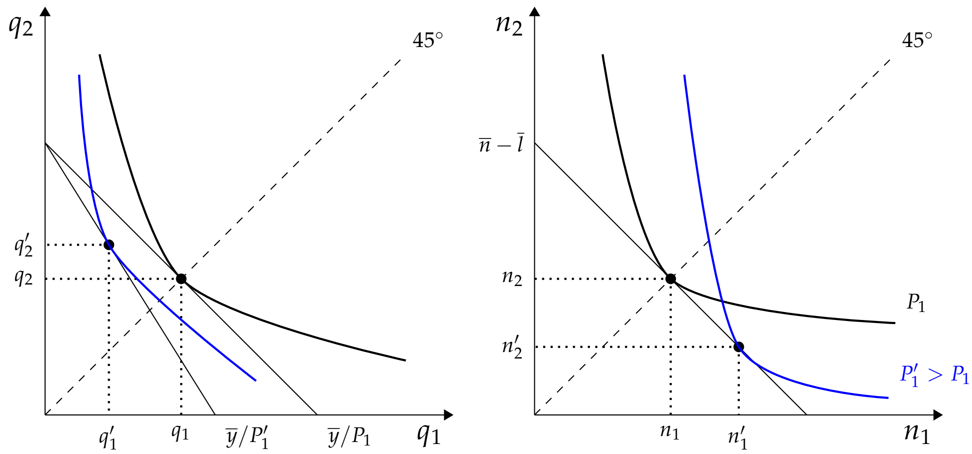


Figure 1.2.: $(\rho < 0)$ — The plot demonstrates how equilibrium outcomes change when $\frac{P_1}{P_2}$ increases relative to baseline unit relative prices, i.e. $P_1 = P_2$. A change in $\frac{P_1}{P_2}$ when final consumption commodities are perfect substitutes induces negative co-movement of $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$. In this sense, when the outputs of home production are complements and consumers have the choice to allocate time toward multiple off-market activities, they can partially insure themselves against adverse price shocks.

1.2.4.2. Wage Effects Under Inelastic Labor $\bar{l}$ and Differing Factor Intensities $\omega_1 \neq \omega_2$

One of the main contributions of our framework is that it provides a possible micro-foundational explanation for why income effects appear in the data: they fundamentally depend on how consumers spend their off-market time engaging in home production activities using different market commodities. We show that income effects can be generated by differences in the labor intensity of in-home activities along with the relative freedom by which consumers are able to divert their resources toward other activities (i.e. the substitution elasticity between home production outputs). We now show that these effects induce non-linear expansion paths in wages when factor shares are different, despite the fact that the underlying utility function is homogeneous of degree one. There is no reason to believe that similar effects would not be generated if the CES or Cobb-Douglas assumptions were relaxed.

How relative market consumption varies in w depends on the sign of the coefficient $\frac{\rho(\omega_2-\omega_1)}{\rho-1}$ which itself depends on whether final goods are complements or substitutes and which production process is more time intensive. Proposition 4 summarizes this.

Proposition 4. Consider the implications of two separate cases and their corresponding subcases:

i. Suppose $\frac{\rho(\omega_2-\omega_1)}{\rho-1} < 0$ so that $\frac{q_1}{q_2}$ is decreasing in w then one and only one of the following must hold:
   a. $\rho < 0$ and $\omega_2 < \omega_1$
   b. $\rho \in (0,1)$ and $\omega_2 > \omega_1$

ii. Suppose $\frac{\rho(\omega_2-\omega_1)}{\rho-1} > 0$ so that $\frac{q_1}{q_2}$ is increasing in w then one and only one of the following must hold:
   a. $\rho < 0$ and $\omega_2 > \omega_1$
   b. $\rho \in (0,1)$ and $\omega_2 < \omega_1$

Proof. See Appendix A.2.3. □

The proof of Proposition 4 is fairly trivial. The main takeaways are as follows. If $\rho < 0$, so that final goods are complements, then consumption and time use shift toward the more time intensive task as w increases. If $\rho \in (0,1)$, so that final goods are substitutes, then consumption and time use shift toward the less time intensive task as w increases. Consider case (i) primarily, so that relative market purchases of q_1 to q_2 fall as wages rise. If w is rising and final goods are complements, so that $\rho < 0$, then consumers scale up purchases of q_2 at a faster rate than q_1 since they can take advantage of q_1's relatively higher factor intensity, ω_1.[15] They thus want relatively more of the factor input associated with the relatively more time intensive process — in this case q_2. n_2 would increase relative to n_1 as well in this case. For case (ii.a.) the same phenomenon occurs, except q_1 is relatively more time intensive, so relative consumption of q_1 to q_2 increases. Returning to case (i), suppose now $\rho \in (0,1)$, so that final goods are substitutes. Then consumers scale up purchases of the good associated with the less time intensive process

[15]Note that q_1 is indeed a "good," so consumption is increasing in w just more slowly than that of q_2.

faster — in this case, q_2. Put another way, they want more of the market good associated with the more goods-intensive process (or less time intensive process). In case (ii.b.) the argument remains the same except consumers want relatively more of q_1 and n_1 as w rises.

Let us place this into anecdotal context. Consider the tradeoff faced by a consumer who is choosing whether to buy more cleaning supplies in order to clean his house or pay someone to do it for him. Denote the former activity — cleaning one's own house using one's own cleaning supplies — as c_1. Denote the latter activity — paying a maid to clean the house — as c_2. c_1 is more time intensive than c_2 which is more dependent on market resources — the services provided by the hired maid, q_2. Suppose the consumer receives a wage increase. If $\rho < 0$ so that the outputs from these activities, c_1 and c_2, are complements, consumers will choose to purchase more cleaning supplies q_1 rather than more cleaning services q_2 as a result of a wage increase. On the other hand, if $\rho \in (0, 1)$ so that c_1 and c_2 are substitutes, consumers will hire more cleaning services q_2 as a result of a wage increase. We remind the reader that the results here hold under the assumption that prices remain fixed. If prices and real wages are simultaneously changing, the value of ρ along with the absolute difference $|\omega_2 - \omega_1|$ will determine whether relative price effects or wage effects dominate.

1.2.4.3. Labor Supply Dependencies on Relative Price Changes

In this section we relax the assumption that labor supply is fixed to analyze how consumers adjust their work hours as a response to changes in the relative price of market goods. Note that in our model there are multiple forces weighing on equilibrium labor supply decisions. Given consumers have multiple choices with respect to how to spend their off-market time, each of which are complimentary to a separate market purchase commodity, changes in the prices of market purchases can impact both the equilibrium distribution of off-market time and labor supply on the intensive margin. These tradeoffs will depend on the underlying time intensities of the home production processes, ω_1 and ω_2, as well as the gross substitutability (or gross complementarity) of final consumption, ρ. As before, assume effective cash on hand $\overline{y}$ is fixed and $\theta_1 = \theta_2$. To illustrate how the intensive margin of labor l depends on prices P_1 and P_2, input elasticities ω_1 and ω_2, and the gross substitutability of final consumption ρ, we derive the equilibrium labor supply function $l(P_1, P_2, \omega_1, \omega_2, \rho, w)$ and plot it for different values of relative prices P_1/P_2 under different parameterizations.[16] We allow P_1/P_2 to vary from 0.1 to 10, which is accomplished by setting $P_2 = 1$ and varying P_1. We also fix $w = 1$ and $\overline{n} = 24$, while letting $\overline{y} = w\overline{n} + save$ where $save = 0.1\,\overline{y}$, giving a rounded value for cash on hand of 26.66667. We choose several combinations of ω_1, ω_2, and ρ presenting l and n_2 supply and demand functions side by side below in four different figures.

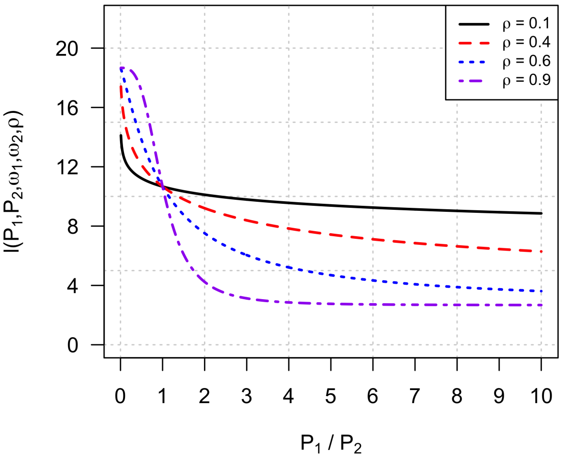
Interactions between home production time intensities governed by ω_1 and ω_2 and gross substitutability governed by ρ lead to some non-monotonic relationships in prices for labor and off-market time use functions. We begin our discussion focussing on cases where $\rho \in (0, 1)$ so that the final outputs of home production are substitutes. Figure 1.3 demonstrates that if the market price associated with the market input for the less time intensive process increases, and home production outputs are substitutes, then off-market time dramatically shifts toward the

[16] A detailed derivation of $l(P_1, P_2, \omega_1, \omega_2, \rho, w)$ can be found in Appendix A.2.2.

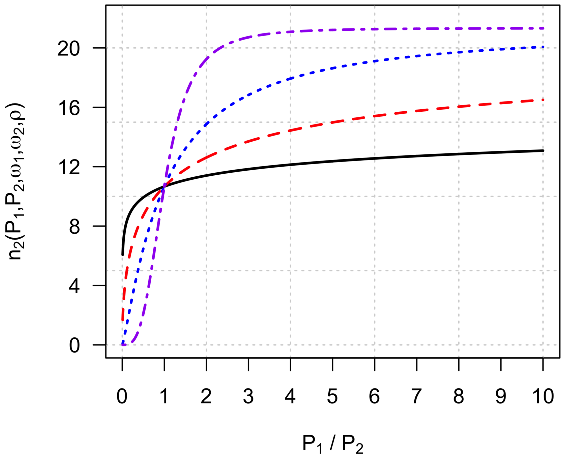
less time intensive process and labor supply falls just as dramatically. Yet this phenomenon does not appear to be symmetric. When the price of the market commodity associated with the more time intensive process increases, consumers do not substitute time away from this process as quickly, in fact increasing time devoted toward this process if the final outputs of home production are only mildly substitutable. This can be seen by noting in Figure 1.4b that n_2 is non-monotonic in changes to P_1 for $0 < \rho \leq 0.6$, so that if P_1 is big enough n_2 and l both fall and consumers spend more time on n_1. Note that q_1 is falling in P_1, but when $0 < \rho \leq 0.6$ it appears that home production time use and market purchase complementarities dominate final commodity substitution effects. As $\rho \to 1$ the substitution effect over final consumption becomes stronger until it is strong enough to induce increases in n_2 and thus corresponding decreases in both n_1 (and q_1), which is evident via the purple line in Figure 1.4b.

In Figures 1.5 and 1.6, we consider supply of l and off-market time allocation when the outputs of home production are complements, $\rho < 0$. Notice that price sensitivity is the same for l as when $\rho \in (0, 1)$: in Figure 1.5a l is sensitive to price changes affecting the market commodity associated with the less time intensive process just as in Figure 1.3a. However, when $\rho \in (0, 1)$ l generally declines in P_1 if consumers must allocate more time to the time intensive process ($i = 2$), while when $\rho < 0$ labor supply increases as P_1 rises since consumers need to allocate more market resources toward q_1 due to both the gross complementarities and home production complementarities. In Figure 1.5 notice that n_2 falls faster than l rises when P_1 increases past $P_1 \approx P_2 = 1$. Thus, for high enough relative prices the increase in P_1 leads to both increases in labor to fund more expensive market purchases and increases in n_1 relative to n_2. Turning to the case where $i = 1$ is more time intensive, labor supply is less sensitive to variation in P_1 than when $i = 2$ is more time intensive, as is seen in Figure 1.6a. This is similar to the flatness of household labor supply when $\rho \in (0, 1)$ observed in Figure 1.4a.

For intuition, suppose in Figure 1.5, $i = 1$ are market services and $i = 2$ are market goods. We expect home production using goods to be more time intensive so $\omega_2 < \omega_1$. This is because services are generally consumed to offset off-market labor: think of purchasing take-out versus making a meal on your own. If the outputs of final consumption are complementary then an increase in the price of services corresponds to an increase in labor and a simultaneous relative increase in time spent using services, n_1. The consumer works more and spends more time using services, presumably because he has less time to spend engaged in laborious home production involving goods.

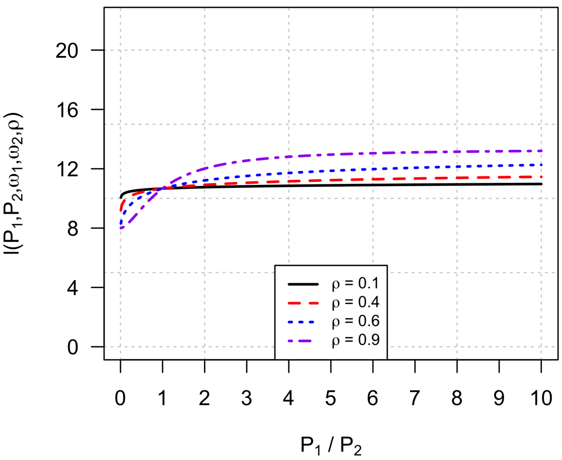


(a) Labor

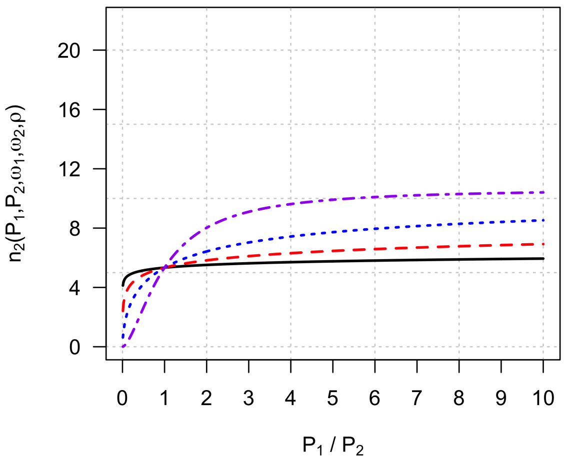


(b) Off-market time n_2

Figure 1.3.: ($0 < \rho < 1$ and $\omega_1 > \omega_2$) — Here, we set $\omega_1 = 0.8$ and $\omega_2 = 0.2$, so that process $i = 2$ is more time intensive. Notice that increasing P_1 relative to numeraire P_2 causes l to fall. As $\rho \to 1$, home production outputs become more substitutable, and the strength of this substitutability dominates complementarities between market purchases and time use.

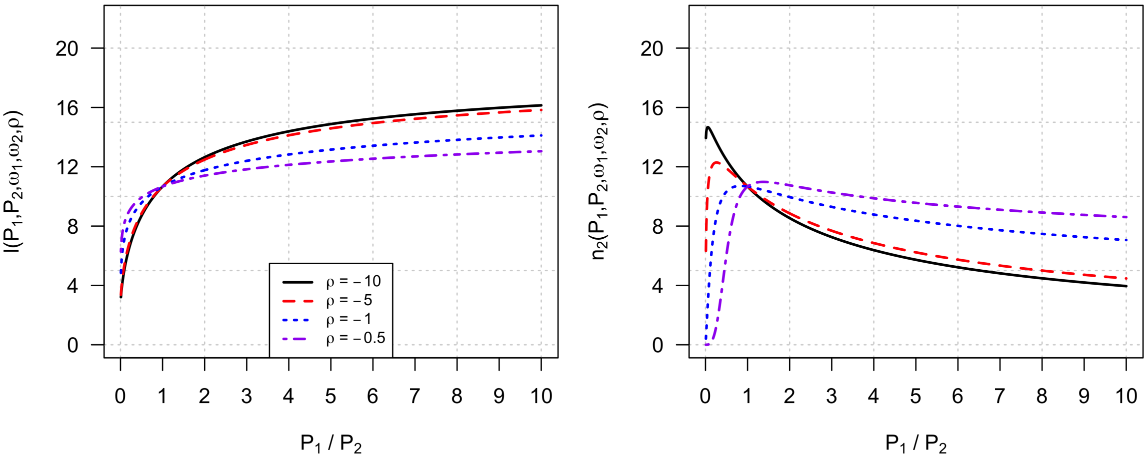


(a) Labor



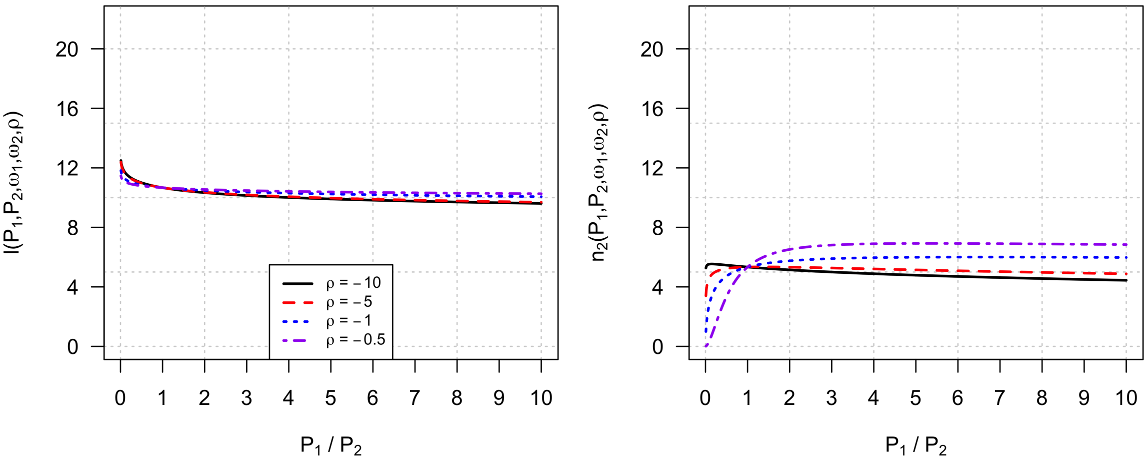
(b) Off-market time n_2

Figure 1.4.: ($0 < \rho < 1$ and $\omega_1 < \omega_2$) — This time we set $\omega_1 = 0.4$ and $\omega_2 = 0.6$, so that process $i = 1$ is more time intensive. As $\rho \to 1$, approaching linear preferences for final consumption, n_2 begins to increase in P_1/P_2 to the point where for $\rho = 0.9$, the n_2 policy function appears monotonic in P_1/P_2. As $\rho \to 0$ consumers compensate for declines in q_1 by spending more time on process $i = 1$.



(a) Labor **(b)** Off-market time n_2

Figure 1.5.: ($\rho < 0$ and $\omega_1 > \omega_2$) — Now final activities are complements. As in Figure 1.3 we set $\omega_1 = 0.8$ and $\omega_2 = 0.2$, so that process $i = 2$ is more time intensive. Labor varies in P_1/P_2 in the opposite way than when $\rho \in (0, 1)$. Further, n_2 declines in P_1/P_2 faster than l rises, so that the remaining time moves to process $i = 1$, the task for which the market input experienced a price increase.



(a) Labor **(b)** Off-market time n_2

Figure 1.6.: ($\rho < 0$ and $\omega_1 < \omega_2$) — As in Figure 1.4 we set $\omega_1 = 0.4$ and $\omega_2 = 0.6$, so that now process $i = 1$ is more time intensive. Labor supply is flat in prices again and n_2 exhibits very small variation, similar to when $\rho \in (0, 1)$. Notice that, though when $\rho \in (0, 1)$ there appears to be a limiting condition in which increases in P_1 lead to changes in the behavior of n_2, this does not occur when $\rho < 0$.

1.2.4.4. Income Effects with Elastic Labor

Now, relax the assumption that w is fixed and consider how labor supply and off-market time use vary in wages. Continue to assume that savings rates are fixed at 10%, so that cash-on-hand is the same as in Section 1.2.4.3 except that it now varies in w. The degree to which income or substitution effects dominate in the context of the classic labor/leisure tradeoff depends on ρ. Let $P_1 = P_2 = 1$. Consider the same parameterizations as above. In Figures 1.7 and 1.8 we plot the equilibrium labor supply functions for the same combinations of ρ and ω as above. Notice that when $\rho \in (0,1)$, so that the outputs of home production are substitutes, the substitution effect appears to dominate as labor supply increases when w rises (Figure 1.7). The opposite is true when $\rho < 0$ in Figure 1.8.

Wage increases have several implications for equilibrium allocations of market consumption and off-market time. First, as Proposition 4 demonstrates, rising wages can indeed lead to changing expenditure shares even if prices are fixed, due to the time use tradeoffs households face. For $\rho \in (0,1)$ and $\omega_2 > \omega_1$ so that $i = 1$ is more time intensive, both $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ will fall as w rises, so that relative consumption and time use shift toward the less time intensive process. This happens as households work more, thus having less time to devote toward their off-market activities. When $\rho < 0$, the more time intensive process receives both more market purchases and more time. However, whether $\frac{q_1}{q_2}$ falls or rises fasters than $\frac{n_1}{n_2}$ falls or rises depends on the relative differences between ω_1 and ω_2. In this sense, for certain values of ω_1 and ω_2 along with certain values of ρ, consumers may shift their off-market time allocation more than their consumption allocation in response to wage increases. The effect of wage increases on the consumption allocation can thus be partially neutralized by households' choosing to re-allocate their off-market time. Failing to account for off-market time use tradeoffs can thus lead to spurious estimates as to the degree to which rising wages are actually responsible for the changing composition of market consumption.

Proposition 5. Let $P_1 = P_2 = 1$ and $z_1 = z_2 = 1$. The following hold:

i. If $\rho \in (0,1)$ then consumption and time use flow toward the less time intensive task, but the consumption allocation changes more than the time use allocation.

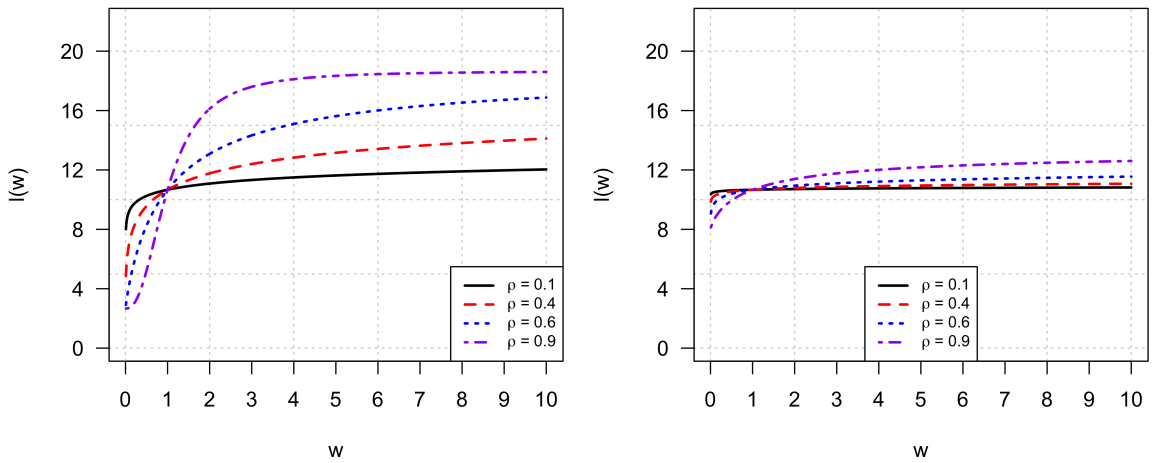
ii. If $\rho < 0$ then consumption and time use flow toward the more time intensive task, but the consumption allocation changes more than the time use allocation.

Proposition 6. Suppose $z_1 = z_2 = 1$, but P_1 and P_2 are free. Consider only non-inflationary relative prices, so that either $P_1 < 1 < P_2$ or $P_2 < 1 < P_1$. Suppose, without loss of generality, that $\omega_2 > \omega_1$. The following must hold:

i. If $\frac{P_1}{P_2} < 1$ then $\frac{q_1}{q_2}$ changes faster than $\frac{n_1}{n_2}$ as w rises.

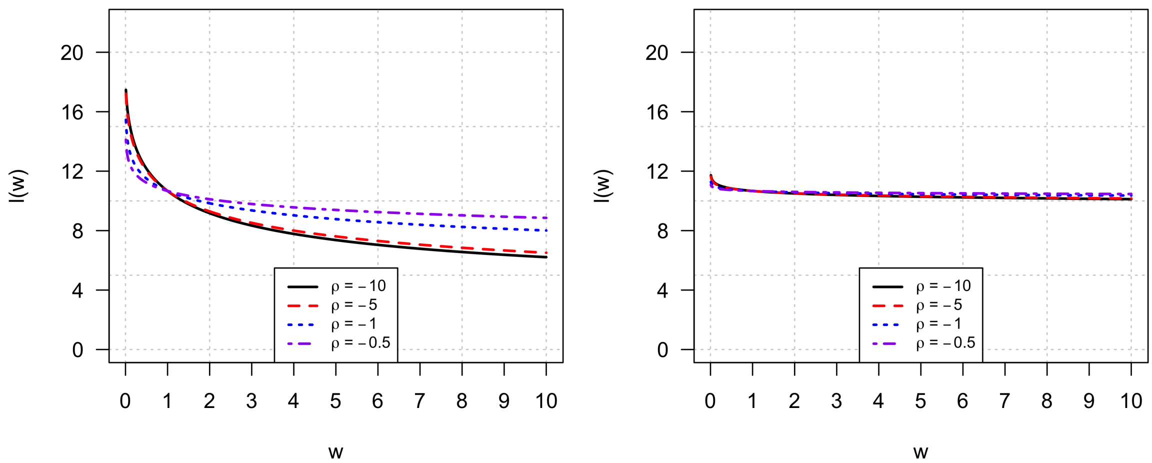
ii. If $\frac{P_1}{P_2} > 1$ then it is ambiguous as to whether $\frac{q_1}{q_2}$ or $\frac{n_1}{n_2}$ changes faster as w rises.

The signs of changes to relative consumption and time use depend on ρ and are the same as those outlined in Proposition 5. For $\omega_1 > \omega_2$, just exchange all index numbers.



(a) $\omega_1 > \omega_2$ **(b)** $\omega_1 < \omega_2$

Figure 1.7.: $(0 < \rho < 1)$ — Here we present household labor supply variation in w for the case when home production outputs are substitutes. In panel (a) $\omega_1 = 0.8$ and $\omega_2 = 0.2$, so that process $i = 2$ is more time intensive. In panel (b) $\omega_1 = 0.4$ and $\omega_2 = 0.6$, so that process $i = 1$ is more time intensive but to a lesser magnitude than as in panel (a). Labor supply appears to be non-decreasing in w for all values of $\rho \in (0, 1)$.



(a) $\omega_1 > \omega_2$ **(b)** $\omega_1 < \omega_2$

Figure 1.8.: $(\rho < 0)$ — Here we present household labor supply variation in w for the case when home production outputs are complements. In panel (a) $\omega_1 = 0.8$ and $\omega_2 = 0.2$, so that process $i = 2$ is more time intensive. In panel (b) $\omega_1 = 0.4$ and $\omega_2 = 0.6$, so that process $i = 1$ is more time intensive but to a lesser magnitude than as in panel (a). Labor supply appears to be non-increasing in w for all values of $\rho < 0$.

Proposition 5 describes how both consumption and time use allocations change as a response to wage changes when there are no price differences. Note that the magnitude of relative market consumption changes always dominates market time use changes, regardless of the value of ρ. Proposition 6 extends the result in Proposition 5 to an environment where $P_1 \neq P_2$ necessarily. Proposition 6 states that the degree to which changes to relative consumption dominate changes to relative time use depends on both relative prices and the relative time intensity of the two home production processes. Note that when the price of the market good associated with the more time intensive process is less than the price of the market good associated with the less time intensive process, it is clear that $\frac{q_1}{q_2}$ changes faster than $\frac{n_1}{n_2}$. However, when the price of the market good associated with the more time intensive process is greater than that of the less time intensive one, the difference in the magnitude of co-movements is ambiguous, depending instead on the underlying degree to which one process is relatively more time intensive than the other. This result demonstrates that the inter-dependencies between time intensity and relative prices determine the magnitude and the sign by which consumption allocations respond to wage increases.

To place these results in a more intuitive context, consider that when $\rho \in (0,1)$, so that the outputs of home production are substitutes, the classical labor/leisure substitution effect dominates. The opportunity cost of engaging in off-market activities increases, so consumers work more and buy more market goods, but they adjust their allocations of market consumption and off-market time toward home production activities that are relatively less time intensive. In this sense, if the outputs of home production are substitutes, consumers can substitute market purchases toward services as wages rise, subsequently spending relatively more of their off-market time using services even though the total amount of time they devote toward off-market activities falls. When $\rho < 0$, so that the outputs of home production are complements, the classical labor/leisure income effect dominates: consumers work less and spend more time on off-market activities. This in turn causes a shift in the allocation of both market resources and time toward the more time intensive process.

Figures 1.9 and 1.10 show the graphical indifference curve analysis as to how the optimal allocations adjust in response to wage increases. Note that the magnitude of budget constraint changes is greater when the home production outputs are substitutes. In that case the time use constraint shifts inward. The quasi-marginal time use indifference curves twist toward n_2, so that n_1 falls but n_2 falls less than n_1. Depending on the values of ω_i, $i \in \{1,2\}$, n_2 could also rise as we demonstrate in Figure 1.9. Regardless of whether n_2 falls or rises, $\frac{n_1}{n_2}$ falls. In Figure 1.10 we show how relative consumption and time use change when home production activities are complements. In the complements case both the budget constraint and off-market time use constraint shift outward, since labor supply falls but wages rise. It is clear, then, that the value of ρ dictates the qualitative nature of labor supply, consumption allocation, and off-market time use allocation responsiveness to wage changes,

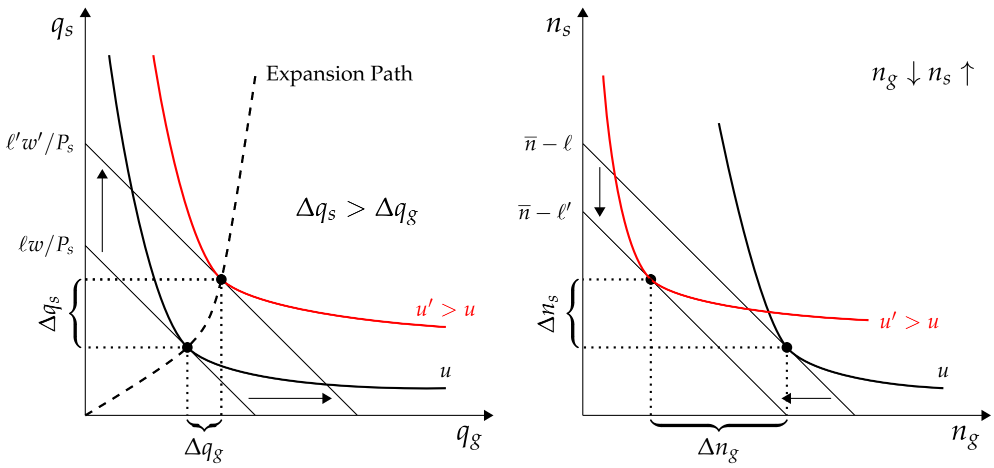


Figure 1.9.: $(0 < \rho < 1)$ — Assume $\omega_2 > \omega_1$ and consider what happens when $w \to w'$ where $w' > w$. Note that $l'(w') > l(w)$ since labor supply is increasing in w when $\rho \in (0,1)$. This implies that $\overline{y}'(l', w') > \overline{y}(l, w)$ and $\overline{n} - l' < \overline{n} - l$. Note that both $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ decline in w. Since off-market time use and consumption are all goods in the classical demand sense, increasing w is welfare improving so $u' > u$.

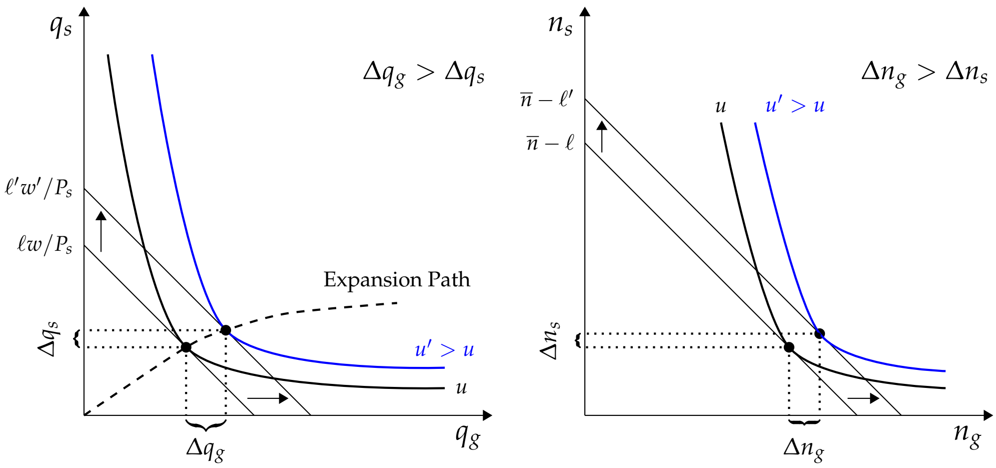


Figure 1.10.: $(\rho < 0)$ — Assume $\omega_2 > \omega_1$ and consider what happens when $w \to w'$ where $w' > w$. Note that $l'(w') < l(w)$ since labor supply is decreasing in w when $\rho < 0$. Since wage increases are welfare improving, the budget constraint shifts out but the magnitude of the shift is not as great as when $\rho \in (0,1)$. Note that both $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ increase in w when $\rho < 0$. Further, the off-market time use constraint shifts out, the opposite of when $\rho \in (0,1)$, but the magnitude of changes to n_1 is greater than changes to n_2.

1.3. Quantitative Application

To understand the empirical implications of our theoretical results, we engage in a quantitative exercise using the model to estimate the degree to which demand-driven income effects versus supply-driven technological effects have contributed to the long-run increase in the services share of United States consumption expenditure. Prior to engaging in this empirical exercise, we first examine several trends in both long-run, aggregate U.S. consumption expenditure and disaggregated spending and time use. After establishing some basic properties with both aggregate and disaggregated data, we then estimate our full general equilibrium model of the U.S. economy assuming a representative household and two representative firms, one which manufactures "goods" (g) and the other a producer of "services" (s). Our estimation procedure is of the Bayesian learning variety where we target a posterior distribution of model parameters conditional upon aggregate U.S. data. To estimate the joint posterior, we use a Hamiltonian Monte Carlo (HMC) routine on a model featuring first differences of our equilibrium equations. With parameter estimates in hand, we then engage in counterfactual exercises fixing relative prices, wages, and productivities separately in order to understand which channel has most affected long-run structural change.

Below, we first outline some empirical regularities in the data in Section 1.3.1. In Section 1.3.2 we present details of the HMC estimation procedure. In Section 1.3.3 we present the estimates from the structural parameters, which we use to engage in counterfactual experiments in Section 1.3.4.

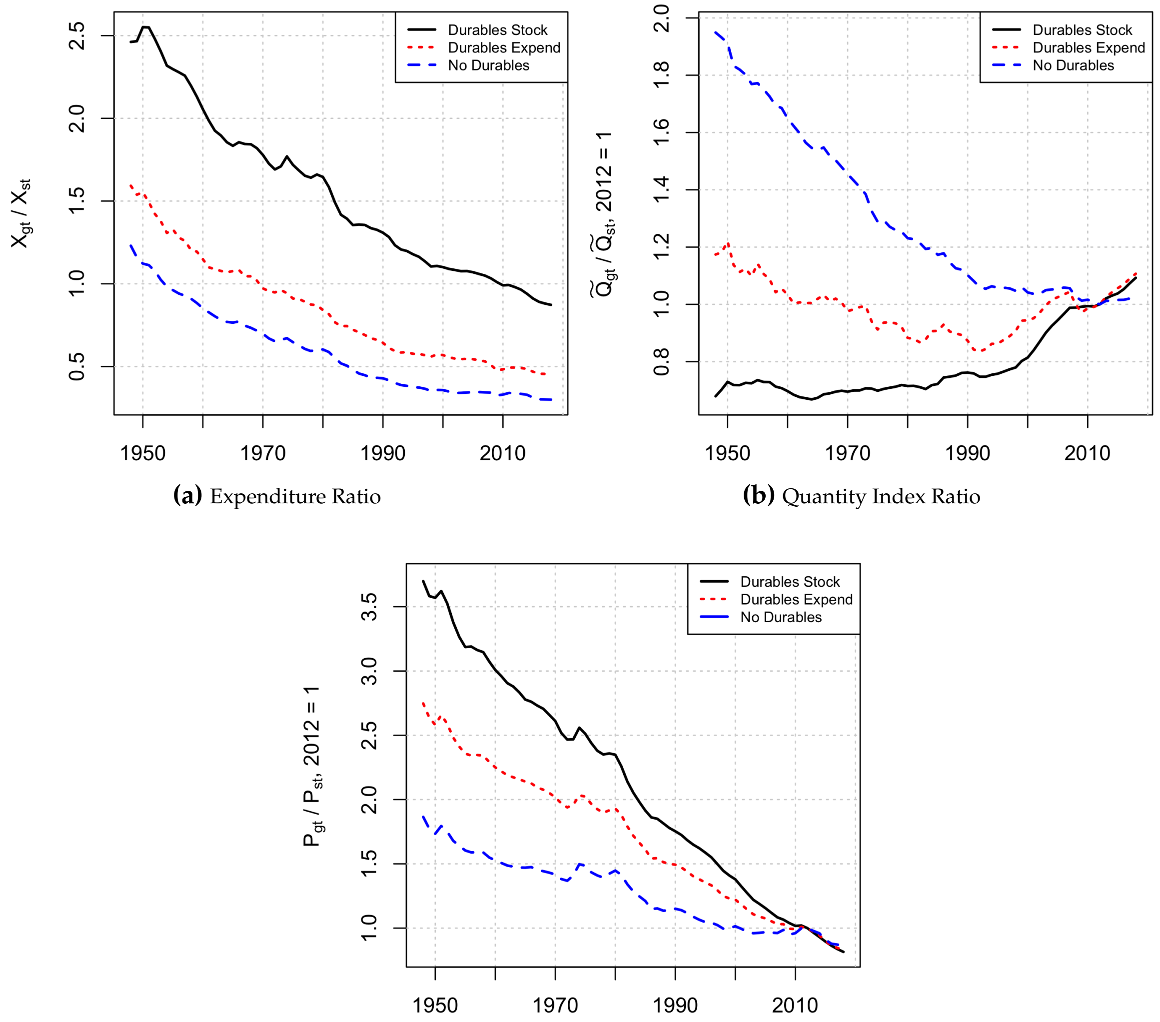
1.3.1. Empirical Regularities

1.3.1.1. Aggregate U.S. Expenditure Data

Our quantitative exercises operate on several well-established long-run trends in U.S. economic activity from 1948-2018: the decline in the aggregate nominal consumption value of goods to services X_{gt}/X_{st}, changes in the aggregate relative quantity indices of goods and services $\widetilde{Q}_{gt}/\widetilde{Q}_{st}$,[17] and the decline in aggregate relative goods to services prices P_{gt}/P_{st}. Both the signs and magnitudes of these changes depend on the degree to which we account for the presence of consumer durables in the various goods series — X_{gt}, $\widetilde{Q}_{gt}$, and P_{gt}. Since durable flows are a non-trivial part of aggregate goods consumption, the failure to properly account for how consumers use accumulated durables in their everyday activities can lead to different estimates as to what degree wage and relative price effects have contributed to structural change. We contend that different market commodities are associated with different tradeoffs between consumption and off-market time use, and these differences are the fundamental causes of income and relative price effects. Since these tradeoffs result from home production complementarities, we must naturally examine data accounting for the value of the entire stock of durables, not just new durables investment. This is because accumulated durable consumption assets, like kitchen appliances for example, contribute to home production output. Assuming the nominal value of the service flows of durables is equal to the aggregate resale value of all durables presently in utilization, the main goods expenditure series we construct will be the sum of non-durable expenditure and the nominal value of all consumer durables. Goods prices are adjusted to accommodate this new series, the details of which are described further in Appendix A.1.

The degree to which the U.S. has undergone structural change from goods to services dominated consumption depends to an extent on what underlying products and activities actually comprise the goods and services expenditure series. The sensitivity of measures of structural

[17] Throughout the quantitative analysis we will distinguish between aggregate quantity indices and actual quantities by denoting quantity indices with a tilde.



(a) Expenditure Ratio

(b) Quantity Index Ratio

(c) Price Index Ratio

Figure 1.11.: Clockwise from top left, we present the ratio of the aggregate nominal value of final goods to services consumption (a), the ratio of goods to services chain-weighted quantity indices with 2012 = 1 (b), and the relative chain-weighted price of goods to services where 2012 = 1 (c). In each plot, we show three data series each constructed to include different measures of consumer durables. The "Durables Stock" plots (solid black line) include the entire stock of existing consumer durables in the goods series. The "Durables Expend" plots (dotted red line) include only new investment in durables. The "No Durables" plots (dashed blue line) only include non-durables in the goods series. All series are annual, 1948-2018.

change can be illustrated by separately examining the long run ratios of goods to services expenditure, goods to services chain-weighted 2012 quantities, and goods to services chain-weighted 2012 prices when the goods data series account for consumer durables in varying degrees. These three different aggregate data series are presented in Figure 1.11, where we plot relative goods to services expenditure, relative quantity indices, and relative price indices separately depending on the degree to which we account for consumer durables. Note that in each of the plots, the services series are taken directly from the National Income and Product Accounts (NIPA) for services consumption expenditure, services chain-weighted 2012 quantity index, and services chain-weighted 2012 price index series. Meanwhile, the goods series in each plot are constructed so as to include the relevant data series associated with non-durable consumption in addition to one of the following: 1) the entire stock of consumer durable assets not including residential housing (solid black line), 2) only investment in new durable assets (dotted red line), 3) no measure of durables at all. Data on the nominal stock of durable assets is taken from the Bureau of Economic Analysis' (BEA) Fixed Asset Tables, while investment in durables comes from NIPA measures of durable consumption expenditure. For details on how our consumption and expenditure data series and their corresponding price and quantity indices are constructed, see Appendix A.1.1. We follow Bernanke (1985), McGrattan, Rogerson, and Wright (1993), and Gomme, Kydland, and Rupert (2001) in constructing a data series of goods expenditure using the durables stock in order to account for the presence of service flows from consumer durables in households' home production activities. In the model, we will not explicitly separate durables and non-durables, so that households can be thought to be purchasing the service flows of durables on the market, hence the need to adjust aggregate goods prices accordingly. This allows us to avoid having to estimate more than one simultaneous demand equation and generates ready comparisons to the literature examining the forces driving the rise of the U.S. services share.

Upon first glance, failure to include durables service flows can lead to biased estimates of the degree to which the value of final consumption in the U.S. has changed over the last half century. In Figure 1.11a notice that the decline in the nominal value of consumption goods, including the durables stock, relative to services is 64.2% (black line) versus a 75.5% decline when durables are totally left out (dashed blue line). Meanwhile, Figure 1.11c shows that when the value of the full durables stock is included, relative market-equivalent prices were over 3.5 times higher in 1948 compared to 2018, but only 1.8 times higher when leaving out durables. When including the full stock of durables, the decline in the nominal ratio appears at first glance to be driven more by strong relative price declines, since relative 2012 chain-weighted quantities of goods to services actually have increased over the last half century (black line in Figure 1.11b).

Looking at the dashed blue lines where durables are excluded, note that relative expenditure, relative quantities, and relative prices are all simultaneously falling, suggesting that something akin to income effects are generally outweighing relative price effects in long run structural transformation, a conclusion consistent with work in Kongsamut, Rebelo, and Xie (2001), Buera and Kaboski (2009), Herrendorf, Rogerson, and Valentinyi (2013), and Boppart (2014). This is because one would expect relative prices to move opposite relative quantities if relative price effects were significant. Either way, even after ignoring the solid black line which includes the full measure of the stock of durables, a natural question to ask is, why might such income effects be driving this trend? If income effects driving structural change do indeed result from consumers substituting market purchases away from time intensive home production tasks toward services, then the entire stock of consumer durables must be included in any goods data series used for quantitative analysis. This is because consumers use durables frequently in home production activities. Thus, inference on structural change using final expenditure cannot be complete without properly accounting for the value of this stock of assets and the ever-changing ways in which consumers use them.

While different empirical assessments of structural change have found that income effects inherent in consumer preferences are important, the definition of what constitutes an income effect has not been consistently deployed. The model in Boppart (2014) accounts for income effects by estimating the expenditure elasticity of demand for goods simultaneously while instrumenting for household income with a fixed-effects regression using a cross-section of consumption expenditure survey (CEX) data. Thus, Boppart (2014) quantifies a pure, classical income effect that can explain both dynamic and cross-sectional variation in expenditure shares. Herrendorf, Rogerson, and Valentinyi (2013), meanwhile, use aggregate data to counterfactually simulate the degree to which aggregate expenditure shares change when relative prices evolve as observed, but total expenditure remains constant. Since total consumption expenditure is historically a constant share of aggregate output, the income effect documented in Herrendorf, Rogerson, and Valentinyi (2013) can be read as an *aggregate* income effect. The measured income effect in Herrendorf, Rogerson, and Valentinyi (2013) is stronger than that in Boppart (2014), though the latter accounts for possible cross-sectional household preference heterogeneity. Herrendorf, Rogerson, and Valentinyi (2013) affirm the strength of the aggregate income effect by comparing the fit of a homothetic model with a model containing non-homotheticicities via Stone-Geary preferences.

Empirically, we examine how changes in relative prices and wages affect long-run expenditure and consumption for goods relative to services. Thus, the "income effect" we estimate is actually a wage effect since we do not account for changes in capital income impacting consumption. Our results and those in the literature are not necessarily comparable one-to-one, since the labor share of income has fallen in the U.S. from approximately 65% in 1948 to just over 58% in 2016.[18] Over the period 1948-2018, aggregate real wages equal to the sum of total labor compensation (employees' hourly wages plus proprietors' income) divided by total hours worked has increased.[19] With our model we exploit the convenient log-linear structure of relative demand to examine wage effects on the composition of aggregate consumption along with price changes. Given the decline in the labor share of income, wage effects will not necessarily be correlated one-to-one with income effects, but instead provide a proxy for us to understand how income effects operate in a structural model of consumption expenditure accounting for both durables utilization and time use complementarities with market purchases.

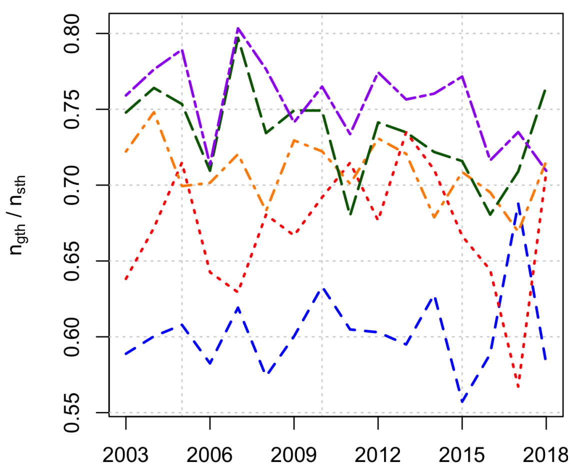
1.3.1.2. Model Support From Micro Data

If relative demand is sensitive to income, or more specifically wages, then we should expect to observe both intertemporal variation, on aggregate, but also cross-sectional variation in demand amongst different consumers in different financial situations. Further, the model provides us with a mechanism describing how demand and time use are linked, if such data can be found. To understand how consumers in different income brackets respond to wage and price variation, we turn to the Consumer Expenditure Survey (CEX) and American Time Use Survey (ATUS).

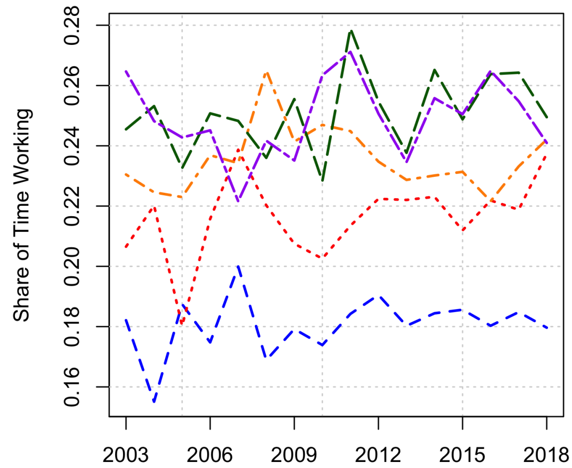
[18]See Giandrea and Sprague (2017).

[19]Nominal wages are equal to total nominal wages divided by total hours. For nominal wages, we use NIPA Table 2.1.1., "Compensation of employees." We then divide this value by non-seasonally adjusted annual total hours from NIPA Tables 6.9B, 6.9C, and 6.9D. See Appendix A.1.1 for how chain-weighted aggregate prices are constructed.

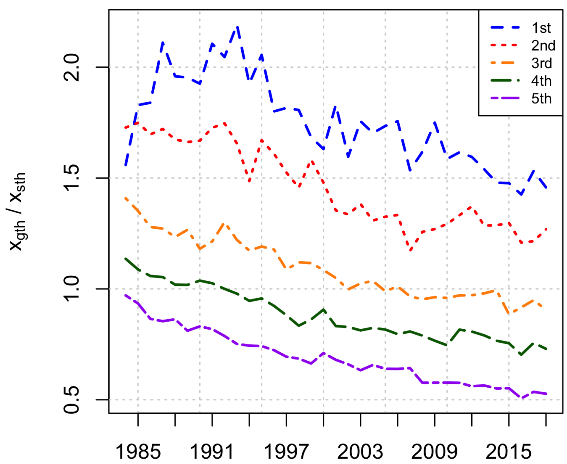
For household level wage data, we look at the annual March release of the Current Population Survey (CPS). We match the CEX summary cross-tabs by income quintile to the ATUS. Then, to construct separate spending series for "goods" and "services," we roughly match CEX spending and ATUS activities to the detailed expenditure categorizations in NIPA Table 2.3.5 — spending by "major type of product."[20]



(a) Relative Off-market Time



(b) Work Share



(c) Expenditure Ratio

Figure 1.12.: Clockwise from top left, we present the ratio of off-market goods to services time utilization from ATUS by income quintile, the share of total time by income quintile devoted to work from ATUS, and the ratio of goods to services expenditure from CEX. ATUS runs from 2003-2018 while the CEX runs from 1984-2018, so the lengths of the series are different. The legend denoting which color and line type scheme correspond to which income quintile is included in the Expenditure Ratio plot.

The ATUS dataset provides a convenient way to distinguish between activities associated with using "goods" versus "services." For example, the dataset contains both a variable that presents the time an individual respondent spent "Interior cleaning" and a separate variable that presents the time that same individual respondent spent "Using interior cleaning services."

[20]For the CEX classification, we apply the same rubric as in Boppart (2014).

In fact, almost all tasks within the survey are classified in this manner. Continuing with the interior cleaning example, a researcher can reasonably assume that an individual engaged in "Interior cleaning" is using his time along with goods like soaps, brushes, vacuums, dusters, etc. to accomplish the task of cleaning, while one engaged in "Using interior cleaning services" could reasonably be thought to be spending time monitoring a maid or housekeeper for whom he pays to perform cleaning services. This is just one example. Almost every household task featured in ATUS is classified in this manner. Thus, the survey structure makes it easy to distinguish between tasks that are complementary to using market goods and those which are complementary to using services.

Figure 1.12 presents a breakdown by income quintile of observed time use and spending behavior from the micro data. Lower income consumers spend relatively more time using services than goods, tend to work less, and spend a larger fraction of their disposable income on goods compared to higher income consumers. The fact that lower income consumers spend relatively more time using services than goods is a reflection of the quality of services they use, such as time intensive public transportation. While all workers spend, on average, more time using services than goods, it appears that the patterns of time use have not changed much since the ATUS was started in 2003. From Figure 1.12c, however, it is clear that relative goods to services expenditure has declined in a consistent manner for all consumers in all income quintiles. To compare the breakdown by income quintile with aggregates, use the dotted red line in Figure 1.11a as a reference point, since new durables expenditure is included in the CEX measures but not the exact value of durable assets owned.

Table 1.1.: Home Production Parameter Estimates From Micro Data

Reduced Form Coefficients		
	Goods	Services
$\ln(w_{th}/P_{gt})$	-0.822	
	(0.013)	
$\ln(w_{th}/P_{st})$		-1.672
		(0.026)
Constant	1.144	3.721
	(0.034)	(0.070)
Observations	80	80
R^2	0.982	0.981
Structural Parameters		
	Goods	Services
ω_i	0.025	0.000006
ν_i	-2.216	-1.598

To assess our structural model's cross-sectional predictions with respect to time utilization and consumption expenditure, we consider a simple linear regression on the equilibrium inframarginal rates of technical substitution for time and market purchases in processes $i \in \{g, s\}$ under CES utility, presented in equation (1.15). All cross-sectional heterogeneity is assumed to be attributed to wage differentials. For the data, we use the reported ATUS hourly wages and link the ATUS and CEX by income quintile. We then multiply both sides of (1.15) by real prices of goods and services in 2012 dollars, take logs, and use OLS on the following:

$$\ln P_{it} n_{ith} - \ln x_{ith} = \frac{1}{\nu_i - 1} \ln \left(\frac{\omega_i}{1-\omega_i} \right) + \frac{1}{\nu_i - 1} \ln \left(\frac{w_{th}}{P_{it}} \right) \quad \forall i \in \{g, s\} \tag{1.35}$$

where h indexes the households by income quintile, $h \in \{1, 2, 3, 4, 5\}$, and t indexes time and separate, independent regressions are run for g and s. The left hand dependent variable is the percent difference in the value of off-market time in minutes per unit of market consumption less the value of market consumption. Note that since ATUS begins in 2003 the regression runs from 2003-2018. Table 1.1 presents the values of the reduced-form and structural coefficients, as well as model fit statistics. The results suggest that income effects via wage differentials are important for characterizing the cross-sectional allocation of time use and consumption. We use the estimated structural parameters ν_i and ω_i from this regression as prior means in our full general equilibrium HMC model in Section 1.3.2.

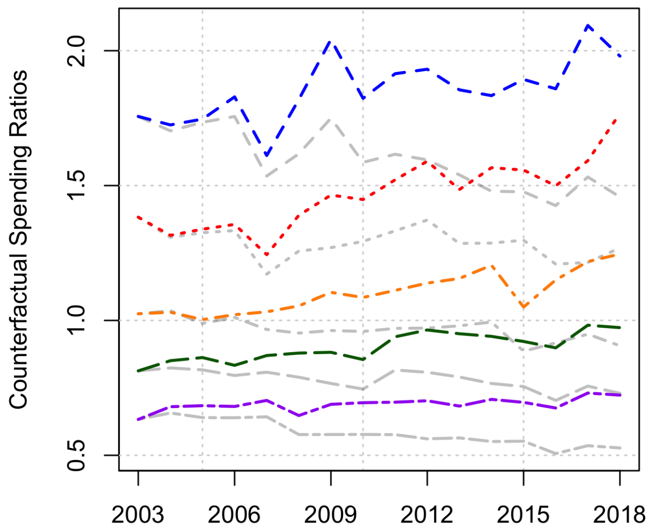
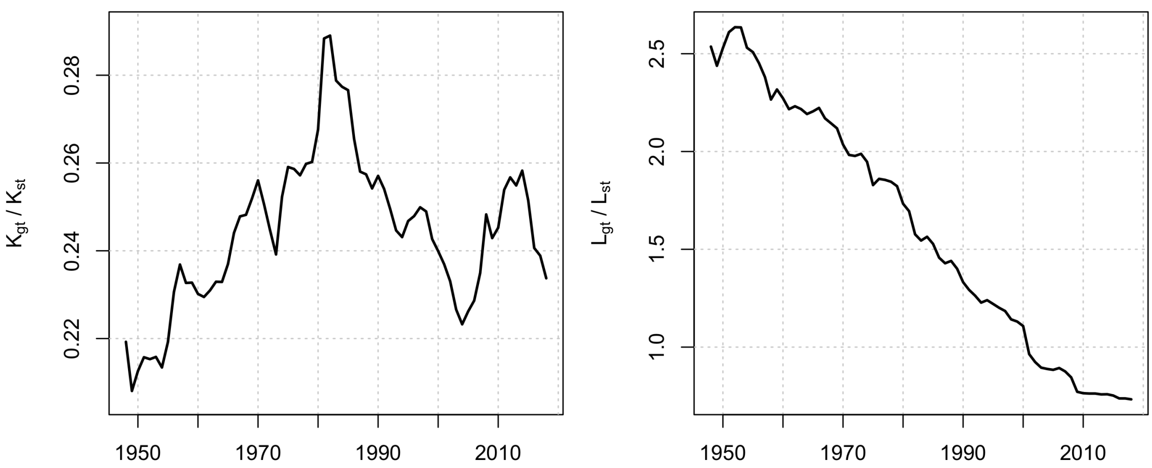


Figure 1.13.: Here we present simulations of the ratio x_{gth}/x_{sth}. In order from top to bottom are the 2003-2018 CEX expenditure ratios of the first, second, third, fourth, and fifth income quintiles. The grey lines represent data while the colored lines represent counterfactual expenditure ratios predicted by the regression in (1.35) when wages are fixed at their 2003 level.

With this incredibly simple picture of how households spend their time and money, it is natural to ask what the OLS regression would predict for consumption expenditure if real wages were constant and fixed at their 2003 levels, yet time use and prices evolved as observed in the data. That is, assume that some external process is causing consumers to change their time use patterns, yet real wages remain fixed: how does predicted expenditure change, simply as a result of price changes? Figure 1.13 shows the 2003-2018 predicted counterfactual expenditure ratios assuming time use and prices evolve as observed in data and wages remain fixed at their 2003 level. Fixing wages and eliminating income growth appears to have an effect on relative spending, suggesting that income effects driving cross-sectional consumption to services are important. This is, of course, only minor evidence for the importance of income effects, as we are limited by the time span of ATUS. Further, the naive regressions fail to account for the substitutability ρ of the outputs of home production associated with using goods g versus services s. Finally, these regressions operate under the strong assumption that wages are uncorrelated with home productivities z_{ith}, which do not factor into the inframarginal condition due to the CES assumption on home production. These assumptions are strong, and may be causing biased estimates of underlying parameters. In Section 1.3.4, after estimating a long run model on aggregate data that accounts for home production substitutability and correlations between average wages and productivities over time, as well as the general equilibrium endogeneity of prices, wage growth appears to have the opposite effect on spending ratios, suggesting a strong upward bias to the estimates presented here from micro-data, as well as possible estimates of income effects in other papers.

1.3.1.3. Capital and Labor Inputs by Sector



(a) Capital Ratio **(b)** Labor Ratio

Figure 1.14.: Here we present relative real capital and labor hours by sector from 1948 to 2018.

For the distribution of capital deployed by each sector, we use the BEA's Fixed Asset Table 3.1ESI and classify the industries into goods and services producing sectors roughly following the categorizations in NIPA Table 2.3.5. Table 3.1ESI provides the current value of fixed assets by industry, so we use the 2012 quantity indices from BEA Fixed Asset Table 3.2ESI to compute price indices for capital following the general price index unwinding procedure in Appendix A.1. To arrive at real values of capital, we deflate the nominal values using these price indices. Finally to put all real indices in the same units of 2012 consumption dollars, we divide all real components by the aggregate consumption price index from NIPA Table 2.3.4. Labor hours by sector are taken from NIPA Tables 6.9B, 6.9C, and 6.9D. We perform the same sector-specific classification for labor as for capital.

Figure 1.14 shows how the distributions of capital and labor hours have changed over time. While labor has shifted toward services consistently since 1948, relative goods to services capital utilization increased until 1980 before falling since. The shift in the labor distribution toward services has been previously observed, inspiring the models of Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008) where relative productivity changes and/or capital deepening contribute to the shift in labor from goods to services production along an aggregate-output balanced growth path.

1.3.2. Estimation Procedure

Our estimation procedure assumes a representative household and operates on aggregate data. The estimation routine over long-run aggregate data operates on three logarithmized equilibrium first-order conditions:

$$\begin{aligned}
&\ln\left(\theta/(1-\theta)\right) + \rho\Big(\ln q_{gt} - \ln q_{st}\Big) + \rho\Big(\ln z_{gt} - \ln z_{st}\Big) \\
&\quad + \ln \omega_g - \ln \omega_s + \left(\frac{\rho - \nu_g}{\nu_g}\right) \ln\left(\omega_g + (1-\omega_g)\left[\frac{W_t \omega_g}{P_{gt}(1-\omega_g)}\right]^{\frac{\nu_g}{\nu_g - 1}}\right) \\
&\quad - \left(\frac{\rho - \nu_s}{\nu_s}\right) \ln\left(\omega_s + (1-\omega_s)\left[\frac{W_t \omega_s}{P_{st}(1-\omega_s)}\right]^{\frac{\nu_s}{\nu_s - 1}}\right) - \ln P_{gt} + \ln P_{st} = 0
\end{aligned} \tag{1.36}$$

$$\ln(1-\alpha_i) + \alpha_i \ln\left(\frac{K_{it}}{L_{it}}\right) + \ln A_{it} - (\alpha_i - 1)\ln \zeta_{it} - \ln\left(\frac{W_t}{P_{it}}\right) = 0 \quad \forall i \in \{g, s\} \tag{1.37}$$

Define the following stacked vector of productivity terms, and assume these terms follow a random walk:

$$\boldsymbol{\xi}_t = \begin{pmatrix} \ln z_{gt} - \ln z_{st} \\ \ln A_{gt} - (\alpha_g - 1)\ln \zeta_{gt} \\ \ln A_{st} - (\alpha_s - 1)\ln \zeta_{st} \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\xi}_{t-1}, \boldsymbol{\Omega}) \tag{1.38}$$

Finally, assume $\theta = \frac{1}{2}$. Note that we cannot identify θ separately from the level change of *relative* home productivities. That is, we can only identify the level change in relative productivities if θ is known. Even then, we cannot identify the level of relative productivities z_{gt}/z_{st} without a normalizing assumption for every period.

As in all DSGE models, the general equilibrium is fully characterized by conditions (1.36) through (1.37) but also the Euler equation, firm marginal products of capital, and market clearing conditions. Note that all structural parameters and exogenous productivity residuals are featured in (1.36) through (1.37) meaning that if we were to form a proper likelihood around these plus any additional equations we would have to introduce either measurement or model error terms. Instead, the equations above, since they must hold in equilibrium, are sufficient to identify the structural parameters under the assumption that productivity residuals are exogenous and both firms and consumers are price takers. Given these assumptions, we capture the dynamic evolution of the system by assuming the productivities described by $\boldsymbol{\xi}_t$ follow a unit-root process. This permits us both to avoid needing to estimate something like an AR(1) or ARIMA shock process on unobserved residuals, while also allowing us to form our likelihood around first differences of (1.36) and (1.37) and subsequently back out structural productivities up to a normalization for the initial period of the sample.

Let Δ be the backwards first-difference operator and consider the regression on first differences:

$$\begin{aligned}
&\rho\Big(\Delta \ln q_{gt} - \Delta \ln q_{st}\Big) + \rho\Big(\Delta \ln z_{gt} - \Delta \ln z_{st}\Big) \\
&\quad + \left(\frac{\rho - \nu_g}{\nu_g}\right) \ln\left(\omega_g + (1-\omega_g)\left[\frac{W_t \omega_g}{P_{gt}(1-\omega_g)}\right]^{\frac{\nu_g}{\nu_g - 1}}\right) \\
&\quad - \left(\frac{\rho - \nu_s}{\nu_s}\right) \ln\left(\omega_s + (1-\omega_s)\left[\frac{W_t \omega_s}{P_{st}(1-\omega_s)}\right]^{\frac{\nu_s}{\nu_s - 1}}\right) \\
&\quad - \left(\frac{\rho - \nu_g}{\nu_g}\right) \ln\left(\omega_g + (1-\omega_g)\left[\frac{W_{t-1} \omega_g}{P_{g,t-1}(1-\omega_g)}\right]^{\frac{\nu_g}{\nu_g - 1}}\right) \\
&\quad + \left(\frac{\rho - \nu_s}{\nu_s}\right) \ln\left(\omega_s + (1-\omega_s)\left[\frac{W_{t-1} \omega_s}{P_{s,t-1}(1-\omega_s)}\right]^{\frac{\nu_s}{\nu_s - 1}}\right) \\
&\quad - \Delta \ln P_{gt} + \Delta \ln P_{st} = \epsilon_t^1
\end{aligned} \tag{1.39}$$

$$\alpha_i(\Delta \ln K_{it} - \Delta L_{it}) + \Delta \ln A_{it} - (\alpha_i - 1)\ln \Delta\zeta_{it} - \Delta \ln W_t + \Delta \ln P_{it} = \epsilon_t^i \quad \forall i \in \{g, s\} \tag{1.40}$$

$\boldsymbol{\epsilon}_t$ is an idiosyncratic model error we assume is multivariate standard normal, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbb{I})$. Note that covariation between equilibrium conditions is captured by $\mathbf{\Omega}$. We make a parameterized assumption on $\mathbf{\Omega}$ which will allow for more variability and flexibility around the predictions than the standard normal assumption on the idiosyncratic error structure above.

All covariates in (1.39) thru (1.40) are endogenous, including the firms' choices of the capital/labor ratio. But information about capital and labor supply decisions by the household are fully encoded in prices and quantities. Thus, the capital/labor ratio is correlated with household decision through wages, since households respond to wages when choosing how much to work. Similarly, quantities of market purchases chosen by the household are linked with firm output only through market prices, which contain information about firm productivity. Thus, prices are endogenous, but the underlying decisions of firms and households are seemingly unrelated except through prices: firms do not think about how much households will adjust their labor supply or capital supply in response to their own decisions, nor do households think about how much firms will raise or lower output in response to the effect of their quantity purchases on prices.

1.3.2.1. Distributional Assumptions

We will run our HMC estimation routine twice, first assuming final goods c_{it} are substitutes so that $\rho \in (0, 1)$, then assuming they are complements so that $\rho < 0$. For each HMC routine, we impose the same priors for parameters ν_i, ω_i, and α_i, $\forall i \in \{g, s\}$ and $\mathbf{\Omega}$, changing only the underlying prior restrictions on ρ. Using the estimates from micro data presented in Section 1.3.1.2 we impose strong priors on ν_i and ω_i, having found that these parameters appear only weakly identified. We impose flat priors on α_i and an LKJ prior distribution on the unit diagonal Cholesky factor, Ξ, of the variance/covariance matrix $\boldsymbol{\Omega}$ to ensure computational tractability in the Hamiltonian Monte Carlo routine implemented with the Stan software package.[21] In Stan, LKJ priors are implemented with a unit diagonal, so we allow the diagonal terms χ_j to take independent half Cauchy priors. The full set of prior distributional assumptions is as follows:

$$-\rho \sim \mathcal{LN}\Big(-\frac{1}{2}, 1\Big) \quad \text{if} \quad \rho < 0 \tag{1.41}$$

$$\rho \sim \text{Beta}(1,1) \quad \text{if} \quad \rho \in (0,1) \tag{1.42}$$

$$\omega_i \sim \text{Beta}(1, \widehat{\omega}_i) \tag{1.43}$$

$$-\nu_i \sim \mathcal{LN}(\widehat{\nu}_i, 1) \tag{1.44}$$

$$\alpha_i \sim \mathcal{U}(0,1) \tag{1.45}$$

$$\text{chol}(\boldsymbol{\Omega}) = \text{diag}(\boldsymbol{\chi}) \cdot \Xi \tag{1.46}$$

$$\chi_j \sim \text{Cauchy}_{(0,\infty)}(0,2) \quad \forall j \in \{1, \ldots, 4\} \tag{1.47}$$

$$\Xi \sim \text{LKJ}(2) \tag{1.48}$$

$\widehat{\omega}_i$ and $\widehat{\nu}_i$ are computed using information from the OLS estimates from Table 1.1. Specifically, we set $\widehat{\omega}_g = 39.684$ and $\widehat{\omega}_s = 1000.0$ so the mean of the Beta prior is centered around micro estimates. We also set $\widehat{\nu}_g = -2.216$ and $\widehat{\nu}_s = -1.598$.

We have now defined all of the objects needed to estimate the Bayesian posterior distribution of parameters conditional on data using a Hamiltonian Monte Carlo regression on first differences. Let $\mathcal{P}$ be the set of structural parameters and $\mathcal{D}$ the data. Let $\pi(\cdot)$ denote an arbitrary density function. We seek to estimate the posterior distribution:

$$\mathcal{P} = \Big\{\rho, \boldsymbol{\omega}, \boldsymbol{\nu}, \boldsymbol{\alpha}, \boldsymbol{\Omega}, \boldsymbol{\chi}, \Xi, \{\boldsymbol{\xi}_t\}_{t\geq 1}\Big\} \tag{1.49}$$

$$\mathcal{D} = \Big\{\{\boldsymbol{q}_t, \boldsymbol{P}_t, w_t, \boldsymbol{K}_i, \boldsymbol{L}_t\}_{t\geq 1}\Big\} \tag{1.50}$$

$$\underbrace{\pi(\mathcal{P} \mid \mathcal{D})}_{\text{Posterior}} \propto \underbrace{\pi(\mathcal{D} \mid \mathcal{P})}_{\substack{\text{Data Given} \\ \text{Parameters}}} \underbrace{\pi(\mathcal{P})}_{\text{Prior}} \tag{1.51}$$

$\pi(\mathcal{D} \mid \mathcal{P})$ represents the probability of realizing the data given the parameterization, i.e. the "likelihood" function:

$$\pi(\mathcal{D} \mid \mathcal{P}) \propto \prod_t \exp\Big\{-\frac{1}{2}\boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t\Big\} \tag{1.52}$$

The full prior representation depends on the whether we initially assume the final goods are complements or substitutes.

[21] For more information on the LKJ distribution see Lewandowski, Kurowicka, and Joe (2009). For more information on Hamiltonian Monte Carlo see the brief summary in Gelman et al. (2013c) or more technical treatments in Neal (2011) and Betancourt and Stein (2011). We also provide a brief overview of how HMC works in Appendix A.3.1. For more information on the Stan software we use to estimate the model, see Stan Development Team (2016) and Gelman et al. (2013b).

1.3.3. Estimation Results

Our estimation strategy is designed to back out structural demand parameters, controlling for price endogeneity using general equilibrium conditions. We then use these parameters to simulate counterfactual allocations of consumption in order to separately assess the degree to which wage changes, relative price changes, and home productivity changes have contributed to the rising services share of expenditure. In two of the estimations we use full consumption data including the service flows from the stock of consumer durables, assuming $\rho \in (0,1)$ in one estimation and $\rho < 0$ in another. In the remaining two estimations we use only data on new investment in durables, again running separate HMC routines for when final home consumption commodities $\boldsymbol{c}_t$ are substitutes or complements. The data series run from 1948-2018 and a representative household is assumed. While this assumption is not necessarily innocuous, since cross-sectional effects are at play as evidence in Section 1.3.1.2 suggests, our goal is to examine the degree to which income effects and relative price effects have impacted aggregate demand when accounting for previously un-modeled off-market time use and consumption complementarities over the long run, as we do.

Table 1.2.: HMC Posterior Results, 1948-2018 Including Durables Stock

$\rho \in (0,1)$	$\widehat{R}$	n_{eff}/N	Mean	S.D.	2.5%	50%	97.5%
ρ	1.001	0.254	0.850	0.113	0.590	0.874	0.994
ν_g	1.001	0.244	-0.927	0.502	-2.058	-0.847	-0.232
ν_s	1.001	0.186	-1.263	1.056	-3.636	-1.070	-0.142
ω_g	1.000	0.445	0.013	0.016	0.000	0.008	0.055
ω_s	1.000	0.616	0.001	0.001	0.000	0.001	0.004
α_g	1.000	0.241	0.102	0.045	0.024	0.100	0.199
α_s	1.000	0.300	0.215	0.042	0.134	0.214	0.300
$U(\mathcal{P})$	1.003	0.160	689.037	2.854	682.480	689.403	693.544
$\rho \in (-\infty,0)$	$\widehat{R}$	n_{eff}/N	Mean	S.D.	2.5%	50%	97.5%
ρ	1.000	0.302	-6.450	4.120	-17.500	-5.332	-2.261
ν_g	1.001	0.146	-1.339	0.995	-3.265	-1.159	-0.588
ν_s	1.001	0.265	-4.755	3.366	-13.556	-3.797	-1.653
ω_g	1.000	0.483	0.027	0.025	0.001	0.020	0.092
ω_s	1.000	0.545	0.001	0.001	0.000	0.001	0.004
α_g	1.000	0.276	0.175	0.064	0.051	0.175	0.302
α_s	1.000	0.316	0.258	0.048	0.167	0.258	0.354
$U(\mathcal{P})$	1.001	0.164	688.692	2.781	682.307	689.051	693.076

The HMC estimation routine operates using the software Stan (Stan Development Team 2016). We use Stan's default tuning parameter values except for the adaptation acceptance rate, which we find from simulations is best picked to be 0.995, as opposed to the default value of 0.8. The HMC is run in parallel on six independent chains each with a sampling space of size $4,000$ for a total posterior sample size of $N = 24,000$. Table 1.2 presents summary statistics for the posterior distribution estimates of reduced-form parameters for the sample accounting

Table 1.3.: HMC Posterior Results, 1948-2018 with Only Durables Investment

$\rho \in (0,1)$	$\widehat{R}$	n_{eff}/N	Mean	S.D.	2.5%	50%	97.5%
ρ	1.001	0.164	0.671	0.164	0.379	0.658	0.974
ν_g	1.000	0.273	-1.253	0.678	-2.806	-1.136	-0.371
ν_s	1.001	0.167	-0.867	1.173	-3.276	-0.584	-0.106
ω_g	1.000	0.438	0.016	0.018	0.000	0.010	0.065
ω_s	1.000	0.506	0.001	0.001	0.000	0.001	0.004
α_g	1.000	0.191	0.225	0.046	0.142	0.223	0.320
α_s	1.000	0.262	0.264	0.047	0.175	0.262	0.360
$U(\mathcal{P})$	1.001	0.148	664.627	2.798	658.288	664.949	669.106

$\rho \in (-\infty,0)$	$\widehat{R}$	n_{eff}/N	Mean	S.D.	2.5%	50%	97.5%
ρ	1.001	0.276	-5.886	4.157	-16.452	-4.748	-1.842
ν_g	1.000	0.262	-3.279	3.099	-11.650	-2.378	-0.862
ν_s	1.000	0.249	-2.688	2.782	-9.303	-1.912	-0.578
ω_g	1.000	0.559	0.027	0.025	0.001	0.020	0.096
ω_s	1.000	0.630	0.001	0.001	0.000	0.001	0.004
α_g	1.000	0.250	0.239	0.067	0.103	0.241	0.365
α_s	1.000	0.267	0.263	0.052	0.161	0.264	0.367
$U(\mathcal{P})$	1.001	0.173	656.137	2.702	650.043	656.461	660.413

for the value of the durables stock used in home production. Table 1.3 presents summary statistics for the estimation routine using just new spending on consumer durables, but not the service flows. The last three columns of each table represent distribution percentiles, with a 95% confidence region buttressed on either side by the 2.5-percentile and 97.5-percentile. The mean and standard deviation of the posterior distributions are also presented, along with the $\widehat{R}$ and effective sample size n_{eff}/N diagnostic statistics. Integration convergence assessments are made by examining $\widehat{R}$ and the effective sample size n_{eff}/N for the log-posterior density, $U(\mathcal{P})$. For all parameters and the density $\widehat{R} \rightarrow_{+} 1$ as the numerical integration converges. Generally speaking, we want $\widehat{R} < 1.1$ and $n_{eff}/N > 0.10$, especially for $U(\mathcal{P})$ (Gelman et al. 2013b). Notice that, under these conditions, it appears our HMC integration scheme has converged and model fit is satisfactory. For a more in-depth discussion about HMC see Appendix A.3.1.

Having run two separate estimation routines each for two separate datasets, we seek a likelihood-based measure of model performance to understand whether $\rho \in (0,1)$ is more likely or $\rho < 0$ more likely. To do this, we compute the posterior odds of the substitutes model being true versus the complements model by using bridge sampling techniques to back out Bayes factors.[22] Let $\mathcal{M}_c$ denote the complements model where $\rho \in (-\infty, 0)$ and let $\mathcal{M}_s$ denote the substitutes model where $\rho \in (0,1)$. Let $\mathcal{D}_{\text{DursStock}}$ denote the dataset with the value of the durables stock included in the goods series and $\mathcal{D}_{\text{NewDurs}}$ denote the dataset where only

[22] The Bayes factor is the ratio of posterior expectations, which each require computation of the marginal likelihood of realizing the observed data given the model under question (see Rossi, Allenby, and McCulloch (2005) for a thorough explanation of model selection using Bayes factors). Bridge sampling is the computational procedure used to compute the marginal likelihood required for Bayes factor analysis (see Meng and Wong (1996) and Gronau et al. (2017) for information regarding the computational procedure).

purchases of new durables are considered. The posterior model odds computed using Bayes factors are

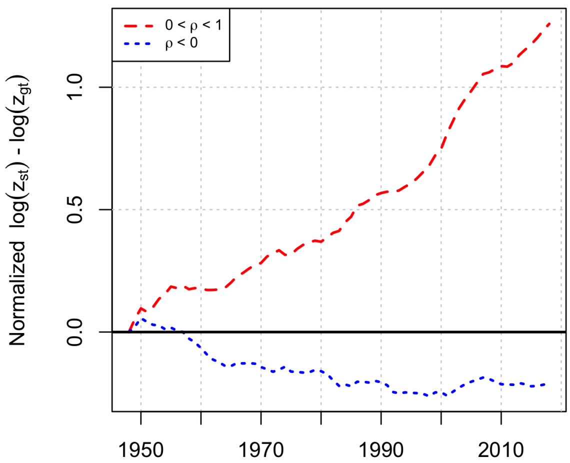
$$\underset{\mathcal{M}_s \text{ vs. } \mathcal{M}_c}{\text{Posterior Odds}} = \frac{\mathbb{P}\big(\mathcal{D}_{\text{DursStock}} \mid \mathcal{M}_s\big)}{\mathbb{P}\big(\mathcal{D}_{\text{DursStock}} \mid \mathcal{M}_c\big)} = 2.571 \quad \Rightarrow \quad \text{Weak Evidence for } \mathcal{M}_s \tag{1.53}$$

$$\underset{\mathcal{M}_s \text{ vs. } \mathcal{M}_c}{\text{Posterior Odds}} = \frac{\mathbb{P}\big(\mathcal{D}_{\text{NewDurs}} \mid \mathcal{M}_s\big)}{\mathbb{P}\big(\mathcal{D}_{\text{NewDurs}} \mid \mathcal{M}_c\big)} = 1346.952 >> 1 \quad \Rightarrow \quad \text{Strong Evidence for } \mathcal{M}_s \tag{1.54}$$

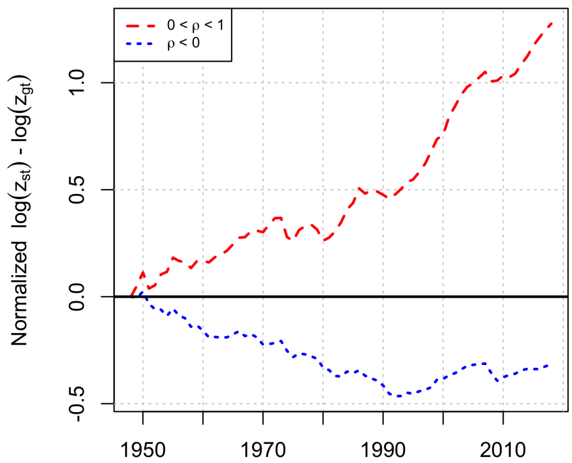
The evidence classification rubric we follow comes from Harold Jeffreys' original scale of evidence as presented in Wasserman (2000). Generally speaking, Bayes factors less than 3 are considered weak and fail to support evidence of one model over the other. The fact that we cannot definitively support one model over the other when the value of all durables in utilization is included suggests that structural parameters are likely co-dependent and only weakly identified. Indeed, focussing on Table 1.2, examination of the values of ρ and $\boldsymbol{\nu}$ in both estimations suggest that the posterior distribution of $\boldsymbol{\nu}$ shifts to accommodate different assumptions on ρ, though values of the log-posterior distribution $U(\mathcal{P})$ between the two models are practically identical. By contrast, turning to Table 1.3, $U(\mathcal{P})$ for $\rho \in (0, 1)$ has a larger posterior peak and does not overlap in value with the $\rho < 0$ case, suggesting that the substitutes model is a significantly better fit of the data, thus leading to a larger Bayes factor estimate.

In terms of identification, values of ν_i appear moderately identified, overcoming their strong priors. Values of ω_i and α_i appear strongly identified as well overcoming their flat unit-interval priors. In the substitutes models, the posterior distribution of ρ is skewed to the right with mass concentrated near $\rho = 1$ suggesting that the outputs of home production are near-perfect substitutes. This substitutability is stronger when the durables services flows are included than when they are left out.

In Figure 1.15 we present the log relative productivity residuals $\ln z_{st} - \ln z_{gt}$ for all four model estimations, normalized so that the log-levels are equal to 0 in the first year of the sample, 1948. We present the normalized residuals, rather than their raw log-levels because our residual estimates are conditional on the identifying assumption that $\theta = \frac{1}{2}$. For different values of the utility weight θ the raw residual time series will shift up or down by a fixed constant, yet their first-order and second-order properties will remain the same. Readers should thus focus on these series first-order properties here. In our counterfactual simulations, we will use the raw residual levels so as to ensure all counterfactual simulations start at the same level of relative consumption. Turning back to Figure 1.15 note that, depending on whether the value of durables is included, inferences on relative productivities differ. When the final outputs of home production are substitutes (red lines), services have increased in relative productivity to goods since 1948, regardless of which data series is used. Our residual estimates suggest the opposite for when the outputs of home production are complements.



(a) Durables Stock



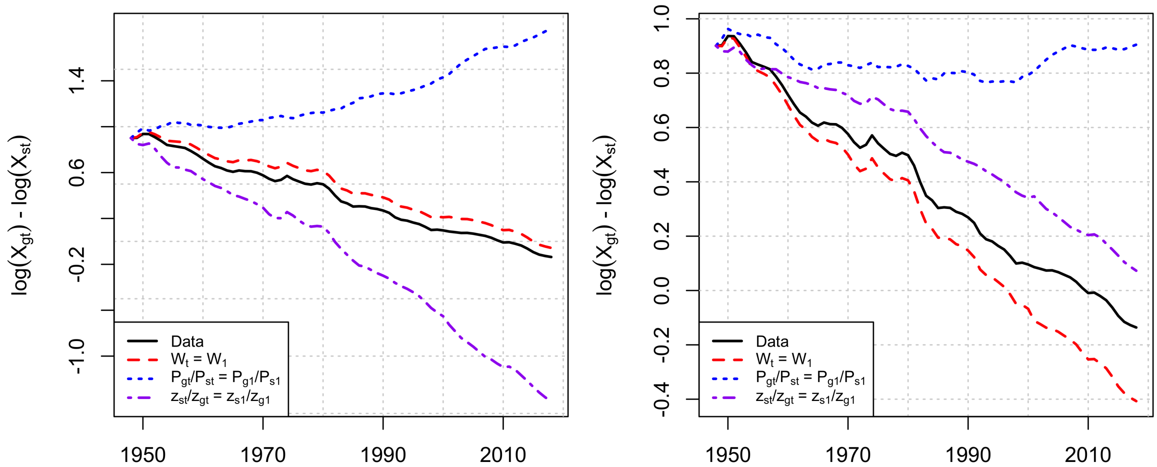
(b) New Durables Only

Figure 1.15.: Here we present the posterior means of the estimated relative home productivities time series as inferred from the HMC estimations for $\rho \in (0, 1)$ (red) and $\rho < 0$ (blue) separately. When relative productivities are greater than zero, services are relatively more productive or quality-efficient in the home production process than goods relative to a 1948 baseline.

1.3.4. Demand-Driven Structural Change with Aggregate Consumption

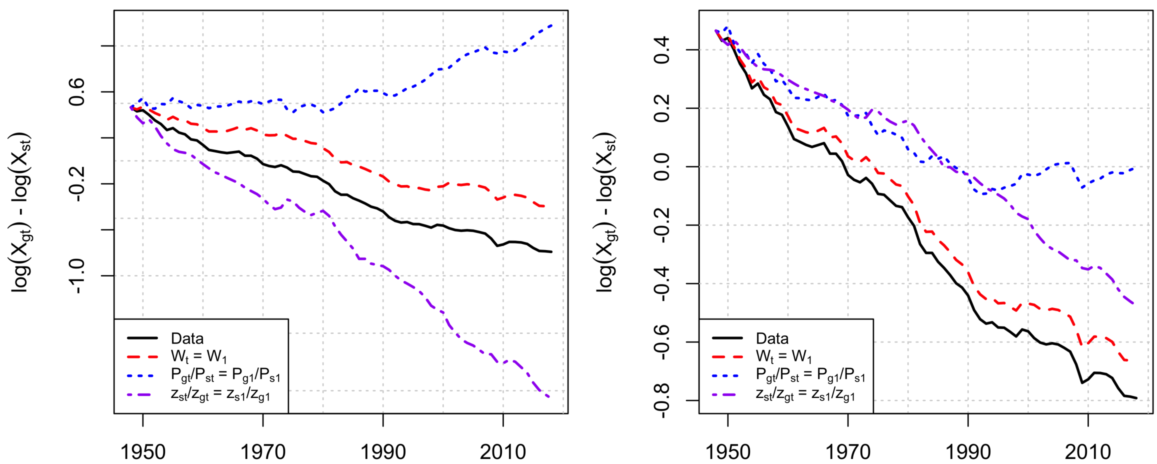
The goal of this paper is to better understand how wage growth and relative price changes affect aggregate demand allocations when controlling for home production complementarities and time to consume. For this reason, we consider three counterfactual simulations for each separate model estimation — $\rho \in (0, 1)$ and $\rho < 0$ — and each dataset — including and not including the value of the stock of durables. Our counterfactuals operate solely on the relative demand equation in (1.36). First, we fix aggregate wages $W_t = W_{1948}$, eliminating the effects of real wage growth by also dividing P_{gt} and P_{st} by an aggregate price index P_t so that the aggregate consumption price level remains fixed as well. In this counterfactual economy, there is thus no real inflation. Second, we allow W_t to evolve as observed in the real data but fix relative prices P_{gt}/P_{st} at their 1948 value. In each of these scenarios we simulate the counterfactual series using the implied relative productivities as estimated from the residuals of either the substitutes or complements models without first differences and presented in Figure 1.15. Third and finally, we assume the relative productivities z_{st}/z_{gt} are fixed at their 1948 levels and do not change. In each scenario, variables that are not explicitly adjusted are taken at their observed data values for the entire time series.

Figure 1.16 compares the counterfactual simulations with the data (black line) for estimations on the dataset featuring the full durables stock. Figure 1.17 shows the same counterfactual simulations using the estimates on the dataset featuring only new spending, not the value of the stock. We normalize all simulations to year one in our sample, 1948, and unwind the first differences equation to plot the counterfactuals and data in log-levels. Note that without wage growth (dashed red lines), long run structural change is slightly weaker than what is observed in the data, though this pails in comparison to the counterfactually estimated relative-price



(a) $\rho \in (0, 1)$ **(b)** $\rho < 0$

Figure 1.16.: We present two counterfactual simulations on the dataset featuring the full durables stock. Under these simulations, we first assume no aggregate wage growth then no aggregate changes to relative prices. Under the red line we assume wages are constant and fixed over time. The blue line presents a counterfactual where real wage growth happens, but relative prices remain constant. Finally, the purple line is a counterfactual simulation where we assume the relative home productivities of services to goods are assumed constant at their inferred 1948 level.



(a) $\rho \in (0, 1)$ **(b)** $\rho < 0$

Figure 1.17.: Here we present results form the same counterfactual simulations describe above, except we use the dataset without accounting for the durables stock and the associated regression parameter estimates.

effect (dotted blue lines). Without wage growth and only relative price changes, we would expect consumers to prefer services to goods slightly *less* than they do now, looking at the substitutes cases. When the durables service flows are included — Figure 1.16a — the wage effect is practically insignificant, though. This calls into question previous findings that fail to account for durables stock utilization in home production but suggest the income effect is the primary driver of structural change (see Herrendorf, Rogerson, and Valentinyi (2013) and Boppart (2014)). Herrendorf, Rogerson, and Valentinyi (2013) come to that conclusion by using a Stone-Geary utility representation that imposes arbitrary non-linearities on Engel curves. Boppart (2014) comes to that conclusion using an equally arbitrary utility specification that also generates non-linearities in the Engel curves. Neither account for time use complementarities in home production or durables services flows, as we do. Further, evidence using all four posterior estimates suggests that shutting off relative price changes while still allowing for real wages to grow (dotted blue line) implies that relative goods to services expenditure would actually have *increased* over the 1948-2018 period. This demonstrates that relative price effects associated with classical substitution effects appear to dominate income effects from wage growth. Our results here are thus the first to show that the income effect story becomes suspect when demand-side home production and time to consume are accounted for.

Focussing in more detail on the results, the data featuring the durables stock show that the ratio of the nominal value of goods to services used by households has fallen by 64.5% since 1948, while when the stock is excluded the fall is 71.5%. Looking first at the model with the full durables stock and assuming $\rho \in (0, 1)$, eliminating counterfactual wages leads to an almost indiscernible difference in the decline of relative consumption, 61.6% to be exact. Allowing for wage growth and fixing relative prices, using the same dataset and looking at the dotted blue line, we actually see a 160% *increase* over the period 1948-2018. Continuing to look specifically at the substitutes estimation, fixing relative prices and simulating the model leads to a 103.8% *increase* with the naive dataset featuring only new investment. When wage growth is included absent relative price changes, we thus see the opposite kind of structural change. On the other hand without wage growth and only relative price changes, we see a mild 57.9% drop in relative consumption, just slightly less than that observed. The qualitative nature of these results does not change when $\rho < 0$. In all of the plots featured in Figures 1.16 and 1.17 we include counterfactual simulations having fixed the relative productivity residuals (purple lines) in case the reader is interested. Again though, we caution readers not to make strong interpretations of these series given the non-identifiability of θ in the CES utility function, as discussed above in Section 1.3.3.

1.3.5. Testing Quantitative Implications Under Cross-Sectional Heterogeneity

Consider the estimated posterior distribution of parameters given data from the general equilibrium model. Suppose that all households face the same productivity evolution, so that z_{gt} and z_{st} are no longer taken to be household-dependent. Under this assumption, we can think of z_{gt} and z_{st} as quality residuals, capturing variation in the relative usefulness of goods and services toward home production over time. Using our parameter and productivity estimates, we now consider how predicted counterfactual expenditure ratios would evolve by income quintile, using wage quintiles from the CPS matched to CEX income quintiles, as described above. Again, we first fix wages at their 1984 levels, the beginning of the CEX sample for which we have data, and watch the spending ratios evolve when influenced solely by relative price changes. We then perform the same counterfactual exercises as above, fixing relative prices and relative productivities separately. The results in terms of data versus posterior counterfactual

Table 1.4.: CEX Percent Change (1984 to 2018) in x_{gth}/x_{sth} — Data vs. Counterfactuals

		$\rho \in (0,1)$		$\rho < 0$	
Income Quintile	Data	$w_{th} = w_{1h}$	$\frac{P_{gt}}{P_{st}} = \frac{P_{g1}}{P_{s1}}$	$w_{th} = w_{1h}$	$\frac{P_{gt}}{P_{st}} = \frac{P_{g1}}{P_{s1}}$
1^{st}	-6.490	33.249	38.066	-60.931	-33.249
2^{nd}	-26.552	34.997	38.624	-60.851	-32.680
3^{rd}	-35.674	35.682	38.735	-60.819	-32.485
4^{th}	-35.750	36.127	39.348	-60.798	-32.605
5^{th}	-45.685	36.677	40.601	-60.772	-32.903

predicted percent changes in spending ratios by income quintile from the beginning of the CEX in 1984 to 2018 are presented in Table 1.4.

The degree to which the outputs of home production are substitutable determines the sign of counterfactual changes. If final outputs are substitutes so that $\rho \in (0, 1)$, we predict a relative increase in goods to services spending both when wages are fixed at their 1984 level and relative prices remain unchanged. However, again, the relative price effect appears stronger than the income effect. If $\rho < 0$ holding relative prices fixed appears to buffer the decline in relative spending for higher income agents, but not lower income ones. Further, when $\rho < 0$ we predict *greater* counterfactual structural change if wages remain fixed, so that when the outputs of home production are complementary, the story whereby increasing income leads to an increasing share of services consumption is suspect.

1.3.6. Empirical Evidence and Model Implications for Labor Supply Responsiveness to Relative Price Changes

We are interested in how the per-worker supply of labor on the intensive margin is affected by relative price changes. Will hours increase in response to certain products in the consumption basket becoming relatively more expensive, for example? This would suggest that consumers may pick up extra work or odd jobs to earn more income and thus better finance certain new consumption baskets if relative prices change but the aggregate price level remains constant.

All of our analyses in this section operate on price data that account for the re-sale value of consumer durables. These are the price data that are used in the estimations where the goods series account for the value of the durables stock used in home production. To understand what the aggregate data say about the relationship between effective full time hours per worker and relative price changes, we run some simple linear regressions presented in Table 1.5 taking the log of hours per worker per week as the left-hand side variable. To compute hours per worker, we divide the total hours data presented in NIPA Tables 6.9B, 6.9C, and 6.9D by the full time equivalent measure of total workers featured in NIPA Tables 6.5B, 6.5C, and 6.5D. Notice that labor hours per worker appear to be decreasing in relative goods to services prices and wages (column 1), though when not controlling for real wage inflation, we would instead expect labor hours per capita to decline as the relative price of services rises (column 2). Indeed, that is what we observe in the data. Labor hours per effective full time worker per week have fallen from a high of 34.43 hours in 1948 to 25.54 hours in 2018, reflecting the fact that a greater share of workers engage in part time work today.

Table 1.5.: Data Dependency of Labor Hours on Relative Prices and Wages

	Dependent variable:		
	$\ln l_t$		
	(1)	(2)	(3)
$\ln P_{gt} - \ln P_{st}$	−0.084	0.196	
	(0.052)	(0.011)	
$\ln W_t$	−0.251		−0.178
	(0.046)		(0.008)
Constant	2.553	3.263	2.758
	(0.130)	(0.008)	(0.029)
Observations	71	71	71
R^2	0.883	0.832	0.879

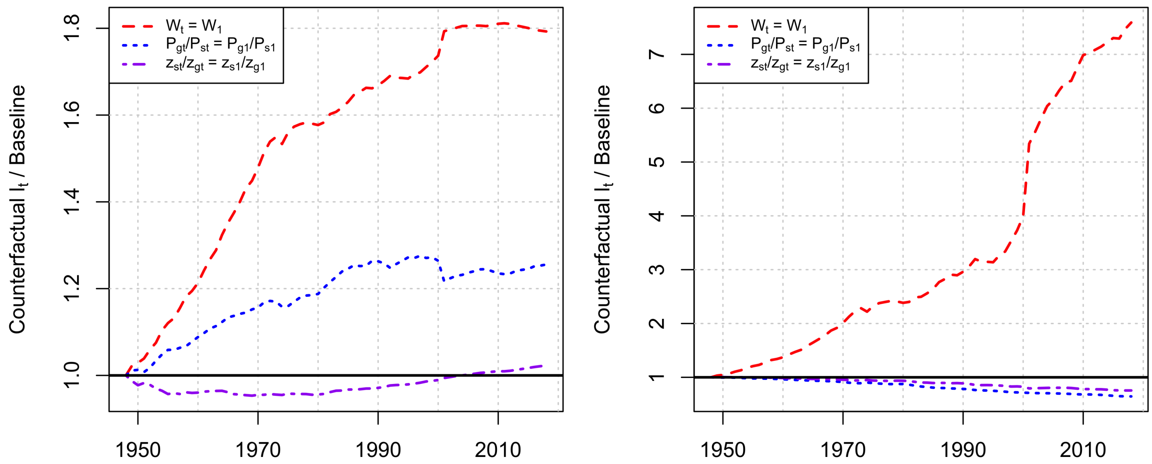
Table 1.6.: Model Predictive Labor Hours Dependency on Relative Prices and Wages

	$\ln \widetilde{l}_t$, *Substitutes or Complements*					
	$\rho \in (0, 1)$			$\rho < 0$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln P_{gt} - \ln P_{st}$	−0.368	0.536		−0.020	1.215	
	(0.101)	(0.026)		(0.052)	(0.027)	
$\ln w_t$	−0.813		−0.491	−1.110		−1.092
	(0.089)		(0.017)	(0.046)		(0.008)
Constant	0.358	2.653	1.258	−1.284	1.851	−1.234
	(0.252)	(0.020)	(0.060)	(0.131)	(0.022)	(0.028)
Observations	71	71	71	71	71	71
R^2	0.939	0.864	0.927	0.996	0.966	0.996

To understand how the qualitative implications for household labor supply under our full model with CES home production, we simulate several counterfactual experiments against a predictive baseline.[23] Note that, while household labor supply is endogenous in the full general equilibrium model, our parameter estimates do not take into direct consideration this choice, operating solely on the equilibrium estimating equations in (1.39) through (1.40). Under the

[23] The full household labor supply function is presented in Appendix A.2.2 after derivative of the Cobb-Douglas home production labor supply function used in our comparative state exercises.

assumption that our model indeed correctly describes the data generating process, the labor supply choice is exactly determined by prices and home productivities, which we estimate from the residuals of the household's relative demand equation. Lacking any additional, identifiable structural productivity components or failing to assume additional measurement or model error, adding the labor supply condition to the estimation equations would yield a stochastically singular system. Since the primary purpose of this paper is to assess which channels — relative price effects or wage effects — have driven the rise in structural change, we do not explicitly consider the labor supply condition in our estimation, but instead use our estimates here to understand the implications for relative price changes and wage growth toward intensive margin labor supply.



(a) $\rho \in (0, 1)$ **(b)** $\rho < 0$

Figure 1.18.: Here we present results form the same counterfactual simulations describe above, except we use the dataset without accounting for the durables stock and the associated regression parameter estimates.

To assess how the posterior predicted labor supply is qualitatively related to relative price and wage changes in data, Table 1.6 presents the same regressions as in Table 1.5 except using posterior predicted labor hours $\widetilde{l}_t$ as the left hand side variable. Columns 1 through 3 of Table 1.5 are the regression estimates for the simulations under the assumption $\rho \in (0, 1)$ and columns 4 through 6 present the regression estimates for the simulations under the assumption $\rho < 0$. The predicted correlations are qualitatively the same, featuring the same signs, though their magnitudes are slightly larger than observed in data.

We use these labor supply simulations as a baseline against which to assess simulations from the following counterfactuals: 1) fixing relative prices, allowing wages and productivities to grow; 2) assuming no wage growth, while still allowing productivities to grow and relative prices to change; 3) fixing productivity growth but not wages or relative prices. If $\underset{\sim}{l}_t$ is a counterfactual simulated labor level and $\widetilde{l}_t$ is the baseline predictive level, Figure 1.18 features the posterior predictive mean time series $\underset{\sim}{l}_t / \widetilde{l}_t$ for each of the aforementioned counterfactuals. These simulated deviations from the baseline are presented in Figure 1.18. Depending on whether final home production outputs are substitutes or complements, we observe markedly different counterfactual behavior for the intensive labor supply when relative prices are left alone (blue lines). Recall that the relative price of goods to services has fallen since 1948. Nonetheless, when $\rho \in (0, 1)$ assuming wages and income grow but relative prices stay the same leads to a relative increase in the intensive margin of labor, but when $\rho < 0$, the labor margin is seemingly unaffected by relative price changes. Eliminating real wage growth (red lines) leads to a relatively higher counterfactual willingness to work, demonstrating the strength of the classical income effect on labor supply. In our final counterfactual, we find that fixing relative productivities at their 1948 value has little impact on labor supply. Thus, the model predicts that the long-run decline in average hours worked per-worker can mostly be attributed to wage increases and, if indeed final home produced commodities are substitutes, relative goods to services price declines.

1.4. Conclusion

Whether you are talking about boats, beds, or restaurant meals, households derive utility from a market purchase by allocating time to consume it. This idea is consistent with early home production models. In this paper, we formalize the concept by allowing households to explicitly choose time to spend consuming individual market goods. We use the resulting estimation of the structural model to understand how accounting for home production affects inference as to the causes of the structural transformation of market demand in the U.S. economy. We show that after controlling for both how consumers derive utility from the service flows of durable goods and how consumers spend their off-market time in home production activities, relative price effects appear to dominate the impact of long run wage growth, calling into question some results in the literature suggesting income effects are strongest.

Future work on macroeconomic trends resulting from household preferences should be careful to consider the importance of home production motivations. Further, since households value and derive utility from the entire stock of durable assets they own, economists should consider incorporating the value of consumer durables into consumption series, lest estimation of underlying utility parameters be biased. Finally, we contend that Beckerian home production models, which feature implicit complementarities between time use and market consumption, are behaviorally reasonable. Considering models where off-market time is split into various tasks in order to better understand how household preferences drive various economic phenomena could lead to future results that call into question other long-held conclusions in economics.

2. The Costs and Benefits of Caring: Aggregate Burdens of an Aging Population

joint with Finn Kydland

2.1. Introduction

For the United States and other developed countries, population aging will increase the absolute number of individuals requiring some form of elder care.[1] Microeconomic evidence suggests that caring for infirm older adults requires substantial resources, both in terms of market-traded services and the off-market time of family members. As an example, the Alzheimer's Association estimates that caring for individuals diagnosed with Alzheimer's and dementia is almost triple the cost of caring for non-diagnosed individuals. While approximately 70% of these costs of care are covered by state and federal social insurance programs, Hurd et al. (2013) estimates that the time-value of informal care provided by family members in 2011 amounted to between \$50 billion and \$106 billion. Recent empirical evidence suggests that many working-age adults spend substantial shares of their available time providing informal care for sick and diseased elders with the average adult who cares for another infirm adult spending 5.18 hours per week doing so. Adults who provide informal care work 1.22 hours per week less and enjoy 3.96 hours less leisure time.[2] The National Institutes of Health, the World Health Organization, and others have warned that the costs of providing assisted-living care for older adults could balloon as the population ages, suggesting that aggregate economic outcomes will be adversely affected by this phenomenon.[3] We show that the ballooning number of elderly people requiring living assistance will have a modest impact on aggregate economic growth independent of the substantial impacts imposed by aging itself. We find that reducing incidence of high cost-of-care old-age diseases can improve welfare for both diseased and healthy agents in a general equilibrium environment.

Our findings also confirm previous studies that showed population aging in general has a large, negative impact on aggregate output growth rates. In our baseline calculation, holding constant the risk rate of acquiring a debilitating, welfare-reducing disease, we project average annual U.S. GDP growth to be 2.2% over the period 2016-2056 and 2.1% over the period

[1]We use the terms "elder care," "informal care," and "assisted-living care" synonymously to refer to any kind of assistance received by diseased elderly individuals to perform day-to-day life functions.

[2]All time-use estimates are population weighted averages over the period 2003-2016 taken from the American Time Use Survey, from here on ATUS (Bureau of Labor Statistics 2017).

[3]See U.S. studies on the implications of aging from the National Institute on Aging (2011), National Research Council (2001), and Knickman and Snell (2002). Also, for costs associated with caring for elderly individuals with dementia and Alzheimer's, see Alzheimer's Association (2011), Hurd et al. (2013), and Lepore, Ferrell, and Wiener (2017).

2016-2096. Eliminating the risk of needing long-term assisted-living care marginally increases projected future average annual growth rates for the United States economy over the period 2016-2056 by 10 basis points relative to the baseline. As in Hubbard, Skinner, and Zeldes (1995) and Prescott (2004) social insurance programs in a pay-as-you-go (PAYGO) structure crowd out investment, reducing long-run growth rates relative to a tax-free environment. In the presence of intergenerational transfers of off-market time from young to old, lifetime welfare increases when social insurance tax rates fall as savings and investment increase. Young agents expect to enjoy being cared for by their offspring when old and plan for this spillover effect when choosing savings. This is because endogenous time transfers from young to old of informal care can help offset the adverse welfare implications of incomplete markets for insurance against old-age welfare shocks. Yet, while reducing social insurance taxes may increase expected lifetime utility, a reduction is not necessarily Pareto improving if the working-age share of adults is low. This is because old agents afflicted with a welfare-reducing disease are made worse off as taxes fall and the number of workers is small enough. In various counterfactual simulations we explore the implications of these trends under different population growth rates and different adverse shock probabilities.

The paper proceeds as follows. In Section 2.2 we discuss the population trends and cost estimates associated with the prevalence of high cost-of-care old-age diseases, while also summarizing available data on the allocation of time to care for infirm elders. In Section 2.3 we outline an OLG model that captures the features discussed in Section 2.2. In Section 2.4, we calibrate this model to match observed data points. In Section 2.5 we simulate counterfactuals to understand how population changes affect long-run economic trends. In Section 2.6 we conclude.

2.2. Background & Discussion

The primary motivation for our undertaking is to understand how population aging affects aggregate economic outcomes when members face ex-post idiosyncratic risk to old-age welfare. While the effects of population aging have been discussed in many contexts, few studies have analyzed general equilibrium outcomes when young people save to insure against idiosyncratic risk that directly impacts old-age consumption utility.[4] The closest study that comes to mind is that of Hall and Jones (2007) who model health risk as endogenously affecting survival rates, along with a health status component in utility.[5] To the best of our knowledge nobody has attempted to place idiosyncratic endogenous health risk into a model where young agents provide informal hospice care to ailing loved ones. Our undertaking thus contextualizes diseases like dementia, including Alzheimer's, and other idiosyncratic old-age welfare shocks within an economic framework that features long-term informal, assisted-living care.

There have been no studies, to our knowledge, that estimate the impacts of providing informal care off-market on general equilibrium economic outcomes. This is important because an aging

[4] We are aware of French and Jones (2011), DeNardi, French, and Jones (2010), Edwards (2008), and Palumbo (1999) who look at financial planning decisions within retirement in a partial equilibrium context.

[5] An unpublished study by Azomahou, Diene, and Soete (2009) models health risk as a shock to a health capital stock, as opposed to a direct change in the utility function, which is our approach.

population will likely lead to higher levels of informal care being provided by young people to old people.[6] Several studies have examined how provisions of informal care impact individual labor force participation and earnings (Muurinen 1986; Carmichael, Charles, and Hulme; Leigh 2010; Van Houtven, Coe, and Skira 2013). Informal caregivers who also participate in the formal labor force work on average 3 to 10 hours less per week than their non-caregiving peers (Van Houtven, Coe, and Skira 2013). Providing informal care can thus lead to considerable earnings losses (Muurinen 1986; Van Houtven, Coe, and Skira 2013). Recent work suggests that substitution rates between formal nursing home care and informal in-home care in the United States depend on individual states' complex Medicaid reimbursement structures (Mommaerts 2016, 2017). Indeed, paid long-term care and unpaid in-home care are imperfect substitutes (Mommaerts 2017). We conjecture that this imperfection is due to trade-offs faced by younger family members who willingly provide informal care to elders. Since providing off-market care requires a time investment, younger family members must weigh the altruistic benefits they receive from caring for older loved ones against the loss in lifetime permanent income due to working less.

Until recently there have been few aggregate data available on the rate at which informal elder care is supplied. In 2011 the American Time Use Survey (ATUS) began asking respondents about time spent engaging in informal care for infirm elders (Bureau of Labor Statistics 2017). These data are available for years 2011 thru 2016, but the number of respondents who participated is small ($N = 1066$). From 2003-2016 ATUS asked respondents how much time they spent caring for or helping adults, not just the elderly, who require assistance. Weighted averages of time use for adults age 25-65, where our primary target variable is "adult care", are presented in Table 2.1.[7][8] Conditional on providing informal care, individuals work less and have less leisure time.

At first glance, the time-use data suggest that the impact of increasing disease prevalence on the intensive margin of labor supply is significant in magnitude as the population ages. For illustration, if providing adult care is perfectly substitutable with working, then for every 1000 people over the age of 65, 3.55 jobs for individuals under the age of 65 would cease to exist. Consider now the effects of such a change: working less results in a reduction in permanent income, resulting in a reduction in investment, resulting in a reduction in aggregate output and social insurance tax receipts. However, our results in Section 2.5 show that young individuals adjust their time use in response to market conditions, including the population distribution, mitigating the aggregate impacts of this disease risk. In fact, changes in the population distribution alone appear to affect the labor supply greatest along the extensive margin. In steady state simulations, we show that young workers increase work time as the relative population of working-age adults to retirees falls, but this increase on the intensive margin does not offset the negative impacts on total labor supply due to a falling extensive margin.

[6]For our purposes, "informal care" encompasses all aspects of care which take place off market. "Formal care" will be used to refer to care paid for on the marketplace. These definitions are consistent with those in Hurd et al. (2013) and Lepore, Ferrell, and Wiener (2017).

[7]We take the denominator in our weekly time-share calculations to be 112 hours, thus allowing individuals 8 hours of non-allocatable personal time per day.

[8]We choose to use the total "adult care" data point rather than the "elder care" data point due to the small sample size of the latter. Empirical tests show that the differences in weighted averages of both data points are not significantly different from zero. More details are available upon request.

Table 2.1.: Per-Capita Time Allocation of Adults $25-65$, (ATUS: 2003-2016)

	Leisure	Labor	Adult Care
Whole Population, $N = 82995$			
Avg. Hrs. per Week	69.642	41.743	0.615
Share of Avg. Total Time*	0.622	0.373	0.005
Provide Positive Off-Market Adult Care, $N = 9937$			
Avg. Hrs. per Week	66.150	40.671	5.179
Share of Avg. Total Time*	0.591	0.363	0.046
Provide No Off-Market Adult Care, $N = 73058$			
Avg. Hrs. per Week	70.113	41.887	0
Share of Avg. Total Time*	0.626	0.374	0

* Total time $7 \cdot (24-8)$.

To understand more broadly how long-run declines in aggregate output are related to population aging, in Figure 2.1 we plot the working-age population ratio ($wapr$), i.e. the ratio of adults age 25-65 to adults age 65 and over, along with the HP-filtered trend of year-on-year aggregate and per-worker GDP growth (g_Y and g_{Y/N_y}, respectively) for the United States economy since 1950.[9] When business cycles are removed, the long-run decline in annual GDP growth appears remarkably correlated with the decline in the working-age population ratio. A regression of $\ln g_{Y,t}$ on $\ln wapr_t$ reveals that the elasticity of the filtered trend in output growth with respect to the working-age population ratio is 1.487, so that a 1% relative increase in workers leads to an approximate 14 basis point increase in the growth rate. Falling $wapr$ accounts for almost 57% of the decline in g_Y since the 1950s. The magnitude of this correlation affirms some of the alarm bells sounded recently in Cooley and Henriksen (2018).

As a motivating example, Alzheimer's disease and dementia impose substantial formal costs on the United States' social insurance system and informal costs on family members tasked with caring for diseased individuals. Total cost estimates for caring for demented elderly individuals range from \$157 to \$215 billion (2010 dollars) depending on the method used to impute time value of informal care (Hurd et al. 2013). Within this range, roughly \$11 billion is covered by Medicare, Medicaid, and Social Security, while the remainder includes both out-of-pocket costs

[9]See Hodrick and Prescott (1997) for more information on the HP filter.

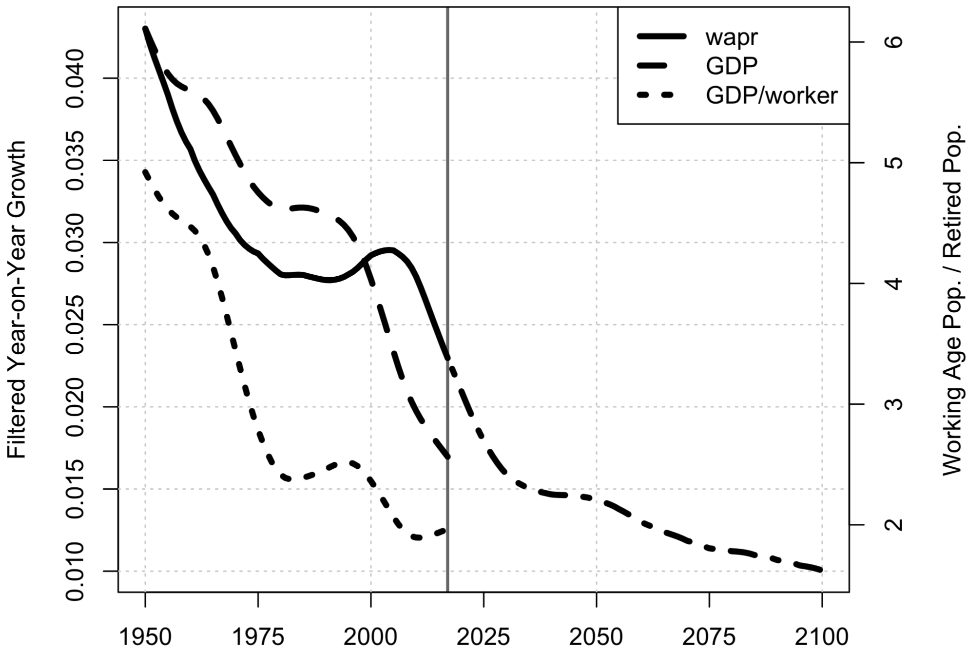


Figure 2.1.: We plot HP-filtered year-on-year net growth: $Y_t/Y_{t-1} - 1$ for the aggregate case or $(Y_t/N_{yt})/(Y_{t-1}/N_{y,t-1}) - 1$ for the per-worker case. $wapr_t$ is unfiltered and downloaded from the United Nations: Department of Economic and Social Affairs (2017). The elasticities of filtered aggregate and per-worker growth with respect to $wapr_t$ are 1.487 and 2.395 respectively. If trends continue, we should expect growth rates to decline throughout the 21st century.

paid by afflicted individuals and their families as well as the time value of unpaid, informal care provided by loved-ones (Hurd et al. 2013). Estimates of total time devoted to informal care for demented persons are not small in magnitude. The Alzheimer's Association estimates that in 2010, 17 billion hours of unpaid care were provided by loved ones to diseased elders, with over 80% of this time burden born by family members. Further, over 90% of those afflicted with Alzheimer's or dementia receive some form of informal care on top of care provided by professional hospice services. The spillover effects on working-age adults of shouldering this burden represent an additional societal cost, the impacts of which have not been directly quantified in past studies. As the population ages and Alzheimer's and dementia prevalences increase, it is reasonable to expect that the quantity of informal care provided by working-age adults to elderly adults will increase.

2.3. Model

Agents live a maximum of two periods, though they may die accidentally after their first period of life.[10] Each period, there exists a population of N_{yt} young households and N_{ot} old

[10]Young households cannot choose to die, nor can individuals, rather the household is thought to "disappear," in that all of its members have perished before becoming old.

households.[11] There is only one type of young household and two types of old households. Old households can be either diseased, $d_t = 1$, or non-diseased, $d_t = 0$. For now, assume the population of young households grows at constant rate g_N so that $N_{yt}/N_{y,t-1} = (1 + g_N)$, and the probability that a young household in period t lives to be an old household in period $t + 1$ is $s_{o,t+1}$.[12] [13]

We model the old agents' consumption process in terms of home production, taking cues from Becker (1965). Young agents can subsidize the home production of diseased old agents' final consumption by supplying them care time h_{yt} outside of formal markets. Diseased individuals thus receive flow utility from final consumption $c_{ot}(d_t = 1)$, which is produced in the home by using inputs of this off-market care time they receive from their children h_{ot} and market resources purchased $x_{ot}(d_t = 1)$.[14] Meanwhile, their healthy peers only use market resources $x_{ot}(d_t = 0)$ for production of final consumption because they do not require additional off-market care time from their children. The home production functions we employ for both diseased and non-diseased old are

$$c_{ot}\big(x_{ot}(d_t = 1), h_{ot}, d_t = 1\big) = x_{ot}(1)^{1-\sigma} h_{ot}^{\sigma} \quad \sigma \in (0, 1) \tag{2.1}$$

$$c_{ot}\big(x_{ot}(d_t = 0), d_t = 0\big) = x_{ot}(0) \tag{2.2}$$

In Equation (2.1), σ is the elasticity of final consumption with respect to informal care time received. Note that both diseased and non-diseased individuals may purchase hospice care or other health services on the formal market. Such a purchase would fall under market consumption, $x_{ot}(d_t)$. Any additional services received by diseased old that are not accounted for on the formal market would fall under informal care-time received, h_{ot}.

Old households have preferences over consumption that depend on health status d_t. The form of an old individuals' utility function is chosen to satisfy several conditions. First, we assume that individuals infected with a disease require more resources, both market and off-market, to care for than those who are not. It would be unreasonable to assume that these individuals, by consuming more, are necessarily better off than their non-diseased peers (after all, they are sick). Let $u_{ot}\big(c_{ot}(d_t), d_t\big)$ denote the flow utility from final consumption for an old individual with disease status d_t. This brings us to an assumption about an old individual's utility function.

Assumption 1. Suppose $c_{ot}(1) = c_{ot}(0) = c$, where c is any feasible level of final consumption. Then $u_{ot}(c, 1) < u_{ot}(c, 0)$. In words, for each level of final consumption, the non-diseased agent always receives higher consumption utility than the diseased agent.

Assumption 1 ensures that it is always better to be non-diseased than diseased. We choose a Stone-Geary flow utility function for diseased old which satisfies this assumption under certain parameter restrictions:

$$u_{ot}\big(c_{ot}(1), d_t = 1\big) = \ln\big(c_{ot}(1) - \nu\big) \tag{2.3}$$

[11] Since agents live only two periods, we use y and o subscripts to denote their ages.

[12] We will relax the constant growth, g_N, assumption in some of our simulations.

[13] The "survival" probability is the probability a young household that enters the economy survives to be an old household next period.

[14] We also refer to h_{ot} as "hospice" care.

$$u_{ot}\big(c_{ot}(0), d_t = 0\big) = \ln c_{ot}(0) \tag{2.4}$$

The flow utility parameterizations in (2.3) and (2.4) lead to two basic lemmas.

Lemma 1. For all $\nu > 0$, Assumption 1 holds.

Proof. See Appendix C.1.1 □

Lemma 2. If $0 < \nu < c_{ot}(1) - c_{ot}(0)$ then the ratio of the marginal utility of non-diseased consumption to diseased consumption is such that $MU_{ot}(0)/MU_{ot}(1) > 1$.

Proof. See Appendix C.1.1 □

Lemma 1 is trivial. Lemma 2 says that for certain combinations of the subsistence parameter ν and consumption policies, non-diseased agents benefit more from additional final consumption than diseased agents. In our calibration we find that the premise of Lemma 2 holds, which is one indicator that in this economy resources are inefficiently allocated in a steady state competitive equilibrium.

Young households use market resources x_{yt} and leisure time l_{yt} to produce a final consumption good c_{yt} according to the home production function:

$$c_{yt} = x_{yt}^{\gamma} \cdot l_{yt}^{\mu} \tag{2.5}$$

Young households additionally supply off-market time h_{yt} to care for their elders, but since this does not affect the final production of their home-produced consumption good, h_{yt} does not enter into Equation (2.5). Rather, young households exhibit imperfect altruism toward their sick elders, discounting the diseased old household's utility at rate η. The flow utility of young households is

$$u_{yt}(c_{yt}, h_{yt}) = \ln c_{yt} + \eta \cdot \ln\big(c_{ot}(h_{yt}, d_t = 1) - \nu\big) \tag{2.6}$$

In addition to consuming and spending time caring for their parents, young agents supply both labor $1 - l_{yt} - h_{yt}$ and invest i_{yt} in the market.[15] Young agents earn a before-tax wage rate w_t and pay social insurance taxes on their income at rate τ_t.

Old agents do not work but earn a gross return on their assets a_{yt} at rate R_t and also receive Social Security and Medicare transfer benefits from the government $T_t(d_t)$ which depend on disease status d_t. Normalize the price of market purchases to 1 each period. For old agents, net outlay must satisfy the budget constraint:

$$x_{ot}(d_t) \leq R_t \cdot a_{yt} + T_t(d_t) \tag{2.7}$$

Old agents die with certainty at the end of their life and will choose to consume the entirety of their available cash-on-hand. At the end of each period, young agents who accidentally and unexpectedly die leave behind total net assets (capital) equivalent to $a_{y,t+1} \cdot (1 - s_{o,t+1}) \cdot N_{yt}$.

[15] Total available time is normalized to 1.

These assets are then distributed evenly and unexpectedly (i.e., "accidentally") as bequests $b_{y,t+1}$ to young agents entering the economy next period according to

$$b_{y,t+1} = a_{y,t+1} \cdot (1 - s_{o,t+1}) \cdot \frac{N_{yt}}{N_{y,t+1}} \tag{2.8}$$

Since returns on investment are compounded at the beginning of the period, young agents earn gross return on these assets $R_t \cdot b_{yt}$. Having fully-described the right-hand side of a young agent's budget constraint, their choices of market purchases x_{yt}, asset-holdings, and labor-supply must satisfy

$$x_{yt} + a_{y,t+1} \leq R_t \cdot b_{yt} + w_t \cdot (1 - \tau_t) \cdot (1 - l_{yt} - h_{yt}) \tag{2.9}$$

Let ψ_t denote the share of old population which is afflicted with a welfare-reducing disease in period t. The supply of hospice care by young equals the total amount of hospice care received by diseased old

$$h_{yt} = \frac{N_{ot} \cdot \psi_t \cdot h_{ot}}{N_{yt}} \tag{2.10}$$

Let $\rho_t = T_t(1)/T_t(0)$ be the ratio of diseased to non-diseased benefits. The government balances Social Security and Medicare transfers and tax receipts:

$$N_{ot} \cdot \big(\psi_t \cdot \rho_t \cdot T_t(0) + (1 - \psi_t) T_t(0)\big) \leq N_{yt} \cdot w_t \cdot \tau_t \cdot (1 - l_{yt} - h_{yt}) \tag{2.11}$$

Young agents do not know whether they will survive to become old and if they do whether they will face a disease that adversely impacts their welfare, but young agents know ψ_t and how it evolves, just as they know the survival rate. Thus, they make their investment choice both with the aim of smoothing consumption and imperfectly insuring themselves against the adverse welfare effects of contracting some kind of disease such as Alzheimer's or dementia, for example. In this model, given young agents in period t know the distribution of diseased agents in $t + 1$, expectations are perfectly rational. Let β be the time discount factor. In competitive equilibrium, utility maximizing young agents seek to smooth expected market consumption over the lifecycle according to the expected intertemporal Euler equation:

$$\underbrace{\frac{\gamma}{x_{yt}}}_{MU_y(x)} = \beta \cdot s_{o,t+1} \cdot R_{t+1} \left[\psi_{t+1} \underbrace{\frac{1-\sigma}{x_{o,t+1}(1)^{1-\sigma} h_{o,t+1}^{\sigma} - \nu} \left(\frac{h_{o,t+1}}{x_{o,t+1}(1)} \right)^{\sigma}}_{MU_o(x(1))} + (1 - \psi_{t+1}) \underbrace{\frac{1}{x_{o,t+1}(0)}}_{MU_o(x(0))} \right] \tag{2.12}$$

Since the model contains only idiosyncratic uncertainty, R_{t+1} is pre-determined by the population distribution which is assumed known. Thus young agents choose investment by equating the marginal utility of present market purchases with the discounted expected marginal utility of future consumption given by weighting diseased and non-diseased marginal utilities by the distribution ψ_{t+1}. The period t choice of labor supply by young depends on the marginal rate of substitution between consumption and leisure and the marginal rate of substitution between leisure and off-market care time:

$$\frac{\mu}{l_{yt}} = \frac{\gamma}{x_{yt}} w_t(1-\tau_t) \tag{2.13}$$

$$\frac{\mu}{l_{yt}} = \eta \cdot \frac{\sigma}{x_{ot}(1)^{1-\sigma} h_{ot}^{\sigma} - \nu} \left(\frac{h_{ot}}{x_{ot}(1)}\right)^{\sigma-1} \frac{N_{yt}}{N_{ot}\psi_t} \tag{2.14}$$

A single firm produces both an investment good I_t and a market good X_t. We assume period t production is Cobb-Douglas $F_t(K_t, L_t) = z_t K_t^{\alpha} L_t^{1-\alpha}$ where K_t and L_t are aggregate capital and labor respectively. Let δ be the net rate of capital depreciation. The rate of return on assets R_t and before-tax wages w_t are determined by the marginal products of capital and labor respectively:

$$R_t = 1 + z_t \cdot \alpha \cdot \left(\frac{L_t}{K_t}\right)^{1-\alpha} - \delta \tag{2.15}$$

$$w_t = z_t \cdot (1-\alpha) \cdot \left(\frac{K_t}{L_t}\right)^{\alpha} \tag{2.16}$$

Each period aggregate capital in the economy is only affected by young agents' choice of investment from last period since surviving old agents consume their entire portfolio.

$$K_t = N_{y,t-1} a_{yt} = \frac{N_{yt}}{1+g_N} a_{yt} \tag{2.17}$$

$$L_t = N_{yt}(1 - l_{yt} - h_{yt}) \tag{2.18}$$

Finally, total factor productivity z_t grows at constant net rate g_z.

2.3.1. Competitive Equilibrium with Transfers

Given g_N, g_z, an exogenously specified sequence of probabilities for disease contraction and survival from young to old, $\{\psi_t, s_{o,t+1}\}_{t\geq 0}$, an initial population level $N_{y0} = 1$, an initial total factor productivity level $z_0 = (1+g_N)^{\alpha}$, and an initial young asset level a_{y0}, a competitive equilibrium with transfers consists of:

i. *Sequences of policies for consumers:* $\Big\{x_{yt}, l_{yt}, h_{yt}, a_{y,t+1}, x_{ot}(1), x_{ot}(0)\Big\}_{t\geq 0}$.

ii. *Sequences of prices* $\{R_t, w_t\}_{t\geq 0}$.

iii. *Government policies* $\{T_t(1), T_t(0), \tau_t\}_{t\geq 0}$.

such that

a. *Young agents' choices satisfy (2.9) and (2.12) thru (2.14).*

b. *Old agent consumption policies satisfy (2.7).*

c. *Asset return rates and wage rates are (2.15) and (2.16).*

d. *Formal and informal markets clear.*

e. The government's budget is balanced.

The beauty of the two period assumption is that it allows us to consider only how changes in the working age to retiree population ratio $wapr_t$ affects equilibrium outcomes. This is illustrated in Proposition 1. As a corollary to Proposition 1, we also demonstrate that in this environment, aggregate output growth g_{Yt} depends only on $wapr_t$, not generational population levels.

Proposition 1. A competitive equilibrium depends only on $wapr_t$, not the population levels.

Proof. Using (2.17) and the normalization $z_0 = (1 + g_N)^\alpha$, we can write the aggregate resource constraint

$$\begin{aligned} &N_{yt} \cdot x_{yt} + N_{ot}\big(\psi_t \cdot x_{ot}(1) + (1 - \psi_t) \cdot x_{ot}(0)\big) + N_{yt} \cdot a_{y,t+1} \\ &\quad \leq (1 - \delta)\frac{N_{yt}}{1 + g_N} a_{yt} + (1 + g_z)^t \cdot N_{yt} \cdot a_{yt}^\alpha (1 - l_{yt} + h_{yt})^{1-\alpha} \end{aligned} \tag{2.19}$$

Dividing both sides by N_{ot}, we can write

$$\begin{aligned} &wapr_t \cdot x_{yt} + \psi_t \cdot x_{ot}(1) + (1 - \psi_t) \cdot x_{ot}(0) + wapr_t \cdot a_{y,t+1} \\ &\quad \leq (1 - \delta)\frac{wapr_t}{1 + g_N} a_{yt} + (1 + g_z)^t \cdot wapr_t \cdot a_{yt}^\alpha \cdot (1 - l_{yt} + h_{yt})^{1-\alpha} \end{aligned} \tag{2.20}$$

Note that consumption decisions depend on R_t and w_t. Using (2.17) and (2.18), R_t and w_t do not depend on population levels except through household policies. In (2.8), under constant g_N, $b_{y,t+1}$ can be written

$$b_{y,t+1} = a_{y,t+1} \cdot (1 - s_o) \cdot \frac{1}{1 + g_N} \tag{2.21}$$

In (2.10) h_{yt} can be written

$$h_{yt} = \frac{\psi_t}{wapr_t} h_{ot} \tag{2.22}$$

The government budget constraint in (2.23) can be written

$$\big(\psi_t \cdot \rho_t \cdot T_t(0) + (1 - \psi_t) T_t(0)\big) \leq wapr_t \cdot w_t \cdot \tau_t \cdot (1 - l_{yt} - h_{yt}) \tag{2.23}$$

Finally, young household policies must satisfy (2.12) thru (2.14). Population levels only enter (2.14), which can be written

$$\frac{\mu}{l_{yt}} = \eta \cdot \frac{\sigma}{x_{ot}(1)^{1-\sigma} h_{ot}^\sigma - \nu} \left(\frac{h_{ot}}{x_{ot}(1)}\right)^{\sigma-1} \frac{wapr_t}{\psi_t} \tag{2.24}$$

□

Corollary 1. Aggregate growth $g_{Y,t}$ depends only on $wapr_t$, not population levels.

Proof. See Appendix C.1.1 □

2.3.2. Steady State and Balanced Growth

To solve for a steady state, we assume a balanced growth path (BGP) and de-trend productivity growth as in Krueger and Ludwig (2007). For now, suppose $\tau_t = \tau$ is exogenously fixed. Along a BGP the population of young agents and productivity grow at constant exogenous rates, g_N and g_z. Proposition 2 demonstrates that if survival rates and disease risk are constant, then $wapr$ is constant across time.

Proposition 2. Assume the survival rate $s_{o,t+1} = s_o$ is constant and the diseased old distribution $\psi_t = \psi$ is stationary. Then along a BGP the working-age population ratio $wapr$ is constant and given by

$$wapr = \frac{1 + g_N}{s_o} \tag{2.25}$$

Proof. See Appendix C.1.1 □

Clearly, the assumption that $wapr$ is constant is unrealistic in practice, as we see that $wapr$ has been falling over time and is projected to continue falling. This fact begs the question as to whether the U.S. economy in the 21st century is in fact on a balanced growth trajectory or rather is exhibiting structural change due to forces such as population aging and potentially associated idiosyncratic welfare risk affecting long-run growth rates. In Section 2.5 we simulate the future path of aggregate output growth to understand the extent to which falling $wapr_t$, coupled with idiosyncratic welfare risk and young agents' altruism, together impact aggregate growth.

2.4. Calibration

For our calibration we set the period length to 40 years and assume young agents enter the economy at age 25 and turn old at age 65. Our calibration assumes the economy is in steady state in 2016, thus taking the 2016 observed population distribution as the initial steady state distribution.[16] We choose parameters to match a set of carefully selected data moments from around 2016. Specifically, we calibrate to leisure, labor, and hospice care average time shares from the 2003-2016 ATUS data, the personal savings rate from the "PSAVERT" time series (BEA, 2016) which measures personal savings as a percentage of disposable income, the ratio of diseased to non-diseased consumption computed from estimates made by Hurd et al. (2013), the 2016 consumption and investment shares of output (BEA, 2016), and the 2015 U.S. labor and capital income shares from Penn World Table 9.0. Table 2.2 presents the data moment targets and their simulated model counterparts, while Table 3.2 presents the calibrated parameter values and their sources.

The calibration requires a couple of assumptions for identification purposes. First, we exogenously set the benefits ratio ρ using estimates by Hurd et al. (2013) and Mommaerts (2016) to get $\rho = 1.923$.[17] Thus, diseased agents receive almost double the benefits from the

[16] Appendix B.2.0.2 presents the steady state equations.
[17] The procedure used to set this parameter is described in detail in Appendix B.2.

government as non-diseased agents. To calibrate to a steady state assuming it has been detrended from a BGP, we only have to pick two of g_N, s_o, and $wapr$, due to Equation (2.25). We set g_N to accommodate growth in the young population since 1976 and $wapr$ to equal the observed population ratio for workers to retirees in 2016. We exogenously fix the parameters μ, γ, and η to reflect the observed ATUS time-use averages from 2003-2016. The output elasticity α is chosen to match the average U.S. capital share since World War II. The risk rate ψ is fixed based on estimates from Hurd et al. (2013). We calibrate the subsistence parameter ν and intensity of hospice care parameter σ to match aggregate data moments, including the ratio of diseased to non-diseased consumption $x_o(1)/x_o(0)$ taken from estimates in Hurd et al. (2013).

Table 2.2.: Calibration Targets

Moment	Model	Data	Source
l_y^*	0.617	0.622	ATUS, 2003-2016 Avg.
$1 - l_y^* - h_y^*$	0.375	0.373	ATUS, 2003-2016 Avg.
h_y^*	0.008	0.005	ATUS, 2003-2016 Avg.
Savings Rate	0.137	0.049	BEA, 2016
$x_o^*(1)/x_o^*(0)$	1.315	1.360	Hurd et al. 2013
X^*/Y^*	0.812	0.680	Consumption Share, 2016
I^*/Y^*	0.188	0.320	Investment Share, 2016
Labor Income Share	0.645	0.600	Penn World Table 9.0, 2015
Capital Income Share	0.355	0.400	Penn World Table 9.0, 2015

Table 2.3.: Calibrated Parameter Values

	Value	Source
g_N	0.660	Growth in Age 25-65 Pop. (1976-2016)
δ	0.9517	40 years of 6% annual depreciation
α	0.35	Post-war avg. capital share (DeJong and Dave 2011)
β	0.4457	Annual discounting of 0.98 over 40 years
γ	0.373	ATUS (2003-2016) avg. work
μ	0.622	ATUS (2003-2016) avg. leisure
η	0.005	ATUS (2003-2016) avg. adult care
$wapr$	3.475	U.S. Working-age pop. ratio 2016 (UN)
τ	0.153	S.S. + Medicare tax rate 2016
ρ	1.923	Ratio of diseased/non-diseased benefits (see Appendix B.2.0.1)
ψ	0.14	Risk of contracting dementia (see Hurd et al. 2013)
ν	0.099	Subsistence of old (calibrated to match data)
σ	0.659	Intensity of h_{ot} in home production (calibrated to match data)

2.5. Findings

2.5.1. Predicted U.S. Aggregate Output Growth

We want to understand how well the model predicts different aggregate growth rates when the population is evolving in ways inconsistent with balanced growth. We compute a transition path from starting steady states with $wapr$ equivalent to those observed in 1950, 1960, 1970, 1980, and 1990, simulating toward a terminal steady state with $wapr = 3.475$ as observed in 2016. We assume period 0 of the model is in one of the old $wapr$, and then the economy suddenly changes to $wapr = 3.475$, allowing 200 simulated periods to facilitate convergence to the new steady state.[18] For each simulation, we set the initial steady state's τ to the actual employee and employer combined Social Security and Medicare tax rate for the given year.[19] Productivity growth is set to accommodate the observed average annual private multi-factor (MFP) productivity growth rates for the periods 1950-2016, 1960-2016, 1970-2016, 1980-2016, and 1990-2016.[20] We compare both predicted productivity re-trended aggregate output growth and per-worker re-trended output growth from the first period after the sudden change in working age population ratio to that observed in the data. These values are presented in Table 2.4 under columns labeled "Model g_Y" and "Model g_{Y/N_y}," where the former describes aggregate growth and the latter growth per working-age adult. Model predictions slightly undershoot aggregate growth rates for all periods, though the aggregate rates are only off by 20 basis points at most.

We run two additional simulations adjusting g_z to match observed g_Y and g_{Y/N_y}. Implied productivity growth from these simulations is presented in columns labeled "Implied g_z" and "Implied g_{z/N_y}" in Table 2.4. In the data, g_{z/N_y} is negative since the 1960s when the denominator we use to compute output per-worker is the entire working-age adult population. To reconcile per-worker growth, our model requires growth in productivity per-worker to exceed growth in aggregate productivity as can be seen by comparing the "Implied" column of the bottom half of Table 2.4 to the top half. The model thus appears to do a decent job of matching aggregate output growth but not output per-worker. This is due to the fact that we assume a 40-year transition period regardless of the starting $wapr$ being associated with the year 1950 or 1990. Comparisons between aggregate numbers are not biased by this fact because the aggregate growth rate does not depend on a scaling with the growth rate of newborns entering the economy, g_N, which must be computed to accommodate the transition from the initial steady state $wapr$ to the terminal one. If we allow for the possibility that perhaps measurements of MFP in the NIPA tables themselves are biased, failing to account for endogeneity due to g_z's dependence on the population distribution, then U.S. productivity growth over the second half of the twentieth century has perhaps been overstated, or at the least misunderstood. At first consideration, it is hard to ignore the positive correlation between measured productivity per-worker and $wapr$. One can think of a number of possible ways in which $wapr$ may affect productivity:

[18] For a thorough explanation of how to accomplish this simulation technique in an overlapping generations model see the endogenous grid point method of Carroll (2006) and Appendix B of Krueger and Ludwig (2007).

[19] The tax rates are as follows: 3% (1950), 6% (1960), 9.6% (1970), 12.26% (1980), 15.3% (1990 and thereafter).

[20] MFP is real value-added output divided by combined inputs — *U.S. Bureau of Labor Statistics, Private Business Sector: Multi-factor Productivity [MFPPBS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/MFPPBS, Accessed: October 23, 2018.*

Table 2.4.: Model Performance: Predicted Avg. Annual Growth to 2016

		Growth in Aggregate Output, (%)			
Data Period	Starting $wapr$	Model g_Y	Data g_Y	Implied g_z	Data g_z
1950-2016	6.111	3.121	3.096	1.177	1.197
1960-2016	5.114	2.924	3.008	1.182	1.107
1970-2016	4.414	2.497	2.742	1.100	0.901
1980-2016	4.075	2.203	2.630	1.268	0.876
1990-2016	4.028	1.963	2.377	1.240	0.906
		Growth in Output Per Working Age Adult, (%)			
Data Period	Starting $wapr$	Model g_{Y/N_y}	Data g_{Y/N_y}	Implied g_{z/N_y}	Data g_{z/N_y}
1950-2016	6.111	1.212	1.952	1.985	0.074
1960-2016	5.114	1.245	1.827	1.661	-0.053
1970-2016	4.414	0.996	1.460	1.404	-0.359
1980-2016	4.075	1.081	1.413	1.194	-0.320
1990-2016	4.028	1.276	1.356	1.030	-0.100

younger workers have more energy and work more in order to build up a nest egg from scratch, for example. In the context of our formulation, a relatively large population of retirees could negatively weigh on aggregate productivity by diverting working-age adults' attention from their jobs because they provide informal care. If this explanation were true, z_t would be an endogenous function of $wapr_t$, and $g_{z,t}$ would vary in time, falling along with $wapr_t$. We do not take a stance on the mechanism by which z_t may partially depend endogenously on $wapr_t$. Rather, the decomposition in Table 2.4 illustrates what the $wapr_t$-independent component of $g_{z,t}$ would need to be in order to match observed output growth under our parameterization. In general, the results of these simulations show that the given model can accurately predict aggregate growth, suggesting researchers should take our future growth estimates presented in Section 2.5.2 seriously, affirming the general spirit of the results in Cooley and Henriksen (2018).

2.5.2. Future Growth Under Different Counterfactual Regimes

One goal of this project is to understand how the welfare risk of contracting a debilitating old-age disease may affect future aggregate output growth while the population is aging. As a baseline, we follow techniques described in Krueger and Ludwig (2007) to simulate a transition path between the calibrated 2016 steady state and a far-off future steady state assuming the population converges after 200 periods to the United Nations predicted, 2096 median-variant population distribution.[21] We then examine projected growth rates and lifetime welfare under the following policy reforms. First, we consider how the economy evolves when the "dynamically ignorant" government suddenly sets $\tau = 0$ one period into the future

[21]See United Nations: Department of Economic and Social Affairs 2017.

and households are surprised by this change, failing to anticipate it.[22] Second, we consider a policy reform where the government decides to fully reimburse working-age adults for their off-market time at the before-tax market wage.[23] Finally, we simulate a dynamic transition path under unexpected changes to the disease risk rate. We consider growth under a cure by 2056 and a cure by 2096, as well as growth under 10%, 20%, 50%, and 100% increases in cross-sectional risk by 2096.[24] These changes are all based on the value of $\psi = 0.14$ used in calibration, taken from estimates of dementia risk for 70 year olds in Hurd et al. (2013). For computational reasons, we assume changes are permanent after 2096 so the economy has some steady state outcome to which to converge.

Table 2.5 presents simulated average annual aggregate output growth rates (g_Y). For the baseline simulations holding τ and ρ at their observed and calibrated 2016 values, we compare the dynamic transition path of an economy aging according to U.N. projections. Our regime-change counterfactuals operate as follows. First, we suppose that the economy is in the initial 2016 steady state, then the regime change occurs suddenly. For all of these changes, in period $t = 2$ right after the 2016 steady state, the economy has changed unexpectedly, but agents have not updated their dynamic plans. We thus simulate the economy for 200 periods to allow for it to converge to the new steady state, which generally happens after only 7-10 model periods anyway. All of our counterfactual simulations occur off a BGP, where the population distribution is evolving exogenously according to U.N. estimates.

From these simulations it is apparent that an aging population substantially reduces growth relative to a BGP where $wapr$ remains constant. In fact this reduction is on the order of 50 basis points annually, leading to compounded aggregate output losses of 17% by 2056 and 39% by 2096 relative to an economy where $wapr$ held constant at 2016 levels. Though perhaps politically unrealistic, it is illustrative that in this economy setting $\tau = 0$ can lead to both Pareto improvements along the dynamic transition path and increases in compounded aggregate output relative to the baseline with population transition — 5% higher by 2056 and 17% higher by 2096. Figure 2.2 presents the fully compounded predicted population baseline growth relative to counterfactual growth projections, including the BGP. Implementing a before-tax reimbursement policy while holding $\tau = 0.153$ fixed yields Pareto improvements but adversely affects compounded growth relative to the population transition baseline in the first period — output is 0.2% lower by 2056 — though growth improves slightly by the end of the century — 1.7% higher by 2096. The predicted baseline falls the most relative to BGP, then the tax-free environment, then finally the full cure. Neither stabilizing the population distribution to achieve

[22] Here, we consider the 2016 $wapr = 3.475$ as "present."

[23] Currently, some U.S. state Medicaid programs reimburse family members for care time they provide to Medicaid recipients, though the rates of reimbursement and restrictions vary substantially across states. Current data on state-level Medicaid policies does not appear to be readily available in a central source.

Under this reform, the young agent's budget constraint is:

$$x_{yt} + a_{y,t+1} \leq R_t \cdot b_{yt} + w_t(1-\tau_t)(1-l_{yt}-h_{yt}) + w_t \cdot h_{yt} \tag{2.26}$$

while the government faces budget constraint:

$$N_{ot} \cdot \big(\psi_t \cdot \rho_t \cdot T_t(0) + (1-\psi_t)T_t(0)\big) \leq N_{yt} \cdot \big(w_t \cdot \tau_t \cdot (1-l_{yt}-h_{yt}) - w_t \cdot h_{yt}\big) \tag{2.27}$$

[24] For a "cure" we consider a situation where ψ drops to 0.0001 to ensure Inada conditions hold.

Table 2.5.: Growth Under Different Regimes, $g_z = 1.4\%$

Model	Pop. Transition?	Predicted Avg. Annual Growth g_Y, (%) 2016-2056	2016-2096	2016-2136
BGP ($\tau = 0.153$, $\psi = 0.14$)*	No	2.693	2.693	2.693
Baseline	Yes	2.228	2.052	1.890
$\tau = 0$	Yes	2.364	2.298	2.096
Reimbursement of h_{yt} at w_t	Yes	2.224	2.074	1.912
$\psi_{2056} = 0.0001$, $\psi_{2096} = 0.0001$	Yes	2.277	2.123	1.947
$\psi_{2056} = 0.07$, $\psi_{2096} = 0.0001$	Yes	2.248	2.096	1.941
$\psi_{2056} = 0.146$, $\psi_{2096} = 0.154$	Yes	2.226	2.047	1.884
$\psi_{2056} = 0.154$, $\psi_{2096} = 0.168$	Yes	2.223	2.042	1.876
$\psi_{2056} = 0.175$, $\psi_{2096} = 0.210$	Yes	2.216	2.026	1.858
$\psi_{2056} = 0.21$, $\psi_{2096} = 0.280$	Yes	2.203	1.997	1.822
End-of-period *wapr* used in simulations:		2.110	1.658	1.658

* $\tau = 0.153$ and/or $\psi = 0.14$ unless otherwise noted.

a BGP nor eliminating Social Security and Medicare are realistically feasible, yet scientists are working to find cures for dementia-like diseases. In the event a cure is found, the model predicts compounded U.S. total output will be 5.4% higher by 2096 and 6.6% higher by 2136 relative to the baseline where $\psi = 0.14$.

One takeaway we wish to emphasize is that achieving a full cure — $\psi = 0.0001$ by 2056 — would have a small impact on growth, increasing g_Y by only about 10 basis points. Long-run growth rates and welfare are decreasing in ψ. The most striking thing about our simulations under different ψ is that changing the risk rate hardly matters for long-run growth prospects. Rather, the population distribution, regardless of the risk rate, has the largest effect on long-run growth, which can be seen by comparing any of the simulations that account for population transitions with the BGP. All counterfactuals result in anywhere from 60 to 90 basis point relative declines in the average annual growth rate by 2136, and 40 to 70 basis point relative declines by 2096. While this result should mitigate concerns that the burdens of old-age care alone will tamp down growth, we confirm recent findings in Cooley and Henriksen (2018) suggesting a long-run "demographic deficit" may be coming to the United States economy.

Yet reducing disease risk, despite having minimal impact on growth, is still welfare improving. Figure 2.3 compares welfare paths for different possible risk rates, relative to a baseline economy where $\psi = 0.14$, as estimated by Hurd et al. (2013). Notice that welfare improves for all agents as $\psi \to 0$, though risk reduction has the most pronounced effect on diseased agents' welfare, leading to a greater than 6% lifetime gain relative to the baseline. Next, young agents enjoy higher expected lifetime utility, but are also hit hardest relative to the baseline when ψ increases. This is because children shoulder the burden of increased numbers of diseased old through

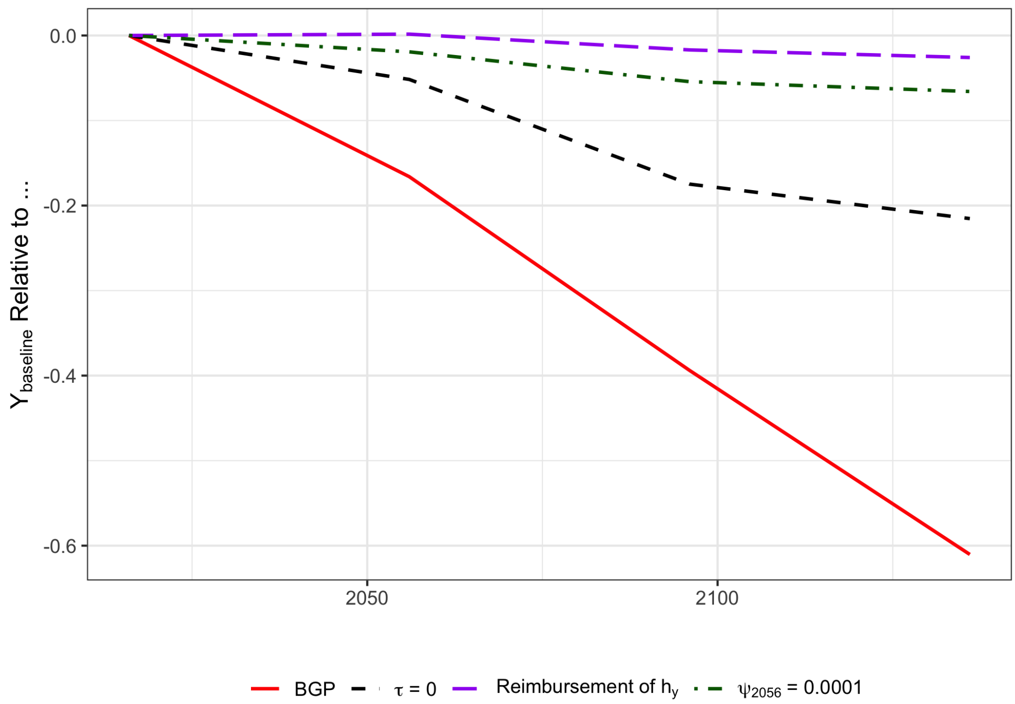


Figure 2.2.: Here we present predicted baseline output relative to various counterfactuals, $(Y_{\text{baseline}} - Y_{\text{counter}})/Y_{\text{counter}}$. A cure for dementia by 2056 ($\psi = 0.0001$; green dashed line) can lead to modest improvements relative to the baseline.

the altruism mechanism. Meanwhile, Figure 2.4 compares welfare paths over the 21st century relative to the population transition baseline for the tax-free environment and an economy with a reimbursement scheme. Lifetime welfare of all agents over the 21st century is improved from baseline under the reimbursement scheme, though again the most notable improvement is for diseased agents. This finding is particularly promising since growth is relatively unaffected by such a scheme, yet all agents are better off. Further, reimbursement schemes are already implemented in certain states. Our results suggest that more adoption of these policies will lead to welfare improvements across the board.

2.5.3. Steady State Comparative Statics

Using the calibrated parameters from Table 3.2 we simulate steady state outcomes under different working-age population ratios and compare them to an economy without a social insurance system. We present selected policies and aggregate outcomes in Figure 2.5. The model predicts that both labor supply and total time spent providing informal care is higher when the social safety net is eliminated. Young people sacrifice leisure time for work because wages are higher and pick up the slack caring for their elders at all levels of $wapr$. All values are monotonic in $wapr$, though the signs of some of the relationships may be surprising. Not surprisingly, labor and output are increasing in $wapr$, but work hours are increasing because wages are decreasing: holding productivity fixed, wages are bid down as the number of workers increases. Steady state savings rates increase in $wapr$ as a response to higher interest rates, driven up by increases in the labor supply forcing firms to acquire more capital to efficiently

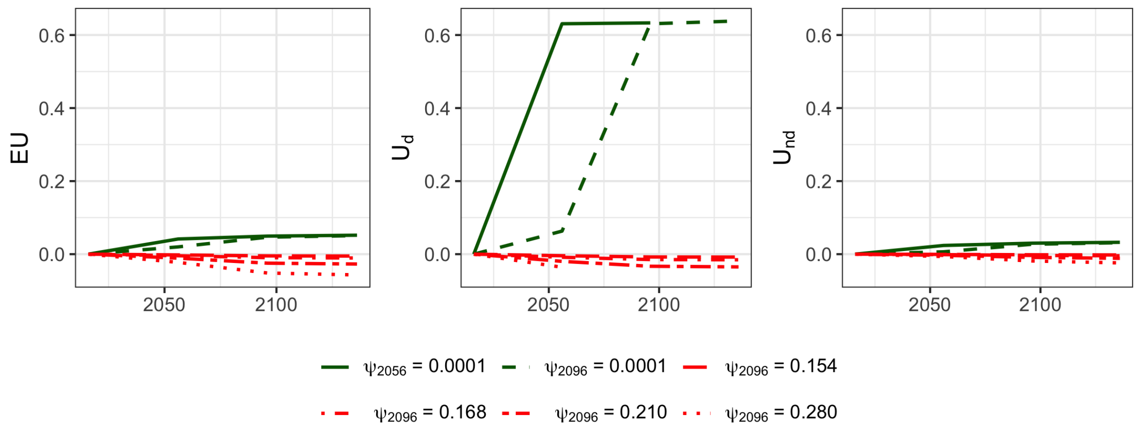


Figure 2.3.: We present welfare as share of the predicted baseline with population transition. Full cures ($\psi \to 0$; black lines) generate Pareto improvements for all types of agents. As ψ increases (red lines), welfare falls relative to the predicted baseline with $\psi = 0.14$.

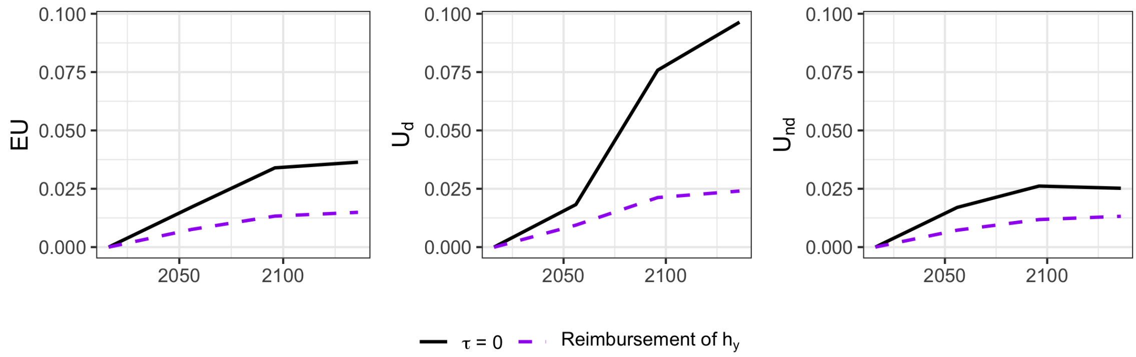


Figure 2.4.: Again, welfare is presented as share of predicted baseline with population transition. Pareto improvements are generated when Social Security and Medicare taxes are unexpectedly eliminated. Reimbursing young agents' time supplying care on the informal market h_{yt} at the market rate w_t also yields Pareto improvements.

utilize the skills of increasing numbers of workers. U.S. personal savings rates have generally fallen since the 1950's, from 11.3% in January 1959 when $wapr$ was at 5.176 to 6.3% in December 2016 with a $wapr$ of 3.475, confirming the validity of the sign of the relationship we observe here.

For each of the $\tau = 0$ and $\tau = 0.153$ case, we simulate expected lifetime utility for a young agent who has not yet realized his old-age disease status as well as realized lifetime utility for both diseased $d = 1$ and non-diseased $d = 0$ old agents.[25] Figure 2.6 presents these welfare values as functions of $wapr$. This exercise demonstrates that for small enough $wapr$, higher social insurance taxes can lead to higher steady state lifetime welfare for diseased agents, though

[25]Let $\mathbb{E}u(d)$ be expected lifetime steady state utility, $u(d = 1)$ be realized lifetime utility for a diseased agent, and $u(d = 0)$ be realized lifetime utility for a non-diseased agent. These values are as follows:

$$\mathbb{E}u(d) = u_y + \beta \cdot [\psi \cdot u_o(1) + (1 - \psi) \cdot u_o(0)] \tag{2.28}$$

$$u(d = 1) = u_y + u_o(1) \tag{2.29}$$

$$u(d = 0) = u_y + u_o(0) \tag{2.30}$$

not non-diseased agents. In steady state, reducing taxes from the 2016 value of $\tau = 0.153$ is Pareto improving as long as $wapr > 2.434$. Why is this? Consider informal care time supplied by young h_{yt}. Figure 2.5 shows that time supplied per-individual is decreasing in $wapr$ though aggregate time supplied is increasing in $wapr$. At a certain threshold, the extensive margin — the total number of young people — dominates the intensive margin — the time supplied by each young person, leading to adverse welfare effects on diseased old.

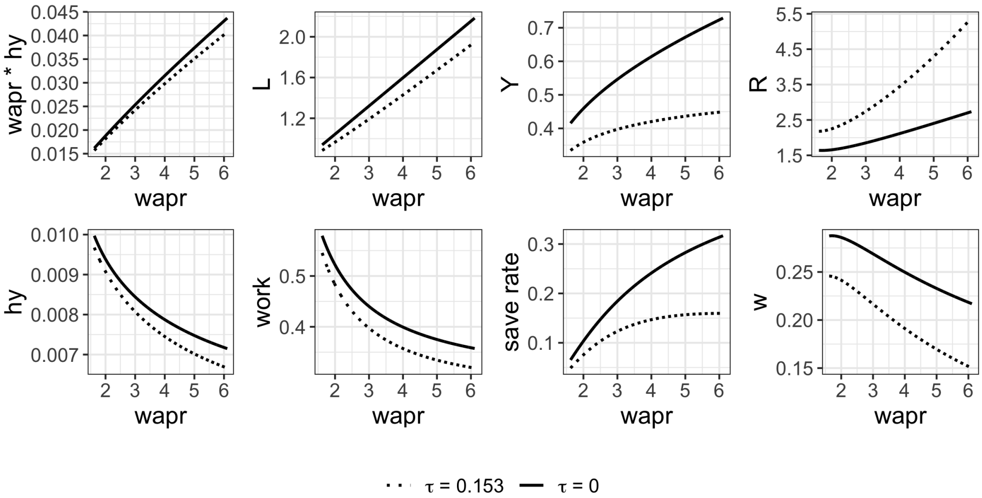


Figure 2.5.: Here we plot steady state outcomes as a function of $wapr$ and different tax rates. Solid lines represent economic variables when the government chooses $\tau = 0$, and dotted lines represent variables under $\tau = 0.153$, the 2016 Social Security and Medicare tax rate.

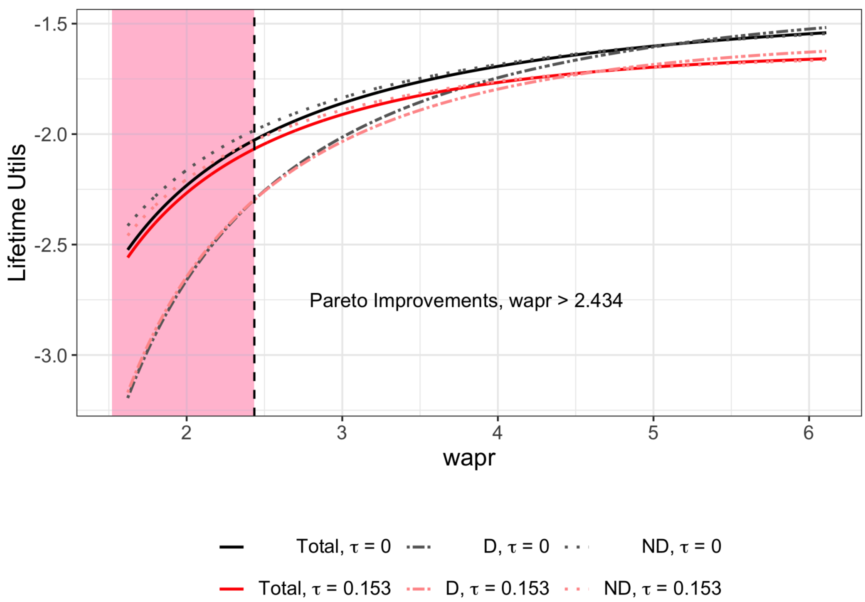


Figure 2.6.: For all $wapr \leq 2.434$ — in the pink box to the left of the dashed vertical line — diseased agents are worse off with $\tau = 0$ than under the baseline 2016 tax rate.

Notice also that diseased old utility falls faster than non-diseased utility as $wapr$ decreases. This is because the decline in the extensive margin drives down total off-market time supplied by young agents as $wapr$ falls, even though every individual young agent is supplying more informal care time on the intensive margin. Meanwhile, as $wapr$ increases, diseased lifetime utility increases faster than non-diseased lifetime utility as the total supply of informal care time increases, allowing diseased agents to supplement their market consumption with increasing amounts of care from their children. Since these are steady state comparisons only, they should be interpreted with caution as such analyses fail to account for productivity gains. We present them to illustrate the co-dependence of lifetime welfare on both $wapr$ and τ.

2.6. Conclusion

Including both idiosyncratic health risk and a motive for young people to engage in informal care of their elders allows the standard, two-period overlapping generations model with production and social insurance taxes to broadly describe the observed decline in aggregate output growth since the 1950s. The model we present qualitatively describes and matches the observed tradeoffs from the ATUS data that agents face when making a decision to provide time on the informal market. These results are important because they should encourage researchers to take seriously the model's predictions about future economic outcomes in an environment with a rapidly aging population. Due to incomplete markets, the rate at which the population ages can adversely impact lifetime welfare of diseased agents when not enough young agents are alive to supply informal care. In counterfactual simulations, reimbursement of informal care time and cures for dementia-like diseases improve both growth and welfare over the U.N.'s medium-variant projected population distribution throughout the 21^{st} century. These results should encourage policy makers to consider how the age-distribution and idiosyncratic risk affect economic aggregates when proposing reforms to address stagnating growth. Aging appears to have broad impacts on long-run GDP growth, regardless of old-age disease risk. Further, finding cures for diseases like Alzheimer's and dementia can lead to modest growth improvements, but more importantly such cures are Pareto improving.

3. Two Stage Budgeting Under Mental Accounting

joint with Alan Montgomery and Christopher Y. Olivola

3.1. Introduction

From week-to-week, consumers' expenditure patterns exhibit substantial heterogeneity. To reconcile these patterns, a model that is both empirically flexible yet consistent with theory is desired. While the literature has documented that consumer spending behavior exhibits departures from standard models of rationality, until recently many behavioral theories have been used only to explain one-off phenomena. To explain high-frequency spending patterns with a flexible, microfounded model, an approach that includes important elements of both neo-classical demand theory and contemporary behavioral theories is needed.

In this paper we construct and estimate a unifying model of two stage budgeting, mental accounting, and sparse maximization to reconcile observed heterogeneity in consumers' week-to-week expenditure patterns. The model yields rich predictions with respect to consumer behavior that affirm both classical and behavioral theories. Specifically, we show that improvements can be made to consumer-level own-price and income elasticity estimates when accounting for endogenous behavioral frictions, like mental accounting and bounded rationality. We also estimate that there is rich heterogeneity across consumers as to the degree to which sparse max rationality constraints bind and the degree to which consumers use mental accounting to discipline spending and saving. As this paper is the first to merge two stage budgeting and mental accounting theories in order to estimate a structural model using high-frequency spending data, this paper should be of broad interest to economists. The results here should be of particular interest to economists and policymakers concerned with household consumption/savings behavior and the estimation of consumer demand models.

The model we build is designed for tractable structural estimation using consumer-level, high-frequency spending, income, and bank balance data. We do not observe consumer budgets or mental accounts, though leaning on two stage budgeting and behavioral theories, we build a model so as to estimate these features. In Deaton and Muellbauer (1980b) two stage budgeting is described as a justification for utility separability to ensure tractable demand estimation, though as has been previously pointed out by Heath and Soll (1996) the original model does not allow for separation between first and second stages, so that budgets are not affected by spending misses but assumed to automatically adjust. Instead, we introduce a timing friction whereby the first stage budgets are formed prior to consumers realizing a price-independent spending or taste shock. Leaning on Gabaix (2014) and the vast literature on narrow choice bracketing (Kahneman and Lovallo 1993; Read, Loewenstein, and Rabin 1999; Rabin and Weizsäcker 2009; Felső and Soetevent 2014; Koch and Nafziger 2016, 2019), consumers are assumed to be sparse maximizers, re-evaluating only a subset of first stage budgets in any given period. The model, however, is flexible, allowing for heterogeneity both across consumers and over time as to the number of first stage budgets they re-optimize in any given period. In this way, we allow for heterogeneous rational inattention, reflecting a broad literature which suggests that there is no reason to assume that all consumers will regard all attributes of a problem as uniformly salient to the same degree (Chetty, Looney, and Kroft 2009; Bordalo, Gennaioli, and Shleifer 2013; Kőszegi and Szeidl 2013; Schwartzstein 2014; Caplin and Dean 2015; Bordalo, Gennaioli, and Shleifer 2017).

Relying on the extensive mental accounting theories described in Shefrin and Thaler (1981), Thaler (1985), Shefrin and Thaler (1988), and Thaler (1990, 1999), we assume that consumers keep track of over and under spending in broad consumption categories. Further, this information informs equilibrium budgeting and spending decisions. As Farhi and Gabaix (2020) note, there is no consensus in the literature as to what exactly constitutes mental accounting. Indeed, the term is used to describe many behaviors while the literature lacks agreement as to which formal, axiomatic foundations necessarily describe an individual engaging in mental accounting. Our formulation relies on several, but not all, behavioral mechanisms described in the literature. First, the model's timing frictions follow from the so-called planner/doer decision structure described in Shefrin and Thaler (1988) and Thaler (1999). Second, consumers have explicit budgets for spending in different categories as in Thaler (1985), however, consumers are assumed to only have utility over the acquisition of final consumption. In Thaler (1985) consumers care both about how much they pay relative to some reference price (transaction utility) and how much they consume (acquisition utility). We take no stand on transaction utility or reference prices. Instead, mental accounting in our model takes the form of consumers keeping track of over and under spending as in Soman (2001). For example, if a consumer faces a \$100 budget for groceries and spends \$110 in week t then, absent a budget update, the consumer encodes his over or under expenditure in a mental account and targets a \$90 level of expenditure on groceries in $t+1$. In the literature mental accounting is also associated with non-fungibilities induced by consumers jointly deciding what to consume and how to pay for it, say by selecting whether to use a credit card or cash for example (Feinberg 1986; Prelec and Loewenstein 1998; Mullainathan 2002). Given our model is designed to estimate parameters from a dataset of underbanked, low-income consumers with regular, observed weekly income using pre-paid debit cards, we abstract from joint expenditure and method-of-payment decisions. We thus assume that underbanked consumers use their pre-paid debit cards for almost all primary consumption.

Our endeavor here falls into a broader conversation about consumption smoothing behavior. A cursory examination of various time series of consumer-level spending in our dataset reveals violations of consumption smoothing over short intervals. The literature has documented violations of consumption smoothing, though most papers rely on spending data aggregated over longer time intervals than we consider (Hall 1978; Zeldes 1989; Souleles 1999; Browning and Collado 2001; Souleles 2002; Japelli and Pistaferri 2010; Kueng 2015; Epper et al. 2020). There are several reasons why we may observe consumption smoothing violations at the weekly level. First, household inventories may take more than one period to draw down. But a model where consumers keep track of unobserved inventories is isomorphic to a mental accounting model where consumers keep track of over and under expenditure. In an inventory model a consumer may only care about his budget for a particular consumption category during the period in which he makes purchases in that category. This would manifest itself as periodic inattentiveness in our formulation with the consumer re-evaluating his budget infrequently. Two mechanisms, then, drive violations of consumption smoothing in our model — budget inattentiveness and mental accounting. The inattention mechanism captures a variety of underlying reasons as to why consumption smoothing is violated. The strength of mental accounting dynamics captures both corrective behavior — consumers consciously wanting to under spend one period after over spending, for example — and inventory effects.

As will be seen, the structural model with endogenous behavioral frictions improves upon estimates of own-price and income elasticities for everyday consumption relative to a naive regression model. Estimates from the naive model suggest that most consumers behave as if most consumption commodities are Giffen goods. In our approach, we assume consumers respond to prices and income either by changing their ex-ante, first stage expenditure budgets or their actual spending in the second stage of the two stage problem. In our sparse max problem consumers do not necessarily re-optimize their budgets every period, but only examine a handful of them at a given time. In our low-income sample consumers only re-evaluate 25-50% of their budgets in a given week. Indeed, in our empirical exercises many consumers appear to engage in persistent over spending relative to optimal budgets over multiple periods.

Analysis of spending patterns, mental accounting tendencies, and budgeting behavior of low-income consumers should be of particular interest to economists and policymakers who are concerned with the savings rates and welfare of the financially marginalized. Our results demonstrate that most consumers behave as if expenditure budgets are relatively strong anchors. When cognitive constraints are randomly relaxed, so that consumers solve their indirect utility problem and optimally update first stage expenditure budgets, both the sign and magnitude of optimal budget updates depend on a consumer's personal sensitivity to over or under expenditure. Consumers that are more sensitive to over expenditure may increase budgets in response to successive periods of high spending in order to offset their disutility from over spending. Loss averse consumers will adjust budgets upward in response to over spending more than they would move them down after the same magnitude of under spending. Under these features, analyzing the behavioral profiles of a sample of low-income consumers can help policymakers understand how these consumers will respond to behavioral nudges designed to increase budget attentiveness and savings. Since low-income households are more likely to face significant financial literacy constraints but also relatively more likely to benefit from increased budget discipline, our structural approach provides both theoretical and quantitative inferences as to how these households may respond and benefit from interventions.

The paper proceeds as follows. In the next section we describe our unique dataset of consumer-level weekly spending and income. We then build a model of two stage budgeting that incorporates features from behavioral economics to explain the variation and heterogeneity in observed weekly spending patterns. After presenting and discussing the theoretical model, we develop a latent inference estimation routine to uncover consumers' unobserved budgeting decisions and mental accounting state variables. Finally, we discuss the results of our estimation and their implications both for inference with regards to classical demand theory, like price and income responsiveness, and behavioral theory, like loss aversion.

3.2. Weekly Expenditure Data

Our goal is to estimate consumer demand parameters while simultaneously accounting for the endogenous link between spending and budgeting decisions using only expenditure data, where control variables to account for unobserved demand shifts for high frequency, weekly expenditure are missing. Further, we want to estimate the model lacking any data about individual spending budgets as well. Incorporating widely-established mental accounting features into a classical two-stage budgeting model accomplishes both of these empirical goals by first placing structure on the underlying budgeting decisions that is consistent with theory and second by allowing for spending to be responsive to an unobserved state variable that encodes consumers' budgeting misses. With such features, we can estimate individual price and income elasticities as well as other demand features using only consumer-level spending data for broad commodity-group aggregates. In this section, we describe our unique consumer-level expenditure dataset and some of its features.

We have access to weekly-aggregated consumption expenditure, income, and running balances for a sample of low-income prepaid debit card users from a North American bank.[1] Since low-income consumers are more likely to be liquidity-constrained and liquidity-constrained households are budget-sensitive and often engage in non-fungible spending behavior for basic necessities like food, studying the spending patterns of low-income households is a natural application for a model with endogenous mental accounting and budgeting features (Gelman et al. 2014; Hastings and Shapiro 2018). Expenditure is categorized by 4-digit Visa merchant category classification codes (MCC).[2] To classify expenditure categories we consider only the first two digits of the MCCs, which give us three explicit consumption categories and one commodity group that collects all other expenditure.[3] The explicit categories we examine, and their 2-digit MCCs are Groceries (54), Gasoline (55), and Food Away from Home (58).[4] The timeline of observations spans from September 2013 to January 2016 with most agents

[1] We have also engaged in analyses of monthly-aggregated expenditure data on the same sample of consumers. The main estimation results are similar to those under weekly aggregation. Monthly results are available upon request.

[2] Within the card industry this code is common across all card processors and is also known as the Standard Industry Code (SIC).

[3] Note that the category for other expenditure includes both classifiable transactions that fall outside the 2-digit codes above and cash withdrawals from the account. This means that some possibly classifiable expenditure may be mis-categorized. Unfortunately, this cannot be reconciled without making some very strong assumptions with regards to how consumers spend their money without evidence.

[4] Our goal with category selection is to look at purchasing behavior in categories whose demand has also been fairly widely analyzed in the literature. For the current analysis, categories have been exogenously defined in order to discipline inference. We have run similar analyses selecting separate categories for each agent using an average budget share ranking criterion. The results are similar to those presented here and also available upon request.

appearing as active users over only the latter part of this time period. After data cleaning and selection, our results operate on transaction profiles for $I = 2{,}509$ customers with 71,859 weeks of observations between them.

We restrict our analysis to a unique sample of consumers who appear to use their prepaid debit card as their only banking product. That is, consumers deposit regular, weekly income to their prepaid card and use the card for consumption expenditure on a regular basis. We acknowledge that restricting our analysis to prepaid debit card users limits possible inferences to only a small subset of the population, namely the underbanked with limited access to other banking products, like checking accounts and credit cards. Yet, given data limitations, we are willing to accept this tradeoff in exchange for higher probability that the income and expenditure profiles we observe represent complete consumption profiles for those agents who use this particular product. Each consumer profile contains at least 16 consecutive weeks of regular income. This ensures that we are observing a regular income process for the consumer. If we observe at least four consecutive weeks of zero expenditure in any of the aforementioned categories, the consumer is dropped from the sample.

Table 3.1.: Summary Statistics Over Agent-Level Means

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
Weeks	16	20	25	28.640	35	39	11.131
Income	23.353	338.855	460.052	510.732	620.131	2,583.379	258.288
Balances	-1,338.279	89.953	204.734	500.744	438.464	59,451.630	1,868.651
Expenditure							
Groceries	2.638	22.282	34.715	42.018	54.195	492.537	30.622
Auto/Gas	2.129	29.191	48.430	60.941	80.198	511.928	46.064
Food Away	2.808	30.179	51.965	68.724	86.684	952.799	61.656
Other	0.666	194.847	283.406	326.124	415.037	2,247.238	195.551

Number of Consumers $I = 2,509$; Total Consumer Weeks $\sum_i \sum_{t_i} = 71,859$

Summary statistics for our entire sample are presented in Table 3.1, where the units are agent-level averages.[5] The median consumer in our sample earns \$460.05 per week after taxes, which aggregates to \$23,922.60 per year in nominal dollars. Median United States household income in 2015 was \$56,516 according to the Census bureau (Proctor, Semega, and Kollar 2015). 96.8% of our sample fall below 2015 median household income. The 2015 poverty thresholds for one, two, three, and four person households where the head was under the age of 65 were \$12,331, \$15,952, \$18,871, and \$24,257. 7.8%, 18.3%, 30.8%, and 51.4% of our sample respectively fall below these corresponding poverty thresholds.[6]

We do not observe purchase prices at the transaction-level, so we turn to the Consumer Price Index for price indices. Price units are \$1982-84 (BLS). All price indices are published at the monthly level, so we use simple linear interpolation to get weekly, aggregate price estimates. For the prices of groceries and food away from home, we use the United States city average

[5]For example, in Table 3.1 "Median" income corresponds to the median agent-level average income and so on.
[6]We do not observe the household size in our dataset.

food and beverage price index (series I.D. “CUUR0000SAF”). For the price of gasoline and other automotive maintenance miscellany, we use the United States city average transportation price index (series I.D. “CUUR0000SAT”). All other expenditure will be deflated using the United States city average all items index (series I.D. “CUUR0000SA0”) when real values are required.

3.2.1. Data Features

Figures 3.1 and 3.2 present the time series of observed expenditures for the median weekly income consumer unit #2,415 and average weekly income consumer unit #1,795 in our sample, respectively.[7] We observe 18 and 29 weeks of expenditure for these two consumers. Notice that weekly expenditure patterns are spiky yet weakly persistent. We thus seek a model, formalized by theory, that can predict spending spikes, persistence, and trend decline or growth in individual consumption expenditure series. Fitting such a model to data will provide quantitative inference as to how consumers make high-frequency consumption expenditure decisions and the tradeoffs they face in doing so.

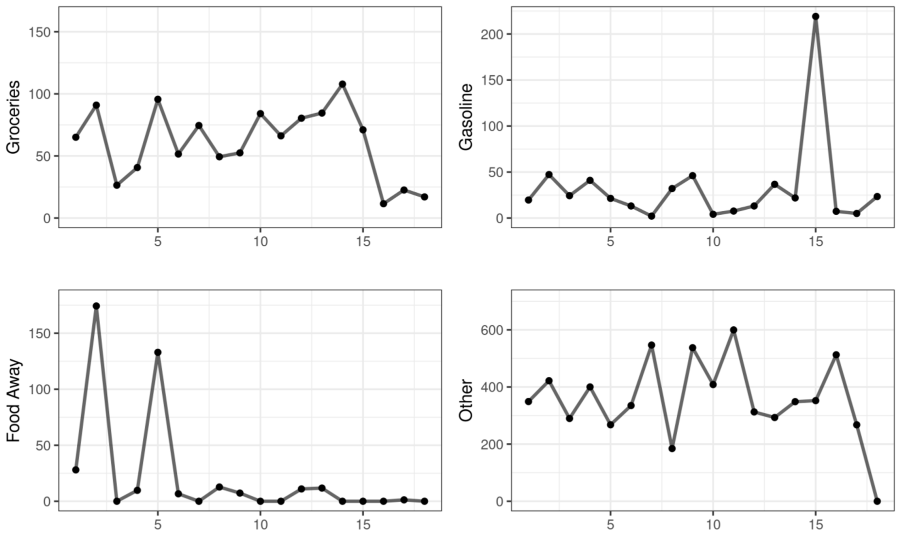


Figure 3.1.: These are the time series of spending for the median income agent #2,415. Mean weekly expenditure in each category is as follows: groceries \$60.67, auto/gasoline \$32.50, food away from home \$21.98, and other expenditure \$357.00. This consumer earns an average of \$460.05 per week and maintains an average weekly card balance of \$115.21 after spending.

3.2.1.1. Naive Price and Income Elasticities

Price and income elasticities that result from estimating a standard, naive model of demand using time series data of this nature lead to some rather strange and dubious conclusions regarding consumer behavior. Let q_{ijt} denote the real quantity of consumption in \$1982-84 for agent i in category j during week t. Let p_{jt} be the price level and l_{it} be income. Consider the following raw price and income elasticity estimates from the simple regressions $\ln q_{ijt} = \beta_{ij0} + \sum_{k=1}^{4} \beta_{ijk} \ln p_{jt} + \beta_{ij5} \ln l_{it} + \eta_{ijt}$ conditional on $q_{ijt} > 0$, where $j = 1$ thru $j = 4$ are

[7] These are the card users whose income is closest to the sample median and sample mean.

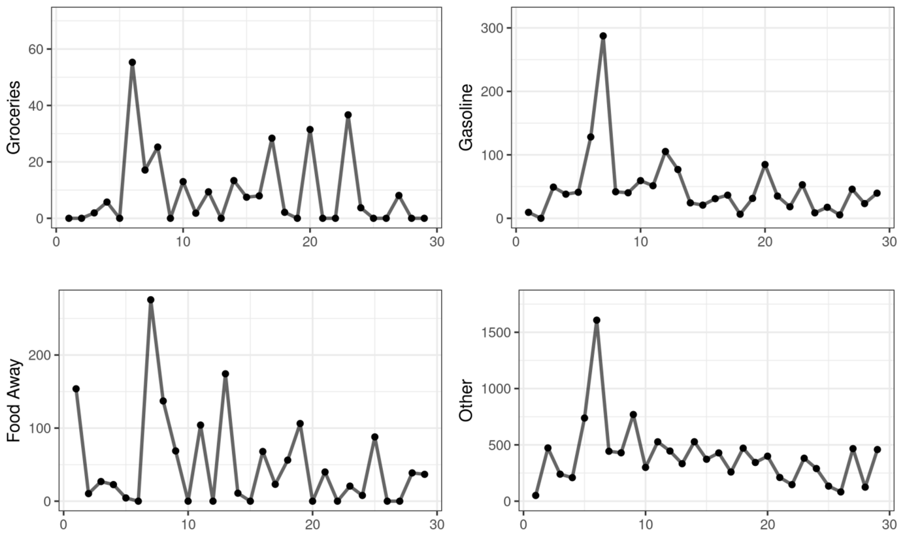


Figure 3.2.: These are the time series of spending for the mean income agent #1,795. Mean weekly expenditure in each category is as follows: groceries \$9.27, auto/gasoline \$48.59, food away from home \$50.83, and other expenditure \$402.15. This consumer earns an average of \$510.80 per week and maintains an average weekly card balance of \$334.28 after spending.

the broad commodity groups: groceries ($j = 1$), auto/gasoline ($j = 2$), food away from home ($j = 3$), and other expenditure ($j = 4$). We specifically do not control for possible demand shifters affecting the error terms η_{ijt}, but rather we just run simple OLS regressions on each equation for each individual.[8] Theoretically, we would need control variables for agent-level demand shifters each period which would account for things like taste shocks and unanticipated expenditures. Lacking such data, we use our structural model formulation to endogenously account for preference shifts resulting from frictions due to two stage budgeting and mental accounting. Why do we even need to consider such possible demand shifters? Consider the own-price elasticity estimates from the aforementioned regressions. For groceries, median and mean agent-level own-price elasticities are -16.554 and 0.617 respectively; for automobile/gasoline, these values are 1.3 and 0.642; for food away from home, these values are -16.501 and 2.318, while for other expenditure we get 7.825 and -45.156. It would appear, at first glance, that for automobile/gasoline expenditure and other expenditure over half of agents treat such commodity groups as Giffen, with positive own-price elasticities (Marshall 1890). Even further, examination of income elasticities suggests that over 25% of consumers appear to treat groceries, automobile/gasoline, and food away from home as inferior goods with negative income elasticities. These results contradict core economic theory, suggesting for some consumers necessities like food and gas are both Giffen and inferior, necessitating a mechanism to both explain and control for changes to demand beyond prices and income responsiveness.

[8] Neither do we control for correlation between η_{ijt} and $\eta_{ij',t}$ for $j \neq j'$, as would be done in a seemingly unrelated regression (SUR) model (Zellner 1962; Deaton and Muellbauer 1980a). We do not control for this correlation because $q_{ijt} = 0$ for different j and t combinations, meaning the length of the qualifying individual time series we can use in each separate first-stage regression varies across consumption categories within an observed agent's sampled time-frame. The SUR procedure is not designed to accommodate this.

While a complex, agent-specific exogenous preference shock process could conceivably reconcile this behavior and generate more plausible price and income elasticity estimates, such a modeling approach is theoretically unsatisfying. Indeed, we would learn very little from a standard consumption model with preferences moving around in some random manner over time. There exists a substantial literature in behavioral economics, however, which provides us with rich theories to help understand why consumption allocations vary idiosyncratically. In the sections that follow, we offer a unifying theory of two stage budgeting and mental accounting, building on the work of Deaton and Muellbauer (1980b) with the theories of Shefrin and Thaler (1981, 1988) and Thaler (1999). We contend that preference shifts result from several features of mental accounting and classic demand theory: 1) consumers infrequently evaluate their first stage expenditure budgets; 2) consumption is referentially dependent, so that spending in $t-1$ informs spending in t;[9] 3) all spending is subject to idiosyncratic, unplanned shocks, akin to taste shocks. Together with two stage budgeting, mental accounting provides us with a mechanism to endogenize the reasons for spiky expenditure, level shifts to spending averages, and apparent changes in spending pattern variances as resulting from a boundedly-rational agent's equilibrium consumption decisions.

3.3. Model

Time is discrete and indexed by t. Column vectors and matrices are denoted with bold font. All variables and functions presented, except market prices $\boldsymbol{p}_t$, are agent-specific, with consumer units indexed by i. Consumers make decisions in a sequentially dynamic environment, choosing savings so as to satisfy utility over liquidity holdings. Further, consumption expenditure and budgeting decisions will be referentially dependent on the previous period's relative over or under expenditure. We proceed by describing the preference and expenditure mechanisms, then consumers' budget updating choices.

3.3.1. Preferences and Expenditure

Let z_{it} be marginal period t savings, m_i be the consumer's borrowing limit, b_{it} available bank balances, and r_t be the gross return on those balances. Account balances evolve according to

$$b_{i,t+1} = r_t b_{it} + z_{it} \tag{3.1}$$

The borrowing limit is such that $m_i > -b_{it}$ always. Consumers have preferences over a J-dimensional vector of real quantities of consumption $\boldsymbol{q}_t$ and their total period t available resources for spending, $z_{it} + m_i + r_t b_{it}$.[10] We assume utility has an additively separable form

[9] See Soman (2001), Kőszegi and Rabin (2006), and Cheema and Soman (2006) for models of reference dependence and mental accounting.

[10] In this formulation the consumer has preferences over money holdings. This condition of our model ensures that consumers never devote all of their available resources to expenditure, desiring instead to carry money forward in time. Since the dynamics of the model are sequential, rather than in terms of continuation values, then if consumers care about balance-matching, as predicted in Gathergood et al. (2019), we expect that they also care about balance holdings and how their behavior contributes to such holdings. For other "money in the utility function" models, refer to the monetary economics literature (Brock 1974; Calvo 1983; Obstfeld and Rogoff 1983; Feenstra 1986; Barnett, Fisher, and Serletis 1992; Walsh 2010).

that allows for zero expenditure:[11]

$$u_{it}(\boldsymbol{q}_{it}, z_{it}) = \sum_{j=1}^{J} \alpha_{ij} \ln(q_{ijt} + 1) + \alpha_{i,J+1} \ln(z_{it} + m_i + r_t b_{it}) \tag{3.2}$$

Let l_{it} be period t income. Let $x_{ijt} = p_{jt}q_{ijt}$ be expenditure on commodity group j. Note that marginal savings must be

$$z_{it} = l_{it} - \sum_{j=1}^{J} x_{ijt} \tag{3.3}$$

z_{it} can indeed be negative if consumers spend more than they earn in income. After substituting the expression for z_{it} into (3.2) it can be shown that the utility function is strictly increasing and strictly concave in q_{ijt} as long as m_i is sufficiently large.[12] Under these conditions, this utility function is well-behaved and results in unique equilibrium demand allocations for any given set of prices $\boldsymbol{p}_t$ and net liquid resources $m_i + l_{it} + r_t b_{it}$.

Under our mental accounting formulation period t expenditure in each commodity group is subject to a separate constraint rather than one single, perfectly linear budget constraint. Coupled with ex-ante, category-specific expenditure uncertainty, this amounts to relaxation of the assumption that resources allocated toward first stage budgets are perfectly fungible between commodity groups. In the standard two stage budgeting model described in Deaton and Muellbauer (1980b), the realized expenditure share for group j is always exactly equal to the first stage budgeting share. In our model timing frictions between budgeting and spending ensure that the ex-ante first stage budgets never exactly equal second stage realized expenditure, except in measure zero cases for each j.

Let θ_{ijt} be a period-specific weighting variable that determines the degree to which changes in income contribute to expenditure changes. Let $a_{ij,t+1}$ be a state variable that encodes the amount a consumer over or under spent relative to his ex-ante period t expenditure expectations for commodity group j. Let ζ_{ijt} be an *iid* idiosyncratic expenditure shock for category j. We assume for each j, $\zeta_{ijt} \perp \zeta_{ij',t}$ where $j \neq j'$, and each shock is mean-zero normally distributed

$$\zeta_{ijt} \underset{iid}{\sim} \mathcal{N}(0, \sigma_{ij}^2)$$

[11] The utility function is a modification of Stone-Geary preferences (Geary 1950; Stone 1954).

[12] Specifically, we require $\sum_{j=1}^{J} \alpha_{ij} = 1$ with $0 \leq \alpha_{ij} \leq 1$ for all $j \leq J$, but we require no explicit restriction on $\alpha_{i,J+1}$, just that any combination of $\alpha_{i,J+1}$ and m_i are such that maximization of u_{it} in $\boldsymbol{q}_{it}$ yields unique equilibrium outcomes. To ensure utility is strictly increasing in q_{ijt}, we must have

$$m_i > \frac{\alpha_{i,J+1}}{\alpha_{ij}} p_{jt}(q_{ijt} + 1) + \sum_{j=1}^{J} p_{jt}q_{ijt} - l_{it} - r_t b_{it}$$

For strict concavity, we must have

$$\frac{\alpha_{i,J+1}}{\alpha_{ij}} p_{jt}^2 (q_{ijt} + 1)^2 < \left(l_{it} - \sum_{j=1}^{J} p_{jt}q_{ijt} + m_i + r_t b_{it} \right)^2$$

These conditions are readily apparent after twice differentiating u_{it} in q_{ijt} following substitution of $z_{it} = l_{it} - \sum_{j=1}^{J} p_{jt}q_{ijt}$.

This shock encodes price-independent unanticipated deviations to the spending plan, due to everything from weather-related shocks that keep consumers from going out to unexpected health shocks. In this sense, ζ_{ijt} is very similar to a taste shock, causing preferences to appear to vary period by period.

The consumer's period t ex-anted expected expenditure budget in commodity group j, is $\mathbb{E}_{it}x_{ijt} = \theta_{ijt}l_{it} + \gamma_i a_{ijt}$, where γ_i is a parameter that governs the degree to which the previous period's expenditure affects the present period's expenditure expectations. Expectations are taken over $\boldsymbol{\zeta}_{it}$, not prices $\boldsymbol{p}_t$, which we assume are ex-ante known. Given our short period length (one week), we argue that it is not unreasonable to suggest that consumers have perfect expectations of the price levels for broad commodity aggregates prior to the week commencing. Sticky menus mean that prices do not change much from week to week anyway. While prices months or years away may be difficult for consumers to forecast, interim prices are likely to be known.[13] Further, Hastings and Shapiro (2013) present evidence that consumers may substitute toward lower quality products in the event of a price increase across the board to a specific class of commodities. In such a situation spending for a broad commodity aggregate may appear constant in the data, despite the fact that the quality of products being purchased has declined. Since we only observe spending and not quantities and thus rely on broad commodity-aggregate CPI price indices to compute the real value of spending, we would not be able to identify the kind of price responsiveness observed in Hastings and Shapiro (2013) anyway. They show that the link between quality and price sensitivity is best explained by a model where consumers budget for specific commodity categories, like ours. Thus, we argue that we can ignore price expectations for two reasons. Since we do not observe specific commodity prices, only indices, we would expect that since such indices barely move over time, consumers would implicitly know their value. Further, even if they do not, if consumers engage in the type of budgeting we posit, responses to such unexpected price changes may not be identifiable anyway.

After realizing ζ_{ijt}, period t ex-post expenditure in commodity group j satisfies

$$x_{ijt} = \theta_{ijt}l_{it} + \gamma_i a_{ijt} + \zeta_{ijt} \tag{3.4}$$

As a_{ijt} induces expenditure reference-dependence, we think of this value as the consumer's "mental account," that is the amount he over or under spent in the previous period which affects his present expenditure. He uses this mental account to discipline expenditure. If $a_{ijt} < 0$ then the consumer over spent in period $t-1$, while if $a_{ijt} > 0$ he under spent relative to expectations. The law of motion for mental account balances is

$$a_{ij,t+1} = \mathbb{E}_{it}x_{ijt} - x_{ijt} = \theta_{ijt}l_{it} + \gamma_i a_{ijt} - x_{ijt} = -\zeta_{ijt} \tag{3.5}$$

$\boldsymbol{a}_{it}$ is $J+1$ dimensional where the $J+1$ component encodes how much a consumer over or

[13]This assertion may seem strong, but it is not. Aggregate price levels for broad commodity groups change very little from week-to-week. A consumer planning his next week's consumption expenditure on a Sunday, having most recently gone to the store on a Thursday, would likely expect to face the same nominal price level.

under saved relative to expectations. This value is

$$a_{i,J+1,t+1} = -\sum_{j=1}^{J} a_{ij,t+1} = \sum_{j=1}^{J} \zeta_{ijt} \tag{3.6}$$

Note that if $a_{i,J+1,t} < 0$ then the consumer has under saved relative to expectations and vice-versa for $a_{i,J+1,t+1} > 0$. By construction $\sum_{j=1}^{J+1} a_{ijt} = 0$ in every period, so that consumers have no cognitive dissonance overall with regards to their mental accounting reference variables. That is, their perception of how much they over or under spent last period in various commodity groups is exact. The mental accounting model in Farhi and Gabaix (2020) has the same feature.

3.3.2. Choices

Consumers enter each period knowing the vector of budget weights ascribed to last period's income $\boldsymbol{\theta}_{i,t-1}$ and the following state variables $\boldsymbol{a}_{it}$, l_{it}, b_{it}, r_t, and $\boldsymbol{p}_{it}$. Herein lies the planner/doer formulation of Thaler (1999): the "planner" chooses his budgets ex-ante and the "doer" engages in expenditure ex-post. The doer's decision is exactly determined by the expenditure constraint. The introduction of uncertainty regarding realized expenditure is what differentiates Thaler's formulation of consumer decisions from Deaton and Muellbauer's two stage budgeting model. We incorporate an additional friction that allows expenditure to exhibit temporary persistence over a few periods. Drawing on psychological evidence regarding how consumers make financial decisions, suppose consumers engage in updates to first stage budgets infrequently and only for a few categories at a time. That is, in any given period, consumers optimally re-evaluate $k_{it} \leq J$ of their budgets for the J different commodity groups. The theoretical motivations for this friction are that consumers face limits with regards to how much numerical information they can cognitively process at any given time, documented in work dating back to Herbert Simon.[14]

The budget re-evaluation process operates as follows. In any given period a consumer may re-evaluate their weighting variable θ_{ijt} for each good or leave it alone. Let ψ_{ij} denote the probability that a consumer re-evaluates their budget for expenditure in commodity category j, and let Γ_{ijt} be an indicator variable that equals 1 when a re-evaluation is made for the budget in j and 0 otherwise. We assume that the re-evaluation decision is Bernoulli distributed

$$\Gamma_{ijt} \underset{iid}{\sim} \text{Bernoulli}(\psi_{ij})$$

with $\Gamma_{ijt} \perp \Gamma_{ij',t}$ for all $j' \neq j$. This independence assumption induces narrow choice bracketing and results in a sparse max equilibrium decision structure, like that discussed in Gabaix (2014).[15].

[14]Lieder, Griffiths, and Goodman (2013) show that people face limits as to how much numerical information they can cognitively process at once. Imposition of an integer constraint on the number of changes that can be made rather than the absolute value of those changes is consistent with findings in Leslie, Gelmand, and Gallistel (2008) who argue that integer representations are innate within individual cognitive processes. Further, the processing capacity of human memory is fairly limited and impacts consumer decisions, as detailed in Malhotra (1982). Findings in Simon (1957) suggest that individuals can consider at most seven choice alternatives at once. Meanwhile, Cowan (2000) says this number is closer to four.

[15]For further discussion of narrow choice bracketing see Kahneman and Lovallo (1993), Read, Loewenstein, and Rabin (1999), Rabin and Weizsäcker (2009), Felső and Soetevent (2014), and Koch and Nafziger (2019)

In the context of our problem, a consumer choosing how to allocate funds for multiple different consumption budgets may only optimally re-evaluate one or two of those budgets in any given period, leaving the remainder fixed. Note a couple things regarding this process. First, we do not assume consumers are choosing whether to re-evaluate a budget, but rather whether or not a re-evaluation is made is an exogenous event that happens to the consumer. In certain periods, this feature allows for spending on certain categories to be more salient to the consumer, so that our model allows for the consumer to exhibit similar attentiveness bias as in Bordalo, Gennaioli, and Shleifer (2013), Kőszegi and Szeidl (2013), and Schwartzstein (2014). While consumers do not choose whether or not to re-evaluate a budget, conditional upon a re-evaluation being made ($\Gamma_{ijt} = 1$) the consumer optimally updates his category j budget by choosing θ_{ijt} to maximize expected indirect utility. Otherwise they set $\theta_{ijt} = \theta_{ij,t-1}$ leaving their planned budget weight for commodity group j alone. Due to this integer constraint consumers may change some budgets but not others. Generally speaking, this is fine, just assume that if only one budget is changed then the implicit budget for savings changes as well. Since the expenditure system and total savings are perfectly collinear, it is sufficient to only specify how consumers alter their expenditure budgets.

Let $k_{it} = \sum_{j=1}^{J} \Gamma_{ijt}$, so that k_{it} describes the total number of expenditure budgets the agent will optimally adjust in period t. The vector $\boldsymbol{\theta}_{it}$ is J-dimensional. Denote $\boldsymbol{\vartheta}_{it}$ as the k_{it}-dimensional vector which holds, in cardinal order, the adjustable budgeting parameters of all commodities j for which $\Gamma_{ijt} = 1$. Note that $\boldsymbol{\vartheta}_{it}$ is a sub-vector of $\boldsymbol{\theta}_{it}$ corresponding to the non-zero indices of $\boldsymbol{\Gamma}_{it}$. Let $\boldsymbol{\vartheta}_{it}^*$ be the optimally chosen analog of this vector. Denote $\mathbb{E}_{it} v_{it}(\boldsymbol{\theta}_{it})$ as expected indirect utility after dividing each component of the expenditure system (3.4) by p_{jt} and substituting into $u_{it}(\boldsymbol{q}_{it}, z_{it})$ along with z_{it}:

$$\mathbb{E}_{it} v_{it}(\boldsymbol{\theta}_{it}) = \mathbb{E}_{it} \left\{ \sum_{j=1}^{J} \alpha_{ij} \ln \left(\frac{\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}}{p_{jt}} + 1 \right) \right.$$
$$\left. + \alpha_{i,J+1} \ln \left(l_{it} - \sum_{j=1}^{J} [\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}] + m_i + r_t b_{it} \right) \right\}$$

Let $\boldsymbol{\iota}_{it}(\boldsymbol{\Gamma}_{it})$ be a vector-valued integer function that maps the index of components of $\boldsymbol{\vartheta}_{it}$ back into the index of components for $\boldsymbol{\theta}_{it}$. This function is such that $\boldsymbol{\iota}_{it}(\boldsymbol{\Gamma}_{it}) : \mathbb{N}^J \to \mathbb{N}^{k_{it}}$ and valid for $k_{it} > 0$. The function outputs in cardinal order the index of the components of $\boldsymbol{\theta}_{it}$ to which we assign the components of $\boldsymbol{\vartheta}_{it}^*$. For example, suppose $J = 4$ and $\boldsymbol{\Gamma}_{it}$. Then this implies $k_{it} = 2$, $\boldsymbol{\Gamma}_{it} = (0, 1, 0, 1)$, and $\boldsymbol{\iota}_{it}(\boldsymbol{\Gamma}_{it}) = (2, 4)$.

Let y index the components of $\boldsymbol{\iota}_{it}(\boldsymbol{\Gamma}_{it})$, so that ι_{iyt} denotes the y^{th} component of the vector $\boldsymbol{\iota}_{it}$.[16] If $k_{it} > 0$ then an optimal choice of ϑ_{iyt}^* must satisfy

$$\mathbb{E}_{it} \frac{\partial v_{it}(\boldsymbol{\theta}_{it})}{\partial \vartheta_{iyt}} = \mathbb{E}_{it} \left\{ \frac{\partial u_{it}}{\partial q_{i,\iota_{iyt},t}} \frac{\partial q_{i,\iota_{iyt},t}}{\partial \theta_{i,\iota_{iyt},t}} + \frac{\partial u_{it}}{\partial z_{it}} \frac{\partial z_{it}}{\partial \theta_{i,\iota_{iyt},t}} \right\} = 0, \quad \forall y > 0 \tag{3.7}$$

Due to the money-in-the-utility-function structure, (3.7) will always be satisfied in equilibrium as consumers equate the expected marginal utility of additional consumption with the expected

[16]In our example above, $\iota_{i1t} = 2$ and $\iota_{i2t} = 4$.

marginal utility of additional liquidity tomorrow.

We can invert (3.7) under our utility parameterization to arrive at an analytical expression for equilibrium values $\boldsymbol{\vartheta}_{it}$. For some y indexing $\boldsymbol{\iota}_{it}$, let $\boldsymbol{\vartheta}^*_{i,-yt}$ denote the vector of optimally chosen budget shares which does not include y. Note that this vector may be empty. The optimal budget share $\vartheta^*_{iyt}(\boldsymbol{\vartheta}^*_{i,-y,t})$ for good ι_{iyt} can be implicitly expressed

$$\begin{aligned}\vartheta^*_{iyt}(\boldsymbol{\vartheta}^*_{i,-y,t}) = \mathbb{E}_{it}\Bigg\{ & \frac{\alpha_{i,\iota_{iyt}} l_{it} - \alpha_{i,\iota_{iyt}} \sum_{j\in\iota_{it}(\Gamma_{it})\, j\neq \iota_{iyt}} (\vartheta^*_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt})}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \\ & - \frac{(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})(\gamma_i a_{i,\iota_{iyt},t} + \zeta_{i,\iota_{iyt},t}) + \alpha_{i,J+1} p_{\iota_{iyt},t}}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \\ & - \frac{\alpha_{i,\iota_{iyt}} \left(\sum_{j=1}^{J} (1-\Gamma_{ijt})(\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}) - m_i - r_t b_{it} \right)}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \Bigg\}\end{aligned} \tag{3.8}$$

where $\theta_{ijt} = \theta_{ij,t-1}$ for $j \notin \boldsymbol{\iota}_{it}(\boldsymbol{\Gamma}_{it})$. For the algebra behind this expression, see Appendix C.1.2. Recall we assume that l_{it} and $\boldsymbol{p}_t$ are ex-ante known. In Section 3.4 we will exploit the independence of the components of $\boldsymbol{\zeta}_{it}$ to iteratively sample the latent shocks Γ_{ijt} and update the components of $\boldsymbol{\vartheta}^*_{it}$ accordingly, taking $\vartheta^*_{i,-y,t}$ as given. This allows for estimation of the mental accounting and budgeting parameters without having to iterate over the component functions of $\boldsymbol{\vartheta}_{it}$ every time an underlying parameter is changed as we progress through the MCMC routine.

3.3.3. Personal Mental Accounting Equilibrium

Let t_{i0} be the period in which an individual consumer enters the economy as an autonomous, decision-making agent. Given a sequence of prices $\{\boldsymbol{p}_t, r_t\}_{t\geq t_{i0}}$, a sequence of income values $\{l_{it}\}_{t\geq t_{i0}}$, a sequence of idiosyncratic expenditure shocks $\{\boldsymbol{\zeta}_{it}\}_{t\geq t_{i0}}$, a sequence of idiosyncratic cognitive shocks $\{\boldsymbol{\Gamma}_{it}\}_{t\geq t_{i0}}$, and initial values for the budget weights, mental account balances, and bank balances $\{\boldsymbol{\theta}_{i,t_{i0}}, \boldsymbol{a}_{i,t_{i0}}, b_{i,t_{i0}}\}$, a personal mental accounting equilibrium consists of:

i. *Sequences of policies:* $\{\boldsymbol{q}_{it}, z_{it}, \boldsymbol{\theta}_{it}\}_{t\geq t_{i0}}$.

ii. *Sequences of balances:* $\{\boldsymbol{a}_{it}, b_{it}\}_{t\geq t_{i0}}$.

such that

a. *Given* $\boldsymbol{\Gamma}_{it}$*,* $\boldsymbol{a}_{it}$*, and* b_{it}*,* $\boldsymbol{\theta}_{it}$ *satisfies the sparse max indirect-utility maximization program.*

b. *Given* $\boldsymbol{\theta}_{it}$*,* l_{it}*,* $\boldsymbol{a}_{it}$*, and* $\boldsymbol{\zeta}_{it}$*,* $\boldsymbol{x}_{it}$ *satisfies (3.4).*

c. *Given* $\boldsymbol{x}_{it}$ *and* l_{it}*,* z_{it} *satisfies (3.3).*

d. *Mental account balances* $\boldsymbol{a}_{it}$ *are updated according to (3.5) and (3.6).*

e. *Bank balances* b_{it} *evolve according to (3.1).*

3.3.4. Implications of Budgeting Frictions

In our formulation the consumer's ex-ante budgeting decisions satisfy the first-order conditions of a sparse max indirect utility optimization problem. By contrast his ex-post expenditure is effectively taken as myopic, subject to stochastic deviations around expected expenditure. In this way the consumer does not proactively choose his real consumption levels but anchors his expenditure around a chosen budget. Through this channel, resources allocated toward first stage budgets are thus not perfectly fungible nor transferable across second stage expenditure. The structure and timing of this decision process has welfare implications relative to how we typically model a neo-classical consumption/savings problem where the consumer chooses his real consumption level and is not subject to budgeting frictions. Indeed, if the consumer was to have one-period-ahead perfect foresight, then our model would collapse into a standard consumption/savings problem with money in the utility function, where the consumer takes prices as given, choosing a vector of real consumption $\boldsymbol{q}_{it}$ and savings z_{it} to maximize utility. Other consequences for the sparse max structure of our optimization problem are the same as those outlined in Gabaix (2014): the Slutsky matrix may be asymmetric, which becomes more clear when we analyze demand elasticities in Section 3.5.1, the weak axiom of revealed preferences may not hold, and Marshallian demand formulations, if taking into consideration the endogenous dependence of θ_{ijt} on prices and income, may not necessarily be homogeneous of degree zero. In this section we focus on our model's unique non-equivalence of choosing budgets ex-ante versus choosing consumption ex-post which is a consequence of the way timing frictions in the sparse max optimization problem.

Ex-post, after realizing $\boldsymbol{\zeta}_{it}$, there is no information lost when plugging (3.3) into (3.1) to get back the standard budget constraint without mental accounting frictions:

$$\sum_{j=1}^{J} x_{ijt} + b_{i,t+1} \leq r_t b_{it} + l_{it} \qquad \text{with} \qquad b_{i,t+1} > -m_i \tag{3.9}$$

Note that this substitution cannot be accomplished prior to realization of $\boldsymbol{\zeta}_{it}$ due to Jensen's inequality. Ex-ante expected expenditure depends on the budgeting decision of the consumer and his expectations over $\boldsymbol{\zeta}_{it}$. When choosing his budgets, he internalizes how balances b_{it} will evolve so that expected expenditure depends on expected balances by way of the budgeting decision.

Proposition 1. Optimally choosing ex-ante budgets is equivalent to optimally choosing ex-post consumption if and only if consumers face perfect foresight over spending shocks ζ_{ijt} and no cognitive frictions ($\Gamma_{ijt} = 1$ for all j).

Proof. See Appendix C.1.1. □

The intuition behind the proof of this Proposition is that ex-ante there exists an expected level of consumption $\mathbb{E}_{it} q_{ijt}$ which also exactly solves the integral in (3.7), in the case wherein ζ_{ijt} is not known. This expected level of consumption is a function of the chosen budget weight, θ_{ijt}. Realizing this expected level of consumption as an ex-post actual consumption level is a measure-zero outcome as long as the measure associated with the distribution of

ζ_{ijt} is absolutely continuous. However, even if ζ_{ijt} is known beforehand, if $\Gamma_{ijt} = 0$, so that consumers face cognitive frictions in making optimal budget updates, then the budget share is $\theta_{ijt} = \theta_{ij,t-1}$. It follows that the ex-ante value of consumption, constrained by a sub-optimal budget weight, will not solve the first order condition in (3.7). In this case, even if ζ_{ijt} is known, the utility maximizing value of q_{ijt} will not be equal to the quantity of consumption associated with the sub-optimal budget weight, except in a measure zero case. Proposition 1 thus shows that the two stage budgeting model with mental accounting collapses into the standard two stage budgeting model of Deaton and Muellbauer (1980b) only when there are no budget-timing or cognition frictions. This Proposition demonstrates that the uncertainty introduced by mental accounting frictions causes consumers to deviate from the optimal consumption allocations associated with fully unbounded rationality, implied under two stage budgeting.

3.3.5. Interaction Between Mental Accounting Sensitivity γ_i and Budgeting

The degree to which budget updates respond to the spending error or mental account balances $\boldsymbol{a}_{it}$ depends on the sensitivity parameter γ_i. Without explicit budgeting data encoding relative over or under expenditure, γ_i will not be identified, so in practice we set this value to either 0 or 1 and estimate the remaining parameters. Analytically, though, we can show that this parameter governs both the sign and magnitude by which a_{ijt} affects budget updates.

Proposition 2. Suppose $\gamma_i > 0$ and $\alpha_{i,J+1} > 0$. For some y indexing $\boldsymbol{\iota}_{it}$, fix $\boldsymbol{\vartheta}^*_{i,-y,t}$ and consider the partial sensitivity of ϑ^*_{iyt} to γ_i under the following conditions:

i. If $a_{i,\iota_{iyt},t} > 0$ and $\sum_{j\neq\iota_{iyt}} a_{ijt} > 0$ then $\frac{\partial\vartheta^*_{iyt}}{\partial\gamma_i} > 0$.

ii. If $a_{i,\iota_{iyt},t} > 0$ and $\sum_{j\neq\iota_{iyt}} a_{ijt} < 0$ then as long as $\alpha_{i,J+1}$ is sufficiently large $\frac{\partial\vartheta^*_{iyt}}{\partial\gamma_i} > 0$.

iii. If $a_{i,\iota_{iyt},t} < 0$ and $\sum_{j\neq\iota_{iyt}} a_{ijt} > 0$ then as long as $\alpha_{i,J+1}$ is sufficiently large $\frac{\partial\vartheta^*_{iyt}}{\partial\gamma_i} < 0$.

iv. If $a_{i,\iota_{iyt},t} < 0$ and $\sum_{j\neq\iota_{iyt}} a_{ijt} < 0$ then $\frac{\partial\vartheta^*_{iyt}}{\partial\gamma_i} < 0$

Proof. See Appendix C.1.1. □

In Proposition 2, cases (i) and (iv) describe how both category-specific and overall under expenditure and over expenditure, respectively, interact with γ_i. In the event of under expenditure in both category ι_{iyt} and over the sum of all the other categories, increasing sensitivity γ_i will lead to increasing optimal budget shares. On the other hand in the event of over expenditure in both category ι_{iyt} and over the sum of all the other categories, increasing sensitivity leads to declining budget shares, demonstrating how the sensitivity parameter governs the degree to which consumers discipline expenditure in the face of previous over spending mistakes. Cases (ii) and (iii) demonstrate that the marginal propensity to save, governed by $\alpha_{i,J+1}$, determines whether budget sensitivity positively affects budget changes or negatively. Specifically, if $a_{ijt} > 0$ the marginal propensity to save must exceed the inverse share of total over spending attributed to category ι_{iyt} less that same category's long-run average expenditure share $\alpha_{i,\iota_{iyt}}$. For cases (ii) and (iii), this threshold is $\alpha_{i,J+1} > -\frac{\sum_{j\neq\iota_{iyt}} a_{ijt}}{a_{i,\iota_{iyt},t}} - \alpha_{i,\iota_{iyt}}$. When $a_{i,\iota_{iyt},t} < 0$ just flip the inequality.

3.3.6. Budgeting Sensitivity to Spending Misses a_{it}

Here, we fix intuition by looking at how equilibrium objects co-vary under different degrees of narrow bracketing in our sparse max structure. Since $\boldsymbol{\vartheta}_{it}^*$ is an implicit function component-wise, the total derivative $\frac{\mathrm{d}\vartheta_{iyt}^*}{\mathrm{d}a_{i,\iota_{iyt},t}}$ includes information regarding how other budgets $\boldsymbol{\vartheta}_{i,-y,t}^*$ respond to variation in the current mental account for commodity ι_{iyt}, $a_{i,\iota_{iyt},t}$. Variation in budget ϑ_{iyt}^* on over or under spending in category $j \neq \iota_{iyt}$ will also inform equilibrium values of ϑ_{iyt}^*. For categories featuring optimal budget updates ($\Gamma_{ijt} = 1$), the responsiveness of ϑ_{iyt}^* to variation in a_{ijt}, where j is one of these optimally-updated budget categories, is systematic and predictable. Yet, over or under spending in outside categories $j \notin \boldsymbol{\iota}_{it}$ induces ambiguous optimal budget responsiveness in the updated categories. In this section only, we will refer to categories for which $\Gamma_{ijt} = 1$ as "inside" the bracket and other categories for which $\Gamma_{ijt} = 0$ as "outside" the bracket. Propositions 3 and 4 summarize how $\boldsymbol{\vartheta}_{it}^*$ responds to over or under spending for categories both inside and outside the bracket.

For categories inside the narrow sparse max bracket, both losses and gains in category y due to respectively over or under spending are fully integrated into the optimal budget update for that same category. In case (i) of Proposition 3, we can specifically see that $\frac{\mathrm{d}\vartheta_{iyt}^*}{\mathrm{d}a_{iyt}} = -\frac{\gamma_i}{l_{it}}$ exactly. For $s \neq y$ where s also indexes a category inside the bracket, i.e. $s \in \boldsymbol{\iota}_{it}$, simultaneous optimal budget updates to ϑ_{ist}^* are independent of the mental account balance in category y, a_{iyt}. This can be seen in case (ii) of Proposition 3. Thus, if the inside bracket is broad enough then cross-category responsiveness is minimized. In fact if the inside bracket consists of all J categories, all over or under expenditure from the previous period is separately but fully integrated into the optimal budget updates for each category.

Proposition 3. Suppose $k_{it} > 0$, $l_{it} > 0$, and $\alpha_{i,J+1} > 0$. Without loss of generality, let $\boldsymbol{\iota}_{it} = (1, 2, 3, \ldots)$ and suppose $\boldsymbol{\iota}_{it}$ is of dimension $J' \leq J$. Consider the total responsiveness of the components of $\boldsymbol{\vartheta}_{it}^*$ to a_{iyt} where $y \in \boldsymbol{\iota}_{it}$.

i. Higher a_{iyt} leads to lower ϑ_{iyt}^*, i.e. $\frac{\mathrm{d}\vartheta_{iyt}^*}{\mathrm{d}a_{iyt}} = -\frac{\gamma_i}{l_{it}}$.

ii. For all $s \in \boldsymbol{\iota}_{it}$ where $s \neq y$, $\frac{\mathrm{d}\vartheta_{ist}^*}{\mathrm{d}a_{iyt}} = 0$.

Proof. See Appendix C.1.1. □

If $k_{it} < J$ consumers take into account how much they over or under spent in categories outside the bracket when making optimal updates to inside categories. Proposition 4, though, demonstrates that the responsiveness of inside categories to a_{ijt} where j is outside the bracket is rather complex. Indeed, there is no hard and fast rule as to which inside bracket categories increase or decrease in response to a_{ijt}. Some categories may see budget increases while others see budget decreases. Further, this responsiveness will be different for different consumers since it depends heavily on the underlying values of $\boldsymbol{\alpha}_i$ and $\alpha_{i,J+1}$. Figure 3.3 shows how $\frac{\mathrm{d}\vartheta_{iyt}^*}{\mathrm{d}a_{ijt}}$, for some $j \notin \boldsymbol{\iota}_{it}$, varies in $\boldsymbol{\alpha}_i$ and $\alpha_{i,J+1}$.[17] In Figure 3.3 the utility weights $\boldsymbol{\alpha}_i$ are increasing from left to right. It appears that categories associated with higher utility weights, such as that in Figure

[17] For the simulations presented, we fix $\gamma_i = l_{it} = 1$ and consider an environment with $J = 4$, $k_{it} = 3$, and $\boldsymbol{\iota}_{it} = (1, 2, 3)$, examining the responsiveness of inside categories to variation in a_{i4t}.

3.3c, more fully integrate perceived gains due to under spending, though this feature greatly depends on the marginal propensity to save. Indeed, the responsiveness is non-monotonic in the savings propensity, which is evident in Figures 3.3a and 3.3b. Thus, it appears that a complex interaction between preferences for consumption for the inside categories and the marginal propensity to save determine the degree to which over or under spending in an outside category affects optimal budget weights.

Proposition 4. Suppose $k_{it} > 0$, $l_{it} > 0$, and $\alpha_{i,J+1} > 0$. Without loss of generality, let $\boldsymbol{\iota}_{it} = (1, 2, 3, \ldots)$ and suppose $\boldsymbol{\iota}_{it}$ is of dimension $J' < J$, strictly. Then both the sign and magnitude of the total responsiveness of the components of $\boldsymbol{\vartheta}^*_{it}$ to a_{ijt} where $j \notin \boldsymbol{\iota}_{it}$ is ambiguous and depends on the underlying values of the utility parameters $\boldsymbol{\alpha}_i$ and $\alpha_{i,J+1}$.

Proof. See Appendix C.1.1. □

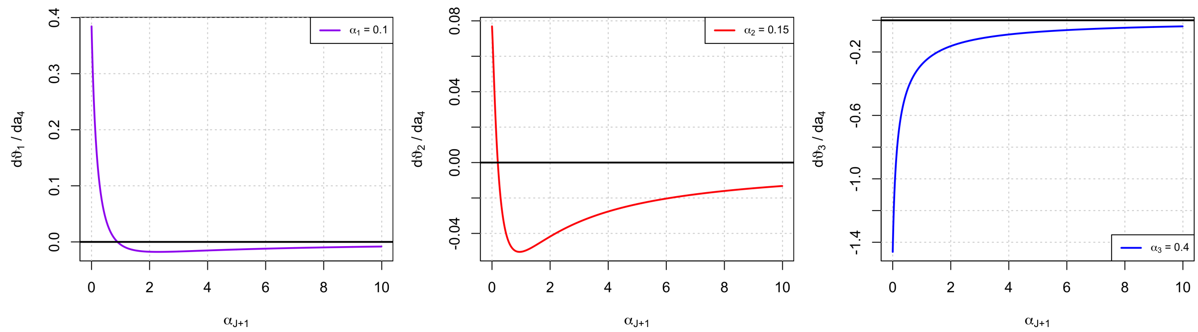


Figure 3.3.: From left to right we show $\frac{d\vartheta^*_{iyt}}{da_{i4t}}$ where $y \in (1, 2, 3)$ as a function of the marginal propensity to save $\alpha_{i,J+1}$. In panel (a) the value of α_{ij} associated with good $y = 1$ is 0.1, in panel (b) it is 0.15, and in panel (c) it is 0.4. Notice that depending on α_{ij} and $\alpha_{i,J+1}$ optimal budget updates respond differently to over or under spending in the non-updated, sparse category $\Gamma_{i4t} = 0$.

3.4. Bayesian MCMC Model Estimation

Since we do not directly observe how individual consumers formulate their budgets each period, nor do we observe their implicit mental account balances, we must estimate budget weights, mental accounts, and cognitive frictions as latent time series variables, $\{\boldsymbol{\theta}_{it}, \boldsymbol{a}_{it}, \boldsymbol{\Gamma}_{it}\}_{t=1}^{T_i}$. Budgeting decisions are thus treated as parameters by the econometrician. Changes to budgets are conditionally identified by level shifts in average spending. The unknown initial mental account balance $\boldsymbol{a}_{i1}$ is conditionally identified by the first period deviation from average spending. Under prior assumptions on the marginal distributions of model parameters, we use Bayesian learning to estimate the posterior distribution of parameters given data, where the integration over uncertainty is performed with a Metropolis within Gibbs Markov Chain Monte Carlo (MCMC) sampling algorithm (Hastings 1970; Geman and Geman 1984).

Note that the data panel is unbalanced, so the time index is technically agent dependent t_i, but we will suppress the dependency of t on i to avoid bogging down our exposition with cluttered notation. Due to the vast degree of observed heterogeneity across agents, most of our estimation routine operates only on agent-level parameters. Because of the need to sample latent time series of consumer decisions, the parameter space is huge, containing over four million parameters.[18] Descriptive lists of all parameters sampled in the model are presented, for ease of reference, in Table 3.2 after we describe our prior distributional assumptions.

3.4.1. Prior Distributional Assumptions

In this section we describe the consumer-specific likelihood function associated with observing given combinations of expenditure, income, and balances. We then describe several prior distributional assumptions on agent-specific parameters and global hyper-parameters. In Section 3.4.2 we discuss the main procedure used to infer latent budget weights, latent mental accounts, and latent cognition shocks.

Given our assumption that the expenditure errors ζ_{ijt} are *iid* and Gaussian normal with agent- and commodity-specific variances σ^2_{ij}, let $\phi(\cdot)$ denote the standard normal probability density function and $\Phi(\cdot)$ denote the associated cumulative distribution function. To accommodate the zero lower bound on expenditure, we assume that the likelihood of realizing agent i's data under the given structural parameterization takes a type I Tobit form (Tobin 1958):

$$\mathcal{L}_i\Big(\underbrace{\{\boldsymbol{x}_{it}, l_{it}\}_t}_{\text{Data}} \mid \underbrace{\{\{\boldsymbol{\theta}_{it}, \boldsymbol{a}_{it}\}_t, \boldsymbol{\sigma}^2_i, \gamma_i\}}_{\text{Structural Parameters}}\Big) = \prod_t \prod_j \underbrace{\left(\frac{1}{\sigma_{ij}}\phi\left(\frac{x_{ijt} - \theta_{ijt}l_{it} - \gamma_i a_{ijt}}{\sigma_{ij}}\right)\right)^{\mathbf{1}(x_{ijt}>0)} \left(1 - \Phi\left(\frac{\theta_{ijt}l_{it} + \gamma_i a_{ijt}}{\sigma_{ij}}\right)\right)^{\mathbf{1}(x_{ijt}\leq 0)}}_{\mathcal{L}_{ijt}(x_{ijt}, l_{it} \mid \theta_{ijt}, a_{ijt}, \sigma^2_{ij}, \gamma_i)}$$

In a slight abuse of notation we denote the agent i, commodity group j, and period t likelihood of observing x_{ijt} and l_{it}, $\mathcal{L}_{ijt}(x_{ijt}, l_{it} \mid \theta_{ijt}, a_{ijt}, \sigma^2_{ij}, \gamma_i)$.[19]

Inference on the agent's budget weights θ_{ijt} and mental account balances a_{ijt} operates on the inverted ex-ante expected indirect utility first-order condition in (3.8). Note that the entire time series of mental account balances $\boldsymbol{a}_{it}$ and by extension budget weights $\boldsymbol{\theta}_{it}$ depends on the initial balances $\boldsymbol{a}_{i1}$ which are unknown parameters. Therefore, we give these initial balances $\boldsymbol{a}_{i1}$ a normal prior structure:

$$a_{ij1} \sim \mathcal{N}(0, \sigma^2_{a_{ij1}})$$

$\sigma^2_{a_{ij1}}$ is fixed for each good and each agent. We set this value to correspond to the expenditure variance for good j over the whole time series of observations divided by average income $\widehat{l_i}$, specifically $\sigma^2_{a_{ij1}} = \text{var}(x_{ijt})/\widehat{l_i}$. Given $\boldsymbol{a}_{i1}$, initial values of θ_{ij1} are then set every iteration of the sampler by passing $\boldsymbol{a}_{i1}$ and data for $t = 1$ through the integral in (3.8). Finally, amongst the budgeting variables, cognitive shocks Γ_{ijt} occur with probability ψ_{ij}, which we give a flat

[18]The exact dimension of the parameter space is 4,664,212 with 4,598,976 of these parameters being the values of latent, time-dependent mental accounting variables.

[19]This notation will come in handy when discussing the sampling steps for latent budgeting parameters.

uniform prior:

$$\psi_{ij} \sim U[0,1]$$

In our various estimation exercises we fix either the mental accounting weight to $\gamma_i = 1$ (baseline estimation) or $\gamma_i = 0$ (estimation with no serial dependence). For the utility weights we require $\sum_{j=1}^{J} \alpha_{ij} = 1$ with $0 \leq \alpha_{ij} \leq 1$ for all $j \leq J$. Following Geary (1950), Stone (1954), and Deaton and Muellbauer (1980b), these values encode the shares of total expenditure devoted toward category j consumption. We fix these values as the mean expenditure shares taken over the entire time series we observe for each consumer: $\widehat{\alpha}_{ij} = \frac{1}{T_i} \sum_t \left(\frac{x_{ijt}}{\sum_j x_{ijt}} \right)$. Note that m_i and $\alpha_{i,J+1}$ cannot be separately identified. We sample $\alpha_{i,J+1}$ explicitly while fixing the consumption borrowing limit m_i to be 20% of an agent's annualized income.[20] From simulations we have inferred that $\alpha_{i,J+1}$ is weakly identified, so we impose a three-tier hierarchical prior structure with empirical priors for the global hierarchical mean and variance:

$$\begin{aligned} &\alpha_{i,J+1} \sim \mathcal{N}(\mu_{i,\alpha_{i,J+1}}, 10^2) \quad \text{with} \quad \mu_{i,\alpha_{i,J+1}} \sim \mathcal{N}(\overline{\mu}_\mu, \overline{\sigma}^2_\mu) \\ &\text{where} \quad \overline{\mu}_\mu = \frac{1}{I} \sum_i \mu_{i,\alpha_{i,J+1}} \quad \& \quad \overline{\sigma}^2_\mu = \frac{1}{I-1} \sum_i (\mu_{i,\alpha_{i,J+1}} - \overline{\mu}_\mu)^2 \end{aligned}$$

We allow the prior variance on the draw of the structural parameter $\alpha_{i,J+1}$ to be large and the same across all agents, inducing flatness for the initial draw. In our simulations, we found that the sampler achieves convergence more efficiently under this structure than one in which the prior variance on $\alpha_{i,J+1}$ is explicitly sampled for each agent.

The orthogonality assumption on error terms such that $\zeta_{ijt} \perp \zeta_{ij',t}$ allows us to impose separate, agent-specific, hierarchical independent gamma priors on the precisions $\frac{1}{\sigma^2_{ij}}$:[21]

$$\frac{1}{\sigma^2_{ij}} \sim \text{Gamma}\left(\tau_{ij}, \frac{1}{\beta_{ij}} \right) \quad \text{with} \quad \tau_{ij} \propto \frac{\beta_{ij}^{\tau_{ij}}}{\Gamma(\tau_{ij}) \sigma_{ij}^{2(\tau_{ij}-1)}} \quad \& \quad \beta_{ij} \sim \text{Gamma}(1,1)$$

The non-conjugate prior on τ_{ij} is discussed in Fink (1997). Its posterior can be easily sampled using a Metropolis step. The rate parameter β_{ij} has a conjugate gamma posterior under a gamma prior.

3.4.2. Estimating Latent Budgeting Variables

Let n index the atomic posterior estimation draws from the Metropolis with Gibbs sampler. Let $\mathcal{P}_{in}$ be an atomic draw of parameters associated with agent i on iteration n, and let $\mathcal{H}_n$ be the

[20] Annualized income is computed by multiplying agent i's weekly average income, $\widehat{l}_i = \frac{1}{T_i} \sum_t l_{it}$, by 52.

[21] Throughout this paper we use the shape-scale parameterization of the gamma distribution such that if $(1/\sigma^2) \sim \text{Gamma}(\tau, 1/\beta)$ then $1/\sigma^2$ has probability density function

$$\frac{\beta^\tau}{\Gamma(\tau)\sigma^{2(\tau-1)}} \exp\left\{ -\frac{\beta}{\sigma^2} \right\}$$

where $\Gamma(\tau)$ is the gamma function.

Table 3.2.: List of Parameters

Parameter	Description
Agent-Level	
γ_i	Governs mental accounting serial dependence
$\boldsymbol{\alpha}_{i,1:J}$	J-dimensional Stone-Geary utility parameters
m_i	Borrowing limit
$\boldsymbol{\sigma}^2_{\boldsymbol{a}_{i1}}$	J-dimensional prior variance for first period $\boldsymbol{a}_{i1}$
$\boldsymbol{\sigma}^2_i$	J-dimensional variance of structural expenditure errors
$\boldsymbol{\tau}_i$	J-dimensional shape parameter for gamma priors on $\boldsymbol{\sigma}^2_i$
$\boldsymbol{\beta}_i$	J-dimensional rate parameter for gamma priors on $\boldsymbol{\sigma}^2_i$
$\alpha_{i,J+1}$	Controls contribution of cash demand to utility
$\mu_{i,\alpha_{i,J+1}}$	Prior mean for $\alpha_{i,J+1}$
$\boldsymbol{\psi}_i$	J-dimensional probability of budget share update
$\{\boldsymbol{\theta}_{it}\}_{t=1}^{T_i}$	Sequence of J-dimensional budget shares
$\{\boldsymbol{\Gamma}_{it}\}_{t=1}^{T_i}$	Sequence of J-dimensional indicators for budget share switches
$\{\boldsymbol{a}_{it}\}_{t=1}^{T_i}$	Sequence of $J+1$-dimensional mental account balances
Global	
$\overline{\mu}_\mu$	Hierarchical mean for $\mu_{i,\alpha_{i,J+1}}$
$\overline{\sigma}^2_\mu$	Hierarchical variance for $\alpha_{i,J+1}$ and $\mu_{i,\alpha_{i,J+1}}$

set of global hierarchical parameters common to all agents.[22] Let $\mathcal{D}_i$ denote the set of agent i's data. These sets are:

$$\mathcal{P}_{in} = \left\{\boldsymbol{\sigma}^2_{in}, \boldsymbol{\tau}_{in}, \boldsymbol{\beta}_{in}, \alpha_{i,J+1,n}, \mu_{i,\alpha_{i,J+1},n}, \boldsymbol{\psi}_{in}, \{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}\right\}$$
$$\mathcal{H}_n = \{\overline{\mu}_{\mu_n}, \overline{\sigma}^2_{\mu_n}\}$$
$$\mathcal{D}_i = \left\{\{\boldsymbol{x}_{it}, l_{it}, r_t, b_{it}, \boldsymbol{p}_t\}_{t=1}^{T_i}\right\}$$

The sampler proceeds in several blocks. Here, we describe the algorithm used to make atomic draws of the budgeting parameters $\{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}$ and leave the description of the remaining Gibbs blocks for Appendix C.2. The algorithm for sampling the cognitive shocks Γ_{ijt} is based on the latent time series level shift identification algorithm from McCulloch and Tsay (1993). While ours is not a standard linear time series model, the spirit of the exercise is the same: to infer whether a change to the budget weight for commodity group j is made from period $t-1$ to t, exploit information provided by data observations in periods t to T_i.

Let $\gamma_i = 1$ and take the non-budgeting parameters $\mathcal{P}_{in} \setminus \left\{\{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}\right\}$ as given, where $\setminus$ is the set exclusion operator. We want to estimate the posterior distribution of $\{\boldsymbol{\Gamma}_{it}\}_t$. Ignoring MCMC iteration indexing for now, consider the sequence $\{\boldsymbol{a}_{it}\}_{t>1}$, which, given $\{\boldsymbol{\theta}_{it}\}_t$ and data, is completely determined by the law of motion (3.5). For each j, recursively substitute out a_{ijt},

[22] Recall, parameters fixed for the entirety of the MCMC sampling routine are γ_i, m_i, α_{ij} for all $j \in \{1, \ldots, J\}$, and $\boldsymbol{\sigma}^2_{a_{i1}}$.

$a_{ij,t-1}, \ldots, a_{ij2}$ to get

$$\mathbb{E}_{it} x_{ijt} = \sum_{s=1}^{t} \theta_{ijs} l_{is} - \sum_{s=1}^{t-1} x_{ijs} + a_{ij1}$$

Thus, for each t, $\mathbb{E}_{it} x_{ijt}$ can be written solely as a function of a_{ij1}. Similarly, ϑ^*_{iyt}, where $\iota_{iyt} = j$ is a function only of the vector of first period mental accounting state variables $\boldsymbol{a}_{i1}$. Re-introducing MCMC indexing, take $\{\boldsymbol{\Gamma}_{it,n-1}\}_{t=1}^{T_i}$ as given. We use a Metropolis step to draw $\boldsymbol{a}_{i1n}$, stepping through from $t = 1$ to T_i to update $\theta_{ijtn} = \vartheta^*_{iytn}(\boldsymbol{a}_{i1n})$ whenever $\Gamma_{ijt,n-1} = 1$ and setting $\theta_{ijtn} = \theta_{ij,t-1,n}$ whenever $\Gamma_{ijt,n-1} = 0$. For each t, we step through, from $j = 1$ to J, and check $\Gamma_{i1t,n-1}, \Gamma_{i2t,n-1}, \ldots, \Gamma_{iJt,n-1}$ evaluating the integral in (3.8) sequentially. Thus, on any given draw, updated values of ϑ^*_{iytn} depend on values $y' < y$ from the current iterate n and potentially $y'' > y$ from the previous iterate $n-1$, as well as past values $\boldsymbol{\theta}_{isn}$, $s < t$:

$$\vartheta^*_{iytn}(\vartheta^*_{i1tn}, \ldots, \vartheta^*_{i,y-1,tn}, \vartheta^*_{i,y+1,t,n-1}, \ldots, \vartheta^*_{i,k_{it},t,n-1}, \boldsymbol{\theta}_{i,t-1,n}, \ldots, \boldsymbol{\theta}_{i1n})$$

Having obtained the candidate sequence $\{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{it,n-1}\}_{t=1}^{T_i}$, we can accept or reject the draw of $\boldsymbol{a}_{i1n}$ using a standard Metropolis acceptance rule. If the draw is rejected, the entire sequence is reset so that $\theta_{ijtn} = \theta_{ijt,n-1}$ and $a_{ijtn} = a_{ijt,n-1}$ for all j and t.[23]

Having performed the Metropolis step for $\boldsymbol{a}_{i1n}$, we proceed to use Gibbs sampling to now jointly draw the sequences $\{\boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}$ and $\{\boldsymbol{\theta}_{itn}\}_{t=1}^{T_i}$. For $t = 1$, we assume $\Gamma_{ij1n} = 1$, so that $k_{i1} = J$ and $\theta_{ij1n} = \vartheta^*_{ij1n}$ always. For $t > 1$, θ_{ijtn} depends on past values $\theta_{ij,t-1,n}$, $\theta_{ij,t-2,n}$, $\ldots$, $\theta_{ij,t-s,n}$ where s is the number of periods that have passed since the last budget update, $\Gamma_{ij,t-s,n} = 1$. Since $\theta_{ij,t-s,n}$ is a function of $\Gamma_{ij,t-s,n}$ then by extension so is θ_{ijtn}. Thus, any new draw of Γ_{ijtn} will affect all future realizations of the budget weights, $\{\theta_{ijsn}\}_{s=t}^{T_i}$. Further, since updating Γ_{ijtn} induces a change to the sequence $\{\theta_{ijsn}\}_{s=t}^{T_i}$, values of $\{a_{ijsn}\}_{s=t}^{T_i}$ are also affected and must be updated to ensure the equilibrium mental account balance law of motion in (3.5) always holds.

Let $\mathcal{D}_{i,t:T_i} \subseteq \mathcal{D}_i$ that only contains data for observed periods t to T_i. Let $\mathcal{P}_{i,t:T_i,n} \subseteq \mathcal{P}_{in}$ similarly represent agent i's parameter set containing only budgeting parameters for periods t to T_i, $\mathcal{P}_{i,t:T_i,n} = \mathcal{P}_{in} \setminus \{\{\boldsymbol{\theta}_{isn}, \boldsymbol{a}_{isn}, \boldsymbol{\Gamma}_{isn}\}_{s=1}^{t-1}\}$. Sampling $\{\Gamma_{ijtn}\}_{t>1}$ proceeds as follows. If $\Gamma_{ijt,n-1} = 0$, then set $\Gamma_{ijtn} = 1$ and compute the sequence of budget weight policies $\{\theta_{ijsn}(\Gamma_{ijtn}, \Gamma_{ij,t+1,n-1}, \ldots, \Gamma_{ijs,n-1})\}_{s \geq t}$, re-computing ϑ^*_{iysn} and setting $\theta_{ijsn} = \vartheta^*_{iysn}$ for $\iota_{iys} = j$ whenever $\Gamma_{ijs,n-1} = 1$ and $s > t$. Future values of θ_{ijsn} will change because they depend on a_{ijsn} which in turn depends on past values of budget weights. After computing the new sequence of implied budget weights, evaluate the

[23] In practice, this Metropolis step involves evaluating a joint draw of all agent-level parameters associated with a conditionally non-conjugate posterior. These parameters are $\boldsymbol{a}_{i1n}$, $\boldsymbol{\tau}_{in}$, and $\alpha_{i,J+1,n}$. Further, we use the adaptive algorithm of Atchade and Rosenthal (2005) to achieve convergence towards a theoretically optimal acceptance rate of 0.234 (Roberts, Gelman, and Gilks 1997). All Metropolis draws are agent-specific, so each agent's parameters are associated with a unique adaptation step and thus a unique acceptance rate. This is described in more detail in Appendix C.2.

posterior probability of observing a budget update:

$$
\begin{aligned}
&\mathbb{P}\Big(\Gamma_{ijtn}=1 \mid \mathcal{D}_{i,t:T_i}, \mathcal{P}_{i,t:T_i,n} \setminus \{\Gamma_{ijtn}\}, \mathcal{H}_n\Big) \\
&= \psi_{ijn} \prod_{s\geq t} \mathcal{L}_{ijsn}\Big(\theta_{ijsn}(\Gamma_{ijtn}=1)\Big)\mathbf{1}\{\theta_{ijsn}\in(-1, m_i/\widehat{l}_i)\}\mathbf{1}\{a_{ijsn}\in(-\widehat{l}_i,\widehat{l}_i)\} \\
&\Bigg/ \Bigg[\psi_{ijn} \prod_{s\geq t} \mathcal{L}_{ijsn}\Big(\theta_{ijsn}(\Gamma_{ijtn}=1)\Big)\mathbf{1}\{\theta_{ijsn}\in(-1, m_i/\widehat{l}_i)\}\mathbf{1}\{a_{ijsn}\in(-\widehat{l}_i,\widehat{l}_i)\} \\
&+(1-\psi_{ijn}) \prod_{s\geq t} \mathcal{L}_{ijsn}\Big(\theta_{ijs,n-1}(\Gamma_{ijt,n-1}=0)\Big)\mathbf{1}\{\theta_{ijs,n-1}\in(-1, m_i/\widehat{l}_i)\}\mathbf{1}\{a_{ijs,n-1}\in(-\widehat{l}_i,\widehat{l}_i)\}\Bigg]
\end{aligned}
\tag{3.10}
$$

This is the probability of simultaneously realizing Γ_{ijt} and θ_{ijt}, where θ_{ijt} is exactly determined by an equilibrium function of data and other parameters. In a slight abuse of notation we index $\theta_{ijs,n-1}(\Gamma_{ijt,n-1}=0)$ and $a_{ijs,n-1}(\Gamma_{ijt,n-1}=0)$ with $n-1$ even though for any $t<s$, θ_{ijsn} and a_{ijsn} will each have been re-evaluated and updated previously to ensure that the equilibrium conditions hold exactly for the entire sequence at each draw of Γ_{ijtn} for each t. Under Gibbs sampling, with some probability we accept that $\Gamma_{ijtn}=1$, or we reject it and set $\Gamma_{ijtn}=\Gamma_{ijt,n-1}=0$, and the budget weight sequence reverts back to that previously computed as a result of following the same procedure to sample $\Gamma_{ij,t-1,n}$. This procedure is repeated, stepping forward in t. In this process θ_{ij1n} is computed once, θ_{ij2n} twice, and θ_{ijtn} t times, once at every draw of Γ_{ijsn} for all $s\leq t$. Further, this process is performed for each j, starting at $j=1$ and on to $j=J$. Note that, despite the fact that ϑ^*_{iytn} depends on $\boldsymbol{\vartheta}^*_{i,-ytn}$ (other optimal budget re-evaluations) and thus, by extension $\boldsymbol{\Gamma}_{i,-jtn}$, we do not re-evaluate the entire vector $\boldsymbol{\vartheta}^*_{itn}$ every time a change is made to Γ_{ijtn} for each particular j. Thus, other values of $\boldsymbol{\theta}_{i,-jtn}$ are treated as parameters and taken as given when re-evaluating θ_{ijtn}. The relevant likelihood function is then only that for commodity group j under the assumption that the components of $\boldsymbol{\zeta}_{it}$ are independent.

3.4.3. Data Variation Identifying Budget Weight Updates

Since our estimation is of the Bayesian learning variety, different parameters may be identified to varying degrees. The key is to assess whether or not prior distributions imposed by the modeler appear different than their posterior estimates. For parameters that are not time-dependent we perform Kolmogorov-Smirnov tests of the null hypothesis that the prior distribution and estimated posterior distributions are the same, H_0 : Prior = Posterior. This assessment is performed at the agent-level when parameters are agent-dependent and at the global level when they are not. The results of these tests are summarized in Appendix C.3. We find that the null hypothesis is rejected for almost all agents and all parameters, suggesting the data provides substantial influence over our parameter estimates so that the results are not driven solely by prior assumptions.

Such a statistical assessment of identification for time-varying mental accounting parameters is more difficult. $\boldsymbol{\theta}_{it}$ and $\boldsymbol{a}_{it}$ are pre-determined by all the other parameters and so are not directly, probabilistically sampled. Suppose for sake of argument, $\boldsymbol{\theta}_{it}$ and $\boldsymbol{a}_{it}$ were given proper prior distributions and probabilistically sampled, not subject to the constraints on their values imposed by the equilibrium modeling conditions. That is, consider the expenditure relation $x_{ijt} = \theta_{ijt} l_{it} + a_{ijt} + \zeta_{ijt}$ strictly in some reduced-form sense, where θ_{ijt} and a_{ijt} are time-varying slopes and intercepts. Then for any given error variance σ^2_{ij}, there exists a thick vein of possible pairs of $(\boldsymbol{\theta}_{it}, \boldsymbol{a}_{it})$ that are all equally likely realizations. Under our model structure however, the likelihood of these realizations are restricted by theory. For any atomic parameter draw, excluding the budget weights in category j, $\mathcal{P}_{in} \setminus \left\{\{\theta_{ijtn}\}_{t=1}^{T_i}\right\} \cup \mathcal{H}_n$, equation (3.8) ensures that there exists a unique sequence $\{\theta_{ijtn}\}_{t=1}^{T_i}$ that conditionally satisfies the personal mental accounting equilibrium. Thus, if we can sample the parameters in $\mathcal{P}_{in}$ and $\mathcal{H}_n$, then we can conditionally identify a unique sequence of budget weights. Note, however, that there is no explicit data variation that identifies the levels of the budget weights, but rather such levels are identified conditional upon the other parameters.

Identifying atomic realizations of the cognitive shocks driving budget updates $\{\Gamma_{ijtn}\}_{t=1}^{T_i}$ relies on exploiting information regarding changes in the average levels of the expenditure time series, conditional on income, over multiple periods, as discussed for more general time series level changes in McCulloch and Tsay (1993). Take σ^2_{ijn}, the likelihood variance, and a_{ij1n}, the initial mental account balance, as given. Also, suppose all future budget-updating indicators $\{\Gamma_{ijs,n-1}\}_{s>t}$ are taken as given from the previous iteration of the MCMC sampler. Recall further that ζ_{it} is mean zero. Suppose for illustration we observe $x_{ijt} >> \theta_{ij,t-1,n} l_{it} + \sum_{s=1}^{t-1}(\theta_{ijsn} l_{is} - x_{ijs}) + a_{ij1n}$, but after this particular deviation of expenditure we observe $x_{ij,t+s'} \approx \sum_{s=t+s'}^{S}(\theta_{ij,t-1,n} l_{is} - x_{ijs}) + \sum_{s=1}^{t-1}(\theta_{ijsn} l_{i,s} - x_{ijs}) + a_{ij1n}$ for some $S \geq t$ and all s' such that $t + s' \leq S$.[24] That is, observations of future expenditure are only small, idiosyncratic deviations from expected expenditure. Then the Gibbs sampler would probabilistically conclude $\Gamma_{ijtn} = 0$, and no switch is necessary, $\theta_{ijtn} = \theta_{ij,t-1,n}$. Now suppose instead for most s', $x_{ij,t+s'} >> \sum_{s=t+s'}^{S}(\theta_{ij,t-1,n} l_{is} - x_{ijs}) + \sum_{s=1}^{t-1}(\theta_{ijsn} l_{is} - x_{ijs}) + a_{ij1n}$. This compounds and leads the period t likelihoods to shift further and further out, so that each successive expenditure realization $x_{ij,t+s'}$ appears less and less likely to the Gibbs sampler. The sampler thus infers that a permanent budget update has been made, i.e. $\Gamma_{ijtn} = 1$ and computes ϑ^*_{iytn} with $\iota_{iyt} = j$ according to (3.8). Joint variation in expenditure, income, and balances, via both the spending constraints and the inversion formula, informs the probabilistic identification of cognitive shocks leading to budget updates.

3.5. Estimation Results

We perform several different passes of the MCMC estimation algorithm and assess and compare the model fit of each in order to understand the degree to which different budgeting frictions and mental accounting features affect predictive inference. First, we estimate the full model with all the bells and whistles described in the sampling algorithm. Measures of fitness for this model are in the first row of Table 3.3, where we allow ψ_{ij} to be interior to the unit interval and $\gamma_i = 1$. Then we turn off different model features one at a time and re-run the estimation routine assuming the following: 1) no reference dependence, $\psi_{ij} \in (0,1)$ and $\gamma_i = 0$; 2) no

[24] Note that it is possible that $s' = 0$ if $\Gamma_{ij,t+1,n-1} = 1$. Further, if S' is the next period at which $\Gamma_{ijS',n-1} = 1$, then $S = S' - 1$, so that it is possible that $S = t$.

cognitive budget updating frictions, $\psi_{ij} = 1$ and $\gamma_i = 1$; 3) constant budget weights, $\psi_{ij} = 0$ and $\gamma_i = 1$; 4) constant budget weights with no reference dependence, $\psi_{ij} = 0$ and $\gamma_i = 0$. The latter estimation routine is designed to best mimic a classical two stage budgeting problem with ex-ante preference uncertainty and constant expenditure share preference weights.

Table 3.3.: Model Performance and Comparisons

				M-H Acceptance Rates		
Model	MAE / $\widehat{l}_i$ [a]	lppd [b]	lppd % Baseline [c]	Mean	Median	Mode
$\psi_{ij} \in (0,1), \gamma_i = 1$	0.236	-208,208,458	—	0.159	0.196	0.231
$\psi_{ij} \in (0,1), \gamma_i = 0$	0.131	-208,209,461	-4.818×10^{-4}	0.220	0.225	0.230
$\psi_{ij} = 0, \gamma_i = 1$	3904.927	-208,208,979	-2.504×10^{-4}	0.060	0.017	0.008
$\psi_{ij} = 0, \gamma_i = 0$	180,077.100	-208,209,690	-6.368×10^{-4}	0.178	0.194	0.215
$\psi_{ij} = 1, \gamma_i = 1$	0.148	-208,209,783	-5.920×10^{-4}	0.199	0.230	0.231

[a] MAE / $\widehat{l}_i = \frac{1}{IJ}\sum_i \sum_j \frac{1}{T_i}\frac{\sum_t |\widehat{x}_{ijt} - x_{ijt}|}{\widehat{l}_i}$ where the predictive mean is $\widehat{x}_{ijt} = \frac{1}{N}\sum_n \widetilde{x}_{ijtn}$ with atomic MCMC predictions $\widetilde{x}_{ijtn}$.

[b] lppd is the log-pointwise predictive density which we take from Gelman et al. (2013a). This number is $\text{lppd} = \sum_i \sum_t \sum_j \ln\left(\frac{1}{N}\sum_n \phi\left(\frac{\widetilde{x}_{ijtn} - x_{ijt}}{\sigma_{ijn}}\right)\right)$. Bigger numbers (the closer to 0 in our case) indicate better model fits.

[c] Since the baseline estimation of the full model with $\psi_{ij} \in (0,1)$ and $\gamma_i = 1$ has the largest lppd, this value represents the percentage loss of information in other models' predictive power relative to the baseline. Negative numbers indicate worse fit than baseline, positive better. Note that all values are negative, so that the full baseline model fits best.

In each estimation, we initialize the MCMC integration scheme with draws from the prior distributions, except for $\boldsymbol{a}_{i1}$ which we set to $\mathbf{0}$ to start. We then proceed to iterate through a chain of length 100,000. That is, we operate on the agent-level Metropolis within Gibbs blocks 100,000 times and the global blocks 100,000/100 = 1,000 times. We set a burnin period of 30,000, keeping every 100^{th} agent-level draw and every global draw thereafter, giving us a total of 701 draws after burnin and trimming. The sampler operates with agent-level blocks parallelized over a 128 core computer, requiring just under 48 hours to completion. For the internal Gaussian quadrature routine on (3.8), we use $4^4 = 256$ quadrature points.

In Table 3.3 we present measures of model performance and fitness.[25] In the full baseline estimation, ψ_{ij} is fully sampled and allowed to vary in $(0,1)$ and agents' expenditure time series are serially dependent via a_{ijt} since $\gamma_i = 1$. Our two main summary statistics comparing model fit are 1) mean absolute predictive error (MAE) as a fraction of each agent's average income $\widehat{l}_i$ and 2) a log-pointwise predictive density (lppd) information criterion as described in Gelman et al. (2013a). As a measure of model convergence and performance we also present the mean, median, and modal Metropolis acceptance rates for the agent-specific Metropolis within Gibbs sampler. Model fit is notably poor when budget weights are assumed constant, $\psi_{ij} = 0$. When budget updating frictions are lifted ($\psi_{ij} = 1$ with $\gamma_i = 1$) and serial dependence is shut off ($\gamma_i = 0$ with $\psi_{ij} \in (0,1)$), MAE as a fraction of average income is lower than in the full model, yet the lppd information criterion suggests the full model is still a better fit. The full baseline model is indeed our preferred specification because of this, so we refer to its parameter estimates

[25] We additionally present a grid of plots for the various agent-level parameter means and global parameter draws from the MCMC routine in Appendix C.3 to visualize the sampler's autocorrelation in n. Given our burnin period length and trimming, autocorrelation is minimal and the sampler appears to have converged.

throughout our empirical analysis and use those estimates in counterfactual simulations.

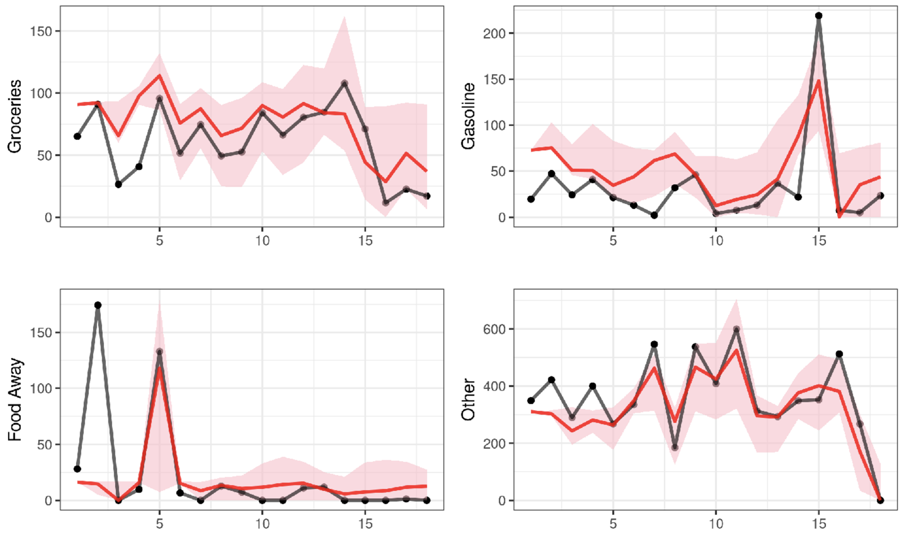


Figure 3.4.: The black line represents weekly spending for the median income agent #2,415, while the red line represents the predictive mean from the baseline model with 95% confidence region in pink.

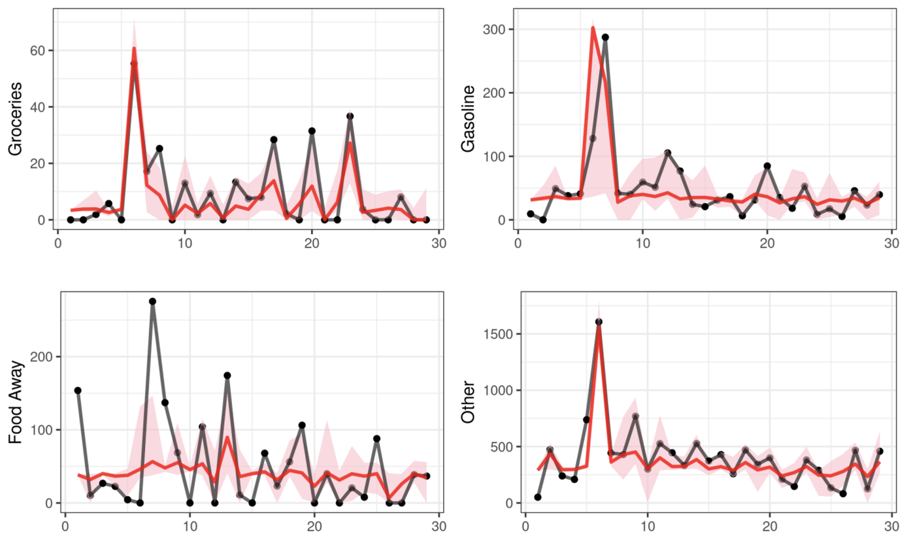


Figure 3.5.: The black line represents weekly spending for the mean income agent #1,795, while the red line represents the predictive mean from the baseline model with 95% confidence region in pink.

Table 3.4.: Posterior Summary Statistics for Agent-Level Means, $\psi_{ij} \in (0, 1)$ & $\gamma_i = 1$

Parameter	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
m_i	242.90	3,524.10	4,784.50	5,311.60	6,449.40	26,867.10	2,686.20
$\widehat{\alpha}_{i1}$	0.008	0.059	0.092	0.106	0.135	0.526	0.064
$\widehat{\alpha}_{i2}$	0.008	0.080	0.130	0.146	0.192	0.686	0.088
$\widehat{\alpha}_{i3}$	0.010	0.079	0.125	0.150	0.197	0.687	0.097
$\widehat{\alpha}_{i4}$	0.042	0.506	0.605	0.598	0.706	0.937	0.148
$\widehat{\sigma}_{a_{i1}}$	0.002	0.009	0.012	0.013	0.015	0.158	0.007
$\widehat{\sigma}_{a_{i2}}$	0.003	0.010	0.013	0.015	0.018	0.271	0.009
$\widehat{\sigma}_{a_{i3}}$	0.003	0.012	0.016	0.018	0.022	0.345	0.011
$\widehat{\sigma}_{a_{i4}}$	0.009	0.025	0.031	0.034	0.038	0.924	0.023
$\widehat{\sigma}_{i1}{}^{\#}$	3.114	17.618	24.948	29.422	35.938	754.703	24.353
$\widehat{\sigma}_{i2}$	2.432	19.863	30.062	45.888	49.646	796.253	54.355
$\widehat{\sigma}_{i3}$	3.478	30.065	46.923	63.459	72.629	1,535.395	63.851
$\widehat{\sigma}_{i4}$	7.934	99.579	153.511	245.444	226.409	98,976.780	2,030.330
$\widehat{\tau}_{i1}$	1.416×10^{-5}	0.124	0.163	0.750	0.303	52.690	2.546
$\widehat{\tau}_{i2}$	2.946×10^{-6}	0.115	0.160	0.997	0.298	35.085	3.340
$\widehat{\tau}_{i3}$	3.362×10^{-5}	0.103	0.139	0.903	0.253	63.582	3.633
$\widehat{\tau}_{i4}$	1.198×10^{-8}	0.081	0.106	0.569	0.189	38.930	2.015
$\widehat{\beta}_{i1}$	0.902	1.112	1.169	1.741	1.303	51.765	2.512
$\widehat{\beta}_{i2}$	0.921	1.102	1.165	1.988	1.293	36.002	3.307
$\widehat{\beta}_{i3}$	0.893	1.087	1.146	1.900	1.264	64.300	3.624
$\widehat{\beta}_{i4}$	0.912	1.064	1.14	1.567	1.202	39.497	2.007
$\widehat{\alpha}_{i,J+1}$	3.319	10.483	11.557	12.166	13.047	52.534	3.336
$\widehat{\mu}_{\alpha_{i,J+1}}$	15.491	15.559	15.576	15.577	15.593	15.751	0.026
$\widehat{\psi}_{i1}$	0.021	0.531	0.704	0.666	0.832	0.963	0.197
$\widehat{\psi}_{i2}$	0.020	0.508	0.674	0.645	0.812	0.965	0.197
$\widehat{\psi}_{i3}$	0.019	0.448	0.599	0.589	0.750	0.955	0.197
$\widehat{\psi}_{i4}$	0.017	0.347	0.540	0.535	0.727	0.971	0.232
$\widehat{\theta}_{i1}{}^{\%}$	0.006	0.057	0.092	0.202	0.142	54.744	1.968
$\widehat{\theta}_{i2}$	0.009	0.079	0.128	0.326	0.200	168.863	4.362
$\widehat{\theta}_{i3}$	0.009	0.081	0.131	0.448	0.217	266.354	6.893
$\widehat{\theta}_{i4}$	−85.897	0.564	0.681	2.158	0.800	1,618.748	40.164
$\widehat{\Gamma}_{i1}$	0.039	0.175	0.307	0.348	0.488	1.000	0.205
$\widehat{\Gamma}_{i2}$	0.036	0.196	0.340	0.369	0.513	1.000	0.205
$\widehat{\Gamma}_{i3}$	0.046	0.260	0.417	0.427	0.575	1.000	0.205
$\widehat{\Gamma}_{i4}$	0.029	0.283	0.476	0.484	0.680	1.000	0.241
$\widehat{a}_{i1}$	−188.904	−4.310	−1.700	−0.305	1.410	1,739.077	36.213
$\widehat{a}_{i2}$	−299.363	−4.045	−0.807	0.420	3.055	1,365.913	33.026
$\widehat{a}_{i3}$	−244.754	−8.967	−3.824	−3.187	1.166	2,727.820	62.525
$\widehat{a}_{i4}$	−427.164.700	−27.907	−8.817	−158.691	6.692	32,189.270	8,556.193
$\widehat{\overline{\mu}}_{\mu}$	0.000	12.351	15.269	15.260	18.036	23.587	3.593
$\widehat{\sigma}_{\mu}$	1.00×10^{-11}	0.537	0.659	0.623	0.708	0.888	0.127

Second subscript corresponds to commodity group — groceries ($j = 1$), auto/gasoline ($j = 2$), food away from home ($j = 3$), and other expenditure ($j = 4$) .

% For time-dependent parameters, θ_{ijt}, Γ_{ijt}, and a_{ijt}, we average over both posterior draws and time for each agent and each commodity category. The statistics presented are summary statistics over these agent-level averages.

To demonstrate pictorially how well the baseline model fits the data, Figures 3.4 and 3.5 present the time series of expenditures along with predictive means (red) from the full model estimation with $\psi_{ij} \in (0,1)$ and $\gamma_i = 1$ for the median and mean income consumers respectively. Summary statistics over agent-level means for parameter estimates are presented in Table 3.4. For parameters that are time dependent, we average both across MCMC draws and over time for each agent. For the global parameters, we simply take summary statistics over the outputted chain. In Table 3.4, note that we present summary statistics for the values of standard deviations $\widehat{\sigma}_{ij}$ and $\widehat{\sigma}_{a_{ij1}}$ as opposed to the variances, so that the reader can assess the likelihood spread in dollar units.

One thing stands out in Table 3.4 amongst the mental accounting parameters. The variances of the posterior distributions of agent-level mental account balance means $\widehat{\boldsymbol{a}}_i$ is rather extreme. Note that some of the likelihood variance estimates $\widehat{\sigma}^2_{ij}$ are also rather large. These high values are not surprising given the spikiness of the various time series. Remember, $a_{ij,t+1,n} = -\zeta_{ijtn}$, so high variance in ζ_{ijt} will lead to high variance in $\widehat{a}_{ij}$ as well. This variability propagates through the model via optimal budget updates ϑ^*_{iytn} leading to some extreme mean outcomes for $\widehat{\theta}_{ij}$ as well with very large budget weights associated with very negative mental account balances. Lacking explicit budgeting data, this posterior variability is to be expected. Yet, while the distributions appear to have fat tails, the first and third quartiles are reasonable and tight, suggesting most values are not so large so as to cast suspicion on the validity of the estimates.

3.5.1. Implied Own-Price and Income Elasticities

In this section we discuss the average own-price and income elasticity estimates generated by the model to understand how well we can improve upon the naive estimates discussed in Section 3.2. Recall that when optimal budgets are set l_{it}, $\boldsymbol{p}_t$, and r_t are known so that all uncertainty acts through $\boldsymbol{\zeta}_{it}$. It can be shown that the commodity group j model-implied own-price ε^p_{ijt} and income ε^l_{ijt} elasticities of demand depend on the responsiveness of consumers' optimal budgeting allocations:

$$\varepsilon^p_{ijt} = \begin{cases} -1 & \Gamma_{ijt} = 0 \\ \frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}\left(\frac{l_{it}}{q_{ijt}}\right) - 1 & \Gamma_{ijt} = 1 \end{cases} \qquad \text{where} \qquad \iota_{iyt} = j$$

$$\varepsilon^l_{ijt} = \begin{cases} \frac{\theta_{ijt} l_{it}}{x_{ijt}} & \Gamma_{ijt} = 0 \\ \frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}\left(\frac{l^2_{it}}{x_{ijt}}\right) + \frac{\vartheta^*_{iyt} l_{it}}{x_{ijt}} & \Gamma_{ijt} = 1 \end{cases} \qquad \text{where} \qquad \iota_{iyt} = j$$

Both ε^p_{ijt} and ε^l_{ijt} are well-defined only when $q_{ijt} > 0$. Exact derivations of these elasticities along with the corresponding total derivatives $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}$ and $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}$ are featured in Appendix C.1.3. Taking the posterior distributions of parameter estimates, for each agent we compute $\widehat{\epsilon}^p_{ij}$ and $\widehat{\epsilon}^l_{ij}$, averaging both across the sampler iterates n and time t. We then compare these results to the constant elasticity estimates from the naive regression. Summary statistics for agent-level means of own-price elasticities are presented in Table 3.5 and income elasticities in Table 3.6.

Turning first to own-price elasticity estimates, note that only optimally chosen ϑ^*_{iyt} with $\iota_{iyt} = j$ depends on p_{jt} (not $\theta_{ijt} = \theta_{ij,t-1}$ in the event $\Gamma_{ijt} = 0$). Engaging in optimal budget

Table 3.5.: Own-Price Elasticities: Full Model vs. Naive Regression

	Full Model [a]						
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
Groceries	-1.815	-1.067	-1.035	-1.052	-1.016	-1.001	0.057
Auto/Gas	-1.905	-1.048	-1.024	-1.038	-1.011	-1.000	0.047
Food Away	-3.162	-1.070	-1.037	-1.058	-1.018	-1.001	0.092
Other	-8.248	-1.016	-1.008	-1.019	-1.004	-1.000	0.160
	Naive Regression [b]						
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
Groceries	-6,378.564	-259.955	-16.554	0.617	280.202	6,028.350	793.806
Auto/Gas	-3,447.366	-51.901	1.300	0.642	63.562	2,177.469	193.927
Food Away	-21,410.260	-322.053	-16.501	2.318	320.747	6,815.177	1,152.328
Other	-5,764.009	-251.189	7.825	-45.156	219.877	9,773.338	823.305

[a] All summary statistics are over agent-level means, conditional upon positive expenditure. That is, we report the distribution of $\frac{\sum_{t_i}\sum_n \epsilon^p_{ijt_i,n}}{N\sum_{t_i}\mathbf{1}(x_{ijt_i}>0)}$.

[b] The regression is $\ln q_{ijt} = \beta_{ij0} + \sum_{k=1}^{4}\beta_{ijk}\ln p_{jt} + \beta_{ij5}\ln l_{it} + \eta_{ijt}$, conditional upon $q_{ijt} > 0$. Summary statistics are over estimated values of β_{ij} for consumption in commodity group j.

adjustments affects price sensitivity in possibly ambiguous ways. Indeed, $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}$ can be positive if k_{it} is close to J and preferences over savings $\alpha_{i,J+1}$ are sufficiently negative, so that consumers either want to borrow or spend down a huge initial balance, b_{i1}. If the marginal propensity to save (MPS) is big enough, where this threshold increases as J increases, then $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}$ will always be negative, and the own-price elasticity of demand will be bounded above at the unit-elastic case. This can be seen by inspecting the parameterized expression for $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}$ in Appendix C.1.3.

Note that $\alpha_{i,J+1}$ is rarely < 0. Observationally, this would occur if an individual consumer were observed to have $z_{it} < 0$ for many periods, which would happen in the event a consumer received a huge one-time windfall gain and used this one-time balance increase to supplement consumption beyond his weekly income.[26] We do not observe that anyone in our dataset exhibits this trait. Looking at posterior estimates of $\widehat{\alpha}_{i,J+1}$, agent-level averages are bounded below by 3.319, so that on average our model predicts for every agent that demand is at least unit elastic in price. This is not necessarily problematic. Levin, Lewis, and Wolak (2017) find that both time and spatial aggregation lead to upwardly biased estimates of the own-price elasticity of demand for gasoline. That is, aggregate estimates are overly inelastic compared to dis-aggregated estimates, which are more likely to be elastic. This jibes with our finding that dis-aggregated own-price elasticity estimates for core commodities tend to be elastic. The timing frictions in our sparse max model have the same implications for price responsiveness as that documented in Gabaix (2014): rational consumers engaging in more budget updates will likely have more elastic ε^p_{ijt}. Table 3.5 shows this empirically, where maximal estimates are very close to -1.000 for all commodity groups. Despite the model's possible downward bias in price estimates, it generates negative elasticities across the board, whereas the naive regression

[26] Note that we take no stand in the model about how one-time windfalls affect consumer behavior.

suggests that over half of consumers treat auto/gas and all other consumption as Giffen. Our model with behavioral frictions can thus uniquely generate price elasticity estimates consistent with neo-classical theory better than a standard reduced-form model that does not control for behavioral budgeting frictions or other demand shifters.

Table 3.6.: Income Elasticities: Full Model vs. Naive Regression

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
				Full Model [a]			
Groceries	0.021	1.250	1.477	1.632	1.824	15.376	0.696
Auto/Gas	0.061	1.216	1.457	1.796	1.917	63.179	1.732
Food Away	-2.683	1.359	1.741	2.252	2.417	60.539	2.365
Other	-261.469	1.288	1.656	3.084	2.423	1,285.233	27.833
				Naive Regression [b]			
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
Groceries	-517.448	-0.307	0.114	0.255	0.519	552.687	16.185
Auto/Gas	-294.083	-0.182	0.167	0.019	0.554	108.915	8.558
Food Away	-307.956	-0.259	0.283	1.866	0.896	4,187.986	84.189
Other	-365.797	0.219	0.631	0.408	1.061	44.310	8.688

[a] Here, we report the distribution of $\frac{\sum_{t_i}\sum_n \epsilon^I_{ijt_i}}{N\sum_{t_i}\mathbf{1}(x_{ijt_i}>0)}$

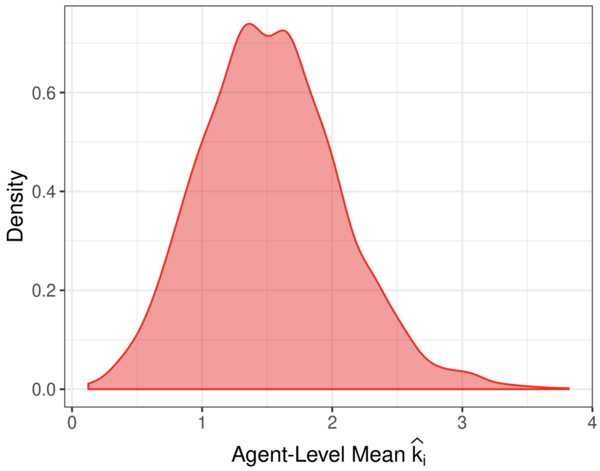
[b] The regression is the same as that used in Table 3.5, except we take summary statistics over the coefficient β_{i5}.

Demand responses to income changes are, by contrast, far more complex. Unlike own-price elasticities, ε^I_{ijt} depends on mental accounting serial dependence a_{ijt}, borrowing limits m_i, balances b_{it}, and the distribution of expenditure shocks ζ_{ijt}, if $\Gamma_{ijt} = 1$ since it depends on the level of ϑ^*_{iyt} which is a function of these underlying objects. Even further, the value of $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}I_{it}}$ depends implicitly on how sensitive optimal budgets are to income in all the other categories $j' \neq j$ for which $\Gamma_{ij',t} = 1$. The full structural model's improvement relative to a naive regression with regards to generating income elasticity estimates consistent with neo-classical theory, is on display in Table 3.6. While over 25% of the sampled agents appear to treat groceries, auto/gasoline, and food away from home as inferior goods in the naive regression ($\widehat{\epsilon}^I_{ij} < 0$), negative elasticities appear only in food away from home and other consumption. In the structural model, one half of consumers appear to elastically respond to income changes, while over 99% still have positive income elasticities, thus treating the commodity groups as neo-classically normal. Together, the model's implied own-price and income elasticity estimates demonstrate the effectiveness of using behavioral theories to account for short-run demand shifts.

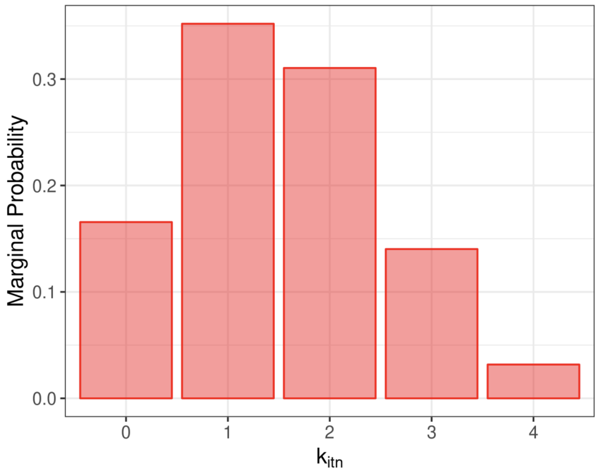
3.5.2. Quantitative Discussion of Mental Accounting Features

Our model estimates can provide empirical inference for both classical and behavioral theories of consumer decision making. Whereas in the previous section we discussed the model's predictions for price and income elasticities and their conformity to classical demand theory, in this section we explore the model's predictions through a behavioral economics lens. In general, agents appear to engage in budget updates around half of the time and approximately half of agents exhibit loss aversion with respect to over spending. We focus on two main results: 1) the full model's predicted number of budget updates k_{it} each period; 2) how the sign and magnitude of budget adjustments $\vartheta^*_{iytn} - \theta_{ij,t-1,n}$, for $\iota_{iyt} = j$, depend on the underlying mental account balance a_{ijtn}.

Figure 3.6a shows that we observe, on average, $\frac{1}{I}\sum_i \widehat{k}_i = 1.53$ budget updates every period for any given consumer. Figure 3.6b shows that the marginal distribution for atomic draws of total budget updates $k_{itn} = \sum_{j=1}^{J} \Gamma_{ijtn}$ is heavily concentrated around $k_{itn} = 1$ and $k_{itn} = 2$. Only about 16.6% of observed weeks across all agents over our entire sample feature atomic draws of $k_{itn} = 0$, i.e. no budget updates, while we infer one to two updates are made over 65% of the time. This estimate reflects the possible presence of cognitive constraints and narrow bracketing. The presence of such constraints, however, does not appear to force agents to stick with budgets for long, extended periods. This suggests agents can exhibit both cognitive processing limitations while not engaging in strongly sticky mental accounting behavior. In fact the posterior average number of weeks between budget updates are 2.42 for groceries, 2.32 for auto/gasoline, 2.11 for food away from home, and 1.88 for all other expenditure.



(a)



(b)

Figure 3.6.: In panel (a) we present the posterior density of $\widehat{k}_i = \frac{1}{N\,T_i}\sum_n \sum_t \sum_j \Gamma_{ijtn}$, i.e. the number of budget changes each period averaged over time and across MCMC draws for each agent, using a smoothed Gaussian kernel with the Silverman rule-of-thumb for choosing the bandwidth (Silverman 1986). In panel (b) we show the marginal posterior probability mass function for the atomic draws of k_{itn}.

To understand the degree to which our model estimates are consistent with various behavioral theories, we assess three basic summary statistics over posterior estimates of budget weights and mental accounting state variables $\{\boldsymbol{\theta}_{it}, \boldsymbol{a}_{it}\}_{t=1}^{T_i}$. In the first two rows of Table 3.7 we present the fractions of our sample of consumers who, on average, appear to decrease their budget weights whenever optimal updates are made, i.e. $\Gamma_{ijtn} = 1$, conditional upon the signs of a_{ijtn}. In row three, we present the fraction of consumers who, on average, appear to always make bigger budget movements, in absolute magnitude, whenever $a_{ijtn} < 0$ versus $a_{ijtn} > 0$. This last

Table 3.7.: Budgeting & Mental Accounting Tendencies

	Groceries	Auto/Gasoline	Food Away	Other	Total [d]
Share Adjust Down, $a_{ijtn} < 0$ [a]	0.522	0.570	0.506	0.462	0.468
Share Adjust Down, $a_{ijtn} > 0$ [b]	0.548	0.582	0.522	0.481	0.479
Share Loss Averse [c]	0.507	0.509	0.508	0.512	0.515

[a] The agent i statistic is $\frac{\sum_n \sum_t (\vartheta^*_{iytn} - \theta_{ij,t-1,n}) \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}<0\right)}{\sum_n \sum_t \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}<0\right)} < 0 \quad \text{for} \quad \iota_{iyt} = j.$

[b] Here, just flip the sign of a_{ijtn}: $\frac{\sum_n \sum_t (\vartheta^*_{iytn} - \theta_{ij,t-1,n}) \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}>0\right)}{\sum_n \sum_t \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}>0\right)} < 0 \quad \text{for} \quad \iota_{iyt} = j.$

[c] For each agent i, the underlying statistic is

$\left|\frac{\sum_n \sum_t (\vartheta^*_{iytn} - \theta_{ij,t-1,n}) \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}>0\right)}{\sum_n \sum_t \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}>0\right)}\right| - \left|\frac{\sum_n \sum_t (\vartheta^*_{iytn} - \theta_{ij,t-1,n}) \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}<0\right)}{\sum_n \sum_t \mathbf{1}\left(\Gamma_{ijtn}=1, a_{ijtn}<0\right)}\right| < 0 \quad \text{for} \quad \iota_{iyt} = j.$

[d] For the "Total" column, we must additionally sum over j.

data point is a measure of loss aversion. We expect consumers' budgets to be more sensitive to $a_{ijtn} < 0$ if they are loss averse, and indeed that is what we find for most consumers in our sample in every consumption category. Note, though, that there is substantial heterogeneity both across consumers and consumption categories for all of these statistics. Consumers are overall more likely to downward adjust both after over and under spending, though on average, slightly more consumers are likely to downward adjust after under spending than over spending, though this difference does not appear significant.

In Table 3.8 we examine how the model-implied explicit budgets, $\theta_{ijt} l_{it}$, change depending on the underlying magnitude and sign of a_{ijt}. These statistics thus describe how budgets change for every dollar of over or under expenditure. Following over spending, explicit income-weighted budgets are more likely to drift upward, whereas after underspending these budgets are more likely to drift downward, suggesting consumers use budget-adjustments to mitigate the disutility of possible future over expenditure, providing evidence that a majority, though not all, of consumers exhibit loss aversion.

We can conclude from our estimation results that consumer expenditure profiles provide no consistent evidence one way or the other that *all* consumers, on the whole, exhibit loss aversion or gain loving behavior, though many consumers indeed do. Rather the model estimation suggests that preferences, budgeting, and mental accounting behavior are all vastly idiosyncratic, which is not surprising given the observed heterogeneity in agent-level expenditure patterns.

3.5.3. Relations Between Mental Accounting, Savings, and Income

To assess whether certain mental accounting features are associated with individual savings and income, we ran some simple, hierarchical regressions. First, we took posterior means of $\widehat{k}_{it}$'s as a left-hand side variable to assess the relations between budgeting and various economic outcomes. Second, we examined the responses of savings z_{it} and savings rates z_{it}/l_{it} to estimated relative over and under saving in the previous period, $\widehat{a}_{i,J+1,t}$. We can better understand a number of aspects of our model under these exercises. If $\widehat{k}_{it}$ is correlated with savings and income, then Γ_{ijt}'s may be endogenous. Further evidence of possible endogeneity

Table 3.8.: Mean Budget Adjustments Relative to Spending Error [a]

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
			Conditional on $a_{ijtn} < 0$				
Groceries	-22.442	-0.221	0.106	0.701	0.745	82.082	3.585
Auto/Gas	-37.973	-0.316	0.156	0.861	1.111	74.082	5.655
Food Away	-144.623	-0.805	0.000	0.714	0.935	202.767	11.914
Other	-882.842	-4.377	0.064	10.746	6.855	2,717.470	86.939
			Conditional on $a_{ijtn} > 0$				
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
Groceries	-35.865	-0.989	-0.243	-0.787	0.073	12.527	2.794
Auto/Gas	-295.611	-1.362	-0.292	-0.869	0.176	72.667	8.512
Food Away	-379.001	-1.342	-0.106	-0.189	0.792	112.495	12.638
Other	-3,828.380	-7.402	-0.485	-6.941	5.435	647.266	109.784

[a] First backwards differences conditional on $\Gamma_{ijtn} = 1$: $\frac{\theta_{ijtn} l_{it} - \theta_{ij,t-1,n} l_{i,t-1}}{|a_{ijt}|}$.

arises if $\widehat{k}_{it}$ is significantly different depending on the sign of $\widehat{a}_{i,J+1,t}$.[27] Assessing how the current period's savings depends on the estimated sign and magnitude of $\widehat{a}_{i,J+1,t}$ can help us understand the degree to which consumers use mental accounting to discipline consumption expenditure by anchoring on their budgets and adjusting them when $\widehat{a}_{i,J+1,t}$ gets too negative, for example.

The exact details of the hierarchical regression specifications we estimate are presented in Appendix C.4. Here, we summarize the findings. First, there appears again to be widespread heterogeneity in terms of how the predicted number of periodic budget updates depend on savings rates and over or under saving in the previous period. In terms of the degree to which over or under saving in the previous period predicts the total number of budget changes in the current period, there is no strong evidence that either have a significant effect on total budget updates. Indeed, this result lends credence to our assumption that the switches themselves may be exogenous, resulting from random increases in attention rather than coming from an additional, higher-level decision process, where consumers assess how much they have over or under spent and then decide to make changes as a result of that assessment.

With regards to how savings levels and savings rates in period t respond to over or under saving in the previous period, our results suggest that consumers exhibit mild myopia with respect to under saving, as they are likely to engage in successive periods of under saving. Following average under saving in period $t-1$, so that $\widehat{a}_{i,J+1,t} > 0$, absolute savings in period t, z_{it}, is predicted to be lower than if over saving had previously occurred. Specifically, under saving by 1% in period $t-1$ leads to an approximate \$18.72 reduction in total savings in period t or a 6.9% reduction in the savings rate as a fraction of income. Consumers thus, on average, fail to immediately adjust to the downward dynamic pressure on savings induced by $\widehat{a}_{i,J+1,t} > 0$. One period of over spending (under saving) can thus result in successive periods of

[27] Recall that over saving in period $t-1$ implies $a_{i,J+1,t} < 0$ (not positive).

over spending as well, a product of myopia due to infrequent budget updates.

Finally, our analyses of the relationships between budget updates, savings, and income, suggest that consumers respond to positive income shocks by increasing savings rates, even though the number of budget weight changes $\widehat{k}_{it}$ is negatively correlated with income. Combined, these two facts reflect an overall tendency for consumers to both, on average, save out of windfall gains and possibly anticipate how such gains may impact their ability to stick to their budgets. In this manner, the negative correlation between $\widehat{k}_{it}$ and income suggests that cognitive shocks may indeed be endogenous, despite the fact that $\widehat{k}_{it}$ appears, on average, uncorrelated with savings behavior. Future work featuring datasets that explicitly record individual attentiveness with respect to their financial profile, say by recording how often people log in via the internet to check their balances in their various accounts, is needed both to understand how such attentiveness is correlated with spending and savings behavior and subsequently account for any possible endogeneity bias attributable to such correlations.

3.5.4. Counterfactual Welfare Under Relaxed Cognitive Frictions

To understand how consumer welfare is affected by relaxing the sparse max constraint on budget updates, we counterfactually simulate consumer expenditure using the posterior distribution of parameters from the full baseline estimation, except we force $\psi_{ij} = 1$ for all i and j combinations.[28] Each period every agent chooses ϑ^*_{it} by maximizing indirect utility subject to the relaxed constraint that $\Gamma_{ijt} = 1$ for all commodities. Under this counterfactual, the average consumer in our sample experiences a 1.8% reduction in posterior average period utility when $\psi_{ij} = 1$ versus when $\psi_{ij} \in (0, 1)$. For 70.4% of consumers, engaging in weekly budget updates for every commodity category does not lead to welfare improvements relative to the posterior predictive baseline.[29]

We also examine counterfactual versus predicted savings rate volatility over time across agents, finding that a majority (64%) of consumers in our sample would experience wider swings in savings rates period-by-period if they engaged in high-frequency budget updating. However, mean savings rates across the whole sample are on average 1.2% higher under the counterfactually imposed budget updates every period, though the distribution's median is 0.028% with a minimum of -0.22%. There is also no apparent difference in average savings rates between those whose welfare is improved versus those with unimproved welfare under the counterfactual. The results on savings rates thus appear mixed, with little counterfactual change to savings rates for approximately half of consumers. This lends credence to the results from the regression analyses in Section 3.5.3, suggesting that the number of budget updates appear on average uncorrelated with savings rates. Still, some extreme behavior with regards to counterfactual savings and balance accumulation is observed. Just under 1% of consumers go bankrupt under the counterfactual simulation with terminal $\widehat{b}_{i,T_i} < -m_i$, while just over 0.5%

[28]Note that this experiment is not the same as the estimation where we force $\psi_{ij} = 1$ and estimate the additional parameters separately. Here, we use the parameter estimates from the main estimation where ψ_{ij} is a free parameter and $\gamma_i = 1$, counterfactually *only* changing ψ_{ij} by setting it to 1.

[29]We compare the counterfactual predictive utility against the posterior predictive utility from the full baseline model, as opposed to actual utility under the posterior parameterization. The reason for this is that we care about how model *predictions* change, given the parameter change. Changing ψ_{ij} in this manner thus amounts to assessing the total variation in model predictive outcomes under variation in ψ_{ij}.

of the sample become hoarders, with $\widehat{b}_{i,T_i} > m_i$, i.e end balances greater than 20% of imputed annual income.

3.6. Conclusion

We have developed a structural model of mental accounting that features both reference dependence and loss aversion using a standard, quasi-concave, monotone, and continuously-differentiable utility function. By incorporating mental accounting into the classical two stage budgeting model, we can endogenously explain idiosyncratic, short term variation in consumption expenditure patterns. Further, by allowing for narrow choice bracketing and reference dependence, we show that our model with all the bells and whistles from behavioral economics generates elasticity estimates consistent with classical demand theory. These results should encourage future work that seeks to unify well-established classical theories with contemporary behavioral ones in order to structurally explain empirical phenomena in consumer decision making.

Finally, the vast heterogeneity we observe with respect both to consumers' spending patterns and the behavioral implications of those patterns generated by our model should encourage policymakers to be cautious in implementing one-size-fits-all behavioral nudges. For example, while savings rates, on average, are counterfactually higher when consumers are attentive to their budgets, there is substantial heterogeneity across consumers as to how they respond to our counterfactual experiment. Push notifications implemented by banks and financial services companies alerting consumers to how their spending deviates from their budgets could encourage higher savings for some individuals but not others. Our results thus underscore the difficulty to creating incentive-compatible behavioral nudges: what might work for one consumer may backfire for another. Since our conclusions are driven simply by counterfactual simulations, future work should attempt to construct field experiments, perhaps using well-established mobile applications to notify individuals of their spending behavior in order to understand how our simulated predictions compare to real world outcomes.

4. Durables, Non-Durables, and a Structural Test of Fungibility

joint with Alan Montgomery and Christopher Y. Olivola

4.1. Introduction

It has been well-established that consumers[1] appear to use credit cards in ways inconsistent with the permanent income hypothesis (PIH) (see for example Prelec and Loewenstein (1998), Prelec and Simester (2001), Gross and Souleles (2002), Huffman and Barenstein (2005), and Quispe-Torreblanca et al. (2019)). There exists strong empirical evidence that consumers treat other sources of liquidity as non-fungible too.[2] Yet despite these previous findings, the distribution of these behavioral tendencies has not been thoroughly investigated, an important undertaking if economists are to understand how violations of fungibility could affect broad economic outcomes. We add to the literature by showing that a dynamic, neo-classical consumption/savings model of the household can be modified to construct a household-level test of fungibility.[3] Allowing for heterogeneity in consumer preferences, we partially confirm the results outlined above, showing that the median/modal household in our sample behaves as if the resources from different liquidity categories are non-fungible both generally and within

[1]Throughout the paper we will use the words "consumer," "individual consumer", and "household" interchangeably. In our data analysis, the unit of measure is "household," though some households consist of multiple individual persons with one account and some of a single individual person with multiple accounts. These different entities are observationally equivalent, so we extrapolate from potential person-level differences and consider the behavior of all persons within one household together, as a collective.

[2]Heath and Soll (1996) wrote one of the first papers to empirically show that consumers behave as if liquidity were non-fungible across different commodity groups. Souleles (1999) shows that non-durable consumption expenditure is excessively sensitive to tax refunds, as opposed to durable consumption expenditure. Hastings and Shapiro (2013) exploit variation in the prices of different grades of gasoline to show that households behave as if budgets for gasoline consumption are non-fungible. Hastings and Shapiro (2017) reject the hypothesis that households respect the fungibility of money by comparing food expenditure using food stamps verses other payment methods.

[3]We use the term "neo-classical" in its broadest sense to encompass any kind of dynamic household consumption/savings model of the form:

$$V_t(w_t) = \max_{\boldsymbol{c}_t, s_t} \quad u_t(\boldsymbol{c}_t) + \beta \cdot \mathbb{E}_t V_{t+1}(w_{t+1}) \tag{4.1}$$

$$\text{s.t.} \quad \boldsymbol{p}_t \cdot \boldsymbol{c}_t + s_t \leq w_t \quad \forall t \tag{4.2}$$

where $\boldsymbol{c}_t$ is a vector of real consumption commodities, $\boldsymbol{p}_t$ is a vector of market prices, s_t is savings, and w_t is available net resources (wealth plus credit). Further, $u_t(\boldsymbol{c}_t)$ is a strictly increasing, strictly concave, at least twice-continuously-differentiable cardinal utility function which takes as its argument *only* the vector of real consumption, but not savings, wealth, or available liquidity. β is the geometric rate of time preference and $V_t(w_t)$ is the agent's optimal value function. Consumers choose sequences of consumption $\boldsymbol{c}_t$ and savings s_t to maximize their period-t utility from consumption plus the discounted expected future value of all future consumption.

durable and non-durable consumption expenditure. Our estimates however also show that the distribution of this behavior across households is disperse: some households treat liquidity as extremely non-fungible while others essentially treat it as fungible.

Despite the degree of heterogeneity, our results affirm that most households treat credit availability and cash as non-fungible, suggesting they make decisions under some degree of mental accounting. Under mental accounting, consumers to care not simply about the sum of their available resources when making consumption expenditure decisions, but the various *components* of these resources (e.g., credit, cash, savings). A consumer who spends as if his available resources were perfectly fungible, absent any liquifying constraints, would respond to the marginal change in one specific class of resources, say the credit limit on his credit card, the same as if a different class of resource, say the balance on his checking account, had experienced the same marginal change. Non-fungibility, by definition, is the result of the disobedience of this stricture. The best theoretical explanation for this behavioral phenomenon was first explained in Thaler (1985) and extended in Shefrin and Thaler (1988) and Thaler (1990, 1999). Yet, few studies have attempted to understand the non-fungibility of liquidity in terms of its broader economic implications. One of the most impactful which comes to mind is Richard Thaler's 1995 paper with Shlomo Benartzi entitled "Myopic Loss Aversion and the Equity Premium Puzzle" where the authors show that the puzzle can be reconciled by a model with mental accounting features (see Benartzi and Thaler (1995)).[4]

Another important paper which models the interaction between method of payment and consumption utility is that of Prelec and Loewenstein (1998). This work is our main anchoring point, as the authors explore how the relative durability of a specific product impacts a consumer's willingness to hold credit card debt after purchasing that product. The authors hypothesize that consumers will be more likely to hold and pay interest on credit card debt for commodities which generate a dynamic flow of utility as opposed to commodities the utility from which is fully experienced at the moment of first consumption. Durable goods, such as furniture, produce a stream of utility over many periods, while non-durable goods like a box of candy are essentially consumed once and thereafter useless.[5] Under the authors' theory, a consumer should be less likely to purchase the candy on credit, and even if they did, they should be less likely to hold debt, and thus have to pay interest, on that credit card purchase. Using their word choices (though the example is our own), paying $1 in interest on a box of candy one has already consumed is "more painful" than $1 in interest on a desk one uses every day. This is because, in the latter situation, as with all durable goods, the pain of paying is offset by the pleasure one gets from continuing usage of the product. Higher levels of durable goods purchases should lead to larger proportions of monthly credit card balances carried forward to the next period, thus accruing interest. This result is empirically confirmed by analyses on U.K. credit card data in Quispe-Torreblanca et al. (2019) who find that consumers are approximately 9% less likely to pay down their full credit card statement balance after making a durable purchase.

[4]More work is needed to explore the broader economic impacts when consumers and investors engage in mental accounting. In the concluding section of this paper, we discuss potential avenues forward exploring potential broader impacts associated with our analysis.

[5]The empty box, of course we must assume, is garbage with no intrinsic value.

Prelec and Loewenstein (1998) provides the springboard for our main analysis where we compare marginal propensities for real consumption out of changes to both available credit and available checking account balances. We do this for both durable and non-durable consumption to argue that the data appear to suggest that the method of payment is coupled with the commodity class within the consumer's decision process. Why focus on the distinction between durables and non-durables? Durables are purchased in the same way as non-durable goods, yet function more like an investment asset so the dynamics behind durable good accumulation are important. Fernández-Villaverde and Krueger (2011) show in fact that durable good expenditure is skewed over the life-cycle toward younger households. In a fairly standard neo-classical model with geometric discounting, Fernández-Villaverde and Krueger (2011) can match the observed hump-shaped profile of durable expenditure along with the tendency of middle-aged households to accumulate more traditional financial assets rather than durable goods. We demonstrate that consumers appear to increase durable consumption out of increases to credit at a faster rate than checking account balance increases, yet durable goods consumption appears less sensitive across consumers to balance changes than that of non-durable goods (see Figure 4.1). The life-cycle effects of durable consumption may be due to younger households' being more likely to be borrowing or liquidity constrained Krueger and Ludwig 2007; Yang 2009; Fernández-Villaverde and Krueger 2011. Unfortunately we lack data on household members' ages, yet our findings complement those of Krueger and Ludwig (2007), Yang (2009), and Fernández-Villaverde and Krueger (2011), since liquidity-constrained households can engage in consumption expenditure out of credit, yet it is more difficult to turn available credit card credit into financial assets.

Our goal is not to explain a specific puzzle, nor show that *all* consumers behave one way or another. Rather we show that the neo-classical model can be used, with only slight modifications, to assess fungibility at the household level. We do not explicitly construct a model of mental accounting, but instead use the neo-classical economic consumption/savings model to empirically test the null hypothesis implied by that model, specifically that credit card and checking account balances are equally fungible when controlling for their different rates of interest and return. The main structural result demonstrates that for the median/modal consumer the null hypothesis can be rejected with $> 99\%$ certainty. This result is a repudiation of fungibility for some consumers within the neo-classical construction, however it is not an absolute repudiation of the neo-classical consumption/savings construction entirely.

The main result we present says that: 1) given individuals face a standard, concave, monotone, at least twice differentiable utility function that is strongly separable over broad commodity groups, *and* 2) that this utility function takes as its arguments consumption commodities *only* and not methods of payment for said commodities, *and* 3) that individual consumers discount the future at a geometric rate, the neo-classical consumption/savings model cannot rationalize, for every consumer, the observed variation in expenditure out of credit cards and checking accounts, absent additional gimmicks. This should be read as evidence in support of alternative models of consumer behavior, foremost mental accounting which allows for consumers to consider the method of payment when making expenditures. The heterogeneity we observe with respect to the degree to which households treat liquidity as fungible opens the door to broader questions which we do not attempt to answer here but address in the concluding section.

This paper proceeds as follows: in Section 4.2 we present some important features of our dataset and discuss their implications for assessment of household-level behavior under mental accounting; in Section 4.3 we specify and estimate the structural model used to assess the degree to which households treat available credit and available checking balances as fungible; in Section 4.4 we conclude and outline potential ways forward.

4.2. Data

We obtain anonymized data from a large U.S. bank which contains credit card, debit card, and checking account transactions (including deposits), matched by household. Our main analysis operates on a subset of this dataset featuring households that use both credit and debit cards. The dataset subsample we use for this analysis and the analyses in the next section features $10,690$ household units each with at least 2 consecutive months of debit/checking account and credit card account transactions and income observations. This gives us a total of $123,112$ unique account/month combinations. The mean household has approximately 11.5 months of observations with a median of 10 observations. The maximal household features 66 consecutive months of observations. The time frame of our data ranges from January 2007 to October 2014. We classify consumption expenditure into durable or non-durable categories by using the descriptions associated with the 4-digit Visa card merchant code and manually matching these descriptions to those for the U.K. credit card data used by Quispe-Torreblanca et al. (2019). These classifications are presented in Table D.1 in Appendix D.1.

Before proceeding to the structural model where we show directly that our data fails to reconcile standard fungibility requirements, we first examine differences in marginal propensity to consume from changes in available credit verses changes in checking account balances. Let nd_{it} denote real non-durable consumption expenditure by the i^{th} household in period t and equivalently denote d_{it} for real durable consumption expenditure. Let c_{it} denote total non-durable plus durable real consumption expenditure. Define the variable $MPND_{it}(j)$ as the change in total non-durable consumption due to a change in the balance of liquidity source j where $j = m$ for credit and $j = z$ for checking account balances, and equivalently $MPD_{it}(j)$ for the change in total durable consumption. The marginal propensity to consume for total non-durable plus durable consumption is $MPC_{it}(j)$. We define available credit as the credit limit b (borrowing limit) less any debts. Since our data are denominated in nominal terms, we deflate non-durable expenditure using the core personal consumption expenditure index (PCE) from the Bureau of Economic Analysis (BEA) and deflate durable expenditure using the PCE for durable goods. Both indices are normalized so that 2009-dollars are the numeraire. The marginal propensities to consume can then be approximated using finite differences where Δ denotes the finite difference operator:

$$MPC_{it}(j) = \frac{\Delta c_{it}}{\Delta j_{it}} \tag{4.3}$$

$$MPND_{it}(j) = \frac{\Delta nd_{it}}{\Delta j_{it}} \tag{4.4}$$

$$MPD_{it}(j) = \frac{\Delta d_{it}}{\Delta j_{it}} \tag{4.5}$$

j_{it} denotes the available balance to consumer i in liquidity-category j entering period t.

Let T_i denote the total number of months of observations for consumer-unit i. For each consumer, we compute the average marginal propensity to consume out of changes in m_{it} and z_{it} for both non-durables, durables, and total consumption which is denoted by C_{it}:

$$\overline{MPC}_i(j) = \frac{1}{T_i}\sum_{t=1}^{T_i} MPC_{it}(j) \tag{4.6}$$

$$\overline{MPND}_i(j) = \frac{1}{T_i}\sum_{t=1}^{T_i} MPND_{it}(j) \tag{4.7}$$

$$\overline{MPD}_i(j) = \frac{1}{T_i}\sum_{t=1}^{T_i} MPD_{it}(j) \tag{4.8}$$

We present summary statistics for these marginal propensities in Appendix D.1, Table D.2.

Figure 4.1 shows a grid of plots featuring the distribution of individual means from (4.6) to (4.8) with both credit and debit card denominators. The distributions in Figure 4.1 have fat tails and a near point-mass at 0. If we difference the consumer-level average marginal propensity to consume from credit cards with the marginal propensity to consume from debit cards within category we get the same shape of distribution. Figure 4.2 shows distributions over consumers of the following differences in means, which we denote with $\mathcal{D}G_i$ where G is some commodity category:

$$\mathcal{D}C_i = \overline{MPC}_i(m) - \overline{MPC}_i(z) \tag{4.9}$$

$$\mathcal{D}ND_i = \overline{MPND}_i(m) - \overline{MPND}_i(z) \tag{4.10}$$

$$\mathcal{D}D_i = \overline{MPD}_i(m) - \overline{MPD}_i(z) \tag{4.11}$$

The distributions in Figure 4.2 appear to have fat tails and be symmetric. If these distributions are centered at 0 then that would tell us that the modal or median consumer treats increases to available credit as equally fungible to increases in cash. Neo-classical theory would suggest that an increase in available credit should actually lead to relatively less consumption than an equivalent increase in available cash, since a rational consumer maximizing expected lifetime utility would rather consume out of cash and avoid high interest payments in the future. On the other hand, a consumer who is more responsive to credit increases than cash increases is borrowing from the future to fund consumption in the present.

To better understand which story fits our data, we assume that each of (4.9) thru (4.11) is distributed according to the fat-tailed, symmetric Cauchy distribution:

$$\mathcal{D}G_i \sim \text{Cauchy}(\mu_G, \sigma_G) \tag{4.12}$$

In the above μ_G is the distribution's location parameter for good G. This parameter is associated with the distribution's median and mode (the Cauchy distribution does not have finite first and second moments and thus does not have a mean). To understand whether the median/modal consumer is more likely to increase consumption due to a credit increase verses a cash increase, we test the following null hypothesis:

$$H_0 : \ \mu_G = 0, \quad \forall G \in \{C, ND, D\} \tag{4.13}$$

The results of maximum likelihood estimation on (4.12) for each of total consumption, non-durable, and durable consumption, along with the results of the associated hypothesis tests on (4.13) are presented below in Table 4.1. The column marked "Wald" tests the hypotheses in (4.13) against a χ^2_1 random variable. All of the median/modal parameters μ_G are significantly different from 0. For each good, the median/modal consumer thus consumes more due to an increase in credit than an increase in cash, though the effect is more pronounced for durable goods than non-durable goods, consistent with the theory of Prelec and Loewenstein (1998) and empirical results of Quispe-Torreblanca et al. (2019).

Table 4.1.: MLE Results for Differences

Model	σ_G	μ_G	Wald	p-value
$\mathcal{D}C_i$	2.212	0.185	42.975	5.545e-11
	(0.033)	(0.028)		
$\mathcal{D}ND_i$	0.734	0.048	28.741	8.273e-08
	(0.012)	(0.009)		
$\mathcal{D}D_i$	2.884	0.315	72.103	0
	(0.042)	(0.037)		

For most consumers, consumption appears to be excessively unresponsive to balance changes and thus excessively smooth over time. Excess smoothness of consumption has been used to justify the PIH (see Campbell and Deaton (1989), Quah (1990), Flavin (1993), Ludvigson and Michaelides (2001), and Attanasio and Pavoni (2011)). However, if individuals engage in two-stage mental accounting and strongly anchor their expenditure on a pre-set budget, then excess consumption smoothness would result. In the former case (PIH), consumers are taken to be *incredibly* forward-looking. In the latter case, consumers are assumed more myopic in terms of dynamic outcomes, instead using rule-of-thumb budgeting to regulate consumption. Observationally, if we look only at marginal propensities, these outcomes are equivalent. Thus, a structural test of fungibility is needed since in the case where PIH holds, we would expect fungibility to also hold. For this, we must take a more rigorous, decision-theoretic approach.

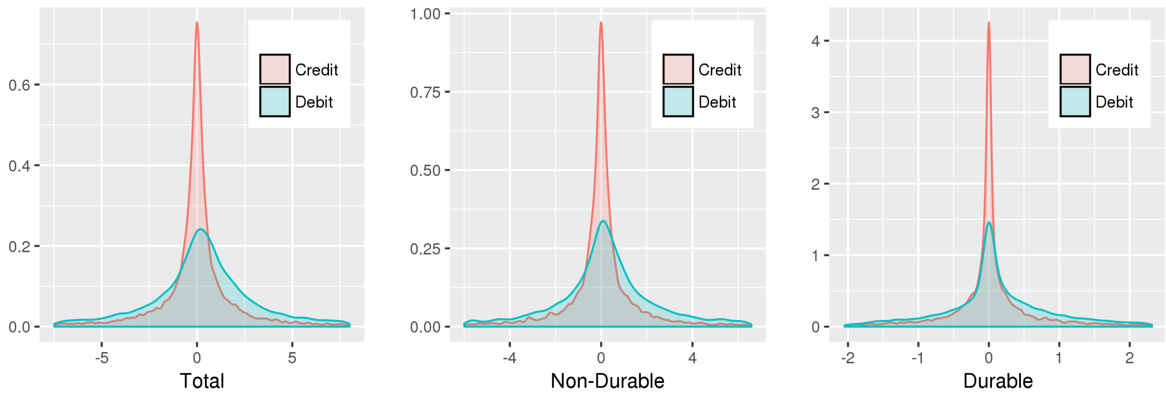


Figure 4.1.: Notice that total consumption (left panel) and non-durable consumption (center panel) are such that the distribution of individual average marginal propensities to consume out of checking account increases appears to dominate the same marginal propensity to consume out of credit increases. This relationship does not appear, at first glance, to hold for durable goods, suggesting consumers again are more comfortable holding credit card debt on durable purchases. (All plots truncated at bottom and top 0.10 quartile.)

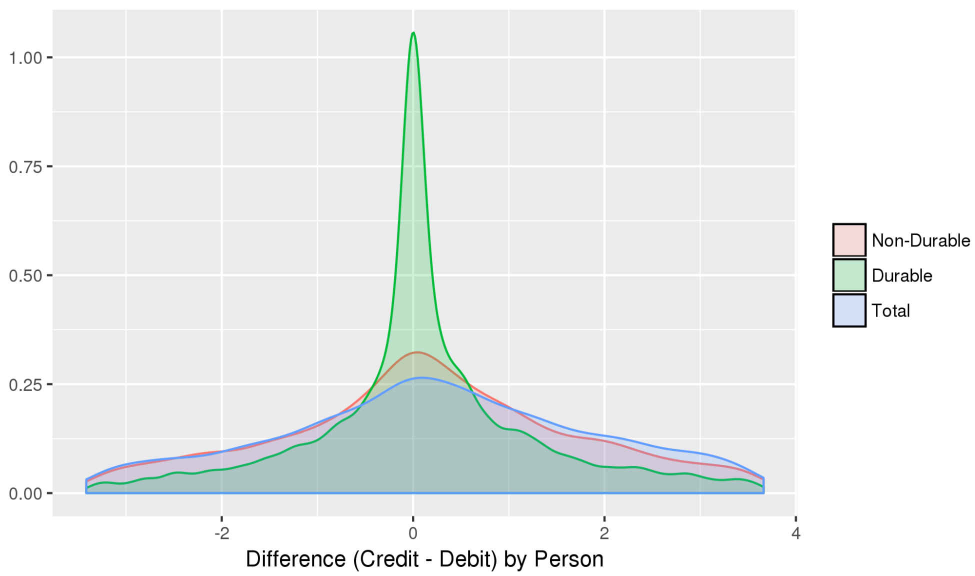


Figure 4.2.: Consumers to the right of the origin increase consumption spending due to a credit increase more than due to a cash increase, while consumers to the left increase consumption spending due to a cash increase more than due to a credit increase. Table 4.1 shows that the modes/medians of these distributions, if we assume a Cauchy specification, are significantly positive and different from 0. (All plots truncated at bottom and top 0.10 quartile.)

4.3. Structural Model

In this section we first define the household's problem in terms of the neo-classical formulation commonly found in consumption/savings models with perfect information, exponential/geometric discounting, and rational expectations. In this initial formulation found in

Section 4.3.1, we describe the important features of this model in which consumers are assumed to treat liquidity as perfectly fungible. In Section 4.3.2 we modify the model to accommodate some of the features of mental accounting by partitioning the budget constraint. We accomplish this without adjusting the preference structure. Finally, we exploit the equilibrium conditions of the mental accounting modification to perform the main, structural fungibility test in Section 4.3.3.

4.3.1. Classical Formulation

Consider a consumer who derives utility from consumption of a non-durable good nd_t and the stock of available durable goods k_t each period. Utility is some strictly increasing, strictly concave, twice-differentiable function $u_t(nd_t, k_t)$ that admits the gross substitutes property. The usefulness of durable goods declines each period due to wear and tear with this depreciation denoted by δ. Each period the consumer chooses how much of his wealth w_t to consume in non-durable goods nd_t, invest in savings s_t, and invest in durable goods d_t by solving the dynamic programming problem

$$V_t(w_t, k_t) = \max_{nd_t, d_t, s_t} \quad u_t(nd_t, k_t) + \beta \cdot \mathbb{E}_t V_{t+1}(w_{t+1}, k_{t+1}) \tag{4.14}$$

$$\text{s.t.} \quad p_t^{nd} \cdot nd_t + p_t^d \cdot d_t + s_t \leq w_t \tag{4.15}$$

$$k_{t+1} \leq (1 - \delta)k_t + d_t \tag{4.16}$$

Here, k_t and w_t are stock variables and nd_t, d_t, and s_t are flow variables. Equation (4.16) describes how the total stock of durable goods in the household evolves from period to period. $V_t(w_t, k_t)$ is a function that describes the value a consumer derives from making an optimal decision, p_t^{nd} is the period t price level for non-durable consumption, p_t^d is the period t price level for durable consumption, and β is a parameter governing the consumer's rate of time preference. In the budget constraint, wealth w_t is a function of the consumer's net assets a_t and his labor income l_t:

$$w_t = r_t \cdot a_t + l_t \tag{4.17}$$

In the above, r_t is the net rate of return on saved assets. Thus in this formulation the consumer values \$1 held in assets the same as \$1 earned from labor income. In the standard formulation, one assumes that asset holdings must satisfy some borrowing limit b_t which usually does not bind:

$$a_{t+1} \geq -b_t \tag{4.18}$$

with $b_t \geq 0$. Thus in this formulation consumers can choose to borrow to finance present consumption but cannot choose to both borrow and save simultaneously. This allows us to express the law of motion for net assets as

$$a_{t+1} \leq a_t + s_t \tag{4.19}$$

Let R_t be the gross rate of return on assets. Now, using (4.19) and (4.17), the budget constraint can be written strictly as a function of asset holdings and labor income:

$$p_t^{nd} \cdot nd_t + p_t^d \cdot d_t + a_{t+1} \leq R_t \cdot a_t + l_t \tag{4.20}$$

The linear structure of the consumer's budget constraint implies that, absent liquidation costs, wealth from investments is perfectly substitutable with wealth derived from labor income. In the forthcoming analysis, we will refer to the linear constraint of (4.20) and assume, for now, that (4.18) does not bind.

To characterize the consumer's optimal decision in period t, let us write out the Lagrangian function using the collapsed budget constraint. Call this function $\mathcal{L}_t(a_t, k_t)$ for the problem described by (4.14) through (4.16) with (4.20) substituting for (4.15). In this formulation agents choose nd_t, d_t, and a_{t+1}, and the state variables are a_t and k_t with l_t introduced as an additional flow variable. Letting λ_t denote the Lagrangian multiplier on (4.20) and η_t the multiplier on (4.16), the period t Lagrangian is:

$$\begin{aligned} \mathcal{L}_t(a_t, k_t) = {} & u_t(nd_t, k_t) + \beta \cdot \mathbb{E}_t V_{t+1}(a_{t+1}, k_{t+1}) \\ & + \lambda_t \cdot (R_t \cdot a_t + l_t - p_t^{nd} \cdot nd_t - p_t^d \cdot d_t - a_{t+1}) \\ & + \eta_t \cdot \big((1-\delta)k_t + d_t - k_{t+1}\big) \end{aligned} \tag{4.21}$$

Differentiating the Lagrangian yields three first-order conditions (plus the budget constraint and law of motion for durable goods) so long as $a_{t+1} > -b_t$:

$$\frac{\partial u_t}{\partial nd_t} = \lambda_t p_t^{nd} \tag{4.22}$$

$$\eta_t = \lambda_t p_t^d \tag{4.23}$$

$$\lambda_t = \beta \cdot \mathbb{E}_t \frac{\partial V_{t+1}}{\partial a_{t+1}} \tag{4.24}$$

The three conditions describe the intratemporal rate of substitution between durable and non-durable consumption and the intertemporal consumption/savings tradeoff a consumer faces. λ_t links the marginal utilities between goods with the discounted expected future marginal value an individual will accrue from his savings. Thus λ_t stands in for the marginal value of period t resources. η_t stands in for the marginal value of an additional unit of durable goods. In equilibrium, the internal household marketplace for present and expected future consumption is cleared when the marginal utilities from present non-durable consumption, present durable consumption, and the expected marginal value of investments next period are both exactly equal to the marginal value of period t wealth. Using the Envelope Theorem, (4.24) can be rewritten:

$$\lambda_t = \beta \cdot \mathbb{E}_t R_{t+1} \lambda_{t+1} \tag{4.25}$$

This condition says that today's marginal value of wealth must equal the discounted expected gross marginal value of resources next period.

Implicitly, this problem says nothing about how consumers value different forms of liquidity, instead focussing on net available resources. Suppose we now allow consumers to choose which forms of available liquidity, such as wealth or credit, out of which to consume and save. Rather than folding borrowing into a single asset-holding decision a_{t+1} where $a_{t+1} < 0$ if a consumer holds net debt, we will now formulate the model to allow consumers to choose the specific forms of available liquidity out of which to consume and invest. Let m_t denote the amount of credit available to the consumer in period t. Let q_t denote the interest rate the consumer must pay if he carries debts into the next period, with Q_t denoting the gross rate associated with q_t. Consumers can choose to pay off, out of assets or labor income, the debts they accrued prior to incurring interest on the balance. Let e_t denote net outlay (expenditures less payments) associated with the debt instrument and denote the credit limit (effective borrowing limit) by b_t. Then available debt evolves according to the law of motion:

$$b_{t+1} - m_{t+1} \leq Q_t(b_t - m_t - e_t) \tag{4.26}$$

Under this formulation, all borrowing happens through choices of e_t not s_t. Thus we can rewrite (4.18) to force $a_{t+1} \geq 0$, so that savings must always be non-negative. This alters the budget constraint. Now consumers must choose s_t and e_t, but these choices depend on the relative rate of return between savings and borrowing. If $e_t > 0$ then the consumer paid down debts more than he borrowed, while if $e_t < 0$ he borrowed more than he paid down debts. Total consumption expenditure, investment, and debt payments must still satisfy the consumer's available net resources, though the budget constraint now additively separates investment from borrowing:

$$p_t^{nd} \cdot nd_t + p_t^d \cdot d_t + s_t + e_t \leq r_t \cdot a_t + l_t \tag{4.27}$$

We can use (4.26) in (4.27) to completely collapse the constraints into a single budget constraint but which now contains an explicit credit feature. The endogenous state variables for the consumer's optimization problem are now a_t, k_t, and m_t, and the recursive optimization problem is

$$V_t(a_t, k_t, m_t) = \max_{nd_t, d_t, a_{t+1}, m_{t+1}} \quad u_t(nd_t, k_t) + \beta \cdot \mathbb{E}_t V_{t+1}(a_{t+1}, k_{t+1}, m_{t+1}) \tag{4.28}$$

$$\text{s.t.} \quad p_t^{nd} \cdot nd_t + p_t^d \cdot d_t + a_{t+1} + \frac{m_{t+1} - b_{t+1}}{Q_t} \leq R_t \cdot a_t + m_t - b_t + l_t \tag{4.29}$$

$$k_{t+1} \leq (1-\delta)k_t + d_t \tag{4.30}$$

Here, the borrowing limits are incorporated directly into the budget constraint, eliminating the additional inequality constraint.

Denote the Lagrangian multiplier on (4.29) as μ_t and that on (4.30) as η_t. The new problem yields four first-order conditions (plus the budget constraint):

$$\frac{\partial u_t}{\partial nd_t} = \mu_t p_t^{nd} \tag{4.31}$$

$$\eta_t = \mu_t p_t^d \tag{4.32}$$

$$\mu_t = \beta \cdot \mathbb{E}_t \frac{\partial V_{t+1}}{\partial a_{t+1}} \tag{4.33}$$

$$\frac{\mu_t}{Q_t} = \beta \cdot \mathbb{E}_t \frac{\partial V_{t+1}}{\partial m_{t+1}} \tag{4.34}$$

Here, the interpretation of μ_t is slightly different than that of λ_t. Whereas λ_t stands in for the marginal value of period-t wealth, μ_t represents the marginal value of period t liquidity, i.e. wealth plus credit. We can combine (4.33) and (4.34) to show that the expected marginal value of having more available credit should equal the expected marginal value of assets weighted by the inverse gross rate of credit interest:

$$\mathbb{E}_t \frac{\partial V_{t+1}}{\partial m_{t+1}} = \left(\frac{1}{Q_t}\right) \cdot \mathbb{E}_t \frac{\partial V_{t+1}}{\partial a_{t+1}} \tag{4.35}$$

Thus as Q_t increases, a consumer can adjust his asset holdings down to offset the additional borrowing cost by paying back his debts faster. Similarly, as Q_t falls, the consumer will borrow more reducing his available credit and increasing his expected marginal value of available credit. These results are a direct consequence of the concavity of the value function inducing diminishing marginal utility.

Despite this small adjustment to the neo-classical consumption/savings model made by explicitly modeling the borrowing decision alongside the savings decision, the consumer by design has the same marginal value for liquidity regardless of the liquidity instrument. In other words, liquidity in this model is "perfectly fungible" across invested assets, income, and credit. This assumption fundamentally contradicts the theory of mental accounting specifically discussed in Shefrin and Thaler (1988) and Thaler (1990, 1999). Other authors like Prelec and Simester (2001) have noted that consumer willingness-to-pay seems directly related to payment method: consumers are willing to spend more to purchase an item when using their credit cards as opposed to their checking accounts. Such a phenomenon could be explained by a liquidity or buffer-stock preference as modeled in Carroll (1997). Even so, a rational consumer who is constrained as described by the budget constraint in (4.29) should seek to ensure that his economic behavior befits the condition in (4.35), and thus price-weighted marginal value of wealth is independent of the liquidity type.

This leads to the natural question: do consumers behave as if the marginal value of liquidity is the same for all liquidity types? To answer it, we will consider consumer behavior under an alternative formulation of the above model where instead of choosing total consumption and investment levels consumers choose shares of liquidity sources out of which to allocate consumption. We will call this formulation the "mental accounting" formulation as it is consistent with the idea that consumers care about the source of funds used to engage in expenditure and/or savings. We accomplish the integration of this modeling feature into the standard, classical utility maximization problem without altering the consumer's preference structure, only his choice set.

4.3.2. Mental Accounting Formulation

Suppose now consumers behave according to the theory of Shefrin and Thaler (1988) whereby ex-ante they set expenditure plans for themselves according to some optimization rule. Such behavior would be consistent with what is known as the "planner's" side of the problem. Consumers then go out and engage in expenditure — the "doer's" side of the problem. Here, we will explicitly focus on the planner's side particularly with respect to the method of payment consumers will use to engage in expenditure.

Consumers have preferences over durable and non-durable consumption given by $u_t(nd_t, k_t)$. Consumers have three sources of liquidity out of which they can consume — investment income $R_t \cdot a_t$, available cash z_t, and credit m_t. The choice process takes a two-step approach. First, consumers decide how much of their expenditure they want to come from the different liquidity sources by choosing goods-dependent shares $\theta_{at}(nd_t)$, $\theta_{zt}(nd_t)$, and $\theta_{mt}(nd_t)$ and $\theta_{at}(d_t)$, $\theta_{zt}(d_t)$, and $\theta_{mt}(d_t)$ where total expenditure of each good in the period is the weighted sum of these shares

$$p_t^{nd} \cdot nd_t = \theta_{at}(nd_t) R_t \cdot a_t + \theta_{zt}(nd_t) z_t + \theta_{mt}(nd_t) m_t \tag{4.36}$$

$$p_t^{d} \cdot d_t = \theta_{at}(d_t) R_t \cdot a_t + \theta_{zt}(d_t) z_t + \theta_{mt}(d_t) m_t \tag{4.37}$$

Here, for example, $\theta_{zt}(nd_t)$ is the fraction of available cash z_t the consumer chooses to spend on non-durable goods nd_t each period.

Given their expenditure choices, consumers then decide how much to save s_t and pay down their credit card debt with payments e_t. The laws of motion change slightly since the payment and savings decisions are fundamentally separate from the expenditure-share decisions. Here, we model the evolution of credit m_t and include an additional state variable z_t which corresponds to the evolution of cash or, equivalently, checking-account holdings[6]:

$$a_{t+1} \leq \big(1 - \theta_{at}(nd_t) - \theta_{at}(d_t)\big) R_t \cdot a_t + s_t \tag{4.38}$$

$$z_{t+1} \leq W_t \big(l_t + (1 - \theta_{zt}(nd_t) - \theta_{zt}(d_t)) z_t - e_t - s_t\big) \tag{4.39}$$

$$b_{t+1} - m_{t+1} \leq Q_t \big(b_t - m_t + (\theta_{mt}(nd_t) + \theta_{mt}(d_t)) m_t - e_t\big) \tag{4.40}$$

The value of cash next period z_{t+1} is the value of cash this period less expenditure, payments, and savings, appreciated at gross rate W_t. Total debt next period $b_{t+1} - m_{t+1}$ is debt this period plus expenditure out of credit this period less payment, appreciated at rate Q_t.

Now assume the consumer chooses these shares $\theta_{at}(nd_t)$, $\theta_{zt}(nd_t)$, $\theta_{mt}(nd_t)$, $\theta_{at}(d_t)$, $\theta_{zt}(d_t)$, and $\theta_{mt}(d_t)$, savings s_t, and payments e_t so as to solve his recursive dynamic programming problem where utility is derived over consumption of non-durable and durable goods. The endogenous state variables for this problem are assets a_t, the stock of durables k_t, cash z_t, and credit m_t. Income l_t is thought to evolve exogenously. Credit lines b_t also evolve exogenously. For notational convenience, denote the vectors of shares where the individual components are the non-durable and durable-specific shares in bold face, (e.g. $\boldsymbol{\theta}_{at}$). The consumer's problem

[6] We use "cash", "debit", and "checking account balances" interchangeably.

under this formulation is:

$$V_t(a_t, k_t, z_t, m_t) = \max_{\substack{\theta_{at}, \theta_{lt}, \theta_{mt} \\ s_t, e_t}} \quad u_t(nd_t, k_t) + \beta \cdot \mathbb{E}_t V_{t+1}(a_{t+1}, k_{t+1}, z_{t+1}, m_{t+1}) \tag{4.41}$$

$$\text{s.t.} \quad a_{t+1} \leq \big(1 - \theta_{at}(nd_t) - \theta_{at}(d_t)\big) R_t \cdot a_t + s_t \tag{4.42}$$

$$z_{t+1} \leq W_t\big(l_t + (1 - \theta_{zt}(nd_t) - \theta_{zt}(d_t))z_t - e_t - s_t\big) \tag{4.43}$$

$$b_{t+1} - m_{t+1} \leq Q_t\big(b_t - m_t + (\theta_{mt}(nd_t) + \theta_{mt}(d_t))m_t - e_t\big) \tag{4.44}$$

$$k_{t+1} \leq (1 - \delta)k_t + d_t \tag{4.45}$$

$$p_t^{nd} \cdot nd_t = \theta_{at}(nd_t) R_t \cdot a_t + \theta_{zt}(nd_t) z_t + \theta_{mt}(c_t) m_t \tag{4.46}$$

$$p_t^{d} \cdot d_t = \theta_{at}(d_t) R_t \cdot a_t + \theta_{zt}(d_t) z_t + \theta_{mt}(d_t) m_t \tag{4.47}$$

Denote the Lagrangian multipliers on (4.42), (4.43), and (4.44) as μ_{at}, μ_{zt}, and μ_{mt}. Let η_t be the multiplier on the durable goods law of motion again. When the agent chooses shares, the first order conditions are:

$$\frac{\partial u_t}{\partial nd_t} = \mu_{at} p_t^{nd} \tag{4.48}$$

$$\frac{\partial u_t}{\partial nd_t} = W_t \mu_{zt} p_t^{nd} \tag{4.49}$$

$$\frac{\partial u_t}{\partial nd_t} = Q_t \mu_{mt} p_t^{nd} \tag{4.50}$$

$$\eta_t = \mu_{at} p_t^{d} \tag{4.51}$$

$$\eta_t = W_t \mu_{at} p_t^{d} \tag{4.52}$$

$$\eta_t = Q_t \mu_{mt} p_t^{d} \tag{4.53}$$

Choices of s_t and e_t provide conditions on the Lagrange multipliers:

$$\mu_{at} = W_t \mu_{zt} \tag{4.54}$$

$$Q_t \mu_{mt} = W_t \mu_{zt} \tag{4.55}$$

Importantly, regardless of the functional form of the utility function chosen, the model says that marginal utility of non-durable consumption must equate with all price-weighted marginal values of liquidity for the different liquidity types. For durables, the marginal value of a one-unit increase to the durable stock η_t must equate with these price-weighted marginal values of liquidity also.

Conditions (4.54) and (4.55) provide natural restrictions to test the validity of this model. (4.54) says that the marginal value of an additional unit of assets is equal to the return-weighted marginal value of an additional unit of cash. Thus, individuals should maintain balances in all of these accounts, including available credit, so that they are roughly proportional to each other. The model thus predicts that individuals will take into consideration the relative interest rates across liquidity sources when making expenditure decisions and choose consumption expenditure in weighted proportions. Note that this is a *result* of the model which holds for all utility functions in which consumption utility is not separable across payment categories. Thus as long as consumers benefit from one unit of consumption from a credit card purchase the same as they benefit from one unit of consumption from a cash purchase, conditions (4.48) thru (4.55) will hold. (4.48) thru (4.55) constitute a system of identifiable simultaneous equations, but whose restrictions force the marginal utilities to relate via (4.54) and (4.55). We can test whether this is a reasonable assumption by empirically examining how the balances of these different liquidity sources covary over time.

4.3.3. A Test of Fungibility

The null hypotheses we are seeking to test is whether the conditions in (4.54) and (4.55) hold with equality. This will tell us whether or not individuals have consistent preferences over different liquidity sources and thus treat available liquidity as fungible, as predicted by the standard neo-classical consumption/savings problem outlined in (4.28) thru (4.30) and modified in (4.41) thru (4.47).

The primary focus of this exercise is to examine whether the classical model predicts that consumers will adjust their expenditure out of both checking and credit card sources at rates consistent with (4.55). Since our data primarily consists of debit card transactions from checking accounts and credit card transactions, we will focus on (4.55) and forego analysis of how individuals consume out of saved assets. Note that μ_{zt} and μ_{mt} denote the rate at which the individual Lagrangian function $\mathcal{L}_{it}$ changes with respect to unit changes in z and m respectively. Let i index individual consumers in our dataset. Using the finite differences operator Δ, we can thus write condition (4.55) at the consumer level:

$$Q_{it}\frac{\Delta\mathcal{L}_{it}}{\Delta m_{it}} = W_{it}\frac{\Delta\mathcal{L}_{it}}{\Delta z_{it}} \tag{4.56}$$

Canceling like terms, re-arranging the equation, then taking an expectation over time, provides us with a testable null hypothesis at the agent level:

$$H_0 : \mathbb{E}_i\left\{Q_{it}\frac{\Delta z_{it}}{\Delta m_{it}} - W_{it}\right\} = 0 \tag{4.57}$$

From Figure 4.3 it appears that the distribution of our test statistic is symmetric with fat tails, much like the distributions in Figure 4.2. Thus suppose (4.57) follows a Cauchy distribution with parameters μ and σ.[7]

$$\mathbb{E}_i\left\{Q_{it}\frac{\Delta z_{it}}{\Delta m_{it}} - W_{it}\right\} \sim \text{Cauchy}(\mu, \sigma) \tag{4.58}$$

Figure 4.3 shows that, even when weighted by the gross monthly interest rate Q_{it}, fluctuations in available credit dominate fluctuations in available cash. Note that available credit can change due to three factors: 1) changes in spending on the credit card, 2) payments made against the credit card balance, or 3) an exogenous increase in the credit limit by the bank. The latter rarely occurs. The fact that the distribution of (4.57) is centered to the left of the origin suggests that (4.55) does not hold with equality. We perform a formal test on this conjecture by estimating μ

[7]The distinction between these parameters from those in (4.12) is that we do not impose subscripts here.

and σ using maximum likelihood estimation on (4.58). The results of the estimation and a Wald test on the hypothesis

$$H_0 : \ \mu = 0 \tag{4.59}$$

shows that $\mu < 0$ almost certainly. This implies that the median/mode of (4.57) is also less than 0, so that for the median/modal consumer in our sample, condition (4.55) does not hold. In fact, this test suggests that most consumers behave as if the marginal return to an additional unit of credit is *greater* than the marginal return to additional cash, a feature which would not be predicted by any standard parameterization of the neo-classical consumption/savings model. The full results of the MLE for (4.58) are presented in Table 4.2 below.

The units of the test statistic are real 2009 U.S. dollars. Essentially (4.57) represents on average how far away a household is from treating liquidity as perfectly fungible. A household with a test statistic value of 2, for example, will on average spend \$2 more on consumption due to a \$1 cash increase than if they had received a \$1 credit increase, while a household with a value of $-$\$2 will on average spend \$2 more on consumption due to a \$1 credit increase than if they had received a \$1 cash increase. Thus, the statistic measures the marginal rate at which households value additional cash verses additional credit.

While the median/modal consumer appears to violate fungibility, Figure 4.3 demonstrates that the distribution of this average behavior across individuals is rather disperse. In fact since the Cauchy distribution itself does not have a finite second moment, the asymptotic variance of (4.57) is effectively infinite. The variance of our sample test statistic, for example, is 6694824685, suggesting that the Cauchy distribution may be a good selection. A small share of households appear to treat liquidity as fungible, though about 59% of households are clearly more sensitive to increases in available credit than increases in cash. To understand the relative shares of households who exhibit behavior "near" fungibility, in Table 4.3 we present estimates of the shares of households whose value of (4.57) resides within \$0.01, \$0.05, \$0.10, \$0.25, \$0.50, \$1, \$2, \$5, and \$10 of 0 for the MLE-fitted Cauchy distribution. About 16.9% of households, for example, reside within \$1 of fungibility, on average. This is not large, but it does suggest that further exploration may be needed to better understand what, if anything, sets more fungible households apart from their less fungible peers.

Whether all of this matters on aggregate remains to be explored. A generous interpretation of these results would say that the neo-classical formulation does not perform *too badly*: a significant share of consumers appear to treat liquidity as near fungible, depending on how one defines "near," and some are even more sensitive to cash increases than credit increases (about 41%). The test reveals what mental accounting theory would expect — most consumers are more sensitive to credit increases than cash increases — and yet the neo-classical model does not appear utterly terrible at reconciling these facts with only slight modifications, true testament to its durability and continuing utility for consumer scientists.

Table 4.2.: Structural Test of Fungibility

σ	μ	Wald	p-value
3.388	-1.029	619.749	0
(0.054)	(0.041)		

Table 4.3.: Shares of Households Within \$$x$ from 0

Absolute $ From 0	0.010	0.050	0.100	0.250	0.500	1	2	5	10
Shares of HH	0.002	0.009	0.017	0.043	0.086	0.169	0.321	0.612	0.790

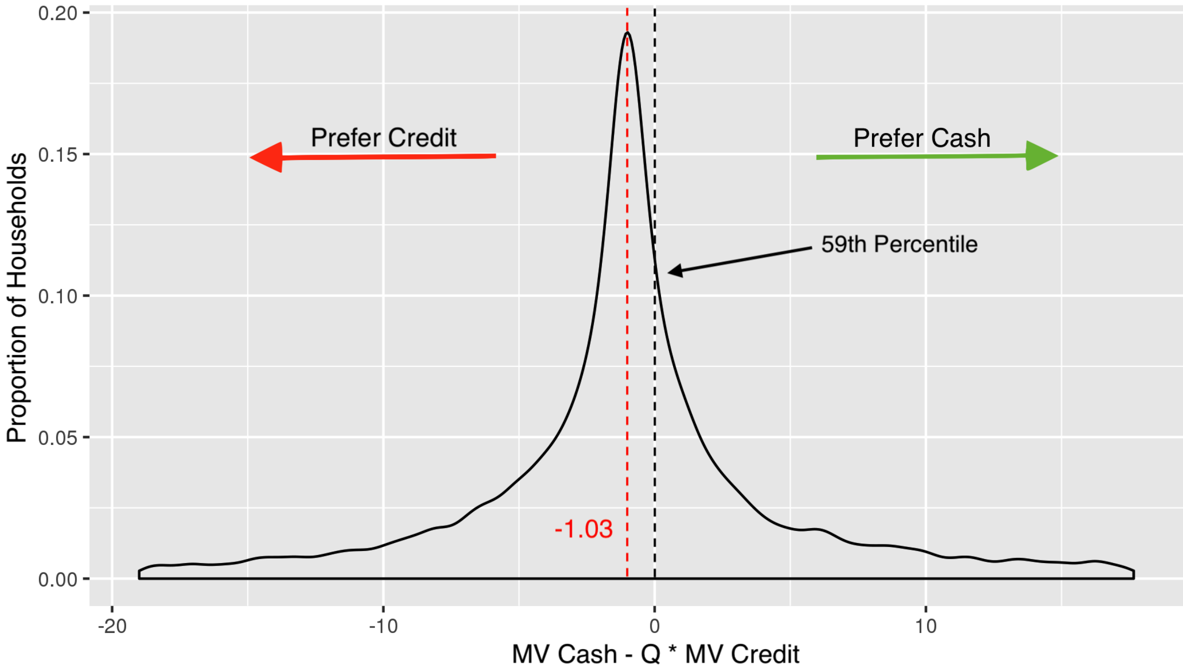


Figure 4.3.: This is the distribution of the individual sample mean analogs of (4.57). Notice that the distribution is centered to the left of the origin, suggesting that the median/modal consumer exhibits behavior that is inconsistent with condition (4.55). (Truncated at bottom and top 0.10 quartile.)

4.4. Conclusion

We have shown that the neo-classical consumption/savings model can be used to perform a consistent, structural test of the fungibility of liquidity at the household level. Most households are credit hungry, increasing consumption more due to credit increases than cash increases, though there is significant heterogeneity in this behavior across households. Further, durable goods consumption exhibits patterns consistent with the theory of mental accounting in Prelec and Loewenstein (1998), suggesting utility from consumption is coupled with method of payment. Broadly speaking, our results provide strong empirical support that consumers engage in various degrees of mental accounting.

But what does this mean for broader economic outcomes and societal impacts? More work is needed to uncover the economic and demographic profiles of households that exhibit highly non-fungible behavior. Understanding how credit hungriness under mental accounting is related on aggregate to broader fluctuations in liquidity and credit availability would be a daunting yet potentially groundbreaking endeavor. Thus our findings here should be used as a springboard for further research that examines how household-level mental accounting behaviors could affect broad economic aggregates, while better informing the design of policies and products that help encourage responsible consumer credit utilization.

A. Appendix: Home Production with Time to Consume

A.1. Data Appendix

A.1.1. Constructing Chain Weighted Price Indices

We will illustrate how to unwind chain-weighted price and quantity indices with an example showing how to combine durables stock data with new consumption expenditure data. Our consumption expenditure data series are taken from the Bureau of Economic Analysis' (BEA) National Income and Product Account (NIPA) tables. Specifically, we take the non-durable goods and services nominal expenditure series from NIPA Table 1.1.5. To construct real data series, we download the chain-type price indices from NIPA Table 1.1.4. and chain type quantity indices from NIPA Table 1.1.3. for non-durable goods and services. To account for the fact that consumers enjoy service flows from durable expenditure over more than one period we turn to BEA Fixed Asset Table 1.1. which gives the current dollar value of the nominal capital stock, including consumer durables. BEA Fixed Asset Table 1.2. provides a corresponding quantity index. From each of these, we use only the "Consumer durable goods" series.

Now in possession of data series for nominal expenditure of non-durable goods and services, the nominal value of the stock of consumer durables and corresponding price and quantity indices where available, we can construct our aggregate "goods" and "services" consumption series in real chained 2012 dollars. Note that construction of the real services consumption series requires no additional steps beyond a standard deflationary procedure dividing the nominal services expenditure series from Table 1.1.5. by the chain type price index from Table 1.1.4. The units of this series should be read as "the real value of services consumption expenditure in 2012 chained dollars."[1] Since "goods" consumption is the sum of non-durable consumption and the consumption of service flows from the net stock of durable assets, we follow the procedure outlined in Online Appendix C of Herrendorf, Rogerson, and Valentinyi (2013) and discussed in Whelan (2002) to construct a measure of real goods consumption in units of 2012 chained dollars. Unfortunately we cannot simply sum expenditure of non-durables and durables and divide this number by the sum of 2012 chain-weighted real consumption since chain-weighted series are generally not additive (Whelan 2000, 2002). Instead we require an aggregate "goods" price index that accounts for changing relative prices of non-durables and the price associated with the stock of all durables. Note that the price index the BEA uses to construct the quantity index associated with the stock of consumer durables from BEA Fixed Asset Table 1.2. is the

[1] See Whelan (2000, 2002) for further discussion of chain-weighted units.

same 2012 chained dollar index for durables *expenditure* presented in NIPA Table 1.1.4.[2] Using this durables expenditure price index, we can construct a total "goods" quantity index that accounts for both aggregate non-durable consumption and service flows derived from the entire stock of consumer durables.

Let $\widetilde{Q}_{gt}$ be a chain-weighted "goods" quantity index.[3] Let the subscript *nd* denote non-durables and *d* durables. Let X_{it} denote current dollar expenditure, P_{it} be a chain-weighted price such that $P_{i,2012} = 100$, and Q_{it} be the real value of consumption in 2012 chained dollars for all $i \in \{nd, d\}$. Note that $Q_{it} = \frac{X_{it}}{P_{it}}$. Set the 2012 base year aggregate quantity index, $\widetilde{Q}_{g,2012} = 1$. We compute

$$\widetilde{Q}_{gt} = \widetilde{Q}_{g,t-1} \sqrt{\frac{(\sum_i P_{it-1} Q_{it})(\sum_i X_{it})}{(\sum_i X_{it-1})(\sum_i P_{it} Q_{it-1})}} \quad \forall t > 2012 \quad \text{where} \quad i \in \{nd, d\} \tag{A.1}$$

$$\widetilde{Q}_{gt} = \widetilde{Q}_{g,t+1} \Bigg/ \sqrt{\frac{(\sum_i P_{it} Q_{it+1})(\sum_i X_{it+1})}{(\sum_i X_{it})(\sum_i P_{it+1} Q_{it})}} \quad \forall t < 2012 \quad \text{where} \quad i \in \{nd, d\} \tag{A.2}$$

Aggregate goods consumption in chained 2012 dollars is then

$$Q_{gt} = \widetilde{Q}_{gt} \sum_{i \in \{nd,d\}} X_{i,2012} \tag{A.3}$$

Finally, the aggregate chain-weighted goods price index with $P_{g,2012} = 1$ is just

$$P_{gt} = \frac{\sum_{i \in \{nd,d\}} X_{it}}{Q_{gt}} \tag{A.4}$$

In addition to a composite goods price index, we also need an aggregate consumption price index that accounts for relative changes in the value of both services and goods over time in order to properly place aggregate wages in the same 2012 chain-weighted units as consumption prices. To construct this series, repeat the above procedure except this time sum over $i \in \{s, nd, d\}$ to get $\widetilde{Q}_t$, an aggregate consumption quantity index in 2012 chain-weighted units. From there, an aggregate consumption price index in 2012 chain-weighted units can be easily derived by first computing real 2012 chain-weighted aggregate consumption $Q_t = \widetilde{Q}_t \sum_{i \in \{s,nd,d\}} X_{i,2012}$, then using that value to get an aggregate price index $P_t = \frac{\sum_{i \in \{s,nd,d\}} X_{it}}{Q_t}$.

[2]Note that the BEA only presents current dollar value X_{dt} and 2012 chain-weighted quantity indices $\widetilde{Q}_{dt}$ for durable stocks in the fixed asset tables, not prices. Nonetheless, it can be confirmed that the price index associated with the stock of durables is the same as that associated with the flow of durables expenditure by performing the following procedure. First, compute the 2012 chain-weighted real value of durables $Q_{dt} = \frac{\widetilde{Q}_{dt} X_{d,2012}}{100}$, since $\widetilde{Q}_{d,2012} = 100$. Then compute $P_{dt} = \frac{X_{dt}}{Q_{dt}}$ and compare this series to the price index for durables expenditure in NIPA Table 1.1.4. They are the same.

[3]Since this appendix describes construction of aggregate expenditure series, household indices *h* are suppressed.

A.2. Mathematical Appendix

A.2.1. Parameterized Household First-Order Conditions

A.2.1.1. CES Utility and Home Production

Here we present detailed, parameterized first-order conditions for the household's problem under CES utility and home production, assuming $\nu_i \neq 0$:

$$\frac{\partial u}{\partial c_{ith}} = u(\boldsymbol{c}_{th})^{1-\rho}\theta_i c_{ith}^{\rho-1} \tag{A.5}$$

$$\frac{\partial \widetilde{f}_{ith}}{\partial q_{ith}} = z_{ith}\left(\omega_i q_{ith}^{\nu_i} + (1-\omega_i)n_{ith}^{\nu_i}\right)^{\frac{1-\nu_i}{\nu_i}} \omega_i q_{ith}^{\nu_i-1} \tag{A.6}$$

$$\frac{\partial \widetilde{f}_{ith}}{\partial n_{ith}} = z_{ith}\left(\omega_i q_{ith}^{\nu_i} + (1-\omega_i)n_{ith}^{\nu_i}\right)^{\frac{1-\nu_i}{\nu_i}} (1-\omega_i) n_{ith}^{\nu_i-1} \tag{A.7}$$

Under this parameterization, the inframarginal rate of technical substitution for process i is

$$\frac{\omega_i}{1-\omega_i}\left(\frac{q_{ith}}{n_{ith}}\right)^{\nu_i-1} = \frac{P_{it}}{w_{th}} \tag{A.8}$$

The marginal rate of substitution for market inputs i and j is

$$\frac{\theta_i}{\theta_j}\left(\frac{c_{ith}}{c_{jth}}\right)^{\rho-1} \frac{z_{ith}\omega_i q_{ith}^{\nu_i-1}}{z_{jth}\omega_j q_{jth}^{\nu_j-1}} \left(\omega_i q_{ith}^{\nu_i} + (1-\omega_i)n_{ith}^{\nu_i}\right)^{\frac{1-\nu_i}{\nu_i}} \left(\omega_j q_{jth}^{\nu_j} + (1-\omega_j)n_{jth}^{\nu_j}\right)^{\frac{\nu_j-1}{\nu_j}} = \frac{P_{it}}{P_{jt}} \tag{A.9}$$

Subbing out $n_{ith}(q_{ith})$ as presented in (1.15) we have:

$$c_{ith}(q_{ith}, n_{ith}(q_{ith})) = q_{ith}\, z_{ith} \left(\omega_i + (1-\omega_i)\left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)}\right]^{\frac{\nu_i}{\nu_i-1}}\right)^{\frac{1}{\nu_i}} \tag{A.10}$$

We can then write the marginal rate of substitution for market inputs i and j as

$$\begin{aligned}\frac{\theta_i\omega_i}{\theta_j\omega_j}\left(\frac{q_{ith}z_{ith}}{q_{jth}z_{jth}}\right)^{\rho} &\left(\omega_i + (1-\omega_i)\left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)}\right]^{\frac{\nu_i}{\nu_i-1}}\right)^{\frac{\rho-\nu_i}{\nu_i}} \\ \times &\left(\omega_j + (1-\omega_j)\left[\frac{w_{th}\omega_j}{P_{jt}(1-\omega_j)}\right]^{\frac{\nu_j}{\nu_j-1}}\right)^{\frac{\nu_j-\rho}{\nu_j}} = \frac{P_{it}}{P_{jt}}\end{aligned} \tag{A.11}$$

Similarly, we can write the marginal rate of substitution for off-market time utilization between processes i and j as

$$\frac{\theta_i}{\theta_j}\left(\frac{c_{ith}}{c_{jth}}\right)^{\rho-1} \frac{z_{ith}(1-\omega_i)n_{ith}^{\nu_i-1}}{z_{jth}(1-\omega_j)n_{jth}^{\nu_j-1}} \left(\omega_i q_{ith}^{\nu_i} + (1-\omega_i)n_{ith}^{\nu_i}\right)^{\frac{1-\nu_i}{\nu_i}} \left(\omega_j q_{jth}^{\nu_j} + (1-\omega_j)n_{jth}^{\nu_j}\right)^{\frac{\nu_j-1}{\nu_j}} = 1 \tag{A.12}$$

Inverting $n_{ith}(q_{ith})$ we get

$$q_{ith}(n_{ith}) = n_{ith}\left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)}\right]^{\frac{1}{1-\nu_i}} \tag{A.13}$$

$$c_{ith}\big(q_{ith}(n_{ith}), n_{ith}\big) = n_{ith}\, z_{ith}\left(\omega_i\left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)}\right]^{\frac{\nu_i}{1-\nu_i}} + 1 - \omega_i\right)^{\frac{1}{\nu_i}} \tag{A.14}$$

which can be subbed into (A.12) to marginalize out the market input quantities:

$$\begin{aligned} &\frac{\theta_i(1-\omega_i)}{\theta_j(1-\omega_j)}\left(\frac{n_{ith}z_{ith}}{n_{jth}z_{jth}}\right)^{\rho}\left(\omega_i\left[\frac{w_{th}\omega_i}{P_{it}(1-\omega_i)}\right]^{\frac{\nu_i}{1-\nu_i}} + 1 - \omega_i\right)^{\frac{\rho-\nu_i}{\nu_i}} \\ &\qquad \times\left(\omega_j\left[\frac{w_{th}\omega_j}{P_{jt}(1-\omega_j)}\right]^{\frac{\nu_j}{1-\nu_j}} + 1 - \omega_j\right)^{\frac{\nu_j-\rho}{\nu_j}} = 1 \end{aligned} \tag{A.15}$$

A.2.1.2. CES Final Utility with Cobb-Douglas Home Production

When $\nu_i = 0$, the marginal utilities in q_{ith} and n_{ith} for the composite utility function $\widetilde{u}_{th}$ are

$$\frac{\partial \widetilde{u}_{th}}{\partial q_{ith}} = \widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th})^{1-\rho}\omega_i\theta_i z_{ith}^{\rho} q_{ith}^{\rho\omega_i} n_{ith}^{\rho-\rho\omega_i}\left(\frac{1}{q_{ith}}\right) \tag{A.16}$$

$$\frac{\partial \widetilde{u}_{th}}{\partial n_{ith}} = \widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th})^{1-\rho}(1-\omega_i)\theta_i z_{ith}^{\rho} q_{ith}^{\rho\omega_i} n_{ith}^{\rho-\rho\omega_i}\left(\frac{1}{n_{ith}}\right) \tag{A.17}$$

We can derive conditions describing the equilibrium relationship between time use and market inputs for a single production process. Dividing (A.16) by (A.17) we get the infra-marginal rate of substitution between time and market goods for process i:

$$\frac{n_{ith}\omega_i}{q_{ith}(1-\omega_i)} = \frac{P_{it}}{w_{th}} \tag{A.18}$$

Using (A.18),we can write the equilibrium choice of q_{it} as an implicit function of n_{it} and vice-versa:

$$q_{ith}(n_{ith}) = \left(\frac{\omega_i}{1-\omega_i}\right)\left(\frac{w_{th}}{P_{it}}\right)n_{ith} \tag{A.19}$$

$$n_i(q_i) = q_i\left(\frac{1-\omega_i}{\omega_i}\right)\left(\frac{P_i}{w}\right) \tag{A.20}$$

Under CES utility for final consumption, and Cobb-Douglas aggregated home production, the relative demand for market good j to market good i can be written

$$\left(\frac{q_{jth}}{q_{ith}}\right) = \left[\frac{\theta_i\omega_i[(1-\omega_i)/\omega_i]^{(1-\omega_i)\rho}}{\theta_j\omega_j[(1-\omega_j)/\omega_j]^{(1-\omega_j)\rho}}\right]^{\frac{1}{\rho-1}} P_{jt}^{\frac{1-\rho+\rho\omega_j}{\rho-1}} P_{it}^{\frac{1-\rho+\rho\omega_i}{1-\rho}} w_{th}^{\frac{\rho(\omega_i-\omega_j)}{\rho-1}}\left[\frac{z_{ith}}{z_{jth}}\right]^{\frac{\rho}{\rho-1}} \tag{A.21}$$

To get this expression, start with the marginal rate of substitution for goods i and j:

$$\frac{\frac{\partial \widetilde{u}_{th}}{\partial q_{ith}}}{\frac{\partial \widetilde{u}_{th}}{\partial q_{jth}}} = \frac{P_{it}}{P_{jt}} \tag{A.22}$$

$$\Rightarrow \quad \frac{\widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th})^{1-\rho}\omega_i\theta_i z_{ith}^{\rho} q_{ith}^{\rho\omega_i} n_{ith}^{\rho-\rho\omega_i}\left(\frac{1}{q_{ith}}\right)}{\widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th})^{1-\rho}\omega_j\theta_j z_{jth}^{\rho} q_{jth}^{\rho\omega_j} n_{jth}^{\rho-\rho\omega_j}\left(\frac{1}{q_{jth}}\right)} = \frac{P_{it}}{P_{jt}} \tag{A.23}$$

Substitute the implicit function $n_{ith}(q_{ith})$ for each process then cancel like terms to get

$$\frac{\omega_i\theta_i z_{ith}^{\rho} q_{ith}^{\rho\omega_i}\left(P_{it}(1-\omega_i)/(w_{th}\omega_i)q_{ith}\right)^{\rho-\rho\omega_i}\left(\frac{1}{q_{ith}}\right)}{\omega_j\theta_j z_{jth}^{\rho} q_{jth}^{\rho\omega_j}\left(P_{jt}(1-\omega_j)/(w_{th}\omega_j)q_{jth}\right)^{\rho-\rho\omega_j}\left(\frac{1}{q_{jth}}\right)} = \frac{P_{it}}{P_{jt}} \tag{A.24}$$

$$\Rightarrow \quad \frac{\omega_i\theta_i z_{ith}^{\rho}\left(P_{it}(1-\omega_i)/(w_{th}\omega_i)\right)^{\rho-\rho\omega_i} q_{ith}^{\rho-1}}{\omega_j\theta_j z_{jth}^{\rho}\left(P_{jt}(1-\omega_j)/(w_{th}\omega_j)\right)^{\rho-\rho\omega_j} q_{jths}^{\rho-1}} = \frac{P_{it}}{P_{jt}} \tag{A.25}$$

Collect like terms and move everything but quantities to the right side:

$$\left(\frac{q_{ith}}{q_{jth}}\right)^{\rho-1} = \left[\frac{\theta_j\omega_j[(1-\omega_j)/\omega_j]^{(1-\omega_j)\rho}}{\theta_i\omega_i[(1-\omega_i)/\omega_i]^{(1-\omega_i)\rho}}\right] P_{jt}^{\rho-\rho\omega_j-1} P_{it}^{1-\rho+\rho\omega_i} w_{th}^{\rho(\omega_i-\omega_j)}\left[\frac{z_{jth}}{z_{ith}}\right]^{\rho} \tag{A.26}$$

Rewrite the left side so that q_{jth} is on top, then raise both sides to $\frac{1}{1-\rho}$ power to get (A.21).

The same procedure can be applied to express the time devoted toward production process j relative to process i as follows:

$$\left(\frac{n_{jth}}{n_{ith}}\right) = \left[\frac{\theta_i(1-\omega_i)[\omega_i/(1-\omega_i)]^{\rho\omega_i}}{\theta_j(1-\omega_j)[\omega_j/(1-\omega_j)]^{\rho\omega_j}}\right]^{\frac{1}{\rho-1}} P_{jt}^{\frac{\rho\omega_j}{\rho-1}} P_{it}^{\frac{\rho\omega_i}{1-\rho}} w_{th}^{\frac{\rho(\omega_i-\omega_j)}{\rho-1}}\left[\frac{z_{ith}}{z_{jth}}\right]^{\frac{\rho}{\rho-1}} \tag{A.27}$$

Again, start with the marginal rate of substitution for time use between processes i and j:

$$\frac{\frac{\partial \widetilde{u}_{th}}{\partial n_{ith}}}{\frac{\partial \widetilde{u}_{th}}{\partial n_{jth}}} = 1 \tag{A.28}$$

$$\begin{aligned}\Rightarrow \quad & \widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th})^{1-\rho}(1-\omega_i)\theta_i z_{ith}^{\rho} q_{ith}^{\rho\omega_i} n_{ith}^{\rho-\rho\omega_i}\left(\frac{1}{n_{ith}}\right) = \\ & \widetilde{u}_{th}(\boldsymbol{q}_{th}, \boldsymbol{n}_{th})^{1-\rho}(1-\omega_j)\theta_j z_{jth}^{\rho} q_{jth}^{\rho\omega_j} n_{jth}^{\rho-\rho\omega_j}\left(\frac{1}{n_{jth}}\right)\end{aligned} \tag{A.29}$$

Substitute (A.19) for each good and cancel like terms to get

$$\begin{aligned}& (1-\omega_i)\theta_i z_{ith}^{\rho}\left(w_{th}\omega_i/[P_{it}(1-\omega_i)]n_{ith}\right)^{\rho\omega_i} n_{ith}^{\rho-\rho\omega_i}\left(\frac{1}{n_{ith}}\right) = \\ & (1-\omega_j)\theta_j z_{jth}^{\rho}\left(w_{th}\omega_j/[P_{jt}(1-\omega_j)]n_{jth}\right)^{\rho\omega_j} n_{jth}^{\rho-\rho\omega_j}\left(\frac{1}{n_{jth}}\right)\end{aligned} \tag{A.30}$$

Collect like terms and isolate relative time use on the left hand side:

$$\left(\frac{n_{ith}}{n_{jth}}\right)^{\rho-1} = \left[\frac{\theta_j(1-\omega_j)[\omega_j/(1-\omega_j)]^{\rho\omega_j}}{\theta_i(1-\omega_i)[\omega_i/(1-\omega_i)]^{\rho\omega_i}}\right] P_{jt}^{-\rho\omega_j} P_{it}^{\rho\omega_i} w_{th}^{\rho(\omega_j-\omega_i)} \left[\frac{z_{jth}}{z_{ith}}\right]^{\rho} \tag{A.31}$$

Rewrite the left side so that n_{jth} is on top, then raise both sides to $\frac{1}{1-\rho}$ power to get (A.27).

A.2.2. Labor Supply Dependency on Prices

A.2.2.1. CES Utility with Cobb-Douglas Home Production

In this section we derive expressions for how the household-level labor supply policy l depends on prices, wages, and productivities for various exercises. First, we derive the function for the case where $\nu_i = 0$ for all i, so that home production is Cobb-Douglas. We consider this home production function in our comparative static exercises on a two-good economy where $z_1 = z_2$. Second, we derive an expression for how labor supply depends on prices, wages, and home productivities when home production functions are CES, $\nu_i \neq 0$.

Deriving an implicit expression for l as a function of prices, wages, and elasticities in the two-good economy we use for comparative statics, recall that cash on hand $\overline{y}$ is fixed, $z_1 = z_2 = 1$, and $\theta_1 = \theta_2$. First, start with the relative time use expression from (24), multiply both sides by n_2, and define the function $\phi^1(P_1, P_2, \omega_1, \omega_2, \rho, w)$ to get an implicit expression for n_1:

$$n_1(n_2) = \underbrace{\left[\frac{(1-\omega_2)[\omega_2/(1-\omega_2)]^{\rho\omega_2}}{(1-\omega_1)[\omega_1/(1-\omega_1)]^{\rho\omega_1}}\right]^{\frac{1}{\rho-1}} P_1^{\frac{\rho\omega_1}{\rho-1}} P_2^{\frac{\rho\omega_2}{1-\rho}} w^{\frac{\rho(\omega_2-\omega_1)}{\rho-1}}}_{\phi^1(P_1,P_2,\omega_1,\omega_2,\rho,w)} n_2 \tag{A.32}$$

Continuing with the relative consumption expression from (23), multiply both sides by q_2, and define the function $\phi^2(P_1, P_2, \omega_1, \omega_2, \rho, w)$ to get an implicit expression for q_1:

$$q_1(q_2) = \underbrace{\left[\frac{\omega_2[(1-\omega_2)/\omega_2]^{(1-\omega_2)\rho}}{\omega_1[(1-\omega_1)/\omega_1]^{(1-\omega_1)\rho}}\right]^{\frac{1}{\rho-1}} P_1^{\frac{1-\rho+\rho\omega_1}{\rho-1}} P_2^{\frac{1-\rho+\rho\omega_2}{1-\rho}} w^{\frac{\rho(\omega_2-\omega_1)}{\rho-1}}}_{\phi^2(P_1,P_2,\omega_1,\omega_2,\rho,w)} q_2 \tag{A.33}$$

Using $q_2(n_2) = \left(\frac{\omega_2}{1-\omega_2}\right)\left(\frac{w}{P_2}\right) n_2$ we can write q_1 as a function of n_2:

$$q_1\big(q_2(n_2)\big) = \phi^2(P_1, P_2, \omega_1, \omega_2, \rho, w)\left(\frac{\omega_2}{1-\omega_2}\right)\left(\frac{w}{P_2}\right) n_2 \tag{A.34}$$

We can now rewrite the budget constraint to get $n_2(P_1, P_2, \omega_1, \omega_2, \rho, w)$:

$$\begin{aligned}&\underbrace{\phi^2(P_1, P_2, \omega_1, \omega_2, \rho, w)\left(\frac{\omega_2}{1-\omega_2}\right)\left(\frac{P_1 w}{P_2}\right) n_2}_{P_1 q_1\left(q_2(n_2)\right)} + \underbrace{\left(\frac{\omega_2}{1-\omega_2}\right) w n_2}_{P_2 q_2(n_2)} \\ &\qquad + \underbrace{w\phi^1(P_1, P_2, \omega_1, \omega_2, \rho, w) n_2}_{w n_1(n_2)} + w n_2 = \overline{y}\end{aligned} \tag{A.35}$$

$$n_2(P_1,P_2,\omega_1,\omega_2,\rho,w) = \frac{\overline{y}}{w\Big[1+\phi^1(P_1,P_2,\omega_1,\omega_2,\rho,w)+\frac{\omega_2}{1-\omega_2}\left(\frac{P_1}{P_2}\right)\phi^2(P_1,P_2,\omega_1,\omega_2,\rho,w)+\frac{\omega_2}{1-\omega_2}\Big]} \tag{A.36}$$

Finally, with the time use constraint we can write $l(P_1,P_2,\omega_1,\omega_2)$ using the objects we just derived:

$$l(P_1,P_2,\omega_1,\omega_2,\rho,w) = \overline{n} - \phi^1(P_1,P_2,\omega_1,\omega_2,\rho,w)n_2(P_1,P_2,\omega_1,\omega_2,\rho,w) - n_2(P_1,P_2,\omega_1,\omega_2,\rho,w) \tag{A.37}$$

For completeness, note that we require $0 \leq l \leq \overline{n}$, a restriction we impose in our numerical simulations in Section 2.2.3, where under certain extreme parameterizations time can be fully devoted to a single, particular task.

A.2.2.2. CES Utility with $\nu_i \neq 0$

Now relax the Cobb-Douglas assumption, allowing $\nu_i < 1$ and $\nu_i \neq 0$, and relax the assumption that $z_1 = z_2$. Again, suppose effective income net of savings (cash on hand) is fixed at $\overline{y}_t = R_t k_t - k_{t+1} + w_t\overline{n}$. Drop t subscripts for now. Note that CES production still allows for us to write $n_1(n_2)$ as a linear function and $q_1(q_2)$ as a linear function:

$$n_1(n_2) = \underbrace{\left(\frac{z_2}{z_1}\right)\left(\frac{\theta_2(1-\omega_2)}{\theta_1(1-\omega_1)}\right)^{\frac{1}{\rho}}\left(\omega_1\left[\frac{w\omega_1}{P_1(1-\omega_1)}\right]^{\frac{\nu_1}{1-\nu_1}}+1-\omega_1\right)^{\frac{\nu_1-\rho}{\rho\nu_1}}\left(\omega_2\left[\frac{w\omega_2}{P_2(1-\omega_2)}\right]^{\frac{\nu_2}{1-\nu_2}}+1-\omega_2\right)^{\frac{\rho-\nu_2}{\rho\nu_2}}}_{\phi^1(P_1,P_2,z_1,z_2,w)} n_2 \tag{A.38}$$

$$q_1(q_2) = \underbrace{\left(\frac{z_2}{z_1}\right)\left(\frac{P_1\theta_2\omega_2}{P_2\theta_1\omega_1}\right)^{\frac{1}{\rho}}\left(\omega_1+(1-\omega_1)\left[\frac{w\omega_1}{P_1(1-\omega_1)}\right]^{\frac{\nu_1}{\nu_1-1}}\right)^{\frac{\nu_1-\rho}{\rho\nu_1}}\left(\omega_2+(1-\omega_2)\left[\frac{w\omega_2}{P_2(1-\omega_2)}\right]^{\frac{\nu_2}{\nu_2-1}}\right)^{\frac{\rho-\nu_2}{\rho\nu_2}}}_{\phi^2(P_1,P_2,z_1,z_2,w)} q_2 \tag{A.39}$$

Using the fact that $q_2(n_2)$ can be written

$$q_2(n_2) = n_2\left[\frac{w\omega_2}{P_2(1-\omega_2)}\right]^{\frac{1}{1-\nu_2}} \tag{A.40}$$

we can write $q_1(n_2)$ like above:

$$q_1(n_2) = \phi^2(P_1,P_2,z_1,z_2,w)\left[\frac{w\omega_2}{P_2(1-\omega_2)}\right]^{\frac{1}{1-\nu_2}} n_2 \tag{A.41}$$

This gives:

$$n_2(P_1, P_2, z_1, z_2, w) = \frac{\overline{y}}{w + w\phi^1(P_1, P_2, z_1, z_2, w) + P_1\phi^2(P_1, P_2, z_1, z_2, w)\left[\frac{w\omega_2}{P_2(1-\omega_2)}\right]^{\frac{1}{1-\gamma_2}} + P_2\left[\frac{w\omega_2}{P_2(1-\omega_2)}\right]^{\frac{1}{1-\gamma_2}}} \tag{A.42}$$

A.2.3. Proofs

Lemma 1. Assume each household is a utility maximizer. Under Assumptions 1, 2, and 3, and under Theorem 1 of Green (1964) attributed to Leontief (1947), we can restrict our analysis to

$$\widetilde{u}_{th}(q_{1th}, \ldots, q_{ith}, \ldots, q_{Ith}, n_{1th}, \ldots, n_{Ith}) \tag{A.43}$$

where q_{ith} is some index that describes the grouping of market goods $\{q_{i1th}, \ldots, q_{ij_ith}, \ldots, q_{iJ_ith}\}$.

Proof. For notational simplicity, denote the marginal final utility for the consumption of market good q_{ij_it} by $MU_t(q_{ij_it})$ which is

$$MU_t(q_{ij_it}) = \frac{\partial u}{\partial c_{it}} \frac{\partial f_{it}}{\partial q_{ij_it}} \tag{A.44}$$

Let j_i and j_i' index two distinct components of $\boldsymbol{q}_{it}$. Theorem 1 of Green (1964) states that the indices comprising grouped market goods must be constructed so that

$$\frac{\partial}{\partial q_{i'j_{i'}t}} \left(\frac{MU_t(q_{ij_it})}{MU_t(q_{ij_i't})} \right) = 0 \quad \forall i' \neq i \tag{A.45}$$

where j_i' indexes a component of $\boldsymbol{q}_{i't}$. Note that under Assumption 3 we can write

$$\left(\frac{MU_t(q_{ij_it})}{MU_t(q_{ij_i't})} \right) = \frac{\frac{\partial u}{\partial c_{it}} \frac{\partial f_{it}}{\partial q_{ij_it}}}{\frac{\partial u}{\partial c_{it}} \frac{\partial f_{it}}{\partial q_{ij_i't}}} \tag{A.46}$$

$$= \frac{\frac{\partial f_{it}}{\partial q_{ij_it}}}{\frac{\partial f_{it}}{\partial q_{ij_i't}}} \tag{A.47}$$

(A.47) only depends on goods in $\boldsymbol{q}_{it}$. By Assumption 2, (A.45) thus holds. Together under Assumption 1, Theorem 1 of Green (1964) is satisfied. □

Lemma 2. The shadow price of activities c_{ith} associated with the consumption of q_{ith} is equal exactly to P_{it} if and only if $\frac{\partial \widetilde{f}_{ith}}{q_{ith}} = 1$.

Proof. (⇐) Suppose $\frac{\partial \widetilde{f}_{ith}}{\partial q_{ith}} = 1$, then clearly $\frac{\partial u}{\partial c_{ith}} = P_{it}\mu_{th}$ from (11).

(⇒) By contrapositive, suppose $\frac{\partial \widetilde{f}_{ith}}{q_{ith}} \neq 1$. Let $\widetilde{P}_{it}$ be a candidate shadow price. We will show this cannot equal P_{it}. Note that

$$\frac{\partial u}{\partial c_{ith}} = \widetilde{P}_{it}\mu_{th} \tag{A.48}$$

$$\text{and} \quad \frac{\partial u}{\partial c_{ith}} = \frac{P_{it}}{\frac{\partial \widetilde{f}_{ith}}{q_{ith}}}\mu_{th} \tag{A.49}$$

$$\Rightarrow \quad \widetilde{P}_{it} = \frac{P_{it}}{\frac{\partial \widetilde{f}_{ith}}{q_{ith}}} \neq P_{it} \tag{A.50}$$

□

Proposition 1. For each i, the value of the production of final good c_{ith} net of the cost of market expenditure (value added) is equal to $w_{th}n_{ith}$, the market value of time spent on task i.

Proof. From Becker (1965), let $\overline{p}_{ith}$ be market inputs per unit of output c_{ith} and let $\overline{w}_{ith}$ be off-market time per unit of c_{ith}. Then we have

$$q_{ith} = \overline{p}_{ith}c_{ith} \tag{A.51}$$

$$n_{ith} = \overline{w}_{ith}c_{ith} \tag{A.52}$$

which implies that implicit prices can be written

$$\overline{p}_{ith} = \frac{q_{ith}}{c_{ith}} = \frac{\frac{\partial f_{ith}}{\partial q_{ith}}\frac{q_{ith}}{c_{ith}}}{\frac{\partial f_{ith}}{\partial q_{ith}}} \tag{A.53}$$

$$\overline{w}_{ith} = \frac{n_{ith}}{c_{ith}} = \frac{\frac{\partial f_{ith}}{\partial n_{ith}}\frac{n_{ith}}{c_{ith}}}{\frac{\partial f_{ith}}{\partial n_{ith}}} \tag{A.54}$$

Now under the homogeneity of degree one assumption, we can apply Euler's Homogeneous Function Theorem to write:

$$c_{ith} = f_{ith}(q_{ith}, n_{ith}) = \frac{\partial f_{ith}}{\partial q_{ith}}q_{ith} + \frac{\partial f_{ith}}{\partial n_{ith}}n_{ith} \tag{A.55}$$

$$\Rightarrow \quad 1 = \frac{\partial f_{ith}}{\partial q_{ith}}\frac{q_{ith}}{c_{ith}} + \frac{\partial f_{ith}}{\partial n_{ith}}\frac{n_{ith}}{c_{ith}} \tag{A.56}$$

Note that the terms on the right hand side of (A.56) are the output elasticities. Denote this value as ω_{ith}, where in the context of this proof the output elasticity is not necessarily time independent. Since $\frac{\partial f_{ith}}{\partial q_{ith}}, \frac{\partial f_{ith}}{\partial n_{ith}} > 0$ and q_{ith}, n_{ith}, and $c_{ith} > 0$, then all terms on the right hand side must be between $(0, 1)$ implying $\omega_{ith} \in (0, 1)$. Thus we can write implicit prices in (A.53) and (A.54) as

$$\overline{p}_{ith} = \frac{\omega_{ith}}{\frac{\partial f_{ith}}{\partial q_{ith}}} \tag{A.57}$$

$$\overline{w}_{ith} = \frac{1 - \omega_{ith}}{\frac{\partial f_{ith}}{\partial n_{ith}}} \tag{A.58}$$

Now consider a version of the budget constraint:

$$\sum_{i=1}^{I} P_{it}q_{ith} \leq w_t\left(\overline{n} - \sum_{i=1}^{I} n_{it}\right) + R_t k_{th} - k_{t+1,h} \tag{A.59}$$

Multiply both sides of (A.51) by the market price P_{it} and both sides of (A.51) by the market wage w_{th} then substitute into the budget constraint to get:

$$\sum_{i=1}^{I} (P_{it}\overline{p}_{it} + w_{th}\overline{w}_{ith})c_{ith} \leq w_{th}\overline{n} + R_t k_{th} - k_{t+1,h} \tag{A.60}$$

Following Becker (1965) the price of final consumption ψ_{ith} is

$$\psi_{ith} = P_{it}\overline{p}_{ith} + w_{th}\overline{w}_{ith} \tag{A.61}$$

Now consider the nominal value of final consumption $\psi_{ith}c_{ith}$. Using (A.57) and (A.58) we can write this as:

$$\psi_{ith}c_{ith} = \left(\frac{P_{it}\omega_{ith}}{\frac{\partial f_{ith}}{\partial q_{ith}}} + \frac{w_{th}(1-\omega_{ith})}{\frac{\partial f_{ith}}{\partial n_{ith}}}\right)c_{ith} \tag{A.62}$$

Applying Euler's theorem as before:

$$\psi_{ith}c_{ith} = \left(\frac{P_{it}\omega_{ith}}{\frac{\partial f_{ith}}{\partial q_{ith}}} + \frac{w_{th}(1-\omega_{ith})}{\frac{\partial f_{ith}}{\partial n_{ith}}}\right)\left(\frac{\partial f_{ith}}{\partial q_{ith}}q_{ith} + \frac{\partial f_{ith}}{\partial n_{ith}}n_{ith}\right) \tag{A.63}$$

$$= q_{ith}P_{it}\omega_{ith} + \frac{\frac{\partial f_{ith}}{\partial n_{ith}}}{\frac{\partial f_{ith}}{\partial q_{ith}}}n_{ith}P_{it}\omega_{ith} + \frac{\frac{\partial f_{ith}}{\partial q_{ith}}}{\frac{\partial f_{ith}}{\partial n_{ith}}}q_{ith}w_{th}(1-\omega_{ith}) + n_{ith}w_{th}(1-\omega_{ith}) \tag{A.64}$$

But note that $\frac{\frac{\partial f_{ith}}{\partial n_{ith}}}{\frac{\partial f_{ith}}{\partial q_{ith}}}$ is just the marginal rate of technical substitution which must be

$$\frac{\frac{\partial f_{ith}}{\partial n_{ith}}}{\frac{\partial f_{ith}}{\partial q_{ith}}} = \frac{w_{th}}{P_{it}} \tag{A.65}$$

Thus (A.64) collapses to $\psi_{ith}c_{ith} = P_{it}q_{ith} + w_{th}n_{ith}$. Then the value added from home production is just the value of the home productive process less the value of market inputs purchased, i.e.:

$$\psi_{ith}c_{ith} - P_{it}q_{ith} = w_{th}n_{ith} \tag{A.66}$$

□

Proposition 2. Fix $P_1 = P_2 = 1$ and z_2. Consider the following cases separately:

i. If the outputs of home production are substitutes so that $\rho \in (0,1)$, then an increase (decrease) in z_1 is welfare improving and results in an increase (decrease) in equilibrium $\frac{q_1}{q_2}$ and an increase (decrease) in $\frac{n_1}{n_2}$.

ii. If the outputs of home production are complements so that $\rho \in (-\infty, 0)$, then an increase (decrease) in z_1 is welfare improving and results in a decrease (increase) in equilibrium $\frac{q_1}{q_2}$ and a decrease (increase) in $\frac{n_1}{n_2}$.

Proof. We present the proofs for the two sub-statements of the proposition together since they each rely on the quasi-Hicksian relative demand functions:

$$q_2(q_1) = \left(\frac{\overline{u} - \theta_1 z_1^{\rho} q_1^{\rho}\left(\frac{1-\omega}{\omega}\right)^{\rho-\rho\omega}P_1^{\rho-\rho\omega}}{\theta_2 z_2^{\rho}\left(\frac{1-\omega}{\omega}\right)^{\rho-\rho\omega}P_2^{\rho-\rho\omega}}\right)^{\frac{1}{\rho}} \tag{A.67}$$

$$n_2(n_1) = \left(\frac{\overline{u} - \theta_1 z_1^{\rho} n_1^{\rho}\left(\frac{\omega}{1-\omega}\right)^{\rho\omega}P_1^{\rho\omega-1}}{\theta_2 z_2^{\rho}\left(\frac{\omega}{1-\omega}\right)^{\rho\omega}P_2^{\rho\omega-1}}\right)^{\frac{1}{\rho}} \tag{A.68}$$

In (A.67) time use has been concentrated out using (21) and in (A.68) market purchases have been concentrated out using (22).

To understand how relative demand changes for any given utility level $\overline{u}$ due to a change in $\frac{z_1}{z_2}$, suppose z_2, q_1, and $\overline{u}$ are fixed, along with prices and elasticities. Partially differentiating (A.67) in z_1 we get:

$$\frac{\partial q_2(q_1)}{\partial z_1} = -q_2(q_1)^{\frac{1-\rho}{\rho}} \left(\frac{P_1}{P_2}\right)^{\rho-\rho\omega} \left(\frac{1}{z_2}\right)^{\rho} q_1^{\rho} z_1^{\rho-1} < 0 \quad \forall \rho \tag{A.69}$$

Thus for all possible ρ we are considering, an upward adjustment in z_1 leads to a downward adjustment in q_2 relative to q_1 at every possible utility level. An identical argument holds for $\frac{\partial n_2(n_1)}{\partial z_1}$. Thus, changes in z_1 affect the curvature of the indifference curves. Specifically, as a result of changes in z_1, relative market purchases and time use either fall dramatically or very little which is readily apparent following from the limiting conditions

$$\lim_{q_1 \to 0} \frac{\partial q_2(q_1)}{\partial z_1}\bigg|_{\rho \in (0,1)} = 0 \tag{A.70}$$

$$\lim_{q_1 \to 0} \frac{\partial q_2(q_1)}{\partial z_1}\bigg|_{\rho < 0} = -\infty \tag{A.71}$$

$$\lim_{q_1 \to \infty} \frac{\partial q_2(q_1)}{\partial z_1}\bigg|_{\rho \in (0,1)} = -\infty \tag{A.72}$$

$$\lim_{q_1 \to \infty} \frac{\partial q_2(q_1)}{\partial z_1}\bigg|_{\rho < 0} = 0 \tag{A.73}$$

The same limiting outcomes occur for $\frac{\partial n_2(n_1)}{\partial z_1}$. This implies that for every $\frac{q_1}{q_2}$ pair, holding q_1 fixed the associated value of q_2 declines, causing the indifference curves for every $\overline{u}$ to shift downward.

The sign of equilibrium changes to $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ depends on how the slope of the indifference curves change as a result of changes to z_1. Note that the composite utility function is homothetic, implying linear expansion paths. □

Proposition 3. Fix $z_1 = z_2 = 1$ and P_2. Consider the following cases separately:

i. If the outputs of home production are substitutes so that $\rho \in (0, 1)$, then an increase (decrease) in $\frac{P_1}{P_2}$ leads to a decrease (increase) in equilibrium $\frac{q_1}{q_2}$ and a decrease (increase) in equilibrium $\frac{n_1}{n_2}$.

ii. If the outputs of home production are complements so that $\rho \in (-\infty, 0)$, then an increase (decrease) in $\frac{P_1}{P_2}$ leads to a decrease (increase) in equilibrium $\frac{q_1}{q_2}$ and an increase (decrease) in $\frac{n_1}{n_2}$.

Proof. Suppose $P_2 = 1$. Consider the two cases separately, and note we need only prove monotonicity:

i. $\rho \in (0, 1)$ — If P_1 increases then the budget constraint shifts inward, as in Figure 1. Refer now to the marginal rate of substitution conditions in (25) and (26) and consider equilibrium choices which must satisfy

$$\frac{q_1}{q_2} = \left(\frac{P_1}{P_2}\right)^{\frac{1-\rho+\rho\omega}{\rho-1}} \tag{A.74}$$

$$\frac{n_1}{n_2} = \left(\frac{P_1}{P_2}\right)^{\frac{\rho\omega}{\rho-1}} \tag{A.75}$$

$\forall \rho, \omega \in (0, 1)$, $\frac{1-\rho+\rho\omega}{\rho-1} < 0$ and $\frac{\rho\omega}{\rho-1} < 0$. Thus as P_1 rises equilibrium $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ fall.

ii. $\rho < 0$ — Note that for $\rho < 0$, $\frac{1-\rho+\rho\omega}{\rho-1} < 0$ as long as $\rho < \frac{1}{1-\omega}$ which holds $\forall \omega \in (0, 1)$. But now the sign of $\frac{\rho\omega}{\rho-1}$ has changed and $\frac{\rho\omega}{\rho-1} > 0$. Thus $\frac{q_1}{q_2}$ and $\frac{n_1}{n_2}$ co-move negatively.

$\square$

Proposition 4. Consider the implications of two separate cases and their corresponding sub-cases:

i. Suppose $\frac{\rho(\omega_2-\omega_1)}{\rho-1} < 0$ so that $\frac{q_1}{q_2}$ is decreasing in w then one and only one of the following must hold:

   a. $\rho < 0$ and $\omega_2 < \omega_1$

   b. $\rho \in (0,1)$ and $\omega_2 > \omega_1$

ii. Suppose $\frac{\rho(\omega_2-\omega_1)}{\rho-1} > 0$ so that $\frac{q_1}{q_2}$ is increasing in w then one and only one of the following must hold:

   a. $\rho < 0$ and $\omega_2 > \omega_1$

   b. $\rho \in (0,1)$ and $\omega_2 < \omega_1$

Proof. We will prove each case separately:

i. Suppose $\frac{\rho(\omega_2-\omega_1)}{\rho-1} < 0$. Clearly, for all feasible $\rho < 1$ the denominator is less than 0. Therefore, we must have $\rho(\omega_2 - \omega_1) > 0$. This happens when $\rho < 0$ and $\omega_2 < \omega_1$ or when $\rho \in (0,1)$ and $\omega_2 > \omega_1$. The converse clearly holds as well.

ii. Suppose $\frac{\rho(\omega_2-\omega_1)}{\rho-1} > 0$. Clearly, for all feasible $\rho < 1$ the denominator is less than 0. Therefore, we must have $\rho(\omega_2 - \omega_1) < 0$. This happens when $\rho < 0$ and $\omega_2 > \omega_1$ or when $\rho \in (0,1)$ and $\omega_2 < \omega_1$. The converse, again, clearly holds.

$\square$

Lemma 3. In a two-good economy with CES utility and Cobb-Douglas home production. Suppose $\omega_2 > \omega_1$. Then

$$\phi^n = \left[\frac{(1-\omega_2)[\omega_2/(1-\omega_2)]^{\rho\omega_2}}{(1-\omega_1)[\omega_1/(1-\omega_1)]^{\rho\omega_1}}\right]^{\frac{1}{\rho-1}} < \left[\frac{\omega_2[(1-\omega_2)/\omega_2]^{(1-\omega_2)\rho}}{\omega_1[(1-\omega_1)/\omega_1]^{(1-\omega_1)\rho}}\right]^{\frac{1}{\rho-1}} = \phi^q \tag{A.76}$$

Proof. Conjecture (A.76) holds. We will show that this is true as long as $\omega_2 > \omega_1$:

$$\left[\frac{(1-\omega_2)[\omega_2/(1-\omega_2)]^{\rho\omega_2}}{(1-\omega_1)[\omega_1/(1-\omega_1)]^{\rho\omega_1}}\right]^{\frac{1}{\rho-1}} < \left[\frac{\omega_2[(1-\omega_2)/\omega_2]^{(1-\omega_2)\rho}}{\omega_1[(1-\omega_1)/\omega_1]^{(1-\omega_1)\rho}}\right]^{\frac{1}{\rho-1}} \tag{A.77}$$

$$\Leftrightarrow \quad \left(\frac{\omega_1}{1-\omega_1}\right)^{1-\rho\omega_1}\left(\frac{\omega_1}{1-\omega_1}\right)^{(\omega_1-1)\rho} < \left(\frac{\omega_2}{1-\omega_2}\right)^{1-\rho\omega_2}\left(\frac{\omega_2}{1-\omega_2}\right)^{(\omega_2-1)\rho} \tag{A.78}$$

$$\Leftrightarrow \quad \left(\frac{\omega_1}{1-\omega_1}\right)^{1-\rho} < \left(\frac{\omega_2}{1-\omega_2}\right)^{1-\rho} \tag{A.79}$$

$$\Leftrightarrow \quad \frac{\omega_1}{\omega_2} < \frac{1-\omega_1}{1-\omega_2} \tag{A.80}$$

$$\Leftrightarrow \quad \omega_2 > \omega_1 \quad \text{since} \quad \omega_1, \omega_2 \in (0,1) \tag{A.81}$$

$\square$

Proposition 5. Let $P_1 = P_2 = 1$ and $z_1 = z_2 = 1$. The following hold:

i. If $\rho \in (0,1)$ then consumption and time use flow toward the less time intensive task, but the consumption allocation changes more than the time use allocation.

ii. If $\rho < 0$ then consumption and time use flow toward the more time intensive task, but the consumption allocation changes more than the time use allocation.

Proof. Assume throughout the proof, WOLOG, that $\omega_2 > \omega_1$. We have

$$\left(\frac{q_1}{q_2}\right) = \underbrace{\left[\frac{\omega_2[(1-\omega_2)/\omega_2]^{(1-\omega_2)\rho}}{\omega_1[(1-\omega_1)/\omega_1]^{(1-\omega_1)\rho}}\right]^{\frac{1}{\rho-1}}}_{\phi^q(\omega_1,\omega_2,\rho)} w^{\frac{\rho(\omega_2-\omega_1)}{\rho-1}} \tag{A.82}$$

$$\left(\frac{n_1}{n_2}\right) = \underbrace{\left[\frac{(1-\omega_2)[\omega_2/(1-\omega_2)]^{\rho\omega_2}}{(1-\omega_1)[\omega_1/(1-\omega_1)]^{\rho\omega_1}}\right]^{\frac{1}{\rho-1}}}_{\phi^n(\omega_1,\omega_2,\rho)} w^{\frac{\rho(\omega_2-\omega_1)}{\rho-1}} \tag{A.83}$$

Note that:

$$\frac{\partial(q_1/q_2)}{\partial w} = \phi^q \frac{\rho(\omega_2-\omega_1)}{\rho-1} w^{\frac{\rho(\omega_2-\omega_1)}{\rho-1}-1} \tag{A.84}$$

$$\frac{\partial(n_1/n_2)}{\partial w} = \phi^n \frac{\rho(\omega_2-\omega_1)}{\rho-1} w^{\frac{\rho(\omega_2-\omega_1)}{\rho-1}-1} \tag{A.85}$$

where $\phi^q, \phi^n > 0$ always.

i. Suppose $\rho \in (0,1)$ and $\omega_2 > \omega_1$. Then $\frac{\rho(\omega_2-\omega_1)}{\rho-1} < 0$ so relative consumption and relative time use both fall, as process $i = 2$ sees relatively more inputs than previously. Process $i = 2$ is the *less* time intensive process. It follows that for $\rho \in (0,1)$, relative consumption and relative time use both flow toward the less time intensive process, but the consumption allocation changes more than the time use allocation since $\phi^n < \phi^q$ by Lemma 3.

ii. Suppose $\rho < 0$ and $\omega_2 > \omega_1$. Then $\frac{\rho(\omega_2-\omega_1)}{\rho-1} > 0$ so relative consumption and relative time use both rise, as process $i = 1$, the *more* time intensive process, sees a relative increase in resources compared to process $i = 2$. Again, the consumption allocation changes more than the time use allocation since $\phi^n < \phi^q$ again by Lemma 3.

□

Proposition 6. Suppose $z_1 = z_2 = 1$, but P_1 and P_2 are free. Consider only non-inflationary relative prices, so that either $P_1 < 1 < P_2$ or $P_2 < 1 < P_1$. Suppose, without loss of generality, that $\omega_2 > \omega_1$. The following must hold:

i. If $\frac{P_1}{P_2} < 1$ then $\frac{q_1}{q_2}$ changes faster than $\frac{n_1}{n_2}$ as w rises.

ii. If $\frac{P_1}{P_2} > 1$ then it is ambiguous as to whether $\frac{q_1}{q_2}$ or $\frac{n_1}{n_2}$ changes faster as w rises.

The signs of changes to relative consumption and time use depend on ρ and are the same as those outlined in Proposition 5. For $\omega_1 > \omega_2$, just exchange all index numbers.

Proof. Define the following intermediate parameters $\beta_1 = \frac{\rho\omega_1}{\rho-1}$ and $\beta_2 = \frac{\rho\omega_2}{1-\rho}$. These are the power coefficients on P_1 and P_2 respectively in the expression for $\frac{n_1}{n_2}$. The coefficients on P_1 and P_2 in the expression for $\frac{q_1}{q_2}$ are $\beta_1 - 1$ and $\beta_2 + 1$ respectively. By Proposition 5, relative consumption and time use both fall when $\rho \in (0,1)$ and rise when $\rho < 0$ as w rises. The same holds when $P_1 \neq P_2$ since P_1 and P_2 enter the expressions for relative consumption and time use positively.

i. Suppose $P_1 < 1 < P_2$. Then for all valid values of $\rho \neq 0$

$$P_1^{\beta_1} < P_1^{\beta_1 - 1} \tag{A.86}$$

$$P_2^{\beta_2} < P_2^{\beta_2 + 1} \tag{A.87}$$

Using the results from Lemma 3, this implies that

$$\phi^n P_1^{\beta_1} P_2^{\beta_2} < \phi^q P_1^{\beta_1 - 1} P_2^{\beta_2 + 1} \tag{A.88}$$

which implies that

$$\left| \frac{\partial (q_1/q_2)}{\partial w} \right| > \left| \frac{\partial (n_1/n_2)}{\partial w} \right| \qquad \forall \rho \in (0,1) \quad \text{and} \quad \forall \rho < 0 \tag{A.89}$$

ii. Suppose $P_2 < 1 < P_1$. Then

$$P_1^{\beta_1} > P_1^{\beta_1 - 1} \tag{A.90}$$

$$P_2^{\beta_2} > P_2^{\beta_2 + 1} \tag{A.91}$$

By Lemma 3, since $\phi^n < \phi^q$, then depending on the magnitudes of the underlying parameters and prices

$$\phi^n P_1^{\beta_1} P_2^{\beta_2} < \phi^q P_1^{\beta_1 - 1} P_2^{\beta_2 + 1} \tag{A.92}$$

$$\text{or} \quad \phi^n P_1^{\beta_1} P_2^{\beta_2} > \phi^q P_1^{\beta_1 - 1} P_2^{\beta_2 + 1} \tag{A.93}$$

The degree to which relative consumption moves more than relative time use is thus ambiguous if $\omega_2 > \omega_1$ and $P_2 < P_1$.

□

A.3. Estimation Details

A.3.1. Description of Hamiltonian Monte Carlo

For a thorough and readable, detailed treatment of Hamiltonian Monte Carlo (HMC), we recommend reading Neal (2011). For an equally thorough, more technical treatment, read Betancourt and Stein (2011). For a brief overview of HMC see Gelman et al. (2013c). Here, we provide the "Reader's Digest" overview of the sampler's properties and benefits since there are very few examples of this particular estimation technique being used in the econometrics literature.[4] This Appendix is not intended to be a full treatment, only an outline.

Recall, in traditional Bayesian MCMC posterior sampling (i.e. Gibbs sampling or Metropolis-Hastings sampling), the state of the sampler at iteration m is the m^{th} draw of the parameter set $\mathcal{P}^m$ from the posterior distribution with density $\pi(\mathcal{P} \mid \mathcal{D}, \mathcal{H})$. Thus a traditional sampler is a Markov Chain operating in a discrete, countable sampling space, with discrete dynamics. For example, the m^{th} draw of a Metropolis-Hastings sampler depends on the $m-1$ draw, and the econometrician knows this distribution has converged when autocorrelation between draws has been sufficiently minimized, depending on the nature of the estimation problem. HMC extends this concept to a sample space with continuous dynamics.[5] In Hamiltonian dynamics commonly employed in physical mechanics, the state of the system depends on both the forward "momentum" of the system $\mathcal{Q}$ and the "position" of the system $\mathcal{P}$, each of

[4]The only published example in the economics literature we were able to find is Burda (2015). Martin Burda also has a recent working paper providing an application of HMC to a discrete choice model (Burda and Daviet 2018).

[5]Recall, "dynamics" in this context refer to the dynamics of the parameter sampler, *not* actual time, as in the modeled structural dynamics of the underlying agent's decision process. One dynamic operates over the sampler dimension m while the other over the data dimension t. This can be confusing when estimating models wrought from a dynamic decision process or which have dynamic time series components.

the same dimension. The Hamiltonian equation associated with this system is the sum of "potential" $U(\cdot)$ and "kinetic" $K(\cdot)$ energy:

$$H(\mathcal{Q},\mathcal{P}) = U(\mathcal{P}) + K(\mathcal{Q}) \tag{A.94}$$

The partial derivatives of this equation will determine how the parameter space $\mathcal{P}$ evolves as the sampler proceeds, as well as the rate of this evolution $\mathcal{Q}$.

Since the "dynamics" of our particular task involve traversing a parameter space over which some posterior density function $\pi(\mathcal{P} \mid \mathcal{D})$ is defined, the potential energy function for HMC is just the negative log of the right hand side of (1.51):

$$U(\mathcal{P}) = -\ln\Big[\pi(\mathcal{D} \mid \mathcal{P})\pi(\mathcal{P})\Big] \tag{A.95}$$

Note that we can sample from this using HMC only if $U(\mathcal{P})$ is continuously differentiable over the entire sample space since HMC operates on Hamilton's equations which require computation of the gradient vectors for both $U(\cdot)$ and $K(\cdot)$. Let # denote the cardinality of a countable set. In practice, the kinetic energy function is defined conditional on $\mathcal{P}$ and taken to be a quadratic of the form:[6]

$$K(\mathcal{Q} \mid \mathcal{P}) = \frac{1}{2}\sum_{k=1}^{\#\mathcal{Q}}\sum_{r=1}^{\#\mathcal{Q}} \mathcal{Q}_k\mathcal{Q}_r\Lambda_{k,r}(\mathcal{P}) - \frac{1}{2}\ln\Big[\det\big(\boldsymbol{\Lambda}(\mathcal{P})\big)\Big] \tag{A.96}$$

where $\boldsymbol{\Lambda}(\mathcal{P})$ is what is called the mass matrix and may be constant. It could also be restricted to the identity matrix or simply be diagonal. At most it is a dense symmetric positive definite matrix that represents the variance/covariance of an underlying conditional Gaussian distribution function for the momentum vector:

$$\mathcal{Q} \mid \mathcal{P} \sim \mathcal{N}\big(\mathbf{0}, \boldsymbol{\Lambda}(\mathcal{P})\big) \tag{A.97}$$

A dense $\boldsymbol{\Lambda}(\mathcal{P})$ can help account for high local non-linearities in $U(\mathcal{P})$, though can be difficult to compute (Betancourt and Stein 2011). In our estimation, we allow $\boldsymbol{\Lambda}(\mathcal{P})$ to be diagonal and tune $\boldsymbol{\Lambda}(\mathcal{P})$ during a warm-up period using Stan's HMC implementation (Stan Development Team 2016).[7]

Having defined the objects on which we operate, we can summarize the algorithm described in detail in Neal (2011). Here is where the kinetic energy component of the Hamiltonian helps greatly speed up convergence and reduce autocorrelation in our sampling routine. After a sufficient warm-up period,[8][9] a sampling step proceeds as follows:

i. Given the position from the previous iteration $\mathcal{P}$, draw a new $\mathcal{Q}$ from (A.97). In a sense (A.97) is thus the HMC analog of what is commonly called a "proposal distribution" for an MCMC Metropolis-Hastings algorithm.

ii. Given $(\mathcal{Q},\mathcal{P})$, iterate on Hamilton's equations using the leapfrog method for L steps to get a pair $(\mathcal{Q}',\mathcal{P}')$.[10] This is your proposed new state.

[6] This is a more flexible form of the kinetic energy function than that in Neal (2011). See Betancourt and Stein (2011) for more details.

[7] Stan is free software for high-performance HMC which utilizes automatic differentiation, LAPACK, and BLAS libraries for computational efficiency and which can be implemented and executed in a number of top-end scientific computing programs like R, Python, Matlab, or executed simply from the shell. See http://mc-stan.org. For examples using the Stan language to execute HMC models, see Gelman et al. (2013b).

[8] This is akin to the "burnin" period for an MCMC operating under discrete dynamics.

[9] This is handled automatically by Stan and discussed thoroughly in the software manual, (Stan Development Team 2016).

[10] See Neal (2011) for a detailed description of this iterative method and how it operates on Hamilton's equations.

iii. Similar to the Metropolis-Hastings algorithm, draw a uniform random deviate $u \sim \mathcal{U}[0,1]$ and accept $(\mathcal{Q}', \mathcal{P}')$ as the next state if

$$u < \exp\left\{ -H(\mathcal{Q}', \mathcal{P}') + H(\mathcal{Q}, \mathcal{P}) \right\} \tag{A.98}$$

HMC is most useful when sampling from $\pi(\mathcal{P} \mid \mathcal{D})$ is computationally burdensome and may require an incredibly long chain of discrete draws in order to achieve convergence. This happens in cases where posterior draws are likely to be highly autocorrelated in a traditional MCMC, especially when using Metropolis-Hastings algorithms where high autocorrelation may also require low acceptance probabilities, thus further giving reason for a long sampling chain.

B. Appendix: Aggregate Burdens of an Aging Population

B.1. Proofs

Lemma 1. For all $\nu > 0$, Assumption 1 holds.

Proof. This proof is trivial, but requires the assumption that $c > \nu$ so that Inada conditions are satisfied and c is thus feasible. Under that assumption, clearly $\ln(c - \nu) < \ln c$. □

Lemma 2. If $0 < \nu < c_{ot}(1) - c_{ot}(0)$ then the ratio of the marginal utility of non-diseased consumption to diseased consumption is such that $MU_{ot}(0)/MU_{ot}(1) > 1$.

Proof. Assume Inada conditions are satisfied such that $c_{ot}(1) > \nu$ and $c_{ot}(0) > 0$. Given Lemma 1, $\nu > 0$. Note that:

$$MU_{ot}(1) = \frac{1}{c_{ot}(1) - \nu} \tag{B.1}$$

$$MU_{ot}(0) = \frac{1}{c_{ot}(0)} \tag{B.2}$$

Rearranging the inequality in the premise we get

$$c_{ot}(0) - \nu < c_{ot}(0) < c_{ot}(1) - \nu \tag{B.3}$$

$$\Rightarrow \quad \frac{1}{c_{ot}(0)} > \frac{1}{c_{ot}(1) - \nu} \tag{B.4}$$

□

Corollary 1. Aggregate growth $g_{Y,t}$ depends only on $wapr_t$, not population levels.

Proof. Define the period t gross output growth rate as $(1 + g_{Y,t}) = \frac{Y_t}{Y_{t-1}}$. Note that

$$Y_t = N_{yt} \cdot (1 + g_z)^t \cdot a_{yt}^{\alpha}(1 - l_{yt} - h_{yt})^{1-\alpha} \tag{B.5}$$

so

$$(1 + g_{Y,t}) = \frac{Y_t}{Y_{t-1}} = (1 + g_N)(1 + g_z)\left(\frac{a_{yt}}{a_{y,t-1}}\right)^{\alpha}\left(\frac{1 - l_{yt} - h_{yt}}{1 - l_{y,t-1} - h_{y,t-1}}\right)^{1-\alpha} \tag{B.6}$$

In Proposition 1 we showed that household policies depend only on $wapr_t$. Thus $g_{Y,t}$ depends only on $wapr_t$. □

Proposition 2. Assume the survival rate $s_{o,t+1} = s_o$ is constant and the diseased old distribution $\psi_t = \psi$ is stationary. Then along a BGP the working-age population ratio $wapr$ is constant and given by

$$wapr = \frac{1 + g_N}{s_o} \tag{B.7}$$

Proof. The population of young agents entering the economy in period t is

$$N_{yt} = (1 + g_N) N_{y,t-1} \tag{B.8}$$

The population of old agents evolves according to

$$N_{ot} = s_o N_{y,t-1} \tag{B.9}$$

Substituting for $N_{y,t-1}$ we can write:

$$\frac{N_{yt}}{N_{ot}} = \frac{1 + g_N}{s_o} \tag{B.10}$$

Note that $\frac{N_{yt}}{N_{ot}}$ is the working-age population ratio $wapr$. The right-hand side of the above does not depend on t. Thus:

$$wapr = \frac{1 + g_N}{s_o} \tag{B.11}$$

□

B.2. Calibration

B.2.0.1. Setting ρ — Ratio of Diseased to Non-Diseased Benefits

We calibrate ρ by using estimates from Hurd et al. (2013) and Mommaerts (2016). Mommaerts (2016) uses RAND Health and Retirement Study (HRS) data to estimate median permanent income (\$14, 157 in 2010 dollars) of sample respondents over age 65 from 1998-2010. We then use the Social Security Administration's rule of thumb permanent income replacement rate (0.4) to compute the implied Social Security benefits for the median retiree:

$$14,157 \cdot 0.4 = 5,662.80 \tag{B.12}$$

Using HRS data, Hurd et al. (2013) estimates that average total annual Medicare spending for demented individuals is \$5, 226. To compute a baseline for total benefits received by diseased agents we add \$5, 226 to Equation (B.12) to get \$10, 888.80. The steady state ratio of diseased to non-diseased benefits is:

$$\rho = \frac{10,888.80}{5,662.80} = 1.923 \tag{B.13}$$

B.2.0.2. Steady State Equations

The de-trended steady state conditions are below. In order they are as follows (B.14) thru (B.17) comprise the young agent's intertemporal consumption condition, intratemporal consumption/leisure condition, intratemporal leisure/hospice care condition, and budget constraint. (B.18) thru (B.19) define the old agents' market goods conditions. (B.20) and (B.21) describe the equilibrium factor prices. (B.22) describes the accidental bequest condition and (B.23) is the government budget constraint. [1] [2]

$$\frac{\gamma}{x_y^*} = \beta \cdot \frac{1+g_N}{wapr} \cdot R^* \left[\psi \frac{1-\sigma}{x_o^*(1)^{1-\sigma}(h_o^*)^{\sigma} - \nu} \left(\frac{h_o^*}{x_o^*(1)} \right)^{\sigma} + (1-\psi) \frac{1}{x_o^*(0)} \right] \tag{B.14}$$

$$\frac{\mu}{l_y^*} = \frac{\gamma}{x_y^*} w^*(1-\tau) \tag{B.15}$$

$$\frac{\mu}{l_y^*} = \eta \cdot \frac{\sigma}{x_o^*(1)^{1-\sigma}(h_o^*)^{\sigma} - \nu} \left(\frac{h_o^*}{x_o^*(1)} \right)^{\sigma-1} \frac{wapr}{\psi} \tag{B.16}$$

$$x_y^* + a_y^* \leq R^* \cdot b_y^* + w^* \cdot (1-\tau) \cdot (1 - l_y^* - h_y^*) \tag{B.17}$$

$$x_o^*(1) \leq R^* \cdot a_y^* + \rho \cdot T^*(0) \tag{B.18}$$

$$x_o^*(0) \leq R^* \cdot a_y^* + T^*(0) \tag{B.19}$$

$$R^* = 1 + (1+g_N)\alpha \left(\frac{(1 - l_y^* - h_y^*)}{a_y^*} \right)^{1-\alpha} + \delta \tag{B.20}$$

$$w^* = (1-\alpha) \left(\frac{a_y^*}{1 - l_y^* - h_y^*} \right)^{\alpha} \tag{B.21}$$

$$b_y^* = a_y^* \frac{1 - \frac{1+g_N}{wapr}}{1+g_N} \tag{B.22}$$

$$T^*(0) = \frac{w^* \cdot \tau \cdot wapr \cdot (1 - l_y^* - h_y^*)}{\rho \cdot \psi + 1 - \psi} \tag{B.23}$$

The steady state does not admit a neat closed form analytical solution due to the non-linearities in (B.16). We solve the steady state using Powell's hybrid method.

[1] Note that we substitute $s_o = \frac{1+g_N}{wapr}$ and take $z_0 = (1+g_N)^{\alpha}$.

[2] The equations presented here are first-order conditions after composing flow utility with the home production functions.

C. Appendix: Two Stage Budgeting Under Mental Accounting

C.1. Mathematical Appendix

C.1.1. Proofs

Proposition 1. Assume l_{it}, $\boldsymbol{p}_t$, and r_t are ex-ante known to the consumer at the beginning of the period when making budgeting decisions. Optimally choosing ex-ante budgets is equivalent to optimally choosing ex-post consumption if and only if consumers face perfect foresight over spending shocks ζ_{ijt} and no cognitive frictions ($\Gamma_{ijt} = 1$ for all j).

Proof. For the following proof, fix l_{it}, $\boldsymbol{a}_{it}$, and $\boldsymbol{p}_{it}$.

($\Leftarrow$) With perfect foresight, by definition $\boldsymbol{\zeta}_{it} = \mathbf{0}$ since there is no uncertainty. If $\Gamma_{ijt} = 1$ for all j then consumers make optimal updates to each budget weight for each j, where it is sufficient to drop the expectations operator since $\boldsymbol{\zeta}_{it} = \mathbf{0}$. Further, note that since z_{it} can be written as a function of q_{it} the equilibrium condition can be written

$$\frac{\partial u_{it}}{\partial q_{ijt}}\frac{\partial q_{ijt}}{\partial \theta_{ijt}} + \frac{\partial u_{it}}{\partial z_{it}}\frac{\partial z_{it}}{\partial \theta_{ijt}} \tag{C.1}$$

$$= \frac{\partial u_{it}}{\partial q_{ijt}}\frac{\partial q_{ijt}}{\partial \theta_{ijt}} + \frac{\partial u_{it}}{\partial z_{it}}\frac{\partial z_{it}}{\partial q_{ijt}}\frac{\partial q_{ijt}}{\partial \theta_{ijt}} \tag{C.2}$$

$$= \frac{\partial u_{it}}{\partial q_{ijt}} + \frac{\partial u_{it}}{\partial z_{it}}\frac{\partial z_{it}}{\partial q_{ijt}} = 0 \tag{C.3}$$

where expectations are dropped since there is no ex-ante uncertainty. Now suppose rather than choosing ex-ante budget weights, consumers choose ex-post consumption. The optimal choice of q_{ijt} must satisfy the equilibrium condition in (C.3). The choices are thus equivalent.

($\Rightarrow$) For this logical direction, we engage in proof by contrapositive. Suppose now that either consumers do not face perfect foresight, so that they do not know $\boldsymbol{\zeta}_{it}$ ex-ante, or there are meaningful cognitive frictions, that is $\sum_{j=1}^{J} \Gamma_{ijt} < J$.

First, let us consider the case where $\boldsymbol{\zeta}_{it}$ is unknown ex-ante but $\sum_{j=1}^{J} \Gamma_{ijt} = J$, so that there are no cognitive frictions. Note that ex-ante budgets θ^*_{ijt} must satisfy the main equilibrium condition. Denote ex-ante expected consumption by $\mathbb{E}_{it} q_{ijt} = \widetilde{q}_{ijt} = \frac{\theta^*_{ijt} l_{it} + \gamma_i a_{ijt}}{p_{jt}}$. Now replace $\frac{\theta^*_{ijt} l_{it} + \gamma_i a_{ijt}}{p_{jt}}$ with $\widetilde{q}_{ijt}$ in the expected indirect utility function. Note that for each j optimal expected consumption $\widetilde{q}^*_{ijt}$ must exactly solve

$$\mathbb{E}_{it}\left\{\frac{\partial u_{it}}{\partial \widetilde{q}_{ijt}} + \frac{\partial u_{it}}{\partial z_{it}}\frac{\partial z_{it}}{\partial \widetilde{q}_{ijt}}\right\} \tag{C.4}$$

$$= \mathbb{E}_{it}\left\{\frac{\alpha_{ij}}{\widetilde{q}_{ijt} + \frac{\zeta_{ijt}}{p_{jt}} + 1} - \frac{p_{jt}\alpha_{i,J+1}}{l_{it} - \sum_{j=1}^{J}[p_{jt}\widetilde{q}_{ijt} + \zeta_{ijt}] + m_i + r_t b_{it}}\right\} = 0 \tag{C.5}$$

Ex-post $q^*_{ijt}(\zeta_{ijt})$ satisfies (C.3) by construction. Since the random variable ζ_{ijt} enters (C.5) non-linearly, by Jensen's inequality $\widetilde{q}^*_{ijt} \neq q^*_{ijt}$ except for a measure-zero realization of ζ_{ijt}. Since $\widetilde{q}^*_{ijt}$ is completely determined by θ^*_{ijt}, l_{it}, a_{ijt}, and p_{jt}, it follows that ex-ante budget choices are not equivalent to choices of ex-post consumption if ζ_{ijt} is not known. Clearly this holds for all j.

Now suppose ζ_{ijt} is known but $\exists j$ such that $\Gamma_{ijt} = 0$, the consumer thus cannot choose an ex-ante budget for category j and sets his budget such that $\theta_{ijt} = \theta_{ij,t-1}$. There is no ex-ante uncertainty, so we can drop expectations. Except for measure zero values of $\theta_{ij,t-1}$,

$$\frac{\partial u_{it}(\theta_{ij,t-1})}{\partial \theta_{ijt}} + \frac{\partial u_{it}}{\partial z_{it}} \frac{\partial z_{it}(\theta_{ij,t-1})}{\partial \theta_{ijt}} \neq 0 \tag{C.6}$$

where the dependencies are to note that the first order condition is evaluated at $\theta_{ij,t-1}$. Now the level of consumption associated with the budget share is $\widetilde{q}_{ijt} = \frac{\theta_{ij,t-1} l_{it} + \gamma_i a_{ijt}}{p_{jt}}$, but this value will not force the left hand side of (C.6) to equal zero. Here, optimally chosen q^*_{ijt} satisfies (C.3) by construction. It follows that $\widetilde{q}_{ijt} \neq q_{ijt}$. □

Proposition 2. Suppose $\gamma_i > 0$ and $\alpha_{i,J+1} > 0$. For some y indexing ι_{it}, fix $\vartheta^*_{i,-y,t}$ and consider the partial sensitivity of ϑ^*_{iyt} to γ_i under the following conditions:

i. If $a_{i,\iota_{iyt},t} > 0$ and $\sum_{j \neq \iota_{iyt}} a_{ijt} > 0$ then $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} > 0$.

ii. If $a_{i,\iota_{iyt},t} > 0$ and $\sum_{j \neq \iota_{iyt}} a_{ijt} < 0$ then as long as $\alpha_{i,J+1}$ is sufficiently large $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} > 0$.

iii. If $a_{i,\iota_{iyt},t} < 0$ and $\sum_{j \neq \iota_{iyt}} a_{ijt} > 0$ then as long as $\alpha_{i,J+1}$ is sufficiently large $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} < 0$.

iv. If $a_{i,\iota_{iyt},t} < 0$ and $\sum_{j \neq \iota_{iyt}} a_{ijt} < 0$ then $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} < 0$

Proof. We will show the sufficient conditions that ensure $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} > 0$. Note from (3.8)

$$\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} = -\frac{\alpha_{i,\iota_{iyt}} \sum_{j=1}^{J} a_{ijt}}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} - \frac{\alpha_{i,J+1} a_{i,\iota_{iyt},t}}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \tag{C.7}$$

where expectations can be dropped due to the linearity of the optimal budget weight in the random variables ζ_{it}. Since $l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1}) > 0$ we can focus on conditions that ensure $\alpha_{i,\iota_{iyt}} \sum_{j=1}^{J} a_{ijt} + \alpha_{i,J+1} a_{i,\iota_{iyt},t} > 0$.

Suppose $a_{i,\iota_{iyt},t} > 0$ then we want conditions that ensure

$$\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1} > -\frac{\sum_{j \neq \iota_{iyt}} a_{ijt}}{a_{i,\iota_{iyt},t}} \tag{C.8}$$

Clearly, if $\sum_{j \neq \iota_{iyt}} a_{ijt} > 0$ then (C.8) always holds ensuring $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} > 0$. If $\sum_{j \neq \iota_{iyt}} a_{ijt} < 0$ then $-\frac{\sum_{j \neq \iota_{iyt}} a_{ijt}}{a_{i,\iota_{iyt},t}} > 0$. Since $\alpha_{i,\iota_{iyt}} \in (0,1)$, by theory, then we require $\alpha_{i,\iota_{iyt}}$ to be sufficiently large in order to ensure that budgeting responsiveness is increasing in the sensitivity parameter γ_i.

Now suppose $a_{i,\iota_{iyt},t} < 0$, so we want conditions that ensure

$$\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1} < -\frac{\sum_{j \neq \iota_{iyt}} a_{ijt}}{a_{i,\iota_{iyt},t}} \tag{C.9}$$

If $\sum_{j \neq \iota_{iyt}} a_{ijt} > 0$ then $-\frac{\sum_{j \neq \iota_{iyt}} a_{ijt}}{a_{i,\iota_{iyt},t}} > 0$. Thus, we need $\alpha_{i,J+1}$ to be sufficiently small relative to total under spending in all categories $j \neq \iota_{iyt}$, $\sum_{j \neq \iota_{iyt}} a_{ijt}$ in order that budget responsiveness is increasing in γ_i.

Finally, when $\sum_{j \neq \iota_{iyt}} a_{ijt} > 0$, it is clear that $-\frac{\sum_{j \neq \iota_{iyt}} a_{ijt}}{a_{i,\iota_{iyt},t}} < 0$, which implies that (C.9) never holds implying $\frac{\partial \vartheta^*_{iyt}}{\partial \gamma_i} < 0$. □

Proposition 3. Suppose $k_{it} > 0$, $l_{it} > 0$, and $\alpha_{i,J+1} > 0$. Without loss of generality, let $\boldsymbol{\iota}_{it} = (1, 2, 3, \ldots)$ and suppose $\boldsymbol{\iota}_{it}$ is of dimension $J' \leq J$. Consider the total responsiveness of the components of $\boldsymbol{\vartheta}^*_{it}$ to a_{iyt} where $y \in \boldsymbol{\iota}_{it}$.

i. Higher a_{iyt} leads to lower ϑ^*_{iyt}, i.e. $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}a_{iyt}} = -\frac{\gamma_i}{l_{it}}$.

ii. For all $s \in \boldsymbol{\iota}_{it}$ where $s \neq y$, $\frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{iyt}} = 0$.

Proof. We begin by differentiating the optimal budget share equilibrium condition in equation (8) in the main text. Note that expectations can be dropped since (8) is linear in unknown random variable $\boldsymbol{\zeta}_{it}$. Consider the system of implicit component-wise total derivatives:

$$\begin{aligned}
\frac{\mathrm{d}\vartheta^*_{i1t}}{\mathrm{d}a_{i1t}} &= -\frac{\gamma_i}{l_{it}} - \sum_{s \neq 1} \Gamma_{ist} \frac{\alpha_{i1}}{\alpha_{i1} + \alpha_{i,J+1}} \frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{i1t}} \\
\frac{\mathrm{d}\vartheta^*_{i2t}}{\mathrm{d}a_{i1t}} &= -\frac{\gamma_i}{l_{it}} \frac{\alpha_{i2}}{\alpha_{i2} + \alpha_{i,J+1}} - \sum_{s \neq 2} \Gamma_{ist} \frac{\alpha_{i2}}{\alpha_{i2} + \alpha_{i,J+1}} \frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{i1t}} \\
&\vdots \\
\frac{\mathrm{d}\vartheta^*_{iJ',t}}{\mathrm{d}a_{i1t}} &= -\frac{\gamma_i}{l_{it}} \frac{\alpha_{iJ'}}{\alpha_{iJ'} + \alpha_{i,J+1}} - \sum_{s \neq J'} \Gamma_{ist} \frac{\alpha_{iJ'}}{\alpha_{iJ'} + \alpha_{i,J+1}} \frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{i1t}}
\end{aligned} \tag{C.10}$$

Note that (C.10) is linear, so after some algebra, we can write this system in canonical form:

$$\boldsymbol{A} \cdot \frac{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}{\mathbf{d}a_{i1t}} = \boldsymbol{b} \tag{C.11}$$

where

$$\boldsymbol{A} = \begin{pmatrix}
1 & \frac{\alpha_{i1}}{\alpha_{i1}+\alpha_{i,J+1}} & \cdots & \cdots & \frac{\alpha_{i1}}{\alpha_{i1}+\alpha_{i,J+1}} \\
\frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} & 1 & \frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} & \cdots & \frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} \\
\vdots & \cdots & \ddots & \cdots & \vdots \\
\vdots & \cdots & \cdots & \ddots & \vdots \\
\frac{\alpha_{iJ'}}{\alpha_{iJ'}+\alpha_{i,J+1}} & \frac{\alpha_{iJ'}}{\alpha_{iJ'}+\alpha_{i,J+1}} & \cdots & \frac{\alpha_{iJ'}}{\alpha_{iJ'}+\alpha_{i,J+1}} & 1
\end{pmatrix} \tag{C.12}$$

$$\frac{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}{\mathbf{d}a_{i1t}} = \begin{pmatrix}
\frac{\mathrm{d}\vartheta^*_{i1t}}{\mathrm{d}a_{i1t}} \\
\frac{\mathrm{d}\vartheta^*_{i2t}}{\mathrm{d}a_{i1t}} \\
\vdots \\
\frac{\mathrm{d}\vartheta^*_{iJ',t}}{\mathrm{d}a_{i1t}}
\end{pmatrix} \tag{C.13}$$

$$\boldsymbol{b} = \begin{pmatrix}
-\frac{\gamma_i}{l_{it}} \\
-\frac{\gamma_i}{l_{it}} \frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} \\
\vdots \\
-\frac{\gamma_i}{l_{it}} \frac{\alpha_{iJ'}}{\alpha_{iJ'}+\alpha_{i,J+1}}
\end{pmatrix} \tag{C.14}$$

Conjecture that a solution to (C.11) is

$$\frac{\widetilde{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}}{\mathbf{d}a_{i1t}} = \begin{pmatrix} -\frac{\gamma_i}{l_{it}} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{C.15}$$

Note that (C.15) is indeed a solution, and it is the only solution since $\boldsymbol{A}$ has full rank J'. It is clear that

$$\boldsymbol{A} \cdot \frac{\widetilde{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}}{\mathbf{d}a_{i1t}} = \boldsymbol{b} \tag{C.16}$$

by inspecting the row-wise dot products on the left hand side of (C.16). This results holds for all values of $y \in \boldsymbol{\iota}_{it}$ indexing a_{iyt}, so that a simple re-arrangement and re-definition of the indices for $y \neq 1$ will yield the same outcome. ◻

Proposition 4. Suppose $k_{it} > 0$, $l_{it} > 0$, and $\alpha_{i,J+1} > 0$. Without loss of generality, let $\boldsymbol{\iota}_{it} = (1, 2, 3, \ldots)$ and suppose $\boldsymbol{\iota}_{it}$ is of dimension $J' < J$, strictly. Then both the sign and magnitude of the total responsiveness of the components of $\boldsymbol{\vartheta}^*_{it}$ to a_{ijt} where $j \notin \boldsymbol{\iota}_{it}$ is ambiguous and depends on the underlying values of the utility parameters $\boldsymbol{\alpha}_i$ and $\alpha_{i,J+1}$.

Proof. We will prove this by considering two numerical parameterizations and showing that both the signs and magnitudes of the total derivatives are different under the different parameterizations. Without loss of generality, let $J = 4$ and $J' = 3$. Let $\gamma_i = l_{it} = 1$. Let $\boldsymbol{\alpha}_i = (0.1, 0.15, 0.4, 0.35)^\top$ and consider two values for $\alpha_{i,J+1}$: $\alpha^1_{i,J+1} = 0.1$ and $\alpha^2_{i,J+1} = 5$. $\boldsymbol{\iota}_{it} = (1, 2, 3)$ so that $j = 4$ is the category for which $\Gamma_{ijt} = 0$. The relevant system of equations is

$$\begin{aligned}
\frac{\mathrm{d}\vartheta^*_{i1t}}{\mathrm{d}a_{i4t}} &= -\frac{\alpha_{i1}}{\alpha_{i1} + \alpha_{i,J+1}} - \sum_{s \neq 1} \Gamma_{ist} \frac{\alpha_{i1}}{\alpha_{i1} + \alpha_{i,J+1}} \frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{i4t}} \\
\frac{\mathrm{d}\vartheta^*_{i2t}}{\mathrm{d}a_{i4t}} &= -\frac{\alpha_{i2}}{\alpha_{i2} + \alpha_{i,J+1}} - \sum_{s \neq 2} \Gamma_{ist} \frac{\alpha_{i2}}{\alpha_{i2} + \alpha_{i,J+1}} \frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{i4t}} \\
\frac{\mathrm{d}\vartheta^*_{i3t}}{\mathrm{d}a_{i4t}} &= -\frac{\alpha_{i2}}{\alpha_{i2} + \alpha_{i,J+1}} - \sum_{s \neq 3} \Gamma_{ist} \frac{\alpha_{i3}}{\alpha_{i3} + \alpha_{i,J+1}} \frac{\mathrm{d}\vartheta^*_{ist}}{\mathrm{d}a_{i4t}}
\end{aligned} \tag{C.17}$$

After some algebra, we get the canonical linear system

$$\boldsymbol{A} \cdot \frac{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}{\mathbf{d}a_{i4t}} = \boldsymbol{b} \tag{C.18}$$

$$\text{where} \quad \boldsymbol{A} = \begin{pmatrix} 1 & \frac{\alpha_{i1}}{\alpha_{i1}+\alpha_{i,J+1}} & \frac{\alpha_{i1}}{\alpha_{i1}+\alpha_{i,J+1}} \\ \frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} & 1 & \frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} \\ \frac{\alpha_{i3}}{\alpha_{i3}+\alpha_{i,J+1}} & \frac{\alpha_{i3}}{\alpha_{i3}+\alpha_{i,J+1}} & 1 \end{pmatrix} \tag{C.19}$$

$$\boldsymbol{b} = \begin{pmatrix} \frac{\alpha_{i1}}{\alpha_{i1}+\alpha_{i,J+1}} \\ \frac{\alpha_{i2}}{\alpha_{i2}+\alpha_{i,J+1}} \\ \frac{\alpha_{i3}}{\alpha_{i3}+\alpha_{i,J+1}} \end{pmatrix} \tag{C.20}$$

Now consider two equilibrium solutions to this system under $\alpha^1_{i,J+1}$ and $\alpha^2_{i,J+1}$. Note that with $\alpha^1_{i,J+1} =$

0.1 we have

$$\frac{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}{\mathbf{d}a_{i4t}} = \begin{pmatrix} 0.227 \\ 0.033 \\ -0.933 \end{pmatrix} \tag{C.21}$$

With $\alpha^1_{i,J+1} = 5$ we have

$$\frac{\mathbf{d}\boldsymbol{\vartheta}^*_{it}}{\mathbf{d}a_{i4t}} = \begin{pmatrix} -0.014 \\ -0.023 \\ -0.073 \end{pmatrix} \tag{C.22}$$

Clearly, depending on the value of $\alpha_{i,J+1}$, ϑ^*_{i1t} either goes up when a_{i4t} goes up or goes down when a_{i4t} goes up. The same holds for ϑ^*_{i2t}. Note that $\alpha_{i1} < \alpha_{i2} < \alpha_{i3}$ suggesting that the magnitudes of $\boldsymbol{\alpha}_i$ and $\alpha_{i,J+1}$ both matter in determining the responsiveness of $\boldsymbol{\vartheta}^*_{it}$ to a_{ijt}, $j \notin \boldsymbol{\iota}_{it}$. □

C.1.2. Inversion of Optimal Budgeting Equation

Suppose there exists at least one j for which $\Gamma_{ijt} = 1$. For some y indexing $\boldsymbol{\iota}_{it}$, let $\boldsymbol{\vartheta}^*_{i,-yt}$ denote the vector of optimally chosen budget shares which does not include y. Note that this vector may be empty. The optimal choice of budget share $\vartheta^*_{iyt}(\boldsymbol{\vartheta}^*_{i,-y,t})$ for good ι_{iyt} can be expressed

$$\begin{aligned} \vartheta^*_{iyt}(\boldsymbol{\vartheta}^*_{i,-y,t}) = \mathbb{E}_{it}\Bigg\{ & \frac{\alpha_{i,\iota_{iyt}} l_{it} - \alpha_{i,\iota_{iyt}} \sum_{j\in\iota_{it}(\Gamma_{it})\, j\neq\iota_{iyt}} (\vartheta^*_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt})}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \\ & - \frac{(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})(\gamma_i a_{i,\iota_{iyt},t} + \zeta_{i,\iota_{iyt},t}) + \alpha_{i,J+1} p_{\iota_{iyt},t}}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \\ & - \frac{\alpha_{i,\iota_{iyt}} \left(\sum_{j=1}^{J} (1-\Gamma_{ijt})(\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}) - m_i - r_t b_{it} \right)}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \Bigg\} \end{aligned} \tag{C.23}$$

where $\theta_{ijt} = \theta_{ij,t-1}$ for $j \notin \boldsymbol{\iota}_{it}(\boldsymbol{\Gamma}_{it})$.

First, we will make a slight transformation to the indirect utility function which will greatly aid with computation of first order conditions. For each additively separable utility component $\ln(q_{ijt} + 1) = \ln(x_{ijt}/p_{jt} + 1) = \ln(x_{ijt} + p_{jt}) - \ln(p_{jt})$. We can thus transform indirect utility to permit direct substitution of x_{ijt} for each j without needing to divide out p_{jt}. This yields the expected indirect utility function

$$\mathbb{E}_{it} v_{it}(\boldsymbol{\theta}_{it}) = \mathbb{E}_{it}\Bigg\{ \sum_{j=1}^{J} \alpha_{ij} \ln(\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt} + p_{jt}) \tag{C.24}$$

$$+ \alpha_{i,J+1} \ln\left(l_{it} - \sum_{j=1}^{J} [\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}] + m_i + r_t b_{it} \right) \Bigg\} - \sum_{j=1}^{J} \ln(p_{jt}) \tag{C.25}$$

where $-\sum_{j=1}^{J} \ln(p_{jt})$ can be safely left outside of the expectation since we assume price levels, entering the week, are known to the consumer. Now under this parameterization, the first-order condition with respect to ϑ_{iyt} is

$$\mathbb{E}_t\left\{\frac{\alpha_{i,\iota_{iyt}} l_{it}}{\vartheta_{iyt} l_{it} + \gamma_i a_{i,\iota_{iyt},t} + \zeta_{i,\iota_{iyt},t} + p_{\iota_{iyt},t}}\right\}$$
$$= \mathbb{E}_t\left\{\frac{l_{it}\alpha_{i,J+1}}{l_{it} - \sum_{y'=1}^{k_t}(\vartheta_{iy',t} l_{it} + \gamma_i a_{i,\iota_{iy',t},t} + \zeta_{i,\iota_{iy',t},t}) - \sum_{j=1}^{J}(1-\Gamma_{ijt})(\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}) + m_i + r_t b_{it}}\right\} \tag{C.26}$$

Ignore expectations for now, divide by l_{it} which is assumed known, and multiply both sides by the terms in the denominator to get

$$\begin{aligned} &\alpha_{i,\iota_{iyt}} l_{it} - \alpha_{i,\iota_{iyt}} l_{it} \vartheta_{iyt} - \alpha_{i,\iota_{iyt}} \sum_{\substack{j \in \iota_{it}(\Gamma_{it}) \\ j \neq \iota_{iyt}}} (\vartheta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}) \\ &- \alpha_{i,\iota_{iyt}}(\gamma_i a_{i,\iota_{iyt},t} + \zeta_{i,\iota_{iyt},t}) - \alpha_{i,\iota_{iyt}} \sum_{j=1}^{J}(1-\Gamma_{ijt})(\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}) + \alpha_{i,\iota_{iyt}}(m_i + r_t b_{it}) \\ &= \alpha_{i,J+1}\vartheta_{iyt} l_{it} + \alpha_{i,J+1}(\gamma_i a_{i,\iota_{iyt},t} + \zeta_{i,\iota_{iyt},t}) + \alpha_{i,J+1} p_{\iota_{iyt},t} \end{aligned} \tag{C.27}$$

Isolate ϑ_{iyt} and take expectations to get (C.23).

C.1.3. Derivations of Elasticities

Assume l_{it}, $\boldsymbol{p}_t$, and r_t are ex-ante known to the consumer at the beginning of the period when making budgeting decisions. Assume $\boldsymbol{\zeta}_{it}$ is unknown and thus the only source of uncertainty. Further, assume $q_{ijt} > 0$, so that the own-price elasticity can be observed. Let ε^p_{ijt} be the ex-post own-price elasticity for good j in period t. This value is

$$\varepsilon^p_{ijt} = \begin{cases} -1 & \Gamma_{ijt} = 0 \\ \frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}\left(\frac{l_{it}}{q_{ijt}}\right) - 1 & \Gamma_{ijt} = 1 \end{cases} \qquad \text{where} \qquad \iota_{iyt} = j \tag{C.28}$$

To see this, start with the case where $\Gamma_{ijt} = 0$, so that category j does not feature an optimal budget update at the beginning of period t. Then $q_{ijt} = \mathbb{E}_{it} q_{ijt} + \frac{\zeta_{ijt}}{p_{jt}} = \frac{\mathbb{E}_{it} x_{ijt}}{p_{jt}} + \frac{\zeta_{ijt}}{p_{jt}}$, where $\mathbb{E}_{it} x_{ijt} = \theta_{ij,t-1} l_{it} + \gamma_i a_{ijt}$ is a constant and invariant in p_{jt} since no optimal budget update was made. This implies

$$\frac{\mathrm{d}q_{ijt}}{\mathrm{d}p_{jt}} = -\frac{\mathbb{E}_{it} x_{ijt} + \zeta_{ijt}}{p_{jt}^2} \tag{C.29}$$

$$\Rightarrow \quad \varepsilon^p_{ijt} = \frac{\mathrm{d}q_{ijt}}{\mathrm{d}p_{jt}}\left(\frac{p_{jt}}{q_{ijt}}\right) = -\frac{\mathbb{E}_{it} x_{ijt} + \zeta_{ijt}}{x_{ijt}} = -1 \tag{C.30}$$

since $\mathbb{E}_{it} x_{ijt} + \zeta_{ijt} = x_{ijt}$. Now consider when $\Gamma_{ijt} = 1$, so that optimally chosen budgets ϑ^*_{iyt} also depend on p_{jt}. By Leibniz's rule for differentiation

$$\frac{\mathrm{d}}{\mathrm{d}p_{jt}}\mathbb{E}_{it} x_{ijt} = \frac{\mathrm{d}}{\mathrm{d}p_{jt}}(\vartheta^*_{iyt}(p_{jt}) l_{it} + \gamma_i a_{ijt}) = l_{it}\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}} \qquad \text{where} \qquad \iota_{iyt} = j \tag{C.31}$$

Putting all of this together and multiplying by $\frac{p_{jt}}{q_{jt}}$ we get (C.28).

To derive the income elasticity, continue operating under the same assumptions regarding the agent's certainty set. We want to derive:

$$\varepsilon^l_{ijt} = \begin{cases} \frac{\theta_{ijt} l_{it}}{x_{ijt}} & \Gamma_{ijt} = 0 \\ \frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}\left(\frac{l^2_{it}}{x_{ijt}}\right) + \frac{\vartheta^*_{iyt} l_{it}}{x_{ijt}} & \Gamma_{ijt} = 1 \end{cases} \qquad \text{where} \qquad \iota_{iyt} = j \tag{C.32}$$

Start again with the case where $\Gamma_{ijt} = 0$, so that category j does not feature an optimal budget update at the beginning of period t. Then $\theta_{ijt} = \theta_{ij,t-1}$ is a constant. With $q_{ijt} = \frac{\theta_{ijt} l_{it} + \gamma_i a_{ijt} + \zeta_{ijt}}{p_{jt}}$ we get

$$\frac{\mathrm{d}q_{ijt}}{\mathrm{d}l_{it}} = \frac{\theta_{ijt}}{p_{jt}} \tag{C.33}$$

$$\Rightarrow \quad \varepsilon^l_{ijt} = \frac{\mathrm{d}q_{ijt}}{\mathrm{d}l_{it}}\left(\frac{l_{it}}{q_{ijt}}\right) = \frac{\theta_{ijt} l_{it}}{x_{ijt}} \tag{C.34}$$

Now with $\Gamma_{ijt} = 1$, by the chain rule we get:

$$\frac{\mathrm{d}q_{ijt}}{\mathrm{d}l_{it}} = \frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}\left(\frac{l_{it}}{p_{jt}}\right) + \frac{\vartheta^*_{iyt}}{p_{jt}} \tag{C.35}$$

$$\Rightarrow \quad \varepsilon^l_{ijt} = \frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}\left(\frac{l^2_{it}}{x_{ijt}}\right) + \frac{\vartheta^*_{iyt} l_{it}}{x_{ijt}} \qquad \text{where} \qquad \iota_{iyt} = j \tag{C.36}$$

where we get the latter expression after multiplying $\frac{\mathrm{d}q_{ijt}}{\mathrm{d}l_{it}}$ by $\frac{l_{it}}{q_{ijt}}$.

Now deriving expressions for $\frac{\mathrm{d}\theta^*_{ijt}}{\mathrm{d}p_{jt}}$ and $\frac{\mathrm{d}\theta^*_{ijt}}{\mathrm{d}l_{it}}$, start with $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{jt}}$ and note that:

$$\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{\iota_{iyt},t}} = \mathbb{E}_{it}\left\{\frac{\partial\vartheta^*_{iyt}}{\partial p_{\iota_{iyt},t}} + \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{iyt}}\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{\iota_{iyt},t}}\right\} \tag{C.37}$$

Since $p_{\iota_{iyt},t}$ and ζ_{ijt} are linearly independent in the expectation in Proposition 1, we can drop the expectation, re-factor, and write:

$$\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}p_{\iota_{iyt},t}} = \frac{\partial\vartheta^*_{iyt}}{\partial p_{\iota_{iyt},t}} \bigg/ \left(1 - \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{iyt}}\right) \tag{C.38}$$

$$\text{where} \quad \frac{\partial\vartheta^*_{iyt}}{\partial p_{\iota_{iyt},t}} = -\frac{\alpha_{i,J+1}}{l_{it}(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \tag{C.39}$$

$$\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}} = -\frac{\alpha_{i,\iota_{iyt}}}{\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1}} \tag{C.40}$$

From (C.40) it is clear that $\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}$ is constant in ϑ^*_{ist} for all s, so that (C.38) varies over time only through $\frac{\partial\vartheta^*_{iyt}}{\partial p_{\iota_{iyt},t}}$ which depends on l_{it}.

The total derivative of $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}$ has a slightly more complicated implicit representation:

$$\begin{aligned}\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}} &= \mathbb{E}_{it}\left\{\frac{\partial\vartheta^*_{iyt}}{\partial l_{it}} + \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\left(\frac{\partial\vartheta^*_{ist}}{\partial l_{it}} + \sum_{j\neq s}\Gamma_{i,\iota_{ijt},t}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{ijt}}\frac{\mathrm{d}\vartheta^*_{ijt}}{\mathrm{d}l_{it}}\right)\right\} \\ &= \mathbb{E}_{it}\left\{\frac{\partial\vartheta^*_{iyt}}{\partial l_{it}} + \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\frac{\partial\vartheta^*_{ist}}{\partial l_{it}} + \right. \\ &\left.\sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\sum_{j\notin\{y,s\}}\Gamma_{i,\iota_{ijt},t}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{ijt}}\frac{\mathrm{d}\vartheta^*_{ijt}}{\mathrm{d}l_{it}} + \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{iyt}}\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}\right\}\end{aligned} \tag{C.41}$$

Again, we can factor the last term on the far right hand side of (C.41) and isolate the total derivative of the budget weight associated with good ι_{iyt} as an implicit function of the total derivatives of the other budget weights:

$$\begin{aligned}\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}} = \mathbb{E}_{it}\Bigg\{&\left(\frac{\partial\vartheta^*_{iyt}}{\partial l_{it}} + \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\frac{\partial\vartheta^*_{ist}}{\partial l_{it}} + \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\sum_{j\notin\{y,s\}}\Gamma_{i,\iota_{ijt},t}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{ijt}}\frac{\mathrm{d}\vartheta^*_{ijt}}{\mathrm{d}l_{it}}\right) \\ &\Bigg/\left(1 - \sum_{s\neq y}\Gamma_{i,\iota_{ist},t}\frac{\partial\vartheta^*_{iyt}}{\partial\vartheta^*_{ist}}\frac{\partial\vartheta^*_{ist}}{\partial\vartheta^*_{iyt}}\right)\Bigg\}\end{aligned} \tag{C.42}$$

$$\begin{aligned}\text{where}\quad \frac{\partial\vartheta^*_{iyt}}{\partial l_{it}} &= -\frac{\alpha_{i,\iota_{iyt}}\sum_{j\in\iota_{it}(\Gamma_{it})}(\gamma_i a_{ijt} + \zeta_{ijt}) + \alpha_{i,J+1}(\gamma_i a_{i,\iota_{iyt},t} + \zeta_{i,\iota_{iyt},t}) + \alpha_{i,J+1}p_{\iota_{iyt},t}}{l_{it}^2(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})} \\ &\quad - \frac{\alpha_{i,\iota_{iyt}}\left(\sum_{j=1}^{J}(1-\Gamma_{ijt})(\gamma_i a_{ijt} + \zeta_{ijt}) - m_i - r_t b_{it}\right)}{l_{it}^2(\alpha_{i,\iota_{iyt}} + \alpha_{i,J+1})}\end{aligned} \tag{C.43}$$

Note that the expectation in (C.42) cannot be ignored since components of $\boldsymbol{\zeta}_{it}$ enter into terms $\frac{\partial\vartheta^*_{iyt}}{\partial l_{it}}$ for all y indexing $\boldsymbol{\iota}_{it}$. To compute $\frac{\mathrm{d}\vartheta^*_{iyt}}{\mathrm{d}l_{it}}$ for each y we iterate over the total derivative indexed by components of $\boldsymbol{\iota}_{it}$.

C.2. Description of Posterior Sampling Algorithm

In Section 4.2 of the main draft, we describe how the latent budgeting decisions and associated cognitive shocks $\{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}$ are sampled. Given $\{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}$, we now describe the Metroplis-Hastings within Gibbs sampling steps for parameters $\mathcal{P}_{in} \setminus \big\{\{\boldsymbol{\theta}_{itn}, \boldsymbol{a}_{itn}, \boldsymbol{\Gamma}_{itn}\}_{t=1}^{T_i}\big\}$.[1] We exploit conjugacy where possible, starting with describing the agent-level draws then moving on to describing the sampling statements for the global, hierarchical parameters. Note that the following agent-level parameters are associated with non-conjugate priors $\{\boldsymbol{a}_{i1n}, \alpha_{i,J+1,n}, \boldsymbol{\tau}_{in}\}$, while the remaining agent-level parameters $\{\boldsymbol{\sigma}^2_{in}, \boldsymbol{\beta}_{in}, \mu_{i,\alpha_{i,J+1},n}, \boldsymbol{\psi}_{in}\}$ can be conditionally sampled directly. In practice, we sample $\{\boldsymbol{a}_{i1n}, \alpha_{i,J+1,n}, \boldsymbol{\tau}_{in}\}$ all together using a random-walk Metropolis step with the adaptive algorithm described in Atchade and Rosenthal (2005).

[1] Again recall that parameters fixed for the entirety of the MCMC sampling routine are γ_i, m_i, α_{ij} for $j \in \{1, \ldots, J\}$, and $\boldsymbol{\sigma}^2_{a_{i1}}$.

Each agent is associated with an adaptive Metropolis step to conditionally draw the column vector

$$\begin{pmatrix} \boldsymbol{a}_{i1n} \\ \alpha_{i,J+1,n} \\ \boldsymbol{\tau}_{in} \end{pmatrix} \Bigg| \mathcal{D}_i, \mathcal{P}_{in} \setminus \{\boldsymbol{a}_{i1n}, \alpha_{i,J+1,n}, \boldsymbol{\tau}_{in}\}, \mathcal{H}_n \propto \mathcal{L}_i(\boldsymbol{a}_{i1n}, \alpha_{i,J+1,n}, \boldsymbol{\tau}_{in})$$
$$\times \left(\prod_j \underbrace{\phi\left(\frac{a_{ij1n}}{\sigma_{a_{ij1n}}}\right)}_{\text{Priors on } a_{ij1n}} \underbrace{\frac{\beta_{ij,n-1}^{\tau_{ijn}}}{\Gamma(\tau_{ijn})\sigma_{ij,n-1}^{2(\tau_{ijn}-1)}} \exp\left\{\frac{\beta_{ij,n-1}}{\sigma^2_{ijn}}\right\}}_{\text{Posterior on } \tau_{ijn} \text{ given } \sigma_{ij,n-1}} \right) \underbrace{\phi\left(\frac{\alpha_{i,J+1,n} - \mu_{i,\alpha_{i,J+1},n-1}}{10}\right)}_{\text{Prior on } \alpha_{i,J+1,n}} \tag{C.44}$$

The adaptation routine is agent-specific, in that each agent's Metropolis step is associated with its own acceptance rate. Let $\boldsymbol{\Lambda}_{in}$ be the column vector of parameters sampled in each agent's Metropolis step, where $\ln \boldsymbol{\tau}_{in}$ is the vector of logged shape parameters. We draw this vector $\boldsymbol{\Lambda}_{in}$ according to

$$\boldsymbol{\Lambda}_{in} = \begin{pmatrix} \boldsymbol{a}_{i1n} \\ \alpha_{i,J+1,n} \\ \ln \boldsymbol{\tau}_{in} \end{pmatrix} \sim \mathcal{MN}(\boldsymbol{\Lambda}_{i,n-1}, \chi_{i,n-1}\, \mathbb{I}_{2J+1}) \tag{C.45}$$

where $\mathcal{MN}$ denotes the multivariate normal distribution, parameterized in terms of the variance/covariance matrix, which is simply the identity matrix scaled by an adaptive scaling factor $\chi_{i,n-1}$. Let $\pi_i(\boldsymbol{\Lambda}_{in})$ be a collect-all expression for the prior densities of Metropolis sampled parameters on the right hand side of (C.44). Let acc_{in} be the algorithm's acceptance rate for agent i at iterate n, and let $\overline{acc} = 0.234$ be the optimally-targeted acceptance rate.[2] At each iterate, χ_{in} is updated according to:

$$\chi_{in} = \chi_{i,n-1} + \frac{2.5}{n}\left(\min\left\{1, \frac{\mathcal{L}_i(\boldsymbol{\Lambda}_{in})\pi_i(\boldsymbol{\Lambda}_{in})}{\mathcal{L}_i(\boldsymbol{\Lambda}_{i,n-1})\pi_i(\boldsymbol{\Lambda}_{i,n-1})}\right\} - \overline{acc}\right) \tag{C.46}$$

We set minimal and maximal values for χ_{in} of 0.0001 and 1000 respectively. Through simulations we find that the adaptation algorithm of Atchade and Rosenthal (2005) performs better in our high-dimensional model with substantial non-linearities and heterogeneity than similar adaptive Metropolis routines described in Haario, Saksman, and Tamminen (1998, 2001).

Our prior specifications allow for conditionally conjugate draws of $\boldsymbol{\sigma}^2_{in}$, $\boldsymbol{\beta}_{in}$, $\mu_{i,\alpha_{i,J+1},n}$, and $\boldsymbol{\psi}_{in}$. Given the independence of ζ_{ijt} and $\zeta_{ij',t}$ we can draw each of the J components of the aforementioned vectors separately. Let $\overline{\zeta}_{ijn}$ be the conditional sample mean of the expenditure error process under the current MCMC draw, and let $\overline{SSQ}_{ijn}$ be the sample sum of squared residuals for good j in the current draw, each defined as follows:

$$\overline{\zeta}_{ijn} = \frac{1}{T_i}\sum_t (x_{ijt} - \theta_{ijtn} l_{it} - a_{ijtn}) \tag{C.47}$$

$$\overline{SSQ}_{ijn} = \sum_t (x_{ijt} - \theta_{ijtn} l_{it} - a_{ijtn} - \overline{\zeta}_{ijn})^2 \tag{C.48}$$

In practice we draw the precisions $\frac{1}{\sigma^2_{ijn}}$ rather than the variances. The conditionally conjugate sampling statements are as follows:

$$\left(\frac{1}{\sigma^2_{ijn}}\right) \mid \mathcal{D}_i, \mathcal{P}_{in} \setminus \{\sigma_{ijn}\}, \mathcal{H}_n \sim \text{Gamma}\left(\tau_{ijn} + \frac{T_i}{2}, \frac{\beta_{ij,n-1}\overline{SSQ}_{ijn} + 2}{2\,\beta_{ij,n-1}}\right) \tag{C.49}$$

[2]See Roberts, Gelman, and Gilks (1997) for discussion of optimal acceptance rates for Metropolis algorithms.

$$\beta_{ijn} \mid \mathcal{D}_i, \mathcal{P}_{in} \setminus \{\beta_{ijn}\}, \mathcal{H}_n \sim \text{Gamma}\left(1 + \tau_{ijn}, \left(1 + \frac{1}{\sigma^2_{ijn}}\right)^{-1}\right) \tag{C.50}$$

$$\mu_{i,\alpha_{i,J+1},n} \mid \mathcal{D}_i, \mathcal{P}_{in} \setminus \{\mu_{i,\alpha_{i,J+1}.n}\}, \mathcal{H}_n \sim \mathcal{N}\left(\left(\frac{\mu_\mu}{\sigma^2_\mu} + \frac{\alpha_{i,J+1,n}}{100}\right)\left(\frac{1}{\frac{1}{\sigma^2_\mu} + \frac{1}{100}}\right), \frac{1}{\frac{1}{\sigma^2_\mu} + \frac{1}{100}}\right) \tag{C.51}$$

$$\psi_{ijn} \mid \mathcal{D}_i, \mathcal{P}_{in} \setminus \psi_{ijn}, \mathcal{H}_n \sim \text{Beta}\left(1 + \sum_t \Gamma_{ijtn}, T_i - \sum_t \Gamma_{ijtn}\right) \tag{C.52}$$

There are only two global parameters common to all agents — μ_μ and σ^2_μ, the mean and variance associated with each agent's $\mu_{i,\alpha_{i,J+1},n}$ parameter. Having placed empirical priors on these parameters, updating them is a matter of simply re-evaluating the formulae in (19) in Section 4.1 of the main text. Given substantial autocorrelation in the agent-level draws due to the complex filtration process needed to uncover the latent mental accounting series, we re-evaluate (19) only once every 100 times through the chain. This means that for each global parameter draw, we proceed through the agent-level sampling steps 100 times for each agent. The idea is to engage in the most computationally efficient sampling scheme without sacrificing inference.

C.3. MCMC Diagnostics

For a visual assessment of model performance and MCMC mixing properties, we plot the entire chain of posterior draws, averaging each draw across agents, for parameters that do not change over time. Specifically, if $\urcorner_{in}$ represents the n^{th} draw after burnin and trimming of parameter $\urcorner$ for agent i, then at each draw we compute $\widehat{\urcorner}_n = \frac{1}{I}\sum_i \urcorner_{in}$. For global parameters μ_μ and σ^2_μ there is no need to average across agents. These chains are presented in Figure C.1

To understand how well non time-dependent model parameters' posterior distributions are identified under our model specification, we run Kolmogorov-Smirnov tests on the MCMC draws against draws from prior distributions for parameters with proper priors. Here we do not average over agents, but rather we test each agent's posterior distribution against his corresponding prior for the set of parameters $\{\boldsymbol{\sigma}^2_i, \boldsymbol{\beta}_i, \alpha_{i,J+1}, \mu_{i,\alpha_{i,J+1}}, \boldsymbol{\psi}_i, \boldsymbol{a}_{i1}\}$. Specifically, for each agent we test the hypothesis

$$H_0: \text{Posterior}_i = \text{Prior}_i \tag{C.53}$$

Each test yields a p-value describing the probability of realizing the estimated posterior distribution under the null hypothesis. For each parameter, we take the average p-value across agents and present this statistic in Table C.1, along with the proportion of our sample for which we can reject the null hypothesis with 99%, 95%, and 90% confidence.

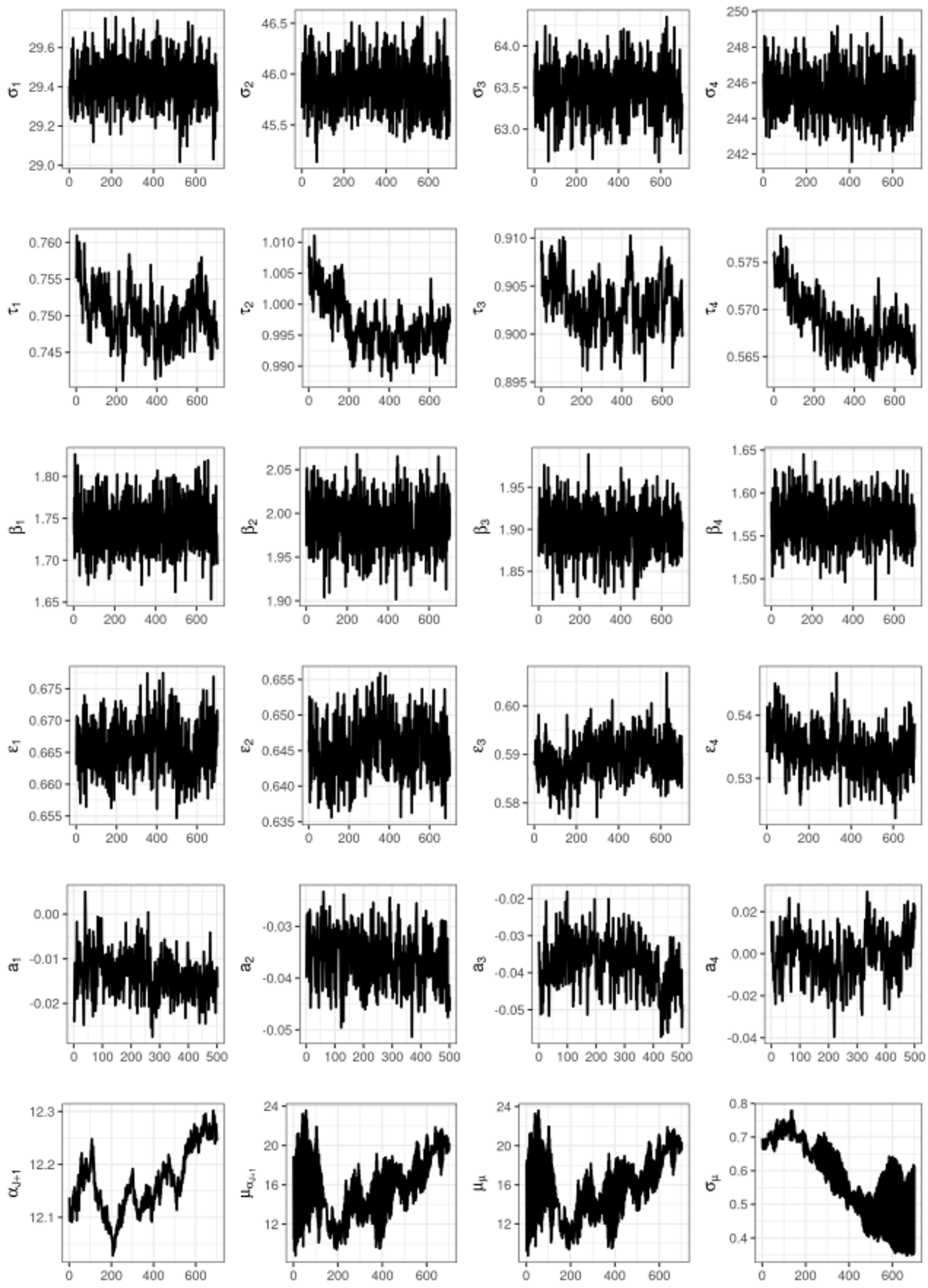


Figure C.1.: On the horizontal axis, 0 corresponds to the first MCMC draw after burnin. Recall the indexing of goods: Groceries ($j = 1$), Auto/Gasoline ($j = 2$), Food Away from Home ($j = 3$), Other ($j = 4$).

Table C.1.: K-S Tests Against Priors

Parameter	Avg. p-value	*Share Rejected* [a] 99%	95%	90%
σ^2_{i1}	0.000	1.000	1.000	1.000
σ^2_{i2}	0.000	1.000	1.000	1.000
σ^2_{i3}	0.000	1.000	1.000	1.000
σ^2_{i4}	0.000	1.000	1.000	1.000
β_{i1}	0.060	0.782	0.853	0.878
β_{i2}	0.069	0.758	0.834	0.864
β_{i3}	0.074	0.696	0.800	0.847
β_{i4}	0.104	0.575	0.707	0.772
$\alpha_{i,J+1}$	0.000	1.000	1.000	1.000
$\mu_{i,\alpha_{i,J+1}}$	0.000	1.000	1.000	1.000
ψ_{i1}	0.000	0.999	1.000	1.000
ψ_{i2}	0.000	0.999	1.000	1.000
ψ_{i3}	0.003	0.998	0.999	1.000
ψ_{i4}	0.002	0.998	0.999	0.999
$a_{i,1,1}$	0.002	0.999	1.000	1.000
$a_{i,2,1}$	0.000	0.999	1.000	1.000
$a_{i,3,1}$	0.003	0.998	0.999	1.000
$a_{i,4,1}$	0.002	0.998	0.999	0.999

The second subscript listed corresponds to the commodity group: groceries ($j = 1$), auto/gasoline ($j = 2$), food away from home ($j = 3$), other ($j = 4$).

[a] H_0 : Prior = Posterior.

C.4. Relations Between Mental Accounting, Savings, and Income — Regression Details

In Section 5.3 of the main draft we discuss the relations between various mental accounting features, savings, and income. Here, we present the details of the hierarchical regression underlying the main discussions. We consider four separate hierarchical regression specifications on posterior estimation outcomes. We estimate these regression models using Gibbs sampling as described in Rossi, Allenby, and McCulloch (2005):

$$\begin{aligned}\ln(\widehat{k}_{it}) = \gamma_{0i} + \gamma_{1i}\ln(l_{it}) + \gamma_{2i}\ln|z_{it}| + \gamma_{3i}\mathbf{1}(z_{it} > 0)\ln|z_{it}| \\ + \gamma_{4i}\ln|\widehat{a}_{i,J+1,t}| + \gamma_{5i}\mathbf{1}(\widehat{a}_{i,J+1,t} > 0)\ln|\widehat{a}_{i,J+1,t}| + \nu_{it}\end{aligned} \tag{C.54}$$

$$\begin{aligned}\ln(\widehat{k}_{it}) = \gamma_{0i} + \gamma_{1i}\ln(l_{it}) + \gamma_{2i}\ln|z_{it}/l_{it}| + \gamma_{3i}\mathbf{1}(z_{it} > 0)\ln|z_{it}/l_{it}| \\ + \gamma_{4i}\ln|\widehat{a}_{i,J+1,t}| + \gamma_{5i}\mathbf{1}(\widehat{a}_{i,J+1,t} > 0)\ln|\widehat{a}_{i,J+1,t}| + \nu_{it}\end{aligned} \tag{C.55}$$

$$z_{it} = \eta_{1i}\ln|\widehat{a}_{i,J+1,t}| + \eta_{2i}\mathbf{1}(\widehat{a}_{i,J+1,t} > 0)\ln|\widehat{a}_{i,J+1,t}| + \eta_{3i}\ln(l_{it}) + \xi_{it} \tag{C.56}$$

$$(z_{it}/l_{it}) = \eta_{1i}\ln|\widehat{a}_{i,J+1,t}| + \eta_{2i}\mathbf{1}(\widehat{a}_{i,J+1,t} > 0)\ln|\widehat{a}_{i,J+1,t}| + \eta_{3i}\ln(l_{it}) + \xi_{it} \tag{C.57}$$

Summary statistics over agent-level parameter estimates from these regressions are presented in Table C.2. The likelihoods are $\nu_{it}, \xi_{it} \sim \mathcal{N}(0, \tau_i)$ where τ_i is the respective likelihood precision with prior $\tau_i \propto 1/\chi^2_K$. Further, $\boldsymbol{\gamma}_i, \boldsymbol{\eta}_i \sim \mathcal{N}(0, \boldsymbol{V})$ with $\boldsymbol{V} \sim \text{InvWish}(K, K \cdot \mathbb{I})$, where K describes the number of right-hand side variables in the corresponding regression. We set the length of the Gibbs samplers to $10,000$ and estimate each model using the "bayesm" package available in R. The first two specifications allow us to decompose the degree to which the average total number of budget changes depends on relative over or under savings in the previous period ($\widehat{a}_{i,J+1,t}$) and income, while the second two specifications demonstrate how absolute savings and savings rates depend on what happened last period.[3] In Table C.2 we present summary statistics over agent-specific posterior mean parameter estimates from the correlative models (not causal) described above. The statistics presented in Table C.2 underly the main talking points and analysis in Section 5.3 of the text.

[3]Notice that the only difference between (C.54) and (C.55) is that savings levels are used in (C.54) and savings rates are used in (C.55).

Table C.2.: Regressions on Relations b/w Savings & Mental Accounting

Eq'n		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	S.D.
(C.54)	$\widehat{\gamma}_{0i}$	-3.763	0.053	0.512	0.507	0.966	4.563	0.350
	$\widehat{\gamma}_{1i}$	-1.442	-0.116	-0.038	-0.038	0.041	2.108	0.064
	$\widehat{\gamma}_{2i}$	-1.362	-0.024	0.014	0.014	0.053	1.590	0.038
	$\widehat{\gamma}_{3i}$	-1.532	-0.038	-0.011	-0.013	0.013	0.977	0.028
	$\widehat{\gamma}_{4i}$	-2.030	-0.042	0.002	0.000	0.044	1.678	0.043
	$\widehat{\gamma}_{5i}$	-1.099	-0.034	-0.007	-0.007	0.021	1.869	0.030
(C.55)	$\widehat{\gamma}_{0i}$	-3.111	0.171	0.603	0.600	1.032	4.103	0.323
	$\widehat{\gamma}_{1i}$	-1.535	-0.119	-0.042	-0.043	0.035	1.386	0.069
	$\widehat{\gamma}_{2i}$	-1.904	-0.033	0.007	0.008	0.049	1.194	0.039
	$\widehat{\gamma}_{3i}$	-1.447	-0.031	0.010	0.012	0.054	2.303	0.039
	$\widehat{\gamma}_{4i}$	-1.844	-0.039	0.004	0.002	0.046	1.635	0.043
	$\widehat{\gamma}_{5i}$	-1.703	-0.037	-0.010	-0.011	0.016	1.472	0.029
(C.56)	$\widehat{\eta}_{1i}$	-41.137	2.562	7.951	8.192	13.253	60.686	3.599
	$\widehat{\eta}_{2i}$	-130.109	-39.070	-25.951	-26.912	-13.248	89.000	10.952
	$\widehat{\eta}_{3i}$	-89.985	15.466	26.674	27.192	38.243	127.757	11.246
(C.57)	$\widehat{\eta}_{1i}$	-1.908	-0.006	0.031	0.035	0.073	2.154	0.036
	$\widehat{\eta}_{2i}$	-2.494	-0.163	-0.097	-0.104	-0.039	1.471	0.063
	$\widehat{\eta}_{3i}$	-2.196	0.034	0.073	0.075	0.115	2.183	0.044

D. Appendix: Durables, Non-Durables, and a Structural Test of Fungibility

D.1. Data Miscellany

In this appendix we present two tables. Table D.1 shows our broad classifications of durable and non-durable goods based on the 4-digit Visa merchant code descriptions. Table D.2 presents summary statistics for the broad marginal propensities to consume and first differences used

Table D.1.: Merchant Category Classification

Category	Durable?
Other svcs. (vet, horticulture, agriculture)	Yes
General contractors	Yes
Contracting services (construction)	Yes
Misc. publishing and printing	No
Specialty cleaning	No
Air travel	No
Car rental	No
Hotels	No
Public transit	No
Freight and courier svcs.	No
Marine and boating svcs.	Yes
Air travel, airports, and air fields	No
Travel agencies and tourism	No
Cable and telephone equipment and svcs.	No
Utilities (electric, gas, sanitary, water)	No
Medical, motor vehicle, office furniture, hardware	Yes
Office supplies, uniforms, books, chemicals	Yes
Home supply warehouses and hardware stores	Yes
Wholesale clubs/dept. Stores	Yes
Groceries	No
Automotive and gasoline	No
Clothing	Yes
Furniture and home electronic sales	Yes
Restaurants, bars, liquor stores	No
Misc.religious products, tents, swimming pools, tobacco, flowers, etc. (day to day and hobbies)	No
Financial svcs.	Yes
Insurance	Yes
Misc. hotels and recreation	No
Misc. svcs. (laundry, dating, massage, escort, photography, health and beauty)	No
Misc. svcs. (advertising, computer prog., information retrieval, bus. Consulting, janitorial, secretarial)	Yes
Automotive repair and cleaning misc.	Yes
HVAC and furniture repair	Yes
Movie theaters and video rentals	No
Bowling, dance halls, golf, club memberships, video gaming, aquariums and other recreation svcs.	No
Medical and dental	Yes
Legal svcs.	Yes
Education (including colleges)	Yes
Child care and charitable svcs.	Yes
Civic, religious, and political organizations	No
Non-medical testing laboratories	Yes
Architectural, account'ing, plus misc. prof. Svcs.	Yes
Court costs, fines, bail and bonds	Yes
Taxes	Yes
Postal svcs. And intra-gov't transactions	No

Table D.2.: Summary Statistics: Marginal Propensities

Statistic	N	Mean	St. Dev.	Min	Max
$MPC_i(z)$	10,690	−0.922	104.359	−9,328.642	1,778.450
$MPND_i(z)$	10,690	0.101	49.516	−2,627.695	1,778.493
$MPD_i(z)$	10,690	−0.750	64.592	−6,222.867	1,039.497
$MPC_i(m)$	10,690	415.052	36,211.200	−259,806.000	3,601,747.000
$MPND_i(m)$	10,690	71.076	10,732.410	−220,949.500	1,074,203.000
$MPD_i(m)$	10,690	3.255	2,577.278	−173,748.800	183,706.500
$\mathcal{D}C_i$	10,690	415.975	36,211.030	−259,806.200	3,601,716.000
$\mathcal{D}ND_i$	10,690	70.975	10,732.540	−220,949.600	1,074,204.000
$\mathcal{D}D_i$	10,690	4.005	2,578.079	−173,749.000	183,704.900

Bibliography

Acemoglu, Daron, and Veronica Guerrieri. 2008. “Capital Deepening and Nonbalanced Economic Growth”. *Journal of Political Economy* 116 (3). (Cit. on pp. 3, 33).

Aguiar, Mark, and Erik Hurst. 2005. “Consumption versus Expenditure”. *Journal of Political Economy* 113 (5). (Cit. on p. 1).

– . 2007. “Life-Cycle Prices and Production”. *The American Economic Review* 97 (5). (Cit. on p. 1).

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. 2013. “Time Use During the Great Recession”. *The American Economic Review* 103 (5). (Cit. on pp. 1, 2).

Alzheimer’s Association. 2011. “Alzheimer’s Association Report: 2011 Alzheimer’s disease facts and figures”. *Alzheimer’s & Dementia* 7:208–244. (Cit. on p. 46).

Atchade, Yves, and Jeffrey Rosenthal. 2005. “On adaptive Markov chain Monte Carlo algorithms”. *Bernoulli* 11 (5): 815–828. (Cit. on pp. 86, 143, 144).

Attanasio, Orazio, and Nicola Pavoni. 2011. “Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption”. *Econometrica* 79 (4). (Cit. on p. 105).

Autor, David, and David Dorn. 2013. “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market”. *The American Economic Review* 103 (5): 1553–1597. (Cit. on p. 3).

Azomahou, Théophile, Bity Diene, and Luc Soete. 2009. “The role of consumption and the financing of health investment under epidemic shocks”. *United Nations University Working Paper Series: #2009-006*. (Cit. on p. 47).

Barnett, William. 1977. “Pollak and Wachter on the Household Production Function Approach”. *Journal of Political Economy* 85 (5). (Cit. on p. 5).

Barnett, William, Douglas Fisher, and Apostolos Serletis. 1992. “Consumer Theory and the Demand for Money”. *Journal of Economic Literature* 30. (Cit. on p. 73).

Becker, Gary. 1965. “A Theory of the Allocation of Time”. *The Economic Journal* 75 (299). (Cit. on pp. ii, 1, 5, 51, 125, 126).

Benartzi, Shlomo, and Richard Thaler. 1995. “Myopic Loss Aversion and the Equity Premium Puzzle”. *The Quarterly Journal of Economics* 110 (1): 73–92. (Cit. on p. 101).

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. “Homework in Macroeconomics: Household Production and Aggregate Fluctuations”. *Journal of Political Economy* 99 (6). (Cit. on p. 1).

Bernanke, Ben. 1985. “Adjustment Costs, Durables, and Aggregate Consumption”. *Journal of Monetary Economics* 15:41–68. (Cit. on p. 27).

Betancourt, Michael, and Leo Stein. 2011. "The Geometry of Hamiltonian Monte Carlo". *arXiv:1112.4118*. (Cit. on pp. 35, 130, 131).

Boerma, Job, and Loukas Karabarbounis. 2019. "Inferring Inequality with Home Production". *Working Paper*. (Cit. on p. 1).

Boppart, Timo. 2014. "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences". *Econometrica* 82 (6). (Cit. on pp. 2, 27–29, 41).

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2017. "Memory, Attention, and Choice". *NBER Working Paper #23,356*. (Cit. on p. 67).

– . 2013. "Salience and Consumer Choice". *Journal of Political Economy* 121 (5). (Cit. on pp. 67, 77).

Bridgman, Benjamin, George Duernecker, and Berthold Herrendorf. 2018. "Structural transformation, marketization, and household production around the world". *Journal of Development Economics* 133:102–126. (Cit. on pp. 1, 2, 8).

Brock, William. 1974. "Money and Growth: The Case of Long Run Perfect Foresight". *International Economic Review* 15 (3). (Cit. on p. 73).

Browning, Martin, and Dolores Collado. 2001. "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates". *The American Economic Review* 91 (3): 681–692. (Cit. on p. 67).

Buera, Francisco, and Joseph Kaboski. 2009. "Can Traditional Theories of Structural Change Fit the Data?" *Journal of the European Economic Association* 7 (2-3): 469–477. (Cit. on pp. 2, 27).

– . 2012. "The Rise of the Service Economy". *The American Economic Review* 102 (6). (Cit. on p. 3).

Burda, Martin. 2015. "Constrained Hamiltonian Monte Carlo in BEKK GARCH with Targeting". *Journal of Time Series Econometrics* 7 (1). (Cit. on p. 130).

Burda, Martin, and Remi Daviet. 2018. "Hamiltonian Sequential Monte Carlo with Application to Consumer Choice Behavior". *University of Toronto, Department of Economics Working Paper #618*. (Cit. on p. 130).

Bureau of Labor Statistics. 2017. "American Time Use Survey (ATUS)". Visited on 11/20/2017. `https://www.bls.gov/tus/home.htm`. (Cit. on pp. 46, 48).

Calvo, Guillermo. 1983. "Staggered Price in a Utility-Maximizing Framework". *Journal of Monetary Economics* 12. (Cit. on p. 73).

Campbell, John, and Angus Deaton. 1989. "Why is Consumption so Smooth?" *Review of Economic Studies* 56:357–374. (Cit. on p. 105).

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition". *American Economic Review* 105 (7): 2183–2203. (Cit. on p. 67).

Carmichael, F., S. Charles, and C. Hulme. "Who will care? Employment participation and willingness to supply informal care". (Cit. on p. 48).

Carroll, Christopher. 1997. "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis". *The Quarterly Journal of Economics* 62 (1). (Cit. on p. 110).

– . 2006. “The method of endogenous gridpoints for solving dynamic stochastic optimization problems”. *Economics Letters* 91. (Cit. on p. 58).

Cheema, Amar, and Dilip Soman. 2006. “Malleable Mental Accounting: The Effect of Flexibility on the Justification of Attractive Spending and Consumption Decisions”. *Journal of Consumer Psychology* 16 (1): 33–44. (Cit. on p. 73).

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. “Salience and Taxation: Theory and Evidence”. *American Economic Review* 99 (4): 1145–1177. (Cit. on p. 67).

Cooley, Thomas, and Espen Henriksen. 2018. “The Demographic Deficit”. *Journal of Monetary Economics* 93. (Cit. on pp. 49, 59, 61).

Cowan, Nelson. 2000. “The magical number 4 in short-term memory: A reconsidration of mental storage capacity”. *Behavioral and Brain Sciences* 24:87–185. (Cit. on p. 76).

Deaton, Angus, and John Muellbauer. 1980a. “An Almost Ideal Demand System”. *The American Economic Review* 70 (3). (Cit. on p. 72).

– . 1980b. *Economics and Consumer Behavior*. Cambridge University Press: New York. (Cit. on pp. 66, 73, 74, 80, 84).

DeJong, David, and Chetan Dave. 2011. *Structural Macroeconometrics*. Second. Princeton University Press. (Cit. on p. 57).

DeNardi, Mariacristina, Eric French, and John Jones. 2010. “Why Do the Elderly Save? The Role of Medical Expenses”. *Journal of Political Economy* 118 (1). (Cit. on p. 47).

Edwards, Ryan. 2008. “Health Risk and Portfolio Choice”. *Journal of Business & Economic Statistics* 26 (4). (Cit. on p. 47).

Epper, Thomas, et al. 2020. “Time Discounting and Wealth Inequality”. *American Economic Review* 110 (4): 1177–1205. (Cit. on p. 68).

Farhi, Emmanuel, and Xavier Gabaix. 2020. “Optimal Taxation with Behavioral Agents”. *American Economic Review* 110 (1): 298–336. (Cit. on pp. 67, 76).

Feenstra, Robert. 1986. “Functional Equivalence Between Liquidity Costs and the Utility of Money”. *Journal of Monetary Economics* 17. (Cit. on p. 73).

Feinberg, Richard. 1986. “Credit Cards as Spending Facilitating Stimuli: A Conditioning Interpretation”. *Journal of Consumer Research* 13 (3): 348–356. (Cit. on p. 67).

Felső, Flóra, and Adriaan Soetevent. 2014. “Broad and narrow bracketing in gift certificate spending”. *European Economic Review* 66:284–302. (Cit. on pp. 67, 76).

Fernández-Villaverde, Jesús, and Dirk Krueger. 2011. “Consumption and Saving Over the Life Cycle: How Important are Consumer Durables?” *Macroeconomic Dynamics* 15:725–770. (Cit. on p. 102).

Fink, Daniel. 1997. *A Compendium of Conjugate Priors*. (Cit. on p. 84).

Flavin, Marjorie. 1993. “The Excess Smoothness of Consumption: Identification and Interpretation”. *Review of Economic Studies* 60. (Cit. on p. 105).

French, Eric, and John Bailey Jones. 2011. “The Effects of Health Insurance and Self-Insurance on Retirement Behavior”. *Econometrica* 79 (3). (Cit. on p. 47).

Gabaix, Xavier. 2014. “A Sparsity-Based Model of Bounded Rationality”. *The Quarterly Journal of Economics* 129 (4): 1661–1710. (Cit. on pp. iii, 67, 76, 79, 93).

Gathergood, John, et al. 2019. “How Do Individuals Repay Their Debt? The Balance-Matching Heuristic”. *American Economic Review* 109 (3): 844–875. (Cit. on p. 73).

Geary, Roy. 1950. “A Note on ‘A Constant-Utility Index of the Cost of Living’”. *Review of Economic Studies* 18 (2): 65–66. (Cit. on pp. 2, 74, 84).

Gelman, Andrew, et al. 2013a. *Bayesian Data Analysis*. Third. Taylor & Francis, CRC Press: Boca Raton, FL. (Cit. on p. 89).

– . 2013b. *Computation in R and Stan*. (Cit. on pp. 35, 37, 131).

– . 2013c. *Computationally efficient Markov chain simulation*. (Cit. on pp. 35, 130).

Gelman, Michael, et al. 2014. “Harnessing naturally occurring data to measure the response of spending to income”. *Science* 345 (6193): 212–215. (Cit. on p. 69).

Geman, Stuart, and Donald Geman. 1984. “Stochastic Relaxation, Gibbs Distributions, and Bayesian Restoration of Images”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6 (6): 721–741. (Cit. on p. 82).

Giandrea, Michael, and Shawn Sprague. 2017. “Estimating the U.S. Labor Share”. *Monthly Labor Review, U.S. Bureau of Labor Statistics*. `https://doi.org/10.21916/mlr.2017.7`. (Cit. on p. 28).

Gomme, Paul, Finn Kydland, and Peter Rupert. 2001. “Home Production Meets Time to Build”. *Journal of Political Economy* 109 (5). (Cit. on pp. 1, 27).

Graham, John, and Carole Green. 1984. “Estimating the Parameters of a Household Production Function with Joint Products”. *The Review of Economics and Statistics* 66 (2). (Cit. on p. 1).

Green, H.A. John. 1964. *The Concept of Functional Separability*. (Cit. on pp. 6, 124).

Gronau, Quentin, et al. 2017. “A Tutorial on Bridge Sampling”. *arXiv:1703.05984v2*. (Cit. on p. 37).

Gronau, Reuben. 1977. “Leisure, Home Production, and Work – the Theory of the Allocation of Time Revisited”. *Journal of Political Economy* 85 (6). (Cit. on p. 1).

Gross, David, and Nicholas Souleles. 2002. “Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data”. *The Quarterly Journal of Economics*. (Cit. on p. 100).

Haario, Heikki, Eero Saksman, and Johanna Tamminen. 1998. “Adaptive proposal distribution for random walk Metropolis algorithm”. *Computational Statistics*. (Cit. on p. 144).

– . 2001. “An adaptive Metropolis algorithm”. *Bernoulli* 7 (2). (Cit. on p. 144).

Hall, Robert. 1978. “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence”. *Journal of Political Economy* 86 (6): 971–987. (Cit. on p. 67).

Hall, Robert, and Charles Jones. 2007. “The Value of Life and the Rise in Health Spending”. *The Quarterly Journal of Economics*. (Cit. on p. 47).

Hastings, Justine, and Jesse Shapiro. 2013. “Fungibility and Consumer Choice: Evidence from Commodity Price Shocks”. *The Quarterly Journal of Economics* 128 (4). (Cit. on pp. 75, 100).

– . 2017. “How are SNAP Benefits Spent? Evidence from a Retail Panel”. *NBER Working Paper*. (Cit. on p. 100).

– . 2018. “How Are SNAP Benefits Spent? Evidence from a Retail Panel”. *American Economic Review* 108 (12): 3493–3540. (Cit. on p. 69).

Hastings, W.K. 1970. “Monte Carlo Sampling Methods Using Markov Chains and Their Applications”. *Biometrika* 57 (1): 97–109. (Cit. on p. 82).

Heath, Chip, and Jack Soll. 1996. “Mental Budgeting and Consumer Decisions”. *Journal of Consumer Research* 23 (1): 40–52. (Cit. on pp. 66, 100).

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2013. “Two Perspectives on Preferences and Structural Transformation”. *American Economic Review* 103 (7). (Cit. on pp. 2, 27, 28, 41, 117).

Hodrick, Robert, and Edward Prescott. 1997. “Postwar U.S. Business Cycles: An Empirical Investigation”. *Journal of Money, Credit, and Banking* 29 (1). (Cit. on p. 49).

Hubbard, R Glenn, Jonathan Skinner, and Stephen Zeldes. 1995. “Precautionary Saving and Social Insurance”. *Journal of Political Economy* 103 (2): 360–399. (Cit. on p. 47).

Huffman, David, and Matias Barenstein. 2005. “A Monthly Struggle for Self-Control? Hyperbolic Discounting, Mental Accounting, and the Fall in Consumption Between Paydays”. *Working Paper*. (Cit. on p. 100).

Hurd, Michael, et al. 2013. “Monetary Costs of Dementia in the United States”. *The New England Journal of Medicine* 368 (14). (Cit. on pp. 46, 48–50, 56, 57, 60, 61, 134).

Ingram, Beth, Narayana Kocherlakota, and N.E. Savin. 1997. “Using theory for measurement: An analysis of the cyclical behavior of home production”. *Journal of Monetary Economics* 40:435–456. (Cit. on p. 1).

Japelli, Tullio, and Luigi Pistaferri. 2010. “The Consumption Response to Income Changes”. *Annual Review of Economics* 2:479–506. (Cit. on p. 68).

Kahneman, Daniel, and Dan Lovallo. 1993. “Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking”. *Management Science* 39 (1). (Cit. on pp. 67, 76).

Kehoe, Timothy, Kim Ruhl, and Joseph Steinberg. 2018. “Global Imbalances and Structural Change in the United States”. *Journal of Political Economy* 126 (1): 761–796. (Cit. on p. 3).

Kőszegi, Botond, and Matthew Rabin. 2006. “A Model of Reference-Dependent Preferences”. *The Quarterly Journal of Economics* 121 (4): 1133–1165. (Cit. on p. 73).

Kőszegi, Botond, and Adam Szeidl. 2013. “A Model of Focusing in Economic Choice”. *The Quarterly Journal of Economics* 128 (1): 53–104. (Cit. on pp. 67, 77).

Knickman, James, and Emily Snell. 2002. “The 2030 Problem: Caring for Aging Baby Boomers”. *HSR: Health Services Research* 37 (4). (Cit. on p. 46).

Koch, Alexander, and Julia Nafziger. 2019. "Correlates of Narrow Bracketing". *The Scandinavian Journal of Economics* 121 (4): 1441–1472. (Cit. on pp. 67, 76).

– . 2016. "Goals and bracketing under mental accounting". *Journal of Economic Theory* 162:305–351. (Cit. on p. 67).

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie. 2001. "Beyond Balanced Growth". *Review of Economic Studies* 68:869–882. (Cit. on pp. 2, 27).

Krueger, Dirk, and Alexander Ludwig. 2007. "On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare". *Journal of Monetary Economics* 54:49–87. (Cit. on pp. 56, 58, 59, 102).

Kueng, Lorenz. 2015. "Explaining Consumption Excess Sensitivity with Near-Rationality: Evidence From Large Predetermined Payments". *NBER Working Paper #21772*. (Cit. on p. 68).

Leigh, Andrew. 2010. "Informal care and labor market participation". *Labour Economics* 17. (Cit. on p. 48).

Leontief, Wassily. 1947. "Introduction to a Theory of the Internal Structure of Functional Relationships". *Econometrica* 15 (4). (Cit. on pp. 6, 124).

Lepore, Michael, Abby Ferrell, and Joshua Wiener. 2017. "Living Arrangements of People with Alzheimer's Disease and Related Dementias: Implications for Services and Supports". *Research Summit on Dementia Care (white paper)*. (Cit. on pp. 46, 48).

Leslie, Alan, Rochel Gelmand, and C.R. Gallistel. 2008. "The generative basis of natural number concepts". *Trends in Cognitive Sciences* 12 (6). (Cit. on p. 76).

Levin, Laurence, Matthew Lewis, and Frank Wolak. 2017. "High Frequency Evidence on the Demand for Gasoline". *American Economic Journal: Economic Policy* 9 (3): 314–347. (Cit. on p. 93).

Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe. 2009. "Generating random correlation matrices based on vines and extended onion method". *Journal of Multivariate Analysis* 100 (9). (Cit. on p. 35).

Lieder, Falk, Thomas Griffiths, and Noah Goodman. 2013. *Advances in Neural Information Processing Systems 25*. Chap. Burn-in, bias, and the rationality of anchoring. Lake Tahoe, USA. (Cit. on p. 76).

Ludvigson, Sydney, and Alexander Michaelides. 2001. "Does Buffer-Stock Saving Explain the Smoothness and Excess Sensitivity of Consumption?" *The American Economic Review*. (Cit. on p. 105).

Malhotra, Naresh. 1982. "Information Load and Consumer Decision Making". *Journal of Consumer Research* 8 (4): 419–430. (Cit. on p. 76).

Marshall, Alfred. 1890. *Principles of Economics*. 8th ed. London: Macmillan / Co., Ltd. (Cit. on p. 72).

Matsuyama, Kiminori. 2009. "Structural Change in an Interdependent World: A Global View of Manufacturing Decline". *Journal of the European Economic Association* 7 (2-3): 478–486. (Cit. on p. 2).

McCulloch, Robert, and Ruey Tsay. 1993. "Bayesian Inference and Prediction for Mean and Variance Shifts in Autoregressive Time Series". *Journal of the American Statistical Association* 88 (423). (Cit. on pp. 85, 88).

McGrattan, Ellen, Richard Rogerson, and Randall Wright. 1993. "Household Production and Taxation in the Stochastic Growth Model". *Federal Reserve Bank of Minneapolis: Staff Report #166*. (Cit. on pp. 1, 27).

Meng, Xiao-Li, and Wing Hung Wong. 1996. "Simulating Ratios of Normalizing Constants Via a Simple Identity: A Theoretical Exploration". *Statistica Sinica* 6. (Cit. on p. 37).

Mommaerts, Corina. 2017. "Are Coresidence and Nursing Homes Substitutes? Evidence from Medicaid Spend-Down Provisions". *Working Paper*. (Cit. on p. 48).

– . 2016. "Long-Term Care Insurance and the Family". *Working Paper*. (Cit. on pp. 48, 56, 134).

Muellbauer, John. 1975. *Review of Economic Studies* 62], Number = 4, Pages = 526-542, Title = Aggregation, Income Distribution, and Consumer Demand. (Cit. on p. 2).

– . 1976. "Community Preferences and the Representative Consumer". *Econometrica* 44 (5): 979–999. (Cit. on p. 2).

Mullainathan, Sendhil. 2002. "A Memory-Based Model of Bounded Rationality". *The Quarterly Journal of Economics* 107 (3): 735–774. (Cit. on p. 67).

Muurinen, Jaana-Marja. 1986. "The Economics of Informal Care: Labor Market Effects in the National Hospice Study". *Medical Care* 24 (11). (Cit. on p. 48).

National Institute on Aging. 2011. *Global Health and Aging*. (Cit. on p. 46).

National Research Council. 2001. *Preparing for an Aging World: The Case for Cross-National Research*. (Cit. on p. 46).

Neal, Radford. 2011. *MCMC using Hamiltonian dynamics*. Ed. by Steve Brooks et al. (Cit. on pp. 35, 130, 131).

Ngai, Rachel, and Christopher Pissarides. 2007. "Structural Change in a Multisector Model of Growth". *American Economic Review* 97 (1). (Cit. on pp. 3, 33).

Obstfeld, Maurice, and Kenneth Rogoff. 1983. "Speculative Hyperinflations in Maximizing Models: Can We Rule Them Out?" *Journal of Political Economy* 91 (4). (Cit. on p. 73).

Palumbo, Michael. 1999. "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle". *The Review of Economic Studies* 66 (2). (Cit. on p. 47).

Pollak, Robert, and Michael Wachter. 1975. "The Relevance of the Household Production Function and Its Implications for the Allocation of Time". *Journal of Political Economy* 83 (2). (Cit. on p. 5).

Prelec, Drazen, and George Loewenstein. 1998. "The Red and the Black: Mental Accounting of Savings and Debt". *Marketing Science* 17 (1): 4–28. (Cit. on pp. 67, 100–102, 105, 115).

Prelec, Drazen, and Duncan Simester. 2001. "Always Leave Home Without It: A Further Investigation of the Credit-Card Effect on Willingness to Pay". *Marketing Letters* 12 (1): 5–12. (Cit. on pp. 100, 110).

Prescott, Edward. 2004. “Why Do Americans Work So Much More Than Europeans?” *Federal Reserve Bank of Minneapolis Quarterly Review* 28 (1). (Cit. on p. 47).

Proctor, Bernadette, Jessica Semega, and Melissa Kollar. 2015. *Income and Poverty in the United States: 2015*. `https://www.census.gov/library/publications/2016/demo/p60-256.html`. Accessed: 2019-05-06. (Cit. on p. 70).

Quah, Danny. 1990. “Permanent and Transitory Movements in Labor Income: An Explanation for "ExcessSmoothness" in Consumption”. *Journal of Political Economy* 98 (3). (Cit. on p. 105).

Quispe-Torreblanca, Edika, et al. 2019. “The Red, the Black, and the Plastic: Paying Down Credit Card Debt for Hotels Not Sofas”. *Management Science* 65 (11). (Cit. on pp. 100, 101, 103, 105).

Rabin, Matthew, and Georg Weizsäcker. 2009. “Narrow Bracketing and Dominated Choices”. *American Economic Review* 99 (4): 1508–1543. (Cit. on pp. 67, 76).

Read, Daniel, George Loewenstein, and Matthew Rabin. 1999. “Choice Bracketing”. *Journal of Risk and Uncertainty* 19 (1): 171–197. (Cit. on pp. 67, 76).

Roberts, Gareth, Andrew Gelman, and Walter Gilks. 1997. “Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms”. *The Annals of Applied Probability* 7 (1): 110–120. (Cit. on pp. 86, 144).

Rossi, Peter, Greg Allenby, and Robert McCulloch. 2005. *Model Choice and Decision Theory*. (Cit. on pp. 37, 148).

Rupert, Peter, Richard Rogerson, and Randall Wright. 1995. “Estimating Substitution Elasticities in Household Production Models”. *Economic Theory* 6 (1). (Cit. on p. 1).

Schwartzstein, Joshua. 2014. “Selective Attention and Learning”. *Journal of the European Economic Association* 12 (6): 1423–1452. (Cit. on pp. 67, 77).

Shefrin, Hersh, and Richard Thaler. 1981. “An Economic Theory of Self-Control”. *Journal of Political Economy* 89 (2): 392–406. (Cit. on pp. 67, 73).

– . 1988. “The Behavioral Life-Cycle Hypothesis”. *Economic Inquiry* 26:609–643. (Cit. on pp. 67, 73, 101, 110, 111).

Silverman, Bernard. 1986. *Density Estimation for Statistics and Data Analysis*. 48. London: Chapman / Hall. (Cit. on p. 95).

Simon, Hebert. 1957. *Models of Man*. John Wiley & Sons: New York. (Cit. on p. 76).

Soman, Dilip. 2001. “Effects of Payment Mechanism on Spending Behavior: The Role of Rehearsal and Immediacy of Payments”. *Journal of Consumer Research* 27:460–474. (Cit. on pp. 67, 73).

Souleles, Nicholas. 2002. “Consumer response to the Reagan tax cuts”. *Journal of Public Economics* 85:99–120. (Cit. on p. 68).

– . 1999. “The Response of Household Consumption to Income Tax Refunds”. *The American Economic Review* 89 (4): 947–958. (Cit. on pp. 67, 100).

Stan Development Team. 2016. *Stan Modeling Language: User's Guide and Reference Manual*. (Cit. on pp. 35, 36, 131).

Stone, Richard. 1954. “Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand”. *The Economic Journal* 64 (255). (Cit. on pp. 2, 74, 84).

Thaler, Richard. 1985. “Mental Accounting and Consumer Choice”. *Marketing Science* 4 (3): 199–214. (Cit. on pp. 67, 101).

– . 1999. “Mental Accounting Matters”. *Journal of Behavioral Decision Making* 12:183–206. (Cit. on pp. iv, 67, 73, 76, 101, 110).

– . 1990. “Saving, Fungibility, and Mental Accounts”. *The Journal of Economic Perspectives* 4 (1): 193–205. (Cit. on pp. 67, 101, 110).

Tobin, James. 1958. “Estimation of Relationships for Limited Dependent Variables”. *Econometrica* 26 (1). (Cit. on p. 83).

United Nations: Department of Economic and Social Affairs. 2017. “World Population Prospects: The 2017 Revision”. Visited on 07/31/2017. `https://esa.un.org/unpd/wpp/Download/Standard/Population/`. (Cit. on pp. 50, 59).

Uy, Timothy, Kei-Mu Yi, and Jing Zhang. 2013. “Structural change in an open economy”. *Journal of Monetary Economics* 60:667–682. (Cit. on p. 2).

Van Houtven, Courtney Harold, Norma Coe, and Meghan Skira. 2013. “The effect of informal care on work and wages”. *Journal of Health Economics* 32. (Cit. on p. 48).

Walsh, Carl. 2010. *Monetary Theory and Policy*. 3rd ed. MIT Press: Cambridge. (Cit. on p. 73).

Wasserman, Larry. 2000. “Bayesian Model Selection and Model Averaging”. *Journal of Mathematical Psychology* 44. (Cit. on p. 38).

Whelan, Karl. 2000. “A Guide to the Use of Chain Aggregated NIPA Data”. *Federal Reserve Board Working Paper*. (Cit. on p. 117).

– . 2002. “A Guide to U.S. Chain Aggregated NIPA Data”. *Review of Income and Wealth* 48 (2): 217–233. (Cit. on p. 117).

Yang, Fang. 2009. “Consumption over the life cycle: How different is housing?” *Review of Economic Dynamics* 12 (3). (Cit. on p. 102).

Zeldes, Stephen. 1989. “Consumption and Liquidity Constraints: An Empirical Investigation”. *Journal of Political Economy* 97 (2). (Cit. on p. 67).

Zellner, Arnold. 1962. “An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias”. *Journal of the American Statistical Association* 57 (298): 348–368. (Cit. on p. 72).