6-30-2022 11:00 AM

Revealed Preference Analysis: Theory and Applications

Charles Gauthier, *The University of Western Ontario*

Supervisor: Aguiar, Victor H., *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

Abstract

This dissertation consists of three chapters studying revealed preference theory and its applications to testing and inference. The first chapter develops a semiparametric revealed preference methodology to analyze the effects of price search on prices paid across income groups. The second chapter derives a novel representation of the exponential discounting model to make inference on the discount factor. The third chapter derives an axiomatization of the exponential discounting model and uses it to propose a nonparametric test of dynamic discrete random utility models.

The first chapter develops a novel semiparametric approach to estimate the impacts of price search on prices paid. My methodology allows for nonparametric preferences, rich heterogeneity, and measurement error in prices. I derive the implications of the model and use the resulting shape constraints to gain additional identification power. Using data on shopping expenditures from the Nielsen Homescan Dataset, I show that the data are consistent with the model and that a doubling of shopping frequency decreases prices paid by about 19%. Moreover, I find that heterogeneity in price search generates within-group consumption inequality and mixed impacts on between-group consumption inequality.

The second chapter derives a novel representation of the exponential discounting model that allows me set identify the discount factor while accounting for measurement error. My approach uses a revealed preference methodology that makes no parametric assumption on the utility function and allows for unrestricted heterogeneity. Using longitudinal data from checkout scanners, I bound household-specific discount factors and assess their sensitivity to measurement error. I find that unobserved heterogeneity is important as observable characteristics fail to capture differences in discounting.

The third chapter generalizes the previous representation and provides an axiomatic characterization of the exponential discounting model. This axiom is used to propose a nonparametric test for dynamic discrete random utility models. The methodology can be used to understand the welfare implications of price changes in demand data when there is no uncertainty for consumers that have preferences over discrete product characteristics or quantities of products.

Keywords: Revealed preference, exponential discounting, inference, testing

Summary for Lay Audience

Revealed preference analysis is an approach that makes minimal assumptions on the consumer or firm behavior. It can be used to test various theories and to study the behavioral response to various policies. This thesis consists of three chapters studying revealed preference theory and its applications to testing and inference.

The first chapter develops a semiparametric revealed preference methodology to analyze the effects of price search on prices paid across income groups, where price search describes the process whereby buyers actively seek to gauge the most favorable prices. Using data on shopping expenditures, I show that the data are consistent with the model and that a doubling of shopping frequency decreases prices paid by about 19%. Moreover, I find that heterogeneity in price search generates within-group consumption inequality and leads to mixed impacts on between-group consumption inequality.

The second chapter derives a novel revealed preference representation of the exponential discounting model to make inference on the discount factor. This endeavor is motivated by the fact that assumptions on the consumer preferences may lead to erroneous conclusions about his time preferences and misrepresent a decision maker's intertemporal choices in a vast range of applications. Using longitudinal data from checkout scanners, I bound household-specific discount factors and assess their sensitivity to measurement error.

The third chapter derives an axiomatization of the exponential discounting model and uses it to propose a nonparametric test of dynamic discrete random utility models. This methodology can be used to understand the welfare implications of a change in prices and to predict the change in the distribution of demand from a change in prices.

Acknowlegements

I wish to express my deepest gratitude to my supervisor Victor Aguiar for his constant support, valuable advice, and useful guidance. I consider myself lucky to have him as a mentor. I also wish to thank my supervisory committee, Roy Allen, Nail Kashaev, and David Rivers, for the many comments and encouragements throughout my PhD. The many interactions I had with them greatly influenced and benefited my work.

I am indebted to every professor that has given me feedback and encouragements within and outside of the University of Western Ontario, including but not limited to, Audra Bowlus, Elizabeth Caucutt, Laurens Cherchye, Rahul Deb, Bram De Rock, Maria Goltsman, Florian Gunsilius, Per Hjertstrand, Igor Livshits, Lance Lochner, Nirav Mehta, Sergio Ocampo Diaz, Gregory Pavlov, John Rehbeck, Baxter Robinson, Bruno Salcedo, Peter Streufert, Susanne Schennach, and Roberto Serrano. I also wish to thank Jessica Lee for her valuable editing work and generosity over the years.

I appreciate the many conversations I have had with fellow graduate students and faculty members. In particular, I extend a special thanks to Bob Millard, Marco Pariguana, Al Slivinski and Tom Trivieri for enjoyable discussions during happy hours.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

In a seminal work, Samuelson (1938) proposed a different approach to consumer theory that does not begin with the specification of the consumer preferences as captured by the utility function. Specifically, he proposed to take the data as primitives and impose restrictions that the data should satisfy if they were generated by a rational consumer. This approach builds on the notion of revealed preference and has the desirable feature to be nonparametric. That is, it does not rely on a parameterization of the consumer preferences.

The goal of this thesis is to show how the nonparametric nature of the revealed preference analysis pioneered by Paul Samuelson may be used for testing and inference in applications of interest. Towards that end, the first chapter develops a novel semiparametric approach to estimate the impacts of price search on prices paid, the second chapter derives a novel representation of the exponential discounting model that can be used to set identify the discount factor without having to specify the consumer preferences, and the third chapter provides an axiom for the exponential discounting model that can be used to obtain a nonparametric test of dynamic discrete random utility models.

Chapter 2 proposes a general model where consumers can pay lower prices by shopping more frequently. I develop a novel semiparametric approach to estimate the impacts of price search on prices paid. My methodology allows for nonparametric preferences, rich heterogeneity, and measurement error in prices. I use a revealed preference approach to derive the nonparametric implications of the model and use the resulting shape constraints to gain additional identification power. Using data on shopping expenditures from the Nielsen Homescan

Dataset, I show that the data are consistent with the model and that a doubling of shopping frequency decreases prices paid by about 19%. My method also provides new insights about the importance of price search for understanding consumption inequality. Specifically, I find that heterogeneity in price search generates within-group consumption inequality and mixed impacts on between-group consumption inequality.

Chapter 3 derives a novel representation of the exponential discounting model that allows me set identify the discount factor while accounting for measurement error. My approach uses a revealed preference methodology that makes no parametric assumption on the utility function and allows for unrestricted heterogeneity. Using longitudinal data from checkout scanners, I bound household-specific discount factors and assess their sensitivity to measurement error. I find that unobserved heterogeneity is important as observable characteristics fail to capture differences in discounting.

Chapter 4 provides a nonparametric test for dynamic discrete random utility models by generalizing the previous representation of exponential discounting. The methodology can be used to understand the effects of a change in prices on welfare or on the distribution of demand in common settings arising in industrial organization.

Chapter 2

Price Search and Consumption Inequality: Robust, Credible, and Valid Inference

2.1 Introduction

Price search describes the process whereby buyers actively seek to gauge the most favorable prices. Its importance has been recognized at least since the seminal paper of Stigler (1961) and has gained strong empirical support over the years.[1] In their influential paper, Aguiar and Hurst (2007) show that the drop in expenditures occurring around retirement is partly due to older households searching more intensively than younger ones. In another paper, Griffith et al. (2009) document savings from various shopping strategies and compare how those savings vary by income group. The ability of consumers to pay lower prices by searching more intensively and more efficiently affects their purchasing power. Thus, price search may be important for understanding consumption inequality and could be used as a mitigation mechanism

[0] Researcher's own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ and Nielsen data are those of the researcher and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[1] See, for example, Sorensen (2000), Brown and Goolsbee (2002), and McKenzie et al. (2011). For a general survey, see Baye et al. (2006).

by policymakers.[2]

The first contribution of this paper is to propose an approach that allows one to statistically test structural models and achieve informative inference despite minimal assumptions on unobservables. The second contribution of this paper is to provide a new revealed preference methodology to characterize a parsimonious model of price search.[3] The model shares key features with the price search literature such as the concavity of the utility function and the log-linearity of the shopping technology. However, to deal with misspecification issues I allow for nonparametric preferences and unrestricted unobserved heterogeneity. The third contribution of this paper is to relax the common homogeneity assumptions on the shopping technology which allows me to explore how it differs across several dimensions such as income, gender, and geographic location.

This paper formalizes the empirical evidence documenting (i) the effects of price search on prices paid (Aguiar and Hurst, 2007), (ii) the use of price search as a mechanism to mitigate adverse income shocks (McKenzie et al., 2011; Nevo and Wong, 2019), and (iii) the wide heterogeneity in prices paid (Kaplan and Menzio, 2015; Kaplan et al., 2019; Hitsch et al., 2019). Additionally, by testing the main assumptions on which the price search literature relies, I provide a foundation for existing models of price search (Aguiar and Hurst, 2007; Pytka, 2017; Arslan et al., 2021) and obtain robust effects of search on prices paid.

While instrumental variable is a standard tool to estimate the causal effect of a treatment on an outcome of interest, it requires the availability of an instrument that only affects the outcome of interest through its impact on the treatment. Unfortunately, it is often difficult to know a priori whether an instrument satisfy this exogeneity condition. Moreover, exogeneity is not always testable when the endogenous variable is continuous (Gunsilius, 2020).[4] Finally, instrumental variable is not immune to potential bias from measurement error. My approach mitigates these issues by exploiting the structure of the model to obtain testable restrictions and by nonparametrically accounting for measurement error.

[2]For example, one may be interested in decreasing search costs of low-income consumers by increasing accessibility to cars. Alternatively, one could improve the shopping technology of consumers in low-income neighborhoods by increasing the number of discount stores.

[3]From a different angle, Tipoe (2021) provides a revealed preference characterization of a model with limited attention to prices à la Gabaix (2014).

[4]It is useful to note that overidentifying restrictions do not allow one to test instrument validity. See, for example, Parente and Silva (2012).

Absent exogenous variation in shopping intensity, I show that the model is set identified in the sense that many elasticities of price with respect to shopping intensity are consistent with the model given the data. To address this problem, I provide a nonparametric characterization of the model and use it to gain additional identification power.[5] The implementation takes advantage of a revealed preference approach (Afriat, 1967; Diewert, 1973; Varian, 1982; Browning, 1989) and its extension to nonlinear budgets (Forges and Minelli, 2009). Accordingly, every consumer is treated separately such that no aggregation assumptions or restrictions on unobserved heterogeneity are imposed.

My empirical application uses the Nielsen Homescan Dataset which is a data set that tracks U.S. households' food purchases on each of their trips to a wide variety of retail outlets. I measure shopping intensity by the number of shopping trips as it captures price variations across stores and price discounts found by frequently visiting stores. The panel structure of the data enables me to set identify the elasticity of price with respect to shopping intensity from individual time-variation in shopping intensity. Under a mild condition on the residual errors, I further show that the true impacts of search on prices paid are recovered. The validity of this assumption can be jointly tested along with the model.

In a validation study of the Nielsen Homescan Dataset, Einav et al. (2010) report severe measurement error in prices and provide information about its structure. The presence of measurement error requires special attention for two reasons. First, the model could be compatible with the true data but incompatible with the observed data, hence leading to the erroneous rejection of the model.[6] Second, measurement error can complicate empirical analyses by obscuring the true behavior of variables such as expenditure.[7] In turn, this can bias estimators in unpredictable ways. For example, measurement error may be nonclassical such that bias could arise even if it appears on the dependent variable in a standard regression setting.

The methodology developed by Schennach (2014) enables me to nonparametrically account for measurement error and statistically test the model. Furthermore, the extension of

[5] This strategy is similar to Blundell et al. (2012) in that it uses shape restrictions from economic theory to improve the precision of an estimate.

[6] Measurement error has been shown to reverse conclusions about the validity of exponential discounting in single-individual households (Aguiar and Kashaev, 2021).

[7] See Attanasio and Pistaferri (2016) for an overview of how measurement error can cloud the evolution of consumption inequality.

Aguiar and Kashaev (2021) allows me to impose the concavity of the utility function without increasing the dimensionality of the problem. I extend the applicability of their framework to allow for nonlinear budget sets and a mixture of parametric and nonparametric components in the model. This is achieved by using a rejection sampling algorithm in the implementation of the methodology.[8] This method has the advantage to be applicable in a broader set of models while remaining computationally tractable.

I find support for price search behavior in the Nielsen Homescan Dataset as the statistical test fails to reject the hypothesis that the model generated the data. Furthermore, I show that heterogeneity in price search explains part of the difference in average price paid between income groups. Nevertheless, my results suggest that the difference in average price paid is mainly driven by other factors such as a difference in average good quality. This finding is consistent with Broda et al. (2009) who document that low-income consumers pay lower prices mainly through the purchase of less expensive goods within product categories.

I document that, on average, low-income consumers shop slightly more frequently than high-income consumers. Additionally, I show that low-income consumers have a slightly better average shopping technology compared to high-income consumers. That is, I find that low-income consumers are better at transforming search intensity into lower prices paid. Since average food consumption levels are comparable between income groups in my data, my results suggest that differences in price search and good quality are successful in reducing average food intake inequality.

Pulling all consumers together, the 95% confidence set on the average elasticity of average price with respect to shopping intensity states that doubling shopping intensity decreases the average price paid by about 19.1% to 19.5%. This effect is well-above the preferred estimate of 7.4% obtained by Aguiar and Hurst (2007) using income as an instrument. Instead, I find support for their estimate of 18.9% obtained using age as an instrument. Thus, my results show that the effects of shopping intensity on prices paid are almost three times larger than previously thought. Moreover, they rationalize the calibration of Arslan et al. (2021) used in a different model of price search.

In a counterfactual exercise, I show that a consumer in the lowest quintile of the shopping

[8] The rejection sampling algorithm exploits GPU parallel computing when applicable.

frequency distribution pays close to 15% more for a good than a consumer in the largest quintile. Assuming that savings from search are spent on the same good, this difference yields within-group consumption inequality of the same order. Furthermore, I show that price search has mixed impacts on between-group consumption inequality: I find that price search mitigates consumption inequality between low-income consumers with high shopping intensity and high-income consumers with low shopping intensity, while exacerbating it for low-income consumers with low shopping intensity and high-income consumers with high shopping intensity.

The rest of the paper is organized as follows. Section 2.2 introduces the model and derives its testable restrictions. Section 2.3 describes the data set. Section 2.4 discusses the environment. Section 2.5 presents the statistical framework. Section 2.6 presents the empirical application. Section 2.7 concludes. The main proofs can be found in Appendix A.

2.2 Description of the Model

This section introduces the notation, presents the model, characterizes the testable implications, and provides the procedure for recovering the identified set.

2.2.1 Notation

The scenario under consideration is that of households making purchases over a certain time window. Let $\mathcal{N} := \{1, \ldots, N\}$, $\mathcal{L} := \{1, \ldots, L\}$, and $\mathcal{T} := \{1, \ldots, T\}$ denote the sets of households, commodities, and periods for which data are observable. For any household $i \in \mathcal{N}$, good $l \in \mathcal{L}$ and time period $t \in \mathcal{T}$, denote prices by $p_{i,l,t}$, consumption by $c_{i,l,t}$, shopping intensity by $a_{i,l,t}$, and search productivity by $\omega_{i,l,t}$. I assume that P, C, $A \subseteq \mathbb{R}^L_{++}$ and $\Omega \subseteq \mathbb{R}^L$, where P, C, A and Ω correspond to the price, consumption, shopping intensity, and search productivity spaces, respectively. An observation on a given household $i \in \mathcal{N}$ at time $t \in \mathcal{T}$ is a triple $(\boldsymbol{p}_{i,t}, \boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \in P \times C \times A$.[9] Accordingly, a data set is written as $\{(\boldsymbol{p}_{i,t}, \boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})\}_{t \in \mathcal{T}}$.

[9] I use bold font to denote vectors and follow the convention that vectors are vector columns.

2.2.2 The Consumer Problem

In every time period, the consumer is assumed to know her realization of search productivity and to choose consumption and shopping intensity accordingly. Formally, a consumer $i \in \mathcal{N}$ behaves as if maximizing her lifetime utility subject to satisfying her budget constraints:

$$\max_{(\boldsymbol{c}_i, \boldsymbol{a}_i) \in C^T \times A^T} \sum_{t=1}^{T} \delta_i^{t-1} u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \tag{2.1}$$

subject to

$$\boldsymbol{p}_i(\boldsymbol{a}_{i,t}, \omega_{i,t})' \boldsymbol{c}_{i,t} + s_{i,t+1} = y_{i,t} + s_{i,t},$$

where $u_i : C \times A \to \mathbb{R}$ is the instantaneous utility function which is continuous, concave, strictly increasing in consumption, and decreasing in shopping intensity; $\delta_i \in [\underline{\delta}, 1]$ is the discount factor, where $\underline{\delta} \in (0, 1]$; $s_{i,t+1}$ is savings in a risk-free asset; $\boldsymbol{p}_i(\boldsymbol{a}, \omega_{i,t})$ is the vector of continuously differentiable good-specific price functions $p_{i,l} : A \times \Omega \to P$; $y_{i,t} > 0$ is income; and $\boldsymbol{p}_{i,t} := \boldsymbol{p}_i(\boldsymbol{a}_{i,t}, \omega_{i,t})$. The econometrician is assumed to only observe the data set $x_i := \{(\boldsymbol{p}_{i,t}, \boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})\}_{t \in \mathcal{T}}$. As explained in the definition below, a data set that can be thought of as stemming from the model (2.1) is said to be rationalized by the model.

Definition 2.1. *A data set x_i is rationalized by the model if there exist a utility function $u_i(\cdot, \cdot)$, a vector of price functions $\boldsymbol{p}_i(\cdot, \cdot)$, an income stream $(y_{i,t})_{t \in \mathcal{T}}$, a savings stream $(s_{i,t+1})_{t \in \mathcal{T}}$, and a discount factor $\delta_i \in [\underline{\delta}, 1]$ such that $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{t \in \mathcal{T}}$ solves the model* (2.1).

The model has two distinctive features. First, the consumer gets utility from consumption and disutility from shopping intensity. The latter captures the opportunity cost of time such as foregone earnings and leisure. Second, the consumer can pay lower prices by shopping more frequently. The extent by which shopping intensity reduces prices paid depends on the consumer's ability to take advantage of sales and other deals such as coupons. The consumer problem boils down to finding the optimal trade-off between utility from consumption and disutility from shopping intensity.

This trade-off is illustrated in Figure 2.1 in the case where there is one good $\mathcal{L} = \{1\}$ and one time period $\mathcal{T} = \{1\}$. The consumer has to choose a bundle that lies within her budget set $\mathcal{B}_i := \{(\boldsymbol{c}, \boldsymbol{a}) \in C \times A : \boldsymbol{p}_i(\boldsymbol{a}, \omega_i)' \boldsymbol{c} \leq y_i + s_i\}$. This set is represented by the shaded area in

Figure 2.1. The affordable bundle that maximizes the consumer utility is (c_i, a_i). At this point, the indifference curve IC_i is tangent to the budget line, hence corresponding to the unique maximizer.[10]

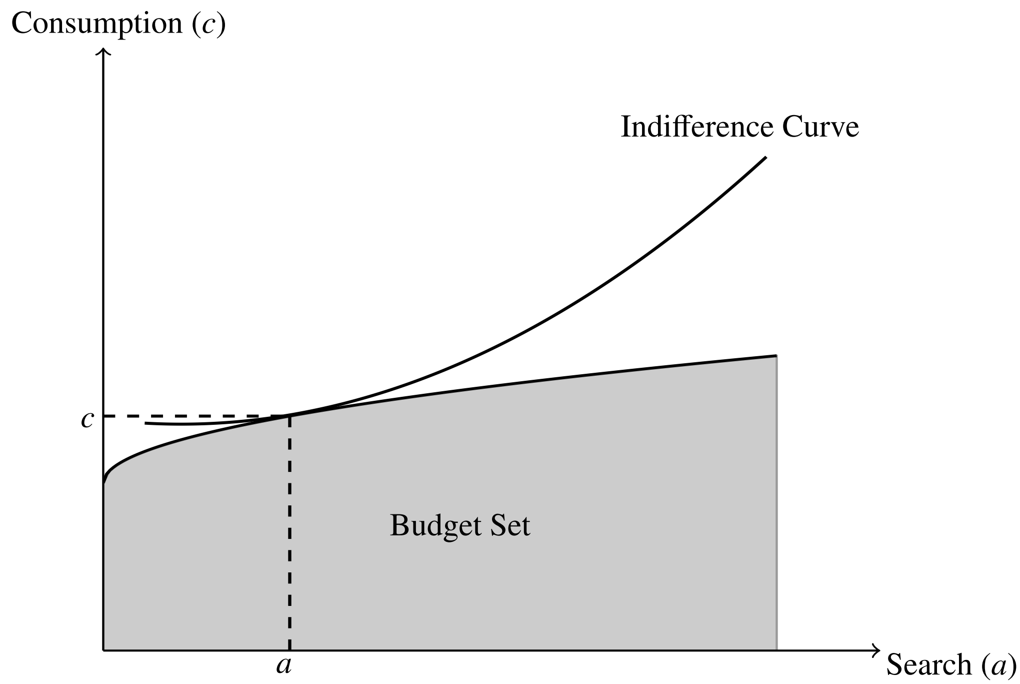

Figure 2.1: Optimal Choice with Price Search

In the remaining of the paper, I follow the price search literature and assume that the price functions are log-linear in shopping intensity.[11]

Assumption 2.1. *For all $l \in \mathcal{L}$, the log price function is given by*

$$\log(p_{i,l}(a_{i,l,t}, \omega_{i,l,t})) = \alpha^0_{i,l} + \alpha^1_{i,l} \log(a_{i,l,t}) - \omega_{i,l,t},$$

where $\alpha^0_{i,l} \in \mathbb{R}$ denotes the intercept and $\alpha^1_{i,l} \leq 0$ denotes the elasticity of price with respect to shopping intensity.

Assumption 2.1 implies that prices paid decrease at a decreasing rate as shopping intensity or search productivity increases. This requirement captures decreasing marginal returns that

[10] In Appendix A.1, I show that my model can be extended to home production and relates it to that of Aguiar and Hurst (2007).

[11] See, for example, Aguiar and Hurst (2007) and Arslan et al. (2021).

arise due to the increasing difficulty of finding discounts surpassing the current best discount.[12] An example of budget set generated by a convex decreasing price function is illustrated in Figure 2.1 in the case of a single good. The figure shows that, while increasing shopping intensity allows a greater consumption level, the marginal benefit is decreasing in shopping intensity.

2.2.3 Identified Set

Given a data set x_i, I am interested in what can be learned about the elasticity $\alpha^1_{i,l}$. For any good $l \in \mathcal{L}$, the first-order conditions with respect to consumption and shopping intensity are

$$\delta^t \nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t} p_{i,l,t} \tag{2.2}$$

$$\delta^t \nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t} \alpha^1_{i,l} a_{i,l,t}^{\alpha^1_{i,l}-1} e^{-(\omega_{i,l,t}-\alpha^0_{i,l})} c_{i,l,t}.^{13} \tag{2.3}$$

Solving for the exponential function in equation (2.3) yields

$$e^{-(\omega_{i,l,t}-\alpha^0_{i,l})} = \frac{\lambda^{-1}_{i,t} \delta^t \nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l}{\alpha^1_{i,l} a_{i,l,t}^{\alpha^1_{i,l}-1} c_{i,l,t}}.$$

This equation can be substituted into the price function equation for good l to get rid of the intercept and search productivity. Further substituting $\lambda^{-1}_{i,t}\delta^t$ by its expression from equation (2.2) then gives

$$\alpha^1_{i,l}(u_i) = \text{MRS}_{i,l} \cdot \frac{a_{i,l,t}}{c_{i,l,t}} \quad \forall l \in \mathcal{L}, \tag{2.4}$$

where $\text{MRS}_{i,l} := \frac{\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l}{\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l} \leq 0$ denotes the marginal rate of substitution and highlights that the consumer would have to receive consumption to increase her shopping intensity.

Equation (2.4) shows that every preference that generates a distinct marginal rate of substitution induces a distinct shopping technology $\alpha^1_{i,l}$. The set of all elasticities that can be sustained by the model is captured by the identified set.

[12] Stigler (1961) shows that the expected value of the minimum price is convex in search, therefore providing a theoretical motivation for this choice.

[13] Strictly speaking, $\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ and $\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ denote a supergradient of $u_i(\cdot,\cdot)$ at $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$. When the utility function is differentiable, the supergradient corresponds to the gradient.

Definition 2.2. *Under Assumption 2.1, the identified set is defined by*

$$\mathcal{I}(x_i) := \{\boldsymbol{\alpha}_i^1(u_i) : x_i \text{ is rationalized by the model}, u_i \in \mathcal{U}\},$$

where $\mathcal{U}$ is the set of utility functions that are continuous, concave, strictly increasing in consumption, and decreasing in shopping intensity.

In other words, the identified set contains every possible elasticity of price with respect to shopping intensity such that data on consumption and shopping intensity can be thought of as maximizers of the consumer problem (2.1) for some well-behaved preferences.

2.2.4 Nonparametric Analysis of the Model

This subsection characterizes the testable implications of the model and provides a feasible procedure to recover the identified set. Let $\odot$ denote the Hadamard product such that $(\boldsymbol{v} \odot \boldsymbol{w})_j = v_j w_j$. The following result characterizes the empirical content of the model.

Theorem 2.1. *Let x_i be a given data set. The statements (i) and (ii) below are related in the following ways: if (i) then (ii), and if the budget sets $\{\mathcal{B}_{i,t}\}_{t\in\mathcal{T}}$ are convex then (ii) implies (i).*

(i) The data set is rationalized by the model, where the utility function $u(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is continuous, concave, strictly increasing in $\boldsymbol{c}_{i,t}$, and decreasing in $\boldsymbol{a}_{i,t}$, and where the vector of price functions $\boldsymbol{p}_i(\cdot,\cdot)$ satisfies Assumption 2.1.

(ii) There exist numbers $u_{i,t}$, $\lambda_{i,t} > 0$, $\alpha_{i,l}^0$, $\alpha_{i,l}^1 \leq 0$, $\omega_{i,l,t}$, and a discount factor $\delta_i \in [\underline{\delta}, 1]$, such that for all $s, t \in \mathcal{T}$, the following system of inequalities is satisfied

$$\begin{aligned} u_{i,s} &\leq u_{i,t} + \lambda_{i,t}\delta_i^{-t}\left[\boldsymbol{p}_{i,t}'(\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) + \boldsymbol{\rho}_{i,t}'(\boldsymbol{a}_{i,s} - \boldsymbol{a}_{i,t})\right] \\ 0 &< \boldsymbol{p}_{i,t} \\ 0 &\geq \boldsymbol{\rho}_{i,t}, \end{aligned}$$

where $\rho_{i,l,t} := \alpha_{i,l}^1 a_{i,l,t}^{\alpha_{i,l}^1 - 1} e^{-(\omega_{i,l,t} - \alpha_{i,l}^0)} c_{i,l,t}$.

The first set of inequalities in Theorem 2.1 (ii) represents the concavity of the utility function, where the numbers $u_{i,t}$ and $\lambda_{i,t} > 0$ can be thought of as utility numbers and marginal utilities of expenditure. The second and third sets of inequalities capture the assumptions that the utility function is strictly increasing in consumption and decreasing in shopping intensity, respectively. Theorem 2.1 states that any data set rationalized by the model must satisfy the inequalities in (*ii*). Moreover, any data set that satisfies these inequalities is rationalized by the model if the budget sets are convex.

Since the model is generally set identified, there may be many solutions to the inequalities in Theorem 2.1 (*ii*). These solutions are observationally equivalent in the sense that the data do not allow one to distinguish one from another. In particular, note that every solution to Theorem 2.1 (*ii*) sustains an elasticity $\alpha^1_{i,l}$. As a consequence, the set of solutions in Theorem 2.1 (*ii*) directly relates to the identified set.

Corollary 2.1. *If the inequalities in Theorem 2.1* (*ii*) *are only necessary for the data to be rationalized by the model, then*

$$\mathcal{I}(x_i) \subset \left\{\boldsymbol{\alpha}_i^1 : x_i \text{ satisfies Theorem 2.1 } (ii)\right\}.$$

If the budget sets $\{\mathcal{B}_{i,t}\}_{t\in\mathcal{T}}$ *are convex, then the inequalities in Theorem 2.1* (*ii*) *are also sufficient for the data to be rationalized by the model and*

$$\mathcal{I}(x_i) = \left\{\boldsymbol{\alpha}_i^1 : x_i \text{ satisfies Theorem 2.1 } (ii)\right\}.$$

Corollary 2.1 states that conservative bounds on the identified set can always be recovered via Theorem 2.1. Moreover, these bounds are sharp whenever the budget sets are convex. Finally, note that given some value of $(\lambda_{i,t})_{t\in\mathcal{T}}$ and δ_i, the existence of a solution can be checked efficiently using linear programming.

Remark Conditional on numbers $(\lambda_{i,t})_{t\in\mathcal{T}} \in \mathbb{R}^L_{++}$ and a discount factor $\delta_i \in [\underline{\delta}, 1]$, the set of solutions $\boldsymbol{\alpha}_i^1$ in Theorem 2.1 is convex. As such, the identified set can be recovered by finding the smallest and largest value of $\boldsymbol{\alpha}_i^1$ for which a solution exists. If the budget sets are not convex, then the procedure returns conservative bounds on the identified set.

2.3 Data

This section presents the data set used in my empirical application and discusses its main source of measurement error.

2.3.1 Sample Construction

For my empirical application, I use the Nielsen Homescan Dataset 2011 (henceforth referred to as the Homescan). This data set contains information on purchases made by a panel of U.S. households in a large variety of retail outlets. The data set is designed to be representative of the U.S. population based on a wide range of annually updated demographic characteristics including age, sex, race, education, and income.

Participating households are provided with a scanner device and instructed to record all of their purchases after each shopping trip. The scanner device first requires participants to specify the date and store associated with each trip. Then, they are prompted to enter the number of units bought. When an item is purchased at a store with point-of-sale data, the average weighted price of the item in that week and store is directly given to Nielsen and recorded as the price paid prior to any coupon. Otherwise, panelists enter the price paid prior to any deal or coupon using the scanner device. In either case, panelists record the amount saved from coupons and the final price paid is the recorded price paid minus coupon discounts.

The Homescan contains information on Universal Product Codes (UPC) belonging to one of 10 departments. In order to mitigate issues associated with stockpiling, I restrict my attention to the following four food departments: dry grocery, frozen foods, dairy, and packaged meat.[14] This selection leaves over a million distinct UPCs representing about 40% of all products in the Homescan. For each household, I also aggregate the data to monthly observations to further reduce stockpiling issues. The resulting UPC prices are calculated as the average UPC prices weighted by quantities purchased.

To obtain regular observations on each good, I aggregate UPCs to their department categories, yielding a total of four "goods". The resulting aggregated prices are calculated as the

[14] This choice implicitly assumes that food is weakly separable from other categories of goods. This assumption is empirically plausible (Cherchye et al., 2015), especially when the presence of measurement error is recognized (Fleissig and Whitney, 2008; Elger and Jones, 2008).

average UPC prices weighted by quantities purchased. Even with this layer of aggregation, some households do not have purchases from each category of goods in every month. Since the model requires price observations in every time period, I discard those households from the analysis.

In addition to the above restriction, I only consider single households. This is motivated by recent evidence that exponential discounting may be rejected because of preference aggregation within a household.[15] Thus, including couple households could lead to the artificial rejection of the model.[16] Finally, I consider consumers that are at least 50 years old to exclude potential online shoppers. The inclusion of younger consumers has no significant impact on the results.

The data set focuses on households that satisfy the above criteria and who participated in the Homescan from April to September of the panel year 2011. The final sample contains 1645 consumers, 4 aggregated goods, and 6 monthly time periods. Additional details about the construction of the data set are provided in Appendix A.2.

2.3.2 Sample Description

To assess whether my constructed sample is in line with the Homescan 2011, Table 2.1 compares demographics between the two samples.

Table 2.1: Summary Demographics

	N	Age	White (%)	Black (%)	Asian (%)	Other (%)	Male (%)	Female (%)	High school or less (%)
Sample	1645	65.45	86.14	11.43	0.55	1.88	34.04	65.96	27.90
Homescan 2011	62 092	57.45	83.77	9.38	3.17	3.99	74.99	90.41	46.21

Notes: N is the sample size. Age is calculated using the year of birth of the head household member. Since households may have many members in the Homescan 2011, I let the head household refer to the male in that case. The percentage of males and females includes both the head male and head female members in the household.

Table 2.1 shows that the average age, proportion of females, and education level are slightly

[15] As Adams et al. (2014) point out, inconsistencies may arise due to negotiation within a couple household. Jackson and Yariv (2015) further show that time inconsistent behavior will appear if individuals in a non-dictatorial household have different discount factors. By accounting for measurement error in survey data, Aguiar and Kashaev (2021) show that single households behave consistently with exponential discounting while couple households do not.

[16] This prediction is confirmed in my empirical application where I find that the model is rejected for couple households.

higher in my sample than in the Homescan. The difference in age is natural given that my sample is restricted to consumers that are at least 50 years old. Despite the higher proportion of females and higher education level in my sample, the demographics are still fairly representative overall.

The fundamental assumption of the model is that consumers can decrease their prices paid by searching more intensively. To provide evidence that price search is an empirical feature of the data, Figure 2.2 displays how log prices vary with log number of shopping trips.

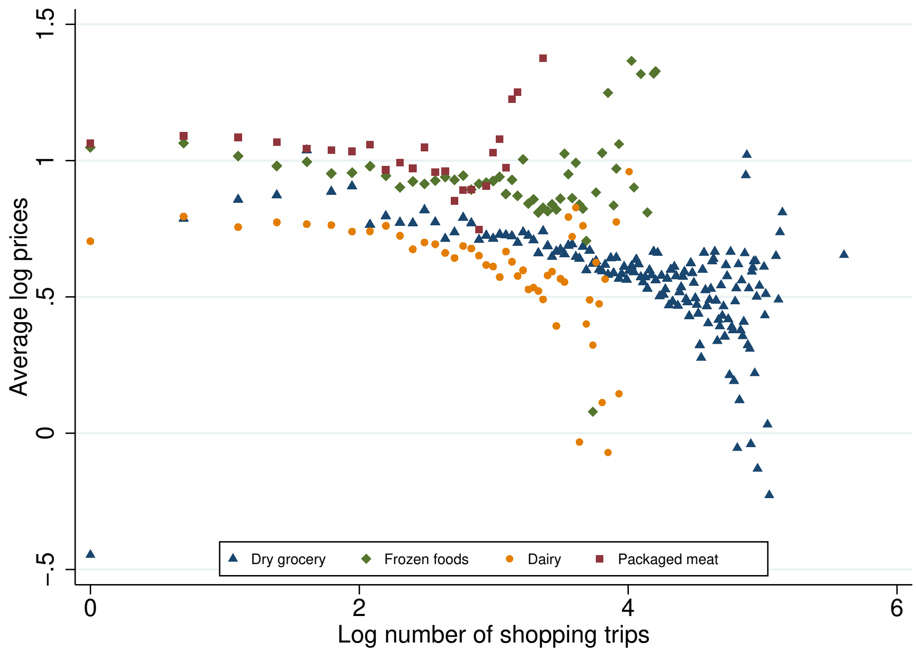

Figure 2.2: Average Log Price by Log Number of Shopping Trips

Note: The vertical axis reports the average log price, where the average is taken across consumers.

Consistent with the main hypothesis of the model, Figure 2.2 shows a negative relationship between prices paid and shopping intensity. That said, the true effect of shopping intensity on prices paid may be quite different from the one displayed in Figure 2.2. Consumers that go on many trips may do so because they do not find satisfactory discounts. This could explain the uptick in prices paid for larger values of shopping trips on frozen foods and packaged meat. Alternatively, those upticks could reflect the purchase of higher quality goods on those shopping trips.

Next, Figure 2.3 compares differences in log prices paid between high- and low-income consumers for each category of goods.

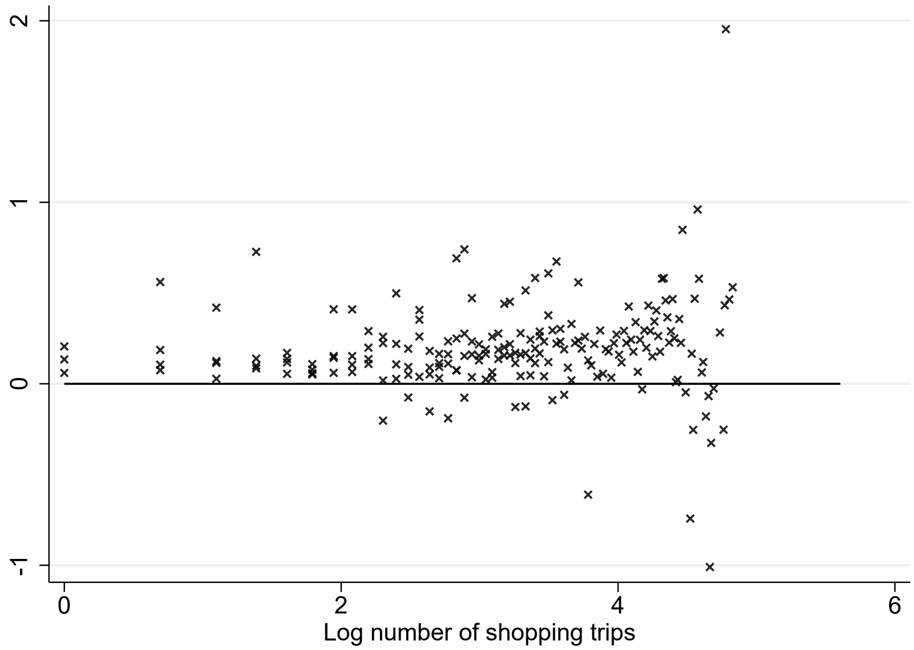

Figure 2.3: Difference in Average Log Price between Income Groups

Notes: The vertical axis reports the difference in average log price between high- and low-income consumers. The average is taken across consumers belonging to the same income group. High-income consumers are defined as those with an income greater than $50,000. Low-income consumers are defined as those with an income lower than $25,000.

The overwhelming trend shown in Figure 2.3 is that, conditional on shopping trips, high-income consumers pay higher prices relative to low-income consumers. Since this feature holds for almost every number of shopping trips, differences in prices paid cannot be attributed to differences in shopping intensity. A plausible reason for the difference could be that high-income consumers purchase goods of higher quality. Yet, a complementary explanation is that low-income consumers have a greater incentive to take advantage of sales and discounts to increase their consumption level. That is, low-income consumers could have a better shopping technology such that shopping intensity would be more effective in decreasing prices paid for them than for high-income consumers.

A possible approach to recover the true effects of shopping intensity on prices paid is to use instrumental variables. However, an issue with this approach is that measurement error in prices could yield inconsistent estimates if measurement error is nonclassical. More generally, instruments could also fail to be exogeneous. In my empirical application, I show that such concerns about instrumental variable are justified in my data.

2.3.3 Measurement Error

The data collection process employed by Nielsen may induce measurement error for three reasons. First, conditional on a shopping trip, entry mistakes may arise as panelists self-report their purchases. Second, when a consumer purchases a UPC at a store that provides Nielsen with point-of-sale data, the price reported (before coupons) is the weighted average price during that week in that particular store. Thus, the reported price will be different from the price paid if the store changed the price during the week. Third, some consumers have loyalty cards whose discounts are not incorporated into the final price paid.

In a validation study of the Homescan 2004, Einav et al. (2010) use transactions from a large retailer in order to document the extent of measurement error. Consistent with the above observations, they find that price is the variable most severely hit by measurement error. Specifically, they find that around 50% of prices are accurately recorded. In contrast, around 90% of UPCs are accurately recorded by panelists on average. This number increases to 99% conditional on the quantity being equal to one. Accordingly, I focus exclusively on measurement error in prices in my application.

Since prices are mismeasured, observed prices $(\boldsymbol{p}_{i,t})_{t\in\mathcal{T}}$ are different from true prices paid by the consumer $(\boldsymbol{p}^*_{i,t})_{t\in\mathcal{T}}$. Let the difference between their logarithms define measurement error: $\boldsymbol{m}_{i,t} := \log(\boldsymbol{p}_{i,t}) - \log(\boldsymbol{p}^*_{i,t})$ for all $t \in \mathcal{T}$.[17] Using price data from a large retailer, Einav et al. (2010) show that the difference between observed and true log prices is centered around zero in the Homescan 2004. Formally, one cannot reject that the difference in sample means of log prices is zero at the 95% confidence level.

As Nielsen's method of data collection has not changed since their study, I take their finding as support for mean zero measurement error in log prices in the Homescan 2011.

Assumption 2.2. *For all $l \in \mathcal{L}$ and $t \in \mathcal{T}$, the following moment condition holds:*

$$\mathbb{E}\left[\log(p_{l,t})\right] = \mathbb{E}\left[\log(p^*_{l,t})\right].$$

Assumption 2.2 says that expected observed log prices and true log prices are the same for

[17] This definition makes no assumption on the way measurement error arises. For example, measurement error could be additive or multiplicative and be correlated across goods or time periods.

each good and time period. Together, they yield a total of $L \cdot T$ moments on measurement error.

2.4 Environment

In this section, I discuss existing assumptions and impose additional structure on the model in preparation of the empirical application.

Conditional on the log-linear specification implied by Assumption 2.1, price functions are otherwise free to vary across goods and consumers. This heterogeneity is important as goods may not be subject to the same discounts and consumers may not have access to the same set of stores.[18] Furthermore, note that the price function for any good $l \in \mathcal{L}$ only depends on the shopping intensity on that good. This precludes complementarities that may naturally arise, for instance, if two goods are in a same aisle in a store. Since goods are aggregated to coarse categories in the data set, this issue should be largely mitigated.

Consistent with the ability of the consumer to transfer income across time, I restrict the marginal utility of expenditure to be constant over time.

Assumption 2.3. *For all $t \in \mathcal{T}$, the marginal utility of expenditure is constant and such that $\lambda_{i,t} = 1$.*

Assumption 2.3 requires that the marginal utility of expenditure be invariant to changes in income. This quasilinearity assumption is justified in my empirical application as the data span only a period of 6 months for which unexpected changes in income should be negligible.[19] Moreover, food tends to be income inelastic such that changes in income should have mild impacts on preferences.[20]

The next assumption imposes the average search productivity to be time-invariant and ensures that price search is refutable. To see why, note that Assumption 2.1 implies that the true average price paid must decrease whenever shopping intensity increases if there is no change

[18]Hence, even if store chains apply a nearly-uniform pricing rule (DellaVigna and Gentzkow, 2019), the shopping technology may still differ across consumers.

[19]Quasilinearity is also used by Echenique et al. (2011) and Allen and Rehbeck (2020a) in a similar scanner data set on food expenditures.

[20]See, for example, Anderson and Blundell (1984), Erdil (2006), and Selvanathan and Selvanathan (2006).

in the average search productivity. Therefore, inconsistencies with price search arise whenever this relationship is violated in the data.[21]

Assumption 2.4. *For all $t \in \mathcal{T}$, $\mathbb{E}[\bar{\omega}_t] = 0$, where $\bar{\omega}_{i,t} := L^{-1} \sum_{l=1}^{L} \omega_{i,l,t}$ denotes the average search productivity across goods.*

Assumption 2.4 allows time-varying search productivity for specific goods as long as the overall search productivity remains constant. Permitting search productivity for a particular good to change over time is important in my application because of the coarse aggregation of the data. Indeed, since a consumer may purchase different baskets of goods in different time periods, prices may vary due to variations in the composition of the baskets of goods.

Aside from the above restriction, Assumption 2.4 is quite general as it does not presume anything about the underlying stochastic process of search productivity. Conditional on the aggregate average search productivity being time-invariant, it allows individual-specific search productivity to vary arbitrarily with both observables and unobservables. In particular, it includes Markovian processes often assumed in the production function literature.[22]

Given the log-linearity of the price functions, Assumption 2.4 implies that average log true prices should be around the mean, conditional on shopping trips. In accordance with this prediction, Figure 2.4 shows that the distribution of observed average log prices is centered around its mean. I report the unconditional distribution for expositional purposes; similar shapes are obtained for the conditional ones.[23]

Lastly, I bound the support of the elasticity of price with respect to shopping intensity to gain further identification power.

Assumption 2.5. *For all $l \in \mathcal{L}$, $\alpha^1_{i,l} \in [-1, 0]$.*

Assumption 2.5 constrains the elasticity of price with respect to shopping intensity to be in $[-1, 0]$ for every good $l \in \mathcal{L}$. In comparison, Aguiar and Hurst (2007) obtain a point-estimate of -0.074 for the elasticity of average price with respect to shopping intensity using

[21] See Appendix A.3 for analytical power results.

[22] See, for example, Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2015), and Gandhi et al. (2020).

[23] Consistent with Assumption 2.1, distributions conditioned on larger values of shopping trips tend to be centered around lower values of log prices.

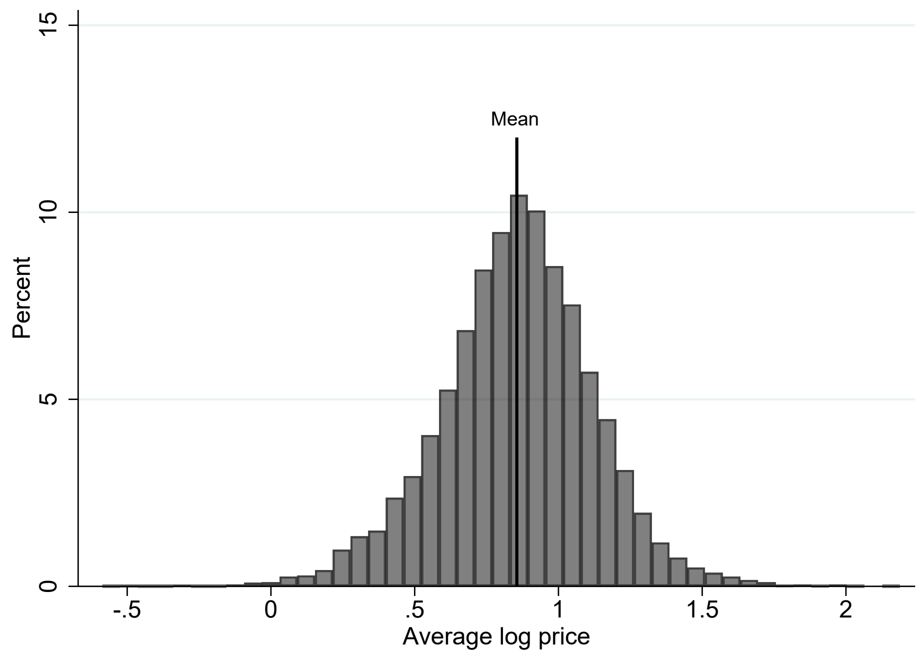

Figure 2.4: Distribution of Average Log Prices

the Homescan 1993-1995.[24] As such, Assumption 2.5 should give enough flexibility for the needs of the data.

Remark Measurement error implicitly accommodates various shocks that may occur outside the model. For example, changes in prices induced by supply shocks would be absorbed by the moments on measurement error provided they satisfy Assumption 2.2. Likewise, exogenous shocks can be absorbed by search productivity provided they satisfy Assumption 2.4. Accordingly, the model is robust to a variety of perturbations.

2.5 Statistical Framework

In this section, I extend the deterministic framework presented in the paper to a statistical one susceptible to testing and inference. I then show that inference on the true expected elasticity can be recovered in the model.

[24]Their estimate is obtained using an instrumental variable approach and is for a single aggregated good.

2.5.1 Characterization of the Model via Moment Functions

From now on, consumer data sets should be viewed as independent and identically distributed draws from some distribution. Let $\mathcal{X} := P \times C \times A$ and $\mathcal{E}|\mathcal{X}$ be the support of the latent random variables conditional on $\mathcal{X}$. Moreover, let $x_i \in \mathcal{X}$ denote the observed random data and $e_i \in \mathcal{E}|\mathcal{X}$ denote the latent random variables $(u_{i,t}, \delta_i, \boldsymbol{\alpha}_i, \boldsymbol{\rho}_{i,t}, \boldsymbol{\omega}_{i,t}, \boldsymbol{m}_{i,t})_{t\in\mathcal{T}}$.

To make the model amenable to statistical testing, I first express it through a set of moments. Therefore, for all $l \in \mathcal{L}$ and $s, t \in \mathcal{T}$, I write the model defined by Assumptions 2.1-2.5 with the following moment functions:

$$
\begin{aligned}
g^u_{i,s,t}(x_i, e_i) &:= \mathbb{1}\left(u_{i,s} - u_{i,t} - \delta_i^{-t}\left[\boldsymbol{p}^{*\prime}_{i,t}(\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) - \boldsymbol{\rho}'_{i,t}(\boldsymbol{a}_{i,s} - \boldsymbol{a}_{i,t})\right] \leq 0\right) - 1,\\
g^p_{i,l,t}(x_i, e_i) &:= \mathbb{1}\left(\log\!\left(p^*_{i,l,t}\right) - \left(\alpha^0_{i,l} + \alpha^1_{i,l}\log(a_{i,l,t}) - \omega_{i,l,t}\right) = 0\right) - 1,\\
g^m_{i,l,t}(x_i, e_i) &:= \log(p_{i,l,t}) - \log\!\left(p^*_{i,l,t}\right),\\
g^\omega_{i,l,t}(x_i, e_i) &:= \bar{\omega}_{i,t},
\end{aligned}
$$

where the first set of functions characterizes the concavity of the utility function, the second the log-linearity of the price functions, the third measurement error, and the last search productivity. The latent random variables satisfy their support constraints: $\delta_i \in [\underline{\delta}, 1]$, $\boldsymbol{\alpha}^1_i \in [-1, 0]$, $\boldsymbol{\rho}_{i,t} \leq 0$ and $\boldsymbol{p}^*_{i,t} > 0$, where $\boldsymbol{\rho}_{i,t}$ further satisfies

$$
\boldsymbol{\rho}_{i,t} := \frac{\partial \boldsymbol{p}_i(\boldsymbol{a}_{i,t}, \boldsymbol{\omega}_{i,t})}{\partial \boldsymbol{a}_{i,t}} \odot \boldsymbol{c}_{i,t}.
$$

This equality constraint implies that $\boldsymbol{\rho}_{i,t}$ is completely determined by the data and latent variables $(u_{i,t}, \delta_i, \boldsymbol{\alpha}_i, \boldsymbol{\omega}_{i,t}, \boldsymbol{m}_{i,t})_{t\in\mathcal{T}}$.

Every consumer has a total of $T^2 + L\cdot T + L\cdot T + T$ moment functions, written as $\boldsymbol{g}_i(x_i, e_i) := (\boldsymbol{g}^u_i(x_i, e_i)', \boldsymbol{g}^p_i(x_i, e_i)', \boldsymbol{g}^m_i(x_i, e_i)', \boldsymbol{g}^\omega_i(x_i, e_i)')'$ for short. Arbitrary combinations of these sets of functions are denoted with their superscripts bundled together. For example, $\boldsymbol{g}^{m,\omega}_i(x_i, e_i)$ is the set of functions on measurement error and search productivity. Note that the moment functions $\boldsymbol{g}_i(x_i, e_i)$ depend on unobservables. As such, the latent variables have to be drawn from some distribution for the moment functions to be evaluated.

Since Assumptions 2.2 and 2.4 only impose centering conditions on the expected measurement error and the expected average search productivity, the latent variables are essentially unrestricted for any particular consumer. Therefore, empirical content is gained by averaging the moment functions $\boldsymbol{g}_i(x_i, e_i)$ across consumers.[25] The distribution of the latent variables that satisfy individual-specific moment functions $\boldsymbol{g}_i^{u,p}(x_i, e_i)$ may then only do so when the expectation of $\boldsymbol{g}_i^{m,\omega}(x_i, e_i)$ deviates from zero.

2.5.2 Statistical Price Search Rationalizability

Let $\mathcal{M}_{\mathcal{X}}$, $\mathcal{M}_{\mathcal{E},\mathcal{X}}$, and $\mathcal{M}_{\mathcal{E}|\mathcal{X}}$ denote the set of all probability measures defined over $\mathcal{X}$, $(\mathcal{E}, \mathcal{X})$, and $\mathcal{E}|\mathcal{X}$, respectively. Moreover, let $\mathbb{E}_{\mu\times\pi}[\boldsymbol{g}_i(x_i, e_i)] := \int_{\mathcal{X}} \int_{\mathcal{E}|\mathcal{X}} \boldsymbol{g}_i(x_i, e_i)\, d\mu\, d\pi$, where $\mu \in \mathcal{M}_{\mathcal{E}|\mathcal{X}}$ and $\pi \in \mathcal{M}_{\mathcal{X}}$. The moment functions previously defined allow me to define the statistical rationalizability of a data set with respect to price search.[26]

Definition 2.3. *Under Assumptions 2.1-2.5, a random data set $x := \{x_i\}_{i=1}^N$ is price search rationalizable (PS-rationalizable) if*

$$\inf_{\mu\in\mathcal{M}_{\mathcal{E}|\mathcal{X}}} ||\mathbb{E}_{\mu\times\pi_0}[\boldsymbol{g}_i(x_i, e_i)]|| = 0,$$

where $\pi_0 \in \mathcal{M}_{\mathcal{X}}$ is the observed distribution of x.

That is, the data are PS-rationalizable if there exists a distribution of the latent random variables conditional on the data such that the expected moment functions are satisfied. In practice, searching over the set of all conditional distributions represents a daunting task. Fortunately, the following result shows that the problem can be greatly simplified without loss of generality.[27]

Theorem 2.2. *The following are equivalent:*

(i) A random data set x is PS-rationalizable.

(ii) $\min_{\boldsymbol{\gamma}\in\mathbb{R}^{L\cdot T+T}} ||\mathbb{E}_{\pi_0}[\tilde{\boldsymbol{h}}_i(x_i; \boldsymbol{\gamma})]|| = 0,$

[25] I show that the model defined by Assumptions 2.1-2.5 is refutable in Appendix A.3.
[26] This definition follows the notion of identified set in Schennach (2014).
[27] See Aguiar and Kashaev (2021) for the weak technical assumptions required for this result to hold.

where

$$\tilde{\boldsymbol{h}}_i(x_i;\boldsymbol{\gamma}) := \frac{\int_{e_i\in\mathcal{E}|\mathcal{X}} \boldsymbol{g}_i^{m,\omega}(x_i,e_i)\exp\Big(\boldsymbol{\gamma}'\boldsymbol{g}_i^{m,\omega}(x_i,e_i)\Big)\mathbb{1}(\boldsymbol{g}_i^{u,p}(x_i,e_i)=0)\,d\eta(e_i|x_i)}{\int_{e_i\in\mathcal{E}|\mathcal{X}} \exp\Big(\boldsymbol{\gamma}'\boldsymbol{g}_i^{m,\omega}(x_i,e_i)\Big)\mathbb{1}(\boldsymbol{g}_i^{u,p}(x_i,e_i)=0)\,d\eta(e_i|x_i)},$$

and where $\eta(\cdot|x_i)$ is an arbitrary user-specified distribution function supported on $\mathcal{E}|\mathcal{X}$ such that $\mathbb{E}_{\pi_0}[\log\Big(\mathbb{E}_\eta[\exp\Big(\boldsymbol{\gamma}'\boldsymbol{g}_i^{m,\omega}(x_i,e_i)\Big)|x_i]\Big)]$ exists and is twice continuously differentiable in $\boldsymbol{\gamma}$ for all $\boldsymbol{\gamma}\in\mathbb{R}^{L\cdot T+T}$.

Proof. See Theorem 2.1 in Schennach (2014) and Theorem 4 in Aguiar and Kashaev (2021).

In words, Theorem 2.2 (*ii*) averages out the unobservables in $\boldsymbol{g}_i(x_i,e_i)$ according to some conditional distribution.[28] The particularity of $\eta(\cdot|x_i)$ is to preserve the set of values that the objective function can take before the latent variables have been averaged out. As such, any minimum achieved under μ can also be achieved under $\eta(\cdot|x_i)$. The dimensionality of the problem is then further reduced by noting that the concavity of the utility function and the log-linearity of the price functions are only restricting the conditional support of the unobservables. Thus, one can draw from the conditional distribution $\tilde{\eta}(\cdot|x_i) := \mathbb{1}(\boldsymbol{g}_i^{u,p}(x_i,\cdot)=0)\eta(\cdot|x_i)$ rather than leaving the moment functions $\boldsymbol{g}_i^{u,p}(x_i,\cdot)$ in the optimization problem.

In most applications, the distribution $\tilde{\eta}(\cdot|x_i)$ may be taken to be proportional to a normal distribution:

$$d\tilde{\eta}(\cdot|x_i) \propto \exp\left(-\|\boldsymbol{g}_i^{m,\omega}(x_i,e_i)\|^2\right),$$

where the value of the mean and variance are inconsequential for the validity of the result. To draw from this distribution, the first step is to obtain latent variables that satisfy the moment functions $\boldsymbol{g}_i^{u,p}(x_i,e_i)$ and can be achieved by rejection sampling. Then, a standard Metropolis-Hastings algorithm can be used to draw from the distribution.[29]

[28] Schennach (2014) shows the existence of an admissible conditional distribution $\eta(\cdot|x_i)$ and gives a generic construction for it.

[29] Additional details about the implementation are given in Appendix A.4.

2.5.3 Statistical Testing

The notion of PS-rationalizability together with Theorem 2.2 provides a feasible way of checking whether the data are consistent with the model. To statistically test the PS-rationalizability of a data set, let

$$\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma}) := \frac{1}{N}\sum_{i=1}^{N}\tilde{\boldsymbol{h}}_i(x_i,\boldsymbol{\gamma})$$

and

$$\hat{\tilde{\boldsymbol{\Omega}}}(\boldsymbol{\gamma}) := \frac{1}{N}\sum_{i=1}^{N}\tilde{\boldsymbol{h}}_i(x_i,\boldsymbol{\gamma})\tilde{\boldsymbol{h}}_i(x_i,\boldsymbol{\gamma})' - \hat{\tilde{\boldsymbol{h}}}_i(\boldsymbol{\gamma})\hat{\tilde{\boldsymbol{h}}}_i(\boldsymbol{\gamma})'$$

denote the sample analogues of $\tilde{\boldsymbol{h}}$ and its variance, respectively. Furthermore, let $\hat{\tilde{\boldsymbol{\Omega}}}^{-}$ denote the generalized inverse of the matrix $\hat{\tilde{\boldsymbol{\Omega}}}$. Schennach (2014) shows that the test statistic

$$\mathrm{TS}_N := N\inf_{\boldsymbol{\gamma}\in\mathbb{R}^{L\cdot T+T}}\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma})'\hat{\tilde{\boldsymbol{\Omega}}}^{-}(\boldsymbol{\gamma})\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma})$$

is stochastically bounded by a χ^2 distribution with $d_m := L\cdot T + T$ degrees of freedom ($\chi^2_{d_m}$).[30] As such, the PS-rationalizability of a data set can be checked by comparing the value of the test statistic against the critical value of the chi-square distribution with d_m degrees of freedom.

2.5.4 Inference

Conditional on the data being consistent with PS-rationalizability, the next step is to make inferences on parameters of interest. First, the next result shows that inference on the true average elasticity of average price is possible in the model.

Proposition 2.1. *The average expected elasticity of average price with respect to shopping intensity is given by*

$$\frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\left[\frac{\partial\overline{\log(p_t^*)}}{\partial\log(a_{l,t})}\right] = \frac{1}{L}\mathbb{E}\left[\bar{\alpha}^1\right],$$

where $\overline{\log(p_t^*)}$ *is the true average log price across goods and* $\bar{\alpha}^1$ *is the average elasticity of*

[30] Aguiar and Kashaev (2021) further show that the test has an asymptotic power equal to one.

price with respect to shopping intensity.

Proposition 2.1 states that the average expected effect of an increase in shopping intensity on the average price can be recovered from data on prices paid and search intensity. The reason why it can be achieved in the model is that Assumption 2.4 restricts the expected average search productivity to be time-invariant. Therefore, any variation in the expected average price must be caused exclusively by variations in shopping intensity.

In the statistical framework previously outlined, inference can be made by adding a moment on a parameter of interest. Proposition 2.1 suggests to choose the following moment:

$$L^{-1}\mathbb{E}[\bar{\alpha}^1] = \theta_0,$$

where $\bar{\alpha}^1$ is the average elasticity of price with respect to shopping intensity and $\theta_0 \in [-0.25, 0]$ is the average expected elasticity of average price with respect to shopping intensity.[31] As before, this condition can be encapsulated in a moment function:

$$g_i^{\alpha}(x_i, e_i) := L^{-1}\bar{\alpha}_i^1 - \theta_0.$$

A conservative 95% confidence set on θ_0 can be obtained by inverting the test statistic:

$$\{\theta_0 \in \Theta : \mathrm{TS}_N(\theta_0) \leq \chi^2_{d_m+1,0.95}\},$$

where $TS_N(\theta_0)$ is the test statistic at a fixed value of θ_0.

2.6 Empirical Application

In this section, I check whether the data are PS-rationalizable, set estimate the shopping technology across multiple demographics, and relate price search to consumption inequality.

[31]The support of θ_0 is given by L^{-1} times the support of the average elasticity of price with respect to shopping intensity $\overline{\alpha}_i = [-1, 0]$.

2.6.1 Price Search

In what follows, I set $\underline{\delta} = 0.95$ such that the support of the discount factor is $[0.95, 1]$. By applying the above methodology to my sample, I find that PS-rationalizability is not rejected by the data at the 95% confidence level. More precisely, I obtain a test statistic of 36.38, which is below the chi-square critical value of 43.77. Since the model is not rejected by the data, I can invert the statistical test to obtain a 95% confidence set on the average expected elasticity of average price with respect to shopping intensity. The results for all consumers, low-income consumers, and high-income consumers are reported in Figure 2.5.

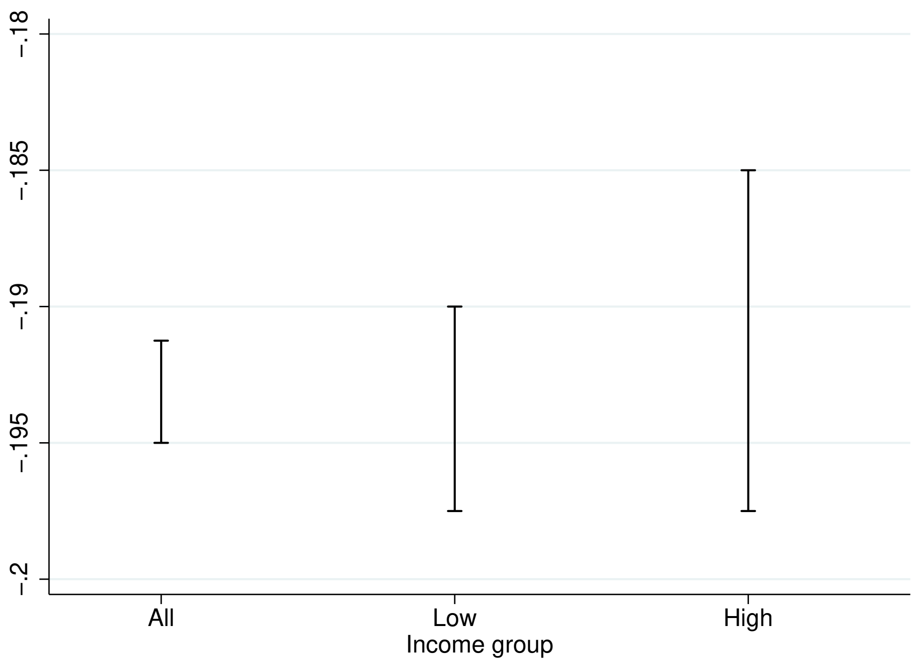

Figure 2.5: Average Expected Elasticity of Average Price with Respect to Shopping Intensity ($\mathbb{E}[\bar{\alpha}]$)

Notes: The vertical axis represents the average expected elasticity of average price with respect to shopping intensity. Low-income consumers are defined as those with an income lower than $25,000. High-income consumers are defined as those with an income greater than $50,000. The sample size is 1645 for all, 660 for low-income consumers, and 348 for high-income consumers.

In this figure, we see that the average expected elasticity of average price with respect to shopping intensity is about -0.1925. That is, doubling shopping intensity decreases the average price paid by about 19.25%. Figure 2.5 also shows that low-income consumers have a slightly better shopping technology compared to high-income consumers. To investigate whether low- and high-income consumers also have different shopping intensities, Figure 2.6 displays the distribution of log number of shopping trips by income group.

Figure 2.6: Log Number of Shopping Trips Density by Department Category

Notes: Low-income consumers are defined as those with an income lower than \$25,000. High-income consumers are defined as those with an income greater than \$50,000. The sample size is 660 for low-income consumers and 348 for high-income consumers.

Overall, Figure 2.6 shows that the distribution of shopping intensity is almost identical between income groups. Consumers earning less than \$25,000 a year are only making 1.65% more shopping trips per month than consumers earning more than \$50,000 a year. Combined with my finding that income groups have a similar shopping technology, it follows that heterogeneity in price search only explains a small portion of the difference in average price paid between low- and high-income consumers.

Lastly, to investigate whether the shopping technology differs across other dimensions than income, I set estimate the shopping technology by gender, education, geographic location, and occupation. These categories are defined as follows: gender separates the sample between males and females; education between consumers that did not graduate college and those that graduated college; geographic location between consumers in rural states and those in urban states; and occupation between workers and retirees.[32]

[32]A rural state is defined as one where the proportion of the urban population is below that of the U.S. population, and an urban state as one where the proportion of the urban population is above that of the U.S. population.

For every demographic, the 95% confidence set is $[-0.1975, -0.19]$, with the exception of workers where the confidence set is $[-0.195, -0.19]$.[33] Since the model allows for sizable heterogeneity, the confidence sets on the average expected elasticity of average price with respect to shopping intensity reflect inherent homogeneity in the average shopping technology rather than induced homogeneity through ex ante restrictions.

2.6.2 Consumption Inequality

To evaluate the impacts of heterogeneity in price search on consumption inequality, I compare how much more a consumer at the $20th$ quantile of the shopping intensity distribution would pay compared to one at the $80th$ quantile. I shall refer to the former type of consumers as infrequent shoppers and the latter type as frequent shoppers. The results are presented in Figure 2.7 and are obtained by using the shopping technologies estimated in Figure 2.5 and the observed distribution of number of shopping trips.

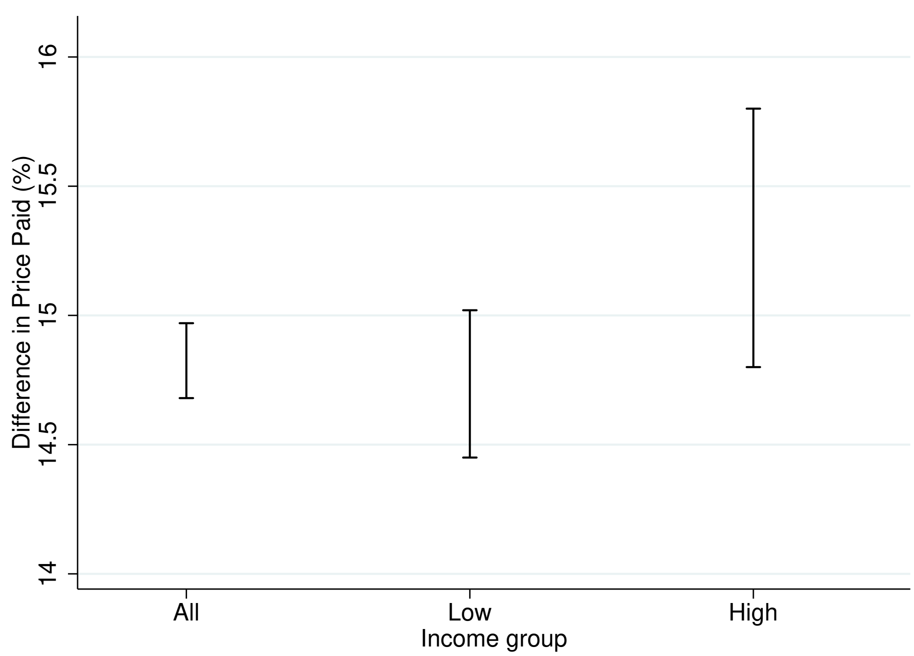

Figure 2.7: Difference in Price Paid due to Heterogeneity in Price Search

Notes: The vertical axis represents how much more consumers at the $20th$ quantile of the distribution of number of shopping trips pay compared to those at the $80th$ quantile.

[33] The 95% confidence set is identical across those demographics at a precision level of 0.0025. Therefore, any difference in the confidence set must be by less than 0.005.

Figure 2.7 shows that infrequent shoppers pay almost 15% more on the goods they purchase than frequent shoppers. Assuming that expenditure is fixed and that savings from search is spent on those same goods, the consumption level of frequent shoppers is 15% larger than that of infrequent shoppers. Figure 2.7 also shows that consumption inequality is larger among high-income consumers than among low-income consumers. Finally, Figure 2.7 implies that price search has two opposite effects on between-group consumption inequality. On one hand, it mitigates consumption inequality between low-income consumers with high shopping frequency and high-income consumers with low shopping frequency. On the other hand, price search exacerbates consumption inequality between low-income consumers with low shopping frequency and high-income consumers with high shopping frequency.

2.6.3 Instrumental Variable

The statistical framework allows me to impose additional moments to test the validity of an instrument within the structure of the model. For conciseness, I consider the validity of income as it is the main instrument used in Aguiar and Hurst (2007). Income is a valid instrument if the exogeneity assumption $\mathbb{E}[y_t \omega_{l,t}] = 0$ holds for all $l \in \mathcal{L}$ and $t \in \mathcal{T}$. Moreover, the instrument would also need to satisfy $\mathbb{E}[y_t m_{l,t}] = 0$ for all $l \in \mathcal{L}$ and $t \in \mathcal{T}$ if measurement error in prices is additive. I find that the model is rejected with a test statistic of 113.69 when further imposing these moments at $t = 1$.[34] In Appendix A.5, I assess the extent of the bias involved with instrumental variable (IV) in a regression setting and find that the estimate is downward biased. Given the severity of measurement error in prices and the evidence towards nonclassical measurement error (Einav et al., 2010), the main issue likely lies in the interaction between the instrument and measurement error.

2.6.4 Discussion

Price search allows consumers to pay lower prices by increasing their shopping intensity. This results in heterogeneity in prices paid such that expenditure may give an erroneous account of

[34] Imposing these moments at additional time periods would increase the test statistic and leave the conclusion unchanged.

consumption and, hence, consumption inequality. In a calibrated model of price search, Arslan et al. (2021) show that consumption inequality is significantly smaller than expenditure inequality in the Homescan 2004. Consistent with their finding, Figure 2.3 and Figure **??** display disparity in prices paid despite similar consumption levels between income groups. Since the difference in expenditures is attributable to lower prices paid by low-income consumers, an analysis based on expenditures would overstate actual consumption inequality in the data.

The conclusion that consumption inequality is negligible between income groups appears somewhat counterintuitive, especially given the overwhelming evidence that consumption inequality has increased over time such as in Aguiar and Bils (2015).[35] This finding can be reconciled by two observations. First, Aguiar and Bils (2015) measure the change in consumption inequality by comparing relative expenditures on luxury versus necessity goods. Since my data set only contains food categories (necessities), the rise in consumption inequality can be explained by changes in expenditures on luxury goods. Second, consumption inequality is typically measured from data on expenditures, thus ignoring differences in prices paid between income groups. As discussed previously, this can lead to an inflated measure of consumption inequality.

My application shows that price search has a strong impact on average prices paid. However, the set estimates in Figure 2.5 show that low-income consumers have only a slightly better average shopping technology compared to high-income consumers. Thus, most of the price differentials between income groups displayed in Figure 2.3 are due to differences in good quality. Using data from the National Health and Nutrition Examination Survey, Wang et al. (2014) document an increase in diet quality inequality between socioeconomic groups.[36] To the extent that lower quality goods are also less healthy, consumption inequality between income groups likely takes the form of diet quality inequality.

[35]See Attanasio and Pistaferri (2016) for a review of the literature investigating the evolution of consumption inequality.

[36]Individuals with low socioeconomic status are defined as those with less than 12 years of education and an income below 130% of the poverty line. Individuals with high socioeconomic status are defined as those with more than 12 years of education and an income above 350% above the poverty line.

2.7 Conclusion

This paper proposes a flexible semiparametric approach to estimate the impacts of shopping intensity on prices paid and consumption inequality. My approach allows one to make informative inferences with minimal data requirements and minimal assumptions on unobservables. Importantly, the econometric framework allows one to statistically test the restrictions of the model. In that general setting, I show that the true average elasticity of average price with respect to shopping intensity can be recovered.

I show that the model I propose is consistent with consumers in the Nielsen Homescan Dataset. Using aggregated data on food expenditures, I find that a doubling of shopping frequency decreases the average price paid by about 19%. Also, by recognizing differences in price search behavior I show that heterogeneity in price search mitigates between-group consumption inequality on average. At the same time, differences in search intensity reveal unequal gains from price search. Specifically, price search reduces consumption inequality between low-income consumers with high shopping frequency and high-income consumers with low shopping frequency but exacerbates it between low-income consumers with low shopping frequency and high-income consumers with high shopping frequency.

The methodology put forward in this paper may be useful in other settings. For example, the shopping technology may be viewed as a function that takes shopping intensity as an input and returns price paid as an output. In this light, the estimation of the shopping technology is related to the estimation of production functions such as in Gandhi et al. (2020).[37] In a firm setting, my centering condition on the error term amounts to assuming that the average productivity across firms is constant over time. This allows firm-specific productivity shocks to follow essentially unrestricted processes. Thus, my methodology could be used to identify the production function without appealing to Markovian processes.

[37] See also Levinsohn and Petrin (2003) for an earlier treatment on the estimation of gross output production functions and Olley and Pakes (1996) and Ackerberg et al. (2015) for the related problem of estimating value-added production functions.

Chapter 3

Robust Inference on Discount Factors

3.1 Introduction

The exponential discounting model is a predominant tool for analyzing dynamic choice in applied work. Its attractiveness rests in that time preferences are summarized by a single parameter—the discount factor. This allows one to tractably analyze a decision maker's intertemporal choices, which is crucial in a vast range of applications. Accordingly, many studies have tried to recover its key time parameter. However, a common feature in this literature is the specification of the consumer's preferences.[1] This constitutes a potentially important limitation as erroneously specifying preferences may lead to spurious estimates of the discount factor.

At its core, the exponential discounting model assumes that the utility function is additively time-separable and stationary. Under these assumptions, the transitivity of preferences can be characterized by the well-known Generalized Axiom of Revealed Preference (GARP). In particular, Afriat (1967) showed that for any finite data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t \in \mathcal{T}}$ of discounted prices and demands, GARP is necessary and sufficient for the existence of a well-behaved utility function that rationalizes the data. The distinctive feature of exponential discounting, though, is the prediction that consumers will be time consistent. Namely, it requires consumers to commit to their initial plan as time unfolds.[2]

[1]For an overview of this large literature, see Frederick et al. (2002).

[2]In experimental settings, a preference reversal occurs when the consumer chooses a sooner-smaller reward over a later-larger one and then switches to the later-larger reward when an equal delay is added to both outcomes. This behavior violates time consistency if the consumer deviates from his plan and chooses the sooner-smaller

I show that the exponential discounting model, which is normally stated as a dynamic maximization problem with an intertemporal budget constraint, may be expressed as a repeated static utility maximization problem without any budget constraint. Specifically, a consumer is an exponential discounter if and only if there exists a locally nonsatiated instantaneous utility function $u : \mathbb{R}^L_+ \to \mathbb{R}$ and a discount factor $\delta \in (0, 1]$ such that

$$\boldsymbol{c}_t \in \arg\max_{\boldsymbol{c}\in\mathbb{R}^L_+} u(\boldsymbol{c}) + \delta^{-t}(y^d_t - \boldsymbol{\rho}'_t\boldsymbol{c}) \quad \forall t \in \mathcal{T},$$

where $y^d_t > 0$ denotes discounted income in period t. Letting $s^d := y^d_t - \boldsymbol{\rho}'_t\boldsymbol{c}$ denote discounted savings and $U_t(\boldsymbol{c}, s^d) := u(\boldsymbol{c}) + \delta^{-t}s^d$, the objective function may be seen as an additively separable time-dependent augmented utility function $U_t : \mathbb{R}^L_+ \times \mathbb{R} \to \mathbb{R}$. The dynamics of the model is captured through the incorporation of savings into the consumer's consideration. Indeed, the amount a consumer is willing to consume in any time period is regulated by his desire to save for future consumption.

My methodology exploits the theory of revealed preference popularized by Afriat (1967) and Varian (1982). This approach obtains sharp conditions that any demand data must satisfy in order to be consistent with utility maximization, and reciprocally, any behavior stemming from utility maximization must satisfy them.[3] In the exponential discounting model, for a given set of observations $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$, these conditions yield a set of linear inequalities that are known up to the discount factor. Since revealed preference conditions are exact, the main requirement maintained in this study is that consumers have perfect foresight. In addition, I impose the marginal utility of discounted expenditure to be constant across time as it is necessary for exponential discounting to have implications beyond GARP (Browning, 1989).

A data set either satisfies or violates the revealed preference inequalities that characterize exponential discounting. This makes the direct implementation of these inequalities of limited applicability as they fail to handle innocuous deviations that may arise in the data. As such, I propose a statistical test that allows for measurement error in variables as in Varian (1985). While this statistical test can be inverted to recover nonparametric bounds on the discount

reward in the future (Halevy, 2015).

[3] Although it is possible to impose additional constraints on the utility function, the revealed preference framework does not require it.

factor, it has the undesirable property to be extremely conservative, thus hindering one's ability to make informative inference. I address this caveat by meaningfully disciplining measurement error in terms of percentage of wasted income.

In my empirical application, I apply my methodology to the checkout scanner panel data set on food expenditures from Echenique et al. (2011). I find that many consumers behave consistently with exponential discounting when measurement error in prices is taken into account. Moreover, I show that bounds on the discount factor get tighter as the extent of measurement error decreases. Finally, I find that observable characteristics such as income, education and age fail to capture heterogeneity in discounting.

The remainder of the paper is organized as follows. Section 3.2 reviews the related literature. Section 3.3 formally defines the exponential discounting model and reviews its testable implications. Section 3.4 obtains a novel representation of the exponential discounting model, generalizes it to a partial efficiency setting, and derives its testable implications. Section 3.5 introduces the statistical test and provides a way to obtain a confidence set on the discount factor. Section 3.6 contains the empirical application and Section 3.7 concludes. The main proofs and supplemental material can be found in Appendix B.

3.2 Related Literature

This paper builds on the exponential discounting characterization of Browning (1989) in order to derive a novel representation of the model in terms of a time-dependent augmented utility function. The use of an augmented utility function has also been used by Deb et al. (2018) in a different framework. They consider the concept of revealed price preference and obtain a consistency condition called the Generalized Axiom of Price Preference (GAPP). The augmented utility function I derive is distinct from theirs as it has the peculiarity of being time-dependent; a notable implication is that exponential discounting can be thought of as a static model with reference-dependent preferences.[4]

This new representation lends itself to a partial efficiency analysis similar to that of Afriat

[4]In this light, my representation relates to the literature on reference-dependent utility functions popularized by the seminal work of Kahneman and Tversky (1979).

(1973) which allows me to measure the severity of a violation from exponential discounting in the data. In this respect, my result relates to existing partial efficiency results such as those for static utility maximization (Halevy et al., 2018; De Clippel and Rozen, 2018), homothetic rationalizability (Heufer and Hjertstrand, 2019), and expected utility maximization (Echenique et al., 2021). I complement these papers by bringing partial efficiency to a dynamic setting. Notably, my extension allows one to use the statistical test of Cherchye et al. (2020) to exponential discounting.

My endeavor is complementary to that of Adams et al. (2014) who extend the analysis of the exponential discounting model for preference heterogeneity and renegotiations within the household. It also relates to models of habit formation such as the one proposed in Crawford (2010) and Demuynck and Verriest (2013) who examine the fit of richer life-cycle models. More generally, my approach is similar to that of Blow et al. (2021) who develop a test for the quasi-hyperbolic model. My work differs from theirs in that I focus on improving the applicability of the standard version of exponential discounting.

My methodology is close to that of Brown and Calsamiglia (2007) who provide conditions for quasilinear utility rationalization, and to Echenique et al. (2020) who provide an axiomatic characterization of exponential discounting for experimental data.[5] Instead, my test is aimed to be applied to survey or scanner data where choices are made over multidimensional consumption bundles.

3.3 Exponential Discounting

In this section, I introduce the notation used throughout the paper, formally define the exponential discounting model, and show how to get nonparametric bounds on the discount factor.

3.3.1 Notation

The typical scenario under consideration is that of purchases made by a consumer over a certain time window. Let $\mathcal{L} \in \{1, ..., L\}$ denote the number of observed commodities and $\mathcal{T} = \{0, ..., T\}$ the periods for which data on consumers are observable. For any good $l \in \mathcal{L}$ and time period

[5] Their test applies to a single good, a case that more naturally occurs in experiments.

$t \in \mathcal{T}$, denote discounted price by $\rho_{l,t} = p_{l,t}/\prod_{i=0}^{t}(1+r_i)$, where $p_{l,t}$ is the spot price and r_i is the interest rate, and denote consumption by $c_{l,t}$.[6] An observation is therefore a pair $(\boldsymbol{\rho}_t, \boldsymbol{c}_t) \in \mathbb{R}^L_{++} \times \mathbb{R}^L_{+}$, and accordingly, a data set is written as $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

3.3.2 Exponential Discounting Rationalizability

The objective function faced by an exponential discounting (ED) consumer at time $\tau \in \mathcal{T}$ is given by

$$U_\tau(\boldsymbol{c}_\tau, ..., \boldsymbol{c}_{T-\tau}) = u(\boldsymbol{c}_\tau) + \sum_{j=1}^{T-\tau} \delta^j u(\boldsymbol{c}_{\tau+j}),$$

where $u(\cdot)$ is the instantaneous utility function and $\delta \in (0, 1]$ is the discount factor. Moreover, consumption satisfies the linear budget constraint

$$\boldsymbol{\rho}_t'\boldsymbol{c}_t + s_t^d = y_t^d + a_t^d \quad \forall t \in \{\tau, \ldots, T\},$$

where s_t^d denotes discounted savings, $y_t^d > 0$ denotes discounted income and a_t is the discounted value of assets held at period t.[7] The assets evolve according to the law of motion: $a_t = (1 + r_t)s_{t-1}$. A data set is consistent with exponential discounting if it can be thought of as stemming from the model.

Definition 3.1. *A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is ED-rationalizable if there exists a locally nonsatiated, continuous, monotonic, and concave instantaneous utility function $u(\cdot)$, an income stream $(y_t^d)_{t\in\mathcal{T}} \in \mathbb{R}^{|\mathcal{T}|}_{++}$, an initial asset level $a_0 \geq 0$, and a discount factor $\delta \in (0, 1]$ such that the consumption stream $(\boldsymbol{c}_t)_{t\in\mathcal{T}}$ solves*

$$\max_{(\boldsymbol{c}_t)_{t\in\mathcal{T}} \in \mathbb{R}_+^{L\times|\mathcal{T}|}} u(\boldsymbol{c}_0) + \sum_{t=1}^{T} \delta^t u(\boldsymbol{c}_t) \ \ s.t.\ \ \boldsymbol{\rho}_0'\boldsymbol{c}_0 + \sum_{t=1}^{T} \boldsymbol{\rho}_t'\boldsymbol{c}_t = y_0 + \sum_{t=1}^{T} y_t^d + a_0.$$

The main requirement in the previous definition of exponential discounting is that consumers have perfect foresight. Consistent with the permanent income hypothesis, I further impose the marginal utility of discounted expenditure to be constant across time (Bewley, 1977).[8]

[6] The interest rate in the first period is set to zero, that is, $r_0 = 0$.

[7] That is, $s_t^d = s_t/\prod_{i=0}^{t}(1+r_i)$, $y_t^d = y_t/\prod_{i=0}^{t}(1+r_i)$ and $a_t^d = a_t/\prod_{i=0}^{t}(1+r_i)$.

[8] In the current framework, this is necessary for exponential discounting to have implications beyond static

This assumption is motivated by the fact that, if the marginal utility of discounted expenditure at period t was higher than period $s \neq t$, then the consumer could move income from s to t such as to increase consumption at t and be better off.

The empirical implications of exponential discounting is captured by the following result due to Browning (1989).

Proposition 3.1. *The following statements are equivalent:*

(i) The data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ *is ED-rationalizable.*

(ii) There exist numbers u_t, $t = 0, \dots, T$, *and a discount factor* $\delta \in (0, 1]$ *such that*

$$u_s \leq u_t + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

(iii) There exists a discount factor $\delta \in (0, 1]$ *such that for any subset of indices* $\tau = \{t_i\}_{i=1}^m$ *with* $t_i \in \mathcal{T}$ *and* $m \geq 2$,

$$0 \leq \delta^{-t_1}\boldsymbol{\rho}_{t_1}'(\boldsymbol{c}_{t_2} - \boldsymbol{c}_{t_1}) + \ldots + \delta^{-t_m}\boldsymbol{\rho}_{t_m}'(\boldsymbol{c}_{t_1} - \boldsymbol{c}_{t_m}). \tag{CM}$$

Proposition 3.1 gives two alternative tests for the exponential discounting model. Conditional on $\delta \in (0, 1]$, condition (*ii*) is a set of linear inequalities and can be solved using linear programming. Turning to (*iii*), note that pairs of indices $s, s+h \in \mathcal{T}$, where $h \geq 1$, provide bounds on the discount factor. Accordingly, I define the greatest lower bound and the least upper bound on the discount factor as

$$glb := \max_{s,s+h\in\mathcal{T}} \left\{ \left(\frac{\boldsymbol{\rho}_{s+h}'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})}{\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})} \right)^{1/h} \right\} \text{ such that } \boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$$

and

$$lub := \min_{s,s+h\in\mathcal{T}} \left\{ \left(\frac{\boldsymbol{\rho}_{s+h}'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})}{\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})} \right)^{1/h} \right\} \text{ such that } \boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) > 0,$$

whenever such bounds exist, and $glb = 0$, $lub = 1$, otherwise.

---

utility maximization (Browning, 1989).

To gain some intuition on these bounds, note that when $\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$, the consumer does *not* reveal a preference for the earlier bundle over the later one.[9] This can only happen if he is somewhat patient and therefore yields a lower bound on the discount factor. In the case where $\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) > 0$, the consumer reveals a preference for the earlier bundle over the later one. In turn, this can only happen if he is somewhat impatient and therefore yields an upper bound on the discount factor.

Furthermore, note that the size of $\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$ gives an indication of how enticing $\boldsymbol{c}_{s+h}$ is compared to $\boldsymbol{c}_s$ at time s. Likewise, $\boldsymbol{\rho}_{s+h}'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$ gives an indication of how enticing $\boldsymbol{c}_{s+h}$ is compared to $\boldsymbol{c}_s$ at time $s + h$. The more enticing $\boldsymbol{c}_{s+h}$ becomes at time $s + h$ relative to time s, the more patient the consumer gets. Intuitively, when $\boldsymbol{c}_{s+h}$ becomes an increasingly better option at time $s + h$ relative to time s, the consumer's willingness to leave $\boldsymbol{c}_{s+h}$ for later strengthens. In other words, the lower bound takes on larger positive values. A similar interpretation holds for upper bounds.

With these bounds in hand, condition (*iii*) allows me to derive necessary conditions that yield additional intuition on the ED model and will prove useful for computational purposes.

Corollary 3.1. *The data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is ED-rationalizable only if GARP holds and*

$$glb \leq lub \ ; \ glb \leq 1 \ ; \ lub > 0. \tag{CD}$$

Corollary 3.1 states that the exponential discounting model has an additional testable implication compared to static utility maximization. As for the latter, GARP captures within-period consistency. That is, it ensures that the bundle chosen at time t is the best among all feasible bundles in that period. In contrast, condition CD represents the dynamic of the model and guarantees that the intertemporal choices of the consumer are pairwise time consistent. However, these conditions are not sufficient for ED-rationalizability, as the following example displays.

Example Consider a bivariate demand ($\mathcal{L} = \{1, 2\}$) with three time periods ($\mathcal{T} = \{0, 1, 2\}$). The consumer has a data set where $(\boldsymbol{\rho}_0, \boldsymbol{c}_0) = ([1, 1]', [4, 3]')$, $(\boldsymbol{\rho}_1, \boldsymbol{c}_1) = ([2, 5]', [1, 2]')$ and $(\boldsymbol{\rho}_2, \boldsymbol{c}_2) = ([4, 2]', [3, 6]')$. It is easy to verify that GARP holds. Indeed, $\boldsymbol{c}_0 R^D \boldsymbol{c}_1$, $\boldsymbol{c}_2 R^D \boldsymbol{c}_0$ and

[9] For any $t \in \mathcal{T}$, a bundle $\boldsymbol{c}_t$ is said to be revealed preferred to a bundle $\boldsymbol{c}$ if $\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t) \leq 0$. See Appendix B.1 for a detailed review of revealed preference concepts.

$\boldsymbol{c}_2 R^D \boldsymbol{c}_1$ so no cycle exists. To see that CD is also satisfied, note that

$$\frac{\boldsymbol{\rho}_1'(\boldsymbol{c}_0 - \boldsymbol{c}_1)}{\boldsymbol{\rho}_0'(\boldsymbol{c}_0 - \boldsymbol{c}_1)} = \frac{11}{4}, \left(\frac{\boldsymbol{\rho}_2'(\boldsymbol{c}_0 - \boldsymbol{c}_2)}{\boldsymbol{\rho}_0'(\boldsymbol{c}_0 - \boldsymbol{c}_2)}\right)^{1/2} = \frac{-2}{-2} = 1, \text{ and } \frac{\boldsymbol{\rho}_2'(\boldsymbol{c}_1 - \boldsymbol{c}_2)}{\boldsymbol{\rho}_1'(\boldsymbol{c}_1 - \boldsymbol{c}_2)} = \frac{-16}{-24} = \frac{2}{3}.$$

Clearly, $glb = 1$ and $lub = 11/4$ so the conditions of CD are met. However, the data set does not satisfy CM since when δ is equal to one[10],

$$f(\delta) := \delta^{-2}\boldsymbol{\rho}_2'(\boldsymbol{c}_1 - \boldsymbol{c}_2) + \delta^{-1}\boldsymbol{\rho}_1'(\boldsymbol{c}_0 - \boldsymbol{c}_1) + \delta^{-0}\boldsymbol{\rho}_0'(\boldsymbol{c}_2 - \boldsymbol{c}_0) = -16 + 11 + 2 < 0.$$

In practice, Corollary 3.1 only involves testing GARP and checking the set of inequalities in CD. As highlighted by Varian (1982), one can use an efficient algorithm called the Floyd-Warshall algorithm to get the transitive closure of the direct revealed preference relation.[11] Importantly, these conditions can be parallelized, thus greatly reducing the computational burden when exponential discounting has to be tested repeatedly.

3.4 Exponential Discounting under Partial Efficiency

This section shows that the exponential discounting model has a time-dependent augmented utility function representation that can be used to account for inconsistent choices in the observed data.

3.4.1 Time-dependent Augmented Utility Function

The main problem with the previous result is that, when a data set is not exactly ED-rationalizable, it becomes impossible to recover bounds on the discount factor. This is highly prohibitive as the observed data are often inconsistent with the model. For example, in the presence of measurement error the observed data could be inconsistent with the model even if the true data are ED-rationalizable.

To remedy this problem, I provide a novel characterization of the exponential discounting model that will allow me to generalize the results introduced in the previous section.

---

[10] It is sufficient to check $\delta = 1$ as the first-order condition of $f(\delta)$ is strictly positive for all $\delta \in (0, 1]$.

[11] See Floyd (1962).

Theorem 3.1. *The following statements are equivalent:*

(i) The data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ *is ED-rationalizable.*

(ii) There exists a locally nonsatiated, continuous, monotonic and concave instantaneous utility function $u(\cdot)$ *and a discount factor* $\delta \in (0, 1]$ *such that for all* $t \in \mathcal{T}$ *and* $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}.$$

A quick comparison of the last condition in Theorem 3.1 with the standard formulation of exponential discounting highlights two major differences. First, there is no budget constraint in the latter. Second, the consumer's problem is much simpler as it only requires solving for optimal consumption bundles rather than the whole consumption stream. To interpret condition (*ii*), it is useful to rewrite it as

$$\boldsymbol{c}_t \in \arg\max_{\boldsymbol{c}\in\mathbb{R}^L_+} u(\boldsymbol{c}) + \delta^{-t}(y^d_t - \boldsymbol{\rho}'_t\boldsymbol{c}) \quad \forall t \in \mathcal{T}.$$

This formulation emphasizes that exponential discounting can be seen as a repeated static utility maximization problem. Letting $s^d := y^d_t - \boldsymbol{\rho}'_t\boldsymbol{c}$ denote savings and $U_t(\boldsymbol{c}, s^d) := u(\boldsymbol{c}) + \delta^{-t}s^d$, the objective function can be interpreted as a time-dependent augmented utility function $U_t : \mathbb{R}^L_+ \times \mathbb{R} \to \mathbb{R}$. It indicates that, in any given time period, the consumer values both current consumption and savings. This compromise between current consumption and savings captures the idea that increasing consumption today leaves a lesser amount of wealth for future periods, thus diminishing future consumption. In the absence of a budget constraint, the mechanism by which an interior solution is achieved therefore relies on the trade-off between the two.

3.4.2 Exponential Discounting under Partial Efficiency

In the revealed preference literature, it is standard to deal with deviations from a given model by slightly relaxing its constraints. Following this approach, I shall adopt the novel representation of Theorem 3.1 for exponential discounting rationalizability under partial efficiency.

Definition 3.2. *Let $e \in (0,1]$. The e-ED model rationalizes the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ if there exists a locally nonsatiated, continuous, monotonic and concave utility function $u(\cdot)$ and a discount factor $\delta \in (0,1]$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$*

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}/e.$$

This definition accounts for digressions from exponential discounting by considering an efficiency level e that rationalizes every choice of a consumer at once.[12] In particular, note that any consumption behavior may be rationalized by the e-ED model for an e arbitrarily close to zero.[13] To see the economic intuition behind e, note that for a given time period t, the expression in the definition may be written as

$$\delta^t(\tilde{u}(\boldsymbol{c}_t) - \tilde{u}(\boldsymbol{c})) \geq e\boldsymbol{\rho}_t'\boldsymbol{c}_t - \boldsymbol{\rho}_t'\boldsymbol{c},$$

where the utilities have been scaled by a factor e. That is, the efficiency level ensures that the discounted benefit from consuming $\boldsymbol{c}_t$ rather than $\boldsymbol{c}$ is greater than the additional cost incurred from purchasing $\boldsymbol{c}_t$ instead of $\boldsymbol{c}$. The difference between the actual cost of acquiring $\boldsymbol{c}_t$ and what it should have been for it to be worthwhile therefore gives a measure of wasted income. Namely, for some $e \in (0,1]$ and period $t \in \mathcal{T}$, the consumer wastes an amount equal to $\boldsymbol{\rho}_t'\boldsymbol{c}_t - e\boldsymbol{\rho}_t'\boldsymbol{c}_t$ or $(1-e)\%$ of his income by making an inefficient choice.[14] The following result extends Proposition 3.1 to a partial efficiency setting.

Proposition 3.2. *For a given $e \in (0,1]$, the following statements are equivalent:*

(i) There exists a locally nonsatiated, continuous, monotonic and concave utility function $u(\cdot)$ and a discount factor $\delta \in (0,1]$ e-ED rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

[12] This choice follows the same suggestion as Afriat (1973) for static utility maximization. Alternatively, one could have an efficiency index for each choice as in Varian (1990), and then consider some aggregator function (Dziewulski, 2020) to determine the overall level of inefficiency. Interestingly, Dziewulski (2020) provides a formal link between efficiency levels and the notion of just-noticeable difference.

[13] That is, e may capture many sources of violation occurring simultaneously, as well as consumption behavior outside of the exponential discounting framework.

[14] Since I consider a common rationalizing efficiency level for all time periods, the consumer wastes up to $(1-e)\%$ of his lifetime income.

(ii) There exist numbers u_t, $t = 0, \ldots, T$, and a discount factor $\delta \in (0,1]$ such that

$$u_s \leq u_t + \delta^{-t} \boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

(iii) There exists a discount factor $\delta \in (0,1]$ such that for any subset of indices $\tau = \{t_i\}_{i=1}^m$ with $t_i \in \mathcal{T}$ and $m \geq 2$,

$$0 \leq \delta^{-t_1} \boldsymbol{\rho}_{t_1}'(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) + \ldots + \delta^{-t_m} \boldsymbol{\rho}_{t_m}'(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}). \qquad \text{(CM}(e))$$

Proposition 3.2 gives a way to gauge the severity of departure from exponential discounting by finding an efficiency index $e \in (0,1]$ e-ED rationalizing the data.[15] Conditional on $(e, \delta) \in (0,1]^2$, the existence of a solution can be checked by solving the set of inequalities in Proposition 3.2 (*ii*) using linear programming. A data set that needs a small efficiency level to be e-ED rationalizable is farther away from exponential discounting than one with a large efficiency level. In particular, if $e = 1$ then the data set is ED-rationalizable.

Remark By imposing $e = \delta = 1$ in Proposition 3.2, I recover the conditions for quasilinear utility maximization from Brown and Calsamiglia (2007). This observation makes clear that quasilinear utility maximization can be viewed as a special instance of exponential discounting. To test quasilinear utility maximization under partial efficiency, it suffices to find a solution to the inequalities in Proposition 3.2 conditional on $\delta = 1$.

3.5 Inference on the Discount Factor

In this section, I introduce measurement error in prices, present the statistical test of Varian (1985) when applied to the exponential discounting model, and propose a constrained statistical test based on e-ED rationalizability. I then show how the test can be inverted to construct a confidence set for the discount factor.

[15] Appendix B.2 discusses efficiency indices of interest such as the largest $e \in (0,1]$ e-ED rationalizing the data.

3.5.1 Statistical Test

Suppose prices are mismeasured such that observed prices $\boldsymbol{\rho}_t$ differ from true prices $\boldsymbol{\rho}_t^*$. Specifically, suppose that

$$\boldsymbol{\rho}_t = \boldsymbol{\rho}_t^*/(1 + \boldsymbol{\epsilon}_t),$$

where $\boldsymbol{\epsilon}_t$ is assumed to be a random vector whose components follow independent normal distributions with mean zero and unknown variance $\sigma^{*2} > 0$.[16] That is, $\epsilon_{l,t} \sim N(0, \sigma^{*2})$ for all $l \in \mathcal{L}$ and all $t \in \mathcal{T}$. It is useful to note that, under this assumption, the test statistic

$$T(\sigma^{*2}) := \sum_{t=0}^{T} \sum_{l=1}^{L} (\rho_{l,t}^*/\rho_{l,t} - 1)^2/\sigma^{*2} \tag{3.1}$$

follows a chi-square distribution. Since true prices are unobservable, the idea consists of obtaining a lower bound on $T(\sigma^{*2})$ by considering the following quadratic programming problem:

$$S(\sigma^{*2}, \delta^*) := \min_{(\boldsymbol{\pi}, \boldsymbol{u}) \in \mathbb{R}_{++}^{L \times |\mathcal{T}|} \times \mathbb{R}^{|\mathcal{T}|}} \sum_{t=0}^{T} \sum_{l=1}^{L} (\pi_{l,t}/\rho_{l,t} - 1)^2/\sigma^{*2}, \tag{3.2}$$

subject to

$$u_s - u_t \leq \delta^{*-t} \boldsymbol{\pi}_t' (\boldsymbol{c}_s - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T},$$

where δ^* is the true discount factor. If the data set is consistent with exponential discounting under true prices, then one can always pick $\boldsymbol{\pi}_t = \boldsymbol{\rho}_t^*$ for all $t \in \mathcal{T}$. Accordingly, $S(\sigma^{*2}, \delta^*) \leq T(\sigma^{*2})$ such that

$$\mathbb{P}\left[S(\sigma^{*2}, \delta^*) \leq \chi^2_{d,\alpha}\right] \geq \mathbb{P}\left[T(\sigma^{*2}) \leq \chi^2_{d,\alpha}\right] = 1 - \alpha,$$

where $\chi^2_{d,\alpha}$ is the critical value of the chi-square distribution with $d = L \cdot |\mathcal{T}|$ degrees of freedom and prespecified confidence level $\alpha \in (0, 1)$.

[16] Alternatively, one can assume an additive error: $\boldsymbol{\rho}_t = \boldsymbol{\rho}_t^* + \boldsymbol{\epsilon}_t$. I consider proportional measurement error as the scale of discounted prices changes significantly across time.

3.5.2 Constrained Statistical Test

The main disadvantage of the previous test is that prices that solve the problem (3.2) may be much closer to observed prices than actual true prices. That is, the test is extremely conservative. To alleviate this limitation, I propose to further restrict the set of allowable true prices to those satisfying e-ED, therefore yielding the following constraints:

$$u_s - u_t \leq \delta^{*-t}\boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \leq \delta^{*-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}, \tag{3.3}$$

where $e \in (0, 1]$, $(u_t)_{t\in\mathcal{T}}$ are real numbers, and $(\boldsymbol{\pi}_t)_{t\in\mathcal{T}}$ are candidate true prices. For a fixed $e \in (0, 1]$, the resulting constrained optimization problem is given by

$$S^C(\sigma^{*2}, \delta^*, e) := \min_{(\boldsymbol{\pi},\boldsymbol{u})\in\mathbb{R}_{++}^{L\times|\mathcal{T}|}\times\mathbb{R}^{|\mathcal{T}|}} \sum_{t=0}^{T}\sum_{l=1}^{L}(\pi_{l,t}/\rho_{l,t} - 1)^2/\sigma^{*2}, \tag{3.4}$$

subject to

$$u_s - u_t \leq \delta^{*-t}\boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \leq \delta^{*-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T},$$

where $(u_t)_{t\in\mathcal{T}}$ and $(\boldsymbol{\pi}_t)_{t\in\mathcal{T}}$ may take different values than those that solve the optimization problem (3.2).

It is useful to note that the constrained optimization problem converges to the unconstrained optimization problem when $e \in (0, 1]$ approaches zero. Thus, the optimization problem (3.4) can be viewed as a generalization of the optimization problem (3.2).

3.5.3 Inference

Let $\sigma^2 \geq \sigma^{*2}$ and $e \in (0, 1]$ be fixed numbers, where σ^2 may be thought of as an upper bound on the variance. Since the true discount factor is unknown, the key to get a feasible test statistic is to set $\delta \in (0, 1]$ and solve

$$V(\delta, e) := \min_{(\boldsymbol{\pi},\boldsymbol{u})\in\mathbb{R}_{++}^{L\times|\mathcal{T}|}\times\mathbb{R}^{|\mathcal{T}|}} \sum_{t=0}^{T}\sum_{l=1}^{L}(\pi_{l,t}/\rho_{l,t} - 1)^2, \tag{3.5}$$

subject to

$$u_s - u_t \leq \delta^{-t}\boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \leq \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

Defining $S^C(\sigma^2, \delta, e) := V(\delta, e)/\sigma^2$, the constrained confidence set is obtained by inverting the constrained test statistic:

$$CS^C := \left\{\delta \in (0,1] : S^C(\sigma^2, \delta, e) \leq \chi^2_{d,\alpha}\right\},$$

and likewise for the unconstrained confidence set:

$$CS := \left\{\delta \in (0,1] : S(\sigma^2, \delta) \leq \chi^2_{d,\alpha}\right\}.$$

Noting that $\lim_{e\to 0} S^C(\sigma^2, \delta, e) = S(\sigma^2, \delta)$, we have

$$\mathbb{P}\left[\lim_{e\to 0} \delta^* \in CS^C\right] = \mathbb{P}\left[\delta^* \in CS\right] \geq \mathbb{P}\left[S(\sigma^{*2}, \delta^*) \leq \chi^2_\alpha\right] \geq 1 - \alpha.$$

In words, the constrained test statistic converges to the unconstrained test statistic when $e \in (0,1]$ approaches zero. In that case, the probability that the constrained confidence set covers the true discount factor becomes at least $1-\alpha$. Conditional on (σ^2, e), the constrained confidence set can be recovered by solving (3.5) for each $\delta \in (0,1]$. The choice of these variables should be chosen from prior knowledge of the data set. For example, it may be possible to set $\sigma^2 = \sigma^{*2}$ if validation data are available.

3.6 Empirical Application

3.6.1 Data

In my empirical analysis, I implement the methodology developed in the previous sections using the Stanford Basket Dataset, which is a panel data set containing expenditures of 494 households between June 1991 and June 1993.[17] Specifically, I use the transformed data set

[17] I treat households as unitary entities even though they may have many members. As such, I refer to a household as a consumer or an individual.

of Echenique et al. (2011). As such, goods for which prices are observed in every week are retained and aggregated by brand for periods of four weeks. This yields a total of 375 distinct goods belonging to one of the following 14 categories: bacon, barbecue sauce, butter, cereal, coffee, crackers, eggs, ice cream, nuts, analgesics, pizza, snacks, sugar and yogurt.

Since none of these categories contain goods purchased for special events (e.g., turkey for Thanksgiving) or products whose quality may change with seasons (e.g., fruits), I expect preferences to be roughly stable over the time window considered. Additionally, due to the focus on food items, I do not expect consumers' purchases to vary considerably in response to changes in income. Finally, aggregation to monthly expenditure should mitigate stockpiling associated with sales.

The data set is prone to measurement error since it contains shelf prices instead of transaction prices. Thus, observed prices differ from actual prices paid whenever a consumer uses discounts such as coupons. As the data do not contain information on interest rates, I include interest rates on personal loans at commercial banks from the Federal Reserve Bank of St. Louis.[18] I report demographic information about households in the data set in Table 3.1. For a comprehensive description of the scanner data set, I refer the reader to Echenique et al. (2011).

Table 3.1: Demographic Variables

[18] Since the data on interest rates are quarterly, I use a linear interpolation to obtain monthly observations.

Variable	Number of Households
Family size:	
Midsize (3,4 members)	187
Large size (> 4 members)	65
Income:	
Mid annual income ($\in$ [\$20k, \$45k])	200
High annual income (> \$45k)	141
Age:[a]	
Middle-aged	201
Old-aged	157
Education:[b]	
High school	197
College	255
Total households	480

[a] Middle-aged households are those in which the average of the spouses' ages is between 30 and 65; in old-aged households, this average exceeds 65.
[b] If both spouses are present in a household, the average education of both spouses is reported.

3.6.2 Specification

In what follows, I restrict the range of the monthly discount factor to $[0.75, 1.0]$ and use a step size of 0.01. This support restriction is essentially without loss of generality as the resulting support of the annualized discount factor becomes approximately $[0.02, 1.0]$.[19] For ease of comparison, I report the confidence sets for the annualized discount factor.

I restrict the analysis to consumers that are e-ED rationalizable for $e \geq 0.85$. This choice is motivated by the fact that the main source of measurement error in the data is from coupons and the broader empirical evidence suggesting that almost no consumer saves more than 15% of their expenditures from sales such as price promotions (Griffith et al., 2009). It follows that a conservative standard deviation of measurement error is $\sigma = 0.15$. Lastly, I set the significance level to $\alpha = 0.05$.[20]

[19]To obtain annualized rates, I raise the monthly discount factor to the power 13. The reason being that data are aggregated to 4-week periods, hence yielding 13 time periods in a year.

[20]For all $l \in \mathcal{L}$ and $t \in \mathcal{T}$, we have $0.85\rho^*_{l,t}c_{l,t} \leq \rho_{l,t}c_{l,t} \leq 1.15\rho^*_{l,t}c_{l,t} \iff 0.85\rho_{l,t}\boldsymbol{c}_{l,t} \leq \rho_{l,t}c_{l,t}(1+\epsilon_t) \leq 1.15\rho_{l,t}c_{l,t} \iff -0.15 \leq \epsilon_t \leq 0.15$. Since $\mu = 0$, then $\sigma \leq 0.15$.

In what follows, a consumer is said to be consistent with the exponential discounting model if there exists a monthly discount factor $\delta \in [0.75, 1.0]$ and an efficiency level $e \in [0.85, 1.0]$ such that $S^C(\sigma^2, \delta, e) \leq \chi^2_{d,0.05}$, where $S^C(\sigma^2, \delta, e)$ is obtained by solving the feasible constrained optimization problem (3.5).[21]

3.6.3 Results

In this subsection, I implement my methodology in the data according to the previous specification. I find that 144 out of the 494 consumers have data sets consistent with exponential discounting. As such, the following analysis focuses on those consumers exclusively.

Figure 3.1 displays how the average constrained confidence set changes by demographic.[22] The average constrained confidence set is obtained by averaging constrained confidence sets across consumers pertaining to a same demographic. The efficiency level e is set to 0.85 such that measurement error is allowed to cause the consumer to waste up to 15% of his expenditure at the observed data.

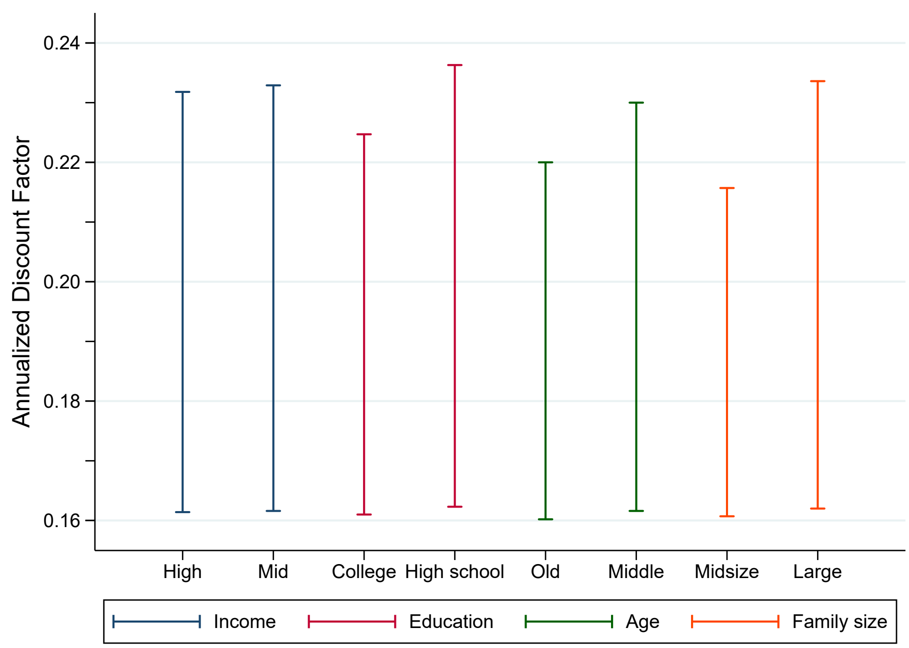

Figure 3.1: Average Constrained Confidence Set by Demographic.

[21] Let R be the number of goods that are never purchased by a consumer. Since changing the price of a good never purchased has no effect on the constraint in (3.3), the number of degrees of freedom is equal to $d = |\mathcal{T}|(L-R)$.

[22] A summary of the demographic variables is given in Table 3.1.

Overall, Figure 3.1 shows no relationship between the discount factor and demographics in the sample. Nevertheless, there may be heterogeneity that is not captured by observable characteristics. Accordingly, I compare the constrained confidence set at various quantiles of the sample in Figure 3.2, where consumers were ordered by the midpoint of their constrained confidence set.

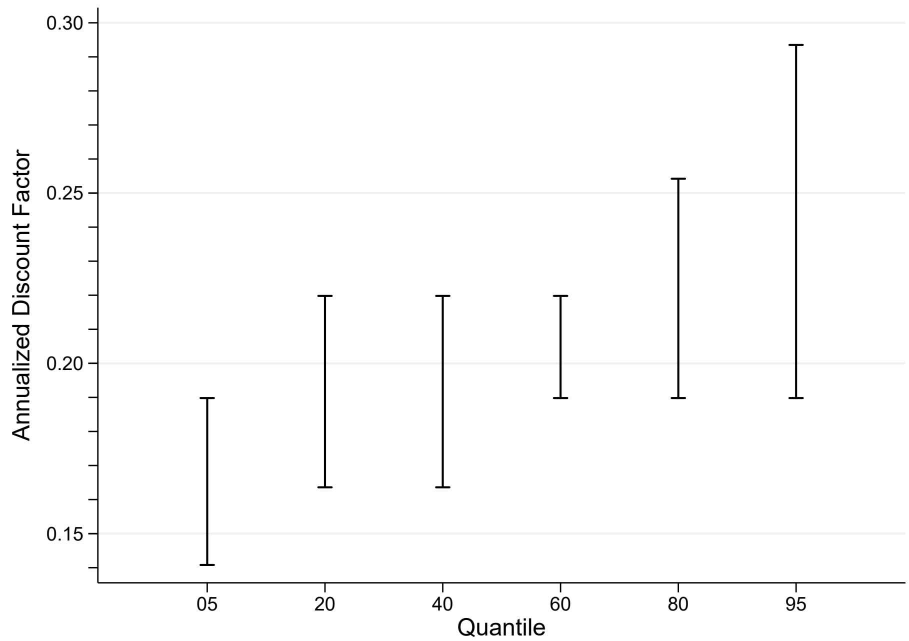

Figure 3.2: Constrained Confidence Set by Quantile.

Contrary to the previous analysis, Figure 3.2 reveals a fair degree of heterogeneity once individual unobserved heterogeneity is fully acknowledged. Furthermore, Figure 3.2 shows that informative confidence sets on the discount factor can be obtained with nonparametric preferences and the presence of measurement error in prices.

3.7 Conclusion

My results allow one to set identify individual discount factors while avoiding the misspecification of preferences. Inference can be made whether or not a data set contains exact information about the variance in measurement error. Once the discount factor is elicited, one could use the revealed preference inequalities to bound a consumer's response to changes in prices. That is, one could undertake a counterfactual analysis in a similar fashion as Blundell et al. (2003, 2008). My methodology could therefore be used to do robust welfare analysis, estimate market power in empirical industrial organization, or determine optimal pricing schemes in marketing.

The annualized discount factors displayed in my application are well below the values usually assumed in the literature. However, the average constrained confidence set is compatible with other studies using analogous data.[23] For example, Ackerberg (2003) estimates a weekly discount factor of 0.98 with scanner data on yogurt, therefore giving an annualized discount factor of 0.35.[24] While I recognize that deviations from exponential discounting may naturally arise due to imperfect measurement, an interesting extension would be to consider a random utility version of exponential discounting to account for changes in preferences.[25]

[23] My results are not directly comparable to those obtained using survey data such as in Blow et al. (2021) as the type of data differs.

[24] See Yao et al. (2012) for further estimates from field data. More generally, see Frederick et al. (2002) for a comprehensive review of the literature.

[25] A general framework to tackle measurement error in utility maximization models is provided by Aguiar and Kashaev (2021).

Chapter 4

Nonparametric Analysis of Dynamic Discrete Random Utility Models

4.1 Introduction

This paper builds on the novel characterization of the exponential discounting model derived in the previous chapter and the revealed price preference framework developed by Deb et al. (2018).

Deb et al. (2018) developed a framework that allows one to make welfare analysis without imposing a quasilinear structure on the consumer preferences. They noted that for a given set of goods, lower prices allow consumers to purchase more of the other (unobserved) goods. Thus, a consumer is better off when his expenditure is smaller. That is, if one observes $\boldsymbol{\rho}_{t_1}' \boldsymbol{c}_{t_2} \leq \boldsymbol{\rho}_{t_2}' \boldsymbol{c}_{t_2}$, then one can infer that the consumer has a preference for prices $\boldsymbol{\rho}_{t_1}$ over prices $\boldsymbol{\rho}_{t_2}$. For welfare comparisons to be meaningful, the previous relation must be acyclic, a condition they dubbed the Generalized Axiom of Price Preference (GAPP). They then show that a consumer satisfies GAPP if and only if he can be thought of as maximizing an expenditure-augmented utility function.

This paper shows that an exponential discounter maximizes a *nonseparable* augmented utility function if and only if he satisfies a slight generalization of GAPP. I then show how these restrictions can be cast into restrictions compatible with the statistical framework of Kitamura and Stoye (2018) developed for random utility models in cross-sectional data.[1]

In related work, Kashaev and Aguiar (2022) extend the approach of Kitamura and Stoye (2018) to longitudinal data which allows them to capture a rich set of time-varying behaviors. My characterization of exponential discounting could also be applied in their framework to analyze dynamic discrete random utility models in panel data. Allen and Rehbeck (2020b) provide a statistical test for a population of approximate quasilinear consumers. The current test generalizes theirs in that exponential discounting includes quasilinearity as a special case while differing in that it uses the distribution of demand rather than just average demand.

4.2 Deterministic Model

In an environment with a time dimension, discounting is essential to make meaningful comparisons. For example, the present value of \$$x$ in t periods is given by:

$$PV = \frac{x}{(1+r)^t},$$

where $r > 0$ is the interest rate. Since the consumer values \$$x$ today more than a year from now, he would choose the former if given the choice. In a similar fashion, a consumer compares the cost of a good at different points in time by appropriately scaling its cost. In particular, exponential discounting means that the cost of a good increases exponentially over time. Thus, the cost of a good at time $t \in \mathcal{T}$ is given by:

$$\boldsymbol{\rho}_t^{pc} := \delta^{-t}\boldsymbol{\rho}_t,$$

where $\delta \in [\underline{\delta}, 1]$ denote the discount factor and $\underline{\delta} \in (0, 1]$. Suppose now that the consumer wants to purchase a bundle $\boldsymbol{c}$. Is there a time at which the consumer would prefer to purchase it? By comparing the cost of the bundle at various periods, the consumer would prefer to

[1] Im and Rehbeck (2021) show that stochastic rationalizability does not imply individual rationality. Hence, welfare predictions may be imperfect at this level of aggregation.

purchase $\boldsymbol{c}$ at period t_1 rather than period t_2 if

$$\delta^{-t_1}\boldsymbol{\rho}_{t_1}\boldsymbol{c} \leq \delta^{-t_2}\boldsymbol{\rho}_{t_2}\boldsymbol{c}.$$

That is, the consumer prefers to purchase $\boldsymbol{c}$ in the period where it is cheapest. Suppose that a consumer were to purchase $\boldsymbol{c}_{t_1}$ in period t_1 and $\boldsymbol{c}_{t_2}$ in period t_2. Further, suppose that

$$\begin{aligned}\delta^{-t_1}\boldsymbol{\rho}_{t_1}\boldsymbol{c}_{t_2} &\leq \delta^{-t_2}\boldsymbol{\rho}_{t_2}\boldsymbol{c}_{t_2}\\ \delta^{-t_2}\boldsymbol{\rho}_{t_2}\boldsymbol{c}_{t_1} &< \delta^{-t_1}\boldsymbol{\rho}_{t_1}\boldsymbol{c}_{t_1},\end{aligned}$$

such that the consumer prefers $\boldsymbol{c}_{t_2}$ in period t_1 and strictly prefers $\boldsymbol{c}_{t_1}$ in period t_2. Then, it is transpiring that the choices of the consumer are suboptimal. Namely, the consumer could have switched the timing of consumption of $\boldsymbol{c}_{t_1}$ and $\boldsymbol{c}_{t_2}$ and be better off. Thus, a consumer that behaves optimally cannot have cyclic preferences.

Example Suppose there are two time periods $\mathcal{T} = \{1,2\}$ and two goods $\mathcal{L} = \{1,2\}$. Further, suppose consumption is $\boldsymbol{c}_1 = \boldsymbol{c}_2 = [1,1]'$ and prices are $\boldsymbol{\rho}_1 = \boldsymbol{\rho}_2 = [2,2]'$. The consumer prefers purchasing $\boldsymbol{c}_2$ at 1 rather than 2 since $\delta^{-1}\boldsymbol{\rho}_1'\boldsymbol{c}_2 \leq \delta^{-2}\boldsymbol{\rho}_2'\boldsymbol{c}_2$. In fact, the consumer would rather get $\boldsymbol{c}_2$ in the first period only because he is impatient, i.e. $\delta \in [\underline{\delta}, 1]$. For the same reason, the consumer does not prefer purchasing $\boldsymbol{c}_1$ in the second period since $\delta^{-2}\boldsymbol{\rho}_2'\boldsymbol{c}_1 \geq \delta^{-1}\boldsymbol{\rho}_1'\boldsymbol{c}_2$.

Formally, $\boldsymbol{\rho}_{t_i}^{pc}$ is said to be *directly revealed preferred* to $\boldsymbol{\rho}_{t_j}^{pc}$ if $\delta^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_j} \leq \delta^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}$. Let R^D denote the direct revealed preference relation and let R denote its transitive closure.[2] When the inequality is strict, $\boldsymbol{\rho}_{t_i}^{pc}$ is said to be *directly revealed strictly preferred* to $\boldsymbol{\rho}_{t_j}^{pc}$ and is denoted P^D. In the case where there is a sequence $\boldsymbol{\rho}_{t_1}^{pc}R^D\boldsymbol{\rho}_{t_2}^{pc}, \boldsymbol{\rho}_{t_2}^{pc}R^D\boldsymbol{\rho}_{t_3}^{pc}, \ldots, \boldsymbol{\rho}_{t_{m-1}}^{pc}R^D\boldsymbol{\rho}_{t_m}^{pc}$ of directly revealed preferences, $\boldsymbol{\rho}_{t_1}^{pc}$ is said to be *revealed preferred* to $\boldsymbol{\rho}_{t_m}^{pc}$. Naturally, if any of those preference relations is strict, then $\boldsymbol{\rho}_{t_1}^{pc}$ is said to be *revealed strictly preferred* to $\boldsymbol{\rho}_{t_m}^{pc}$. I can now define an axiom for exponential discounting that rules out cyclic preferences.

Definition 4.1. *A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfies GAPP(δ) if for all $t_i, t_j \in \mathcal{T}$, there exists a discount factor $\delta \in [\underline{\delta}, 1]$ such that $\delta^{-t_i}\boldsymbol{\rho}_i R \delta^{-t_j}\boldsymbol{\rho}_{t_j}$ implies not $\delta^{-t_j}\boldsymbol{\rho}_{t_j}P^D\delta^{-t_i}\boldsymbol{\rho}_{t_i}$.*

[2] The transitive closure R of a relation R^D is the smallest relation containing R^D satisfying transitivity.

It is useful to note that GAPP(δ) is a straightforward generalization of GAPP in which time preferences are introduced.

Define an augmented utility function as a function $U : \mathbb{R}_+^L \times \mathbb{R}_- \to \mathbb{R}$ that is continuous, strictly increasing, and concave. The consumer picks $\boldsymbol{c}_t$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in C \subset \mathbb{R}_+^L$

$$U(\boldsymbol{c}_t, -\delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t) \geq U(\boldsymbol{c}, -\delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}).$$

The interpretation is that the consumer dislikes expenditure as it removes money that could be used for future consumption. In particular, minus expenditure may be viewed as the outside good as the consumer always has the option to save up money and spend it in the future. Note that this problem is a slight generalization of Definition 3.2 where consumption and expenditure are nonseparable and $e = 1$.

The next theorem shows that this utility maximization problem is equivalent to GAPP(δ).

Theorem 4.1. *For a given data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$, the following are equivalent:*

(i) The data are rationalized by a continuous, strictly increasing, and concave time-dependent augmented utility function.

(ii) There exists $\delta \in [\underline{\delta}, 1]$ such that the data satisfy GAPP(δ).

This result states that to check whether a data set is consistent with exponential discounting, one only has to verify that the revealed preference relation is acyclic for some $\delta \in [\underline{\delta}, 1]$. Conditional on the discount factor, the latter is a well-known problem and existing algorithms for computing the transitive closure can be used directly.

The methodology proposed in this paper may be applied in any finite discrete consumption space. Some special cases are of particular interest due to their wide popularity in the empirical industrial organization literature.

Example Suppose products have L characteristics taking value of at most x such that the consumption space is $C = \{0, 1, \ldots, x\}^L$.

Example Suppose the consumer can choose at most 1 unit of at most one good such that the consumption space is $C = \{\mathbf{0}_{-l}\}_{l=1}^L$, where $\mathbf{0}_{-l}$ is a vector of zeros with its lth element equal to 1.

In either of the above examples, GAPP(δ) provides a necessary and sufficient condition that the consumer choices must satisfy to be an exponential discounter.

Remark It is informative to realize that since GAPP(δ) is necessary for the maximization of a nonseparable augmented utility function, it also is for the maximization of a separable utility function. Theorem 1 of the previous chapter therefore implies that GAPP(δ) is necessary for a consumer to be a standard exponential discounter as defined in Definition 3.1.

4.3 Stochastic Rationalizability

This section builds on the individual analysis of exponential discounting behavior to define stochastic rationalizability in a population of exponential discounters.

4.3.1 Definition

Conditional on a discount factor $\delta \in [\underline{\delta}, 1]$, the consumer picks $\boldsymbol{c}_t$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in C$

$$U(\boldsymbol{c}_t, -\delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t) \geq U(\boldsymbol{c}, -\delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}).$$

Let the consumption space $C \subset \mathbb{N}^L$ be finite.[3] Then, there is only a finite number of options $\mathcal{I}_t$ that the consumer can choose from in any given time period $t \in \mathcal{T}$. As such, there is also only finitely many choice profiles defined over the T choice situations. Let r index each such profile and denote a choice profile by $\boldsymbol{a}^r = (a^r_{1,1}, a^r_{1,2}, \ldots, a^r_{T,I_T})$, where $a^r_{t,i} = 1$ if option $\boldsymbol{c}_i$ is chosen at t and $a^r_{t,i} = 0$ otherwise.[4] The set of rational choice profiles $\mathcal{R}$ is the set of all profiles r for which there exists a utility function U^r and a discount factor $\delta \in [\underline{\delta}, 1]$ such that for all $t \in \mathcal{T}$

$$a^r_{t,i} = 1 \text{ if and only if } \boldsymbol{c}_i \in \arg\max_{c \in C} U^r(\boldsymbol{c}, -\delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}).$$

Let $P_{\mathcal{R}}$ be a probability distribution over all rational choice profiles, and p_r be the probability of a given profile. Define the set $\mathcal{R}_{t,i}$ as the subset of $\mathcal{R}$ such that $r \in \mathcal{R}_{t,i}$ if and only if

[3]One could allow for a continuous consumption space as it is always possible to transform the problem into a discrete one (Deb et al., 2018).

[4]The vector $\boldsymbol{a}^r$ is another way to represent a consumption stream. The distinction is superficial in this paper but nevertheless introduced to be consistent with the literature.

$a^r_{t,i} = 1$. In other words, $\mathcal{R}_{t,i}$ is the set of rational choice profiles that choose $\boldsymbol{c}_i$ at time t.

Suppose that we observe the choices made by a population of consumers and let $\pi_{t,i}$ denote the probability that option i is chosen at time t. Stochastic rationalizability requires that there exists a probability distribution $P_{\mathcal{R}}$ such that, summed over all rational choice types r, the probability of choosing option $\boldsymbol{c}_i$ at t equals $\pi_{t,i}$. Letting $\boldsymbol{\pi} = (\pi_{1,1}, \pi_{1,2}, \ldots, \pi_{T,I_T})$ denote the vector of choice probabilities, the following definition formalizes the notion of stochastic rationalizability.

Definition 4.2. *The choice probabilities $\boldsymbol{\pi}$ are stochastically rationalizable if and only if there exists a distribution $P_{\mathcal{R}}$ over rational choice profiles such that*

$$\sum_{r \in \mathcal{R}_{t,i}} p_r = \pi_{t,i} \quad \forall t \in \mathcal{T}, \forall \boldsymbol{c}_i \in C.$$

That is, a population of consumers is stochastically rationalizable if choice probabilities can be viewed as stemming from some combination of rational choice profiles.

4.3.2 Rational Choice Profiles

The previous section defined what it means for a population of consumers to be stochastically rationalizable. Namely, one simply needs to find a distribution over rational choice profiles that matches choice probabilities.

The main hurdle is that such endeavor assumes prior knowledge of the set of all rational choice profiles $\mathcal{R}$. Theorem 4.1 tells us that the set of all rational choice types $\mathcal{R}$ can be found by finding every vector $\boldsymbol{a}^r$ that satisfies GAPP(δ) for some $\delta \in [\underline{\delta}, 1]$. Thus, recovering $\mathcal{R}$ in its entirety would involve checking each choice profile for every possible value of the discount factor. The limitation of this approach is that the discount factor is a continuous variable.

The next result solves this complication by providing a different characterization of a rational choice profile that only involves finitely many computations. Let $(\boldsymbol{c}^r_t)_{t \in \mathcal{T}}$ be consumption associated with the choice profile $\boldsymbol{a}^r$.

Theorem 4.2. *Let $(\boldsymbol{\rho}_t)_{t \in \mathcal{T}}$ be a given set of prices. A choice profile $\boldsymbol{a}^r$ is rational if and only if*

$glb^r \leq lub^r$ and $glb^r \leq 1$, where

$$lub^r := \min_{(t_i,t_j)\in T^+} \left(\frac{\boldsymbol{\rho}'_{t_j}\boldsymbol{c}^r_{t_j}}{\boldsymbol{\rho}'_{t_i}\boldsymbol{c}^r_{t_j}}\right)^{\frac{1}{t_j-t_i}}; \quad glb^r := \max_{(t_i,t_j)\in T^-} \left(\frac{\boldsymbol{\rho}'_{t_j}\boldsymbol{c}^r_{t_j}}{\boldsymbol{\rho}'_{t_i}\boldsymbol{c}^r_{t_j}}\right)^{\frac{1}{t_j-t_i}},$$

$(t_i, t_j) \in T^+$ if $\delta^{-t_i}\boldsymbol{\rho}_{t_i} P \delta^{-t_j}\boldsymbol{\rho}_{t_j}$, $(t_i, t_j) \in T^-$ if not $\delta^{-t_i}\boldsymbol{\rho}_{t_i} P \delta^{-t_j}\boldsymbol{\rho}_{t_j}$, $T^+ \cap T^- = \emptyset$, and $lub = 1$ if $T^+ = \emptyset$.

Theorem 4.2 shows that every rational choice profile $r \in \mathcal{R}$ is characterized by simple inequalities that can be viewed as logical restrictions on the set of discount factors. They require the greatest lower bound on the discount factor to be smaller than the least upper bound and the greatest lower bound to be smaller than one.

Conditional on prices and a choice profile $\boldsymbol{a}^r$, the bounds are deduced from the revealed preferences. In particular, the revealed preferences must be acyclic ($T^+ \cap T^- = \emptyset$) as any rational choice profile satisfies GAPP(δ) for some $\delta \in [\underline{\delta}, 1]$. Since the revealed preferences are unknown, the inequalities cannot be checked directly. However, there are only finitely many acyclic revealed preferences. Thus, one can identify the set of all acyclic revealed preferences, construct the sets T^+ and T^-, and check that the inequalities are satisfied. This approach therefore gives a feasible procedure to recover the set of all rational choice profiles.

Corollary 4.1. $\mathcal{R} = \{\boldsymbol{a}^r : glb^r \leq lub^r,\ glb^r \leq 1\}$.

Corollary 4.1 states that the set of all rational choice profiles $\mathcal{R}$ can be identified by finding every choice profile $\boldsymbol{a}^r$ that induces logically feasible bounds on the discount factor for some acyclic revealed preferences. Importantly, the set $\mathcal{R}$ can be recovered in a finite number of steps; a significant improvement compared to using Theorem 1. An efficient depth-first search algorithm to identify $\mathcal{R}$ is provided in Appendix C.

4.3.3 Example

Let the consumption space be given by $C = \{0, 1\}^2$. Suppose that prices are $\boldsymbol{\rho}_1 = [2, 3]'$ and $\boldsymbol{\rho}_2 = [3, 2]'$ in period 1 and 2, respectively. Also, let $\underline{\delta} = 0.9$ such that the support of the discount factor is $[0.9, 1]$. Table 4.1 displays the 16 choice profiles and identifies which ones are rational.[5]

Type	Choice profile	$\boldsymbol{c}_1$	$\boldsymbol{c}_2$	Rational
1	$\boldsymbol{a}^1$	(0,0)	(0,0)	Yes
2	$\boldsymbol{a}^2$	(0,0)	(1,0)	No
3	$\boldsymbol{a}^3$	(0,0)	(0,1)	Yes
4	$\boldsymbol{a}^4$	(0,0)	(1,1)	Yes
5	$\boldsymbol{a}^5$	(1,0)	(0,0)	Yes
6	$\boldsymbol{a}^6$	(1,0)	(1,0)	Yes
7	$\boldsymbol{a}^7$	(1,0)	(0,1)	Yes
8	$\boldsymbol{a}^8$	(1,0)	(1,1)	Yes
9	$\boldsymbol{a}^9$	(0,1)	(0,0)	No
10	$\boldsymbol{a}^{10}$	(0,1)	(1,0)	No
11	$\boldsymbol{a}^{11}$	(0,1)	(0,1)	Yes
12	$\boldsymbol{a}^{12}$	(0,1)	(1,1)	Yes
13	$\boldsymbol{a}^{13}$	(1,1)	(0,0)	Yes
14	$\boldsymbol{a}^{14}$	(1,1)	(1,0)	Yes
15	$\boldsymbol{a}^{15}$	(1,1)	(0,1)	Yes
16	$\boldsymbol{a}^{16}$	(1,1)	(1,1)	Yes

Table 4.1: Choice Profiles.

To see that Type 2 is not rational, note that prices in period 2 are weakly revealed preferred to those in period 1 since $0 = \delta^{-2}\boldsymbol{\rho}_2'\boldsymbol{c}_1 \leq \delta^{-1}\boldsymbol{\rho}_1'\boldsymbol{c}_1 = 0$ and prices in period 1 are strictly revealed preferred to those in period 2 for any discount factor $\delta \in [0.9, 1]$ since $\delta^{-1}\boldsymbol{\rho}_1'\boldsymbol{c}_2 < \delta^{-2}\boldsymbol{\rho}_2'\boldsymbol{c}_2 \iff \delta < \frac{\rho_2' c_2}{\rho_1' c_2} = 3/2$, which is always the case. Similar calculations can be done to derive the entire last column of Table 4.1.

4.3.4 Statistical Test

The statistical test of this section was proposed by Kitamura and Stoye (2018) for static random utility models. Theorem 4.1 and 4.2 of this paper allow me to directly use their methodology.

Let $\hat{\boldsymbol{\pi}}$ be an empirical estimate for the choice probabilities $\boldsymbol{\pi}$. Kitamura and Stoye (2018)

[5]This way of labeling and summarizing deterministic types given a set of budgets follows Im and Rehbeck (2021).

propose to use the test statistic J_N given by:

$$\underset{p_r,\, s_{t,i}}{\text{Minimize}} \qquad J_N = N \sum_{t=1}^{T} \sum_{i=1}^{I_t} s_{t,i}^2 \tag{4.1}$$

subject to

$$\sum_{r \in \mathcal{R}_{t,i}} p_r + s_{t,i} = \hat{\pi}_{t,i} \qquad \forall t \in \mathcal{T}, \forall \boldsymbol{c}_i \in C \tag{4.2}$$

$$p_r \geq 0 \qquad \forall r \in \mathcal{R}. \tag{4.3}$$

For each period t and every choice $\boldsymbol{c}_i$, this problem finds the closest distance $s_{t,i}$ between a linear combination of the rational choice types and the estimated choice probability $\hat{\pi}_{t,i}$. In particular, note that $J_N = 0$ if and only if the $\hat{\boldsymbol{\pi}}$ is stochastically rationalizable.

Kitamura and Stoye (2018) propose to compute the critical value used to determine if the test statistic is rejected via a bootstrap procedure. I refer the reader to Kitamura and Stoye (2018) for the details and properties of this approach. I also refer the reader to Smeulders et al. (2021) for an efficient implementation through a column generation method.

4.4 Conclusion

This note characterizes the empirical implications of nonparametric dynamic discrete random utility models and cast them into a finite set of restrictions compatible with the statistical methodology of Kitamura and Stoye (2018). An efficient algorithm to compute the set of all rational choice types is provided. The approach can be used for statistical testing and robust welfare analysis in discrete consumption spaces such as those commonly used in industrial organization.

Bibliography

Ackerberg, Daniel A. 2003. "Advertising, learning, and consumer choice in experience good markets: an empirical examination." *International Economic Review*, 44 (3), 1007–1040.

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica*, 83 (6), 2411–2451.

Adams, Abi, Laurens Cherchye, Bram De Rock, and Ewout Verriest. 2014. "Consume now or later? Time inconsistency, collective choice, and revealed preference." *American Economic Review*, 104 (12), 4147–83.

Afriat, Sydney N. 1967. "The construction of utility functions from expenditure data." *International economic review*, 8 (1), 67–77.

——— 1973. "On a system of inequalities in demand analysis: an extension of the classical method." *International economic review*, 460–472.

Aguiar, Mark and Mark Bils. 2015. "Has consumption inequality mirrored income inequality?" *American Economic Review*, 105 (9), 2725–56.

Aguiar, Mark and Erik Hurst. 2007. "Life-cycle prices and production." *American Economic Review*, 97 (5), 1533–1559.

Aguiar, Victor H and Nail Kashaev. 2021. "Stochastic revealed preferences with measurement error." *The Review of Economic Studies*, 88 (4), 2042–2093.

Allen, Roy and John Rehbeck. 2020a. "Satisficing, aggregation, and quasilinear utility." *Available at SSRN 3180302*.

——— 2020b. “Satisficing, aggregation, and quasilinear utility.” *Available at SSRN 3180302*.

Anderson, Gordon and Richard Blundell. 1984. “Consumer non-durables in the UK: a dynamic demand system.” *The Economic Journal*, 94, 35–44.

Arslan, Yavuz, Bulent Guler, and Temel Taskin. 2021. “Price Search, Consumption Inequality And Expenditure Inequality Over The Life-Cycle.” *International Economic Review*, 62 (1), 295–320.

Attanasio, Orazio P and Luigi Pistaferri. 2016. “Consumption inequality.” *Journal of Economic Perspectives*, 30 (2), 3–28.

Baye, Michael R, John Morgan, Patrick Scholten et al. 2006. “Information, search, and price dispersion.” *Handbook on economics and information systems*, 1, 323–375.

Becker, Gary S. 1965. “A Theory of the Allocation of Time.” *The economic journal*, 75 (299), 493–517.

Bewley, Truman. 1977. “The permanent income hypothesis: A theoretical formulation.” *Journal of Economic Theory*, 16 (2), 252–292, http://dx.doi.org/https://doi.org/10.1016/0022-0531(77)90009-6rm https://doi.org/10.1016/0022-0531(77)90009-6.

Blow, Laura, Martin Browning, and Ian Crawford. 2021. “Non-parametric Analysis of Time-Inconsistent Preferences.” *The Review of Economic Studies*, 88 (6), 2687–2734.

Blundell, Richard, Martin Browning, and Ian Crawford. 2008. “Best nonparametric bounds on demand responses.” *Econometrica*, 76 (6), 1227–1262.

Blundell, Richard, Joel L Horowitz, and Matthias Parey. 2012. “Measuring the price responsiveness of gasoline demand: Economic shape restrictions and nonparametric demand estimation.” *Quantitative Economics*, 3 (1), 29–51.

Blundell, Richard W, Martin Browning, and Ian A Crawford. 2003. “Nonparametric Engel curves and revealed preference.” *Econometrica*, 71 (1), 205–240.

Broda, Christian, Ephraim Leibtag, and David E Weinstein. 2009. “The role of prices in measuring the poor’s living standards.” *Journal of economic Perspectives*, 23 (2), 77–97.

Brown, Donald J and Caterina Calsamiglia. 2007. "The nonparametric approach to applied welfare analysis." *Economic Theory*, 31 (1), 183–188.

Brown, Jeffrey R and Austan Goolsbee. 2002. "Does the Internet make markets more competitive? Evidence from the life insurance industry." *Journal of political economy*, 110 (3), 481–507.

Browning, Martin. 1989. "A nonparametric test of the life-cycle rational expections hypothesis." *International Economic Review*, 979–992.

Cherchye, Laurens, Thomas Demuynck, Bram De Rock, and Per Hjertstrand. 2015. "Revealed preference tests for weak separability: an integer programming approach." *Journal of econometrics*, 186 (1), 129–141.

Cherchye, Laurens, Thomas Demuynck, Bram De Rock, and Joshua Lanier. 2020. "Are consumers rational? Shifting the burden of proof." *FEB Research Report Department of Economics*.

Crawford, Ian. 2010. "Habits revealed." *The Review of Economic Studies*, 77 (4), 1382–1402.

De Clippel, Geoffroy and Kareen Rozen. 2018. "Consumer theory with misperceived tastes." *Available at SSRN 3223654*.

Deb, Rahul, Yuichi Kitamura, John K-H Quah, and Jörg Stoye. 2018. "Revealed price preference: theory and empirical analysis." *arXiv preprint arXiv:1801.02702*.

DellaVigna, Stefano and Matthew Gentzkow. 2019. "Uniform pricing in us retail chains." *The Quarterly Journal of Economics*, 134 (4), 2011–2084.

Demuynck, Thomas and Ewout Verriest. 2013. "I'll never forget my first cigarette: a revealed preference analysis of the "habits as durables" model." *International Economic Review*, 54 (2), 717–738.

Diewert, W Erwin. 1973. "Afriat and revealed preference theory." *The Review of Economic Studies*, 40 (3), 419–425.

Dziewulski, Paweł. 2020. “Just-noticeable difference as a behavioural foundation of the critical cost-efficiency index.” *Journal of Economic Theory*, 188, 105071.

Echenique, Federico, Taisuke Imai, and Kota Saito. 2020. “Testable implications of models of intertemporal choice: Exponential discounting and its generalizations.” *American Economic Journal: Microeconomics*, 12 (4), 114–43.

Echenique, Federico, Sangmok Lee, and Matthew Shum. 2011. “The money pump as a measure of revealed preference violations.” *Journal of Political Economy*, 119 (6), 1201–1223.

Echenique, Federico, Kota Saito, and Taisuke Imai. 2021. “Approximate expected utility rationalization.” *arXiv preprint arXiv:2102.06331*.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo. 2010. “Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?” *QME*, 8 (2), 207–239.

Elger, Thomas and Barry E Jones. 2008. “Can rejections of weak separability be attributed to random measurement errors in the data?” *Economics Letters*, 99 (1), 44–47.

Erdil, Erkan. 2006. “Demand systems for agricultural products in OECD countries.” *Applied Economics Letters*, 13 (3), 163–169.

Fleissig, Adrian R and Gerald A Whitney. 2008. “A nonparametric test of weak separability and consumer preferences.” *Journal of Econometrics*, 147 (2), 275–281.

Floyd, Robert W. 1962. “Algorithm 97: shortest path.” *Communications of the ACM*, 5 (6), 345.

Forges, Francoise and Enrico Minelli. 2009. “Afriat’s theorem for general budget sets.” *Journal of Economic Theory*, 144 (1), 135–145.

Fostel, Ana, Herbert Scarf E, and Michael Todd J. 2004. “Two new proofs of afriat’s theorem.” *Economic Theory*, 211–219.

Frederick, Shane, George Loewenstein, and Ted O’donoghue. 2002. “Time discounting and time preference: A critical review.” *Journal of economic literature*, 40 (2), 351–401.

Gabaix, Xavier. 2014. "A sparsity-based model of bounded rationality." *The Quarterly Journal of Economics*, 129 (4), 1661–1710.

Gandhi, Amit, Salvador Navarro, and David A Rivers. 2020. "On the identification of gross output production functions." *Journal of Political Economy*, 128 (8), 2973–3016.

Geanakoplos, John D. 2013. "Afriat from maxmin." *Economic Theory*, 54 (3), 443–448.

Griffith, Rachel, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo. 2009. "Consumer shopping behavior: how much do consumers save?" *Journal of Economic Perspectives*, 23 (2), 99–120.

Gunsilius, Florian F. 2020. "Non-testability of instrument validity under continuous treatments." *Biometrika*.

Halevy, Yoram. 2015. "Time consistency: Stationarity and time invariance." *Econometrica*, 83 (1), 335–352.

Halevy, Yoram, Dotan Persitz, and Lanny Zrill. 2018. "Parametric recoverability of preferences." *Journal of Political Economy*, 126 (4), 1558–1593.

Heufer, Jan and Per Hjertstrand. 2019. "Homothetic efficiency: Theory and applications." *Journal of Business & Economic Statistics*, 37 (2), 235–247.

Hitsch, Günter J, Ali Hortacsu, and Xiliang Lin. (2019) "Prices and promotions in us retail markets: Evidence from big data."Technical report, National Bureau of Economic Research.

Im, Changkuk and John Rehbeck. 2021. "Non-rationalizable Individuals, Stochastic Rationalizability, and Sampling." *Available at SSRN 3767994*.

Jackson, Matthew O and Leeat Yariv. 2015. "Collective dynamic choice: the necessity of time inconsistency." *American Economic Journal: Microeconomics*, 7 (4), 150–78.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect theory: An analysis of decisions under risk." 263–292.

Kaplan, Greg and Guido Menzio. 2015. "The morphology of price dispersion." *International Economic Review*, 56 (4), 1165–1206.

Kaplan, Greg, Guido Menzio, Leena Rudanko, and Nicholas Trachter. 2019. "Relative price dispersion: Evidence and theory." *American Economic Journal: Microeconomics*, 11 (3), 68–124.

Kashaev, Nail and Victor H. Aguiar. 2022. "Nonparametric Analysis of Dynamic Random Utility Models." https://arxiv.org/abs/2204.07220, 1–33.

Kitamura, Yuichi and Jörg Stoye. 2018. "Nonparametric analysis of random utility models." *Econometrica*, 86 (6), 1883–1909.

Levinsohn, James and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *The review of economic studies*, 70 (2), 317–341.

McKenzie, David, Ernesto Schargrodsky, and Guillermo Cruces. 2011. "Buying Less but Shopping More: The Use of Nonmarket Labor during a Crisis [with Comment]." *Economia*, 11 (2), 1–43.

Nevo, Aviv and Arlene Wong. 2019. "The elasticity of substitution between time and market goods: Evidence from the Great Recession." *International Economic Review*, 60 (1), 25–51.

Olley, G Steven and Ariel Pakes. 1996. "The dynamics of productivity in the telecommunications equipment industry." *Econometrica*, 64 (6), 1263–1297.

Parente, Paulo MDC and JMC Santos Silva. 2012. "A cautionary note on tests of overidentifying restrictions." *Economics Letters*, 115 (2), 314–317.

Pytka, Krzysztof. 2017. "Shopping Effort in Self-Insurance Economies."

Reny, Philip J. 2015. "A characterization of rationalizable consumer behavior." *Econometrica*, 83 (1), 175–192.

Samuelson, Paul A. 1938. "A note on the pure theory of consumer's behaviour." *Economica*, 5 (17), 61–71.

Schennach, Susanne M. 2014. “Entropic latent variable integration via simulation.” *Econometrica*, 82 (1), 345–385.

Selvanathan, Saroja and EA Selvanathan. 2006. “Consumption patterns of food, tobacco and beverages: a cross-country analysis.” *Applied Economics*, 38 (13), 1567–1584.

Smeulders, Bart, Laurens Cherchye, and Bram De Rock. 2021. “Nonparametric analysis of random utility models: computational tools for statistical testing.” *Econometrica*, 89 (1), 437–455.

Sorensen, Alan T. 2000. “Equilibrium price dispersion in retail markets for prescription drugs.” *Journal of Political Economy*, 108 (4), 833–850.

Staiger, Douglas and James H Stock. 1997. “Instrumental variables regression with weak instruments.” *Econometrica: journal of the Econometric Society*, 557–586.

Stigler, George J. 1961. “The economics of information.” *Journal of political economy*, 69 (3), 213–225.

Tipoe, Eileen. 2021. “Price inattention: A revealed preference characterisation.” *European Economic Review*, 134, 103692.

Varian, Hal R. 1982. “The nonparametric approach to demand analysis.” *Econometrica: Journal of the Econometric Society*, 945–973.

——— 1985. “Non-parametric analysis of optimizing behavior with measurement error.” *Journal of Econometrics*, 30 (1-2), 445–458.

——— 1990. “Goodness-of-fit in optimizing models.” *Journal of Econometrics*, 46 (1-2), 125–140.

Wang, Dong D, Cindy W Leung, Yanping Li, Eric L Ding, Stephanie E Chiuve, Frank B Hu, and Walter C Willett. 2014. “Trends in dietary quality among adults in the United States, 1999 through 2010.” *JAMA internal medicine*, 174 (10), 1587–1595.

Yao, Song, Carl F Mela, Jeongwen Chiang, and Yuxin Chen. 2012. “Determining consumers’ discount rates with field studies.” *Journal of Marketing Research*, 49 (6), 822–841.

Appendix A

Appendices for Chapter 2

A.1 Relationship with Models of Household Production

Although the focus of this paper is on the price function, the framework of the model is consistent with one of household production similar in spirit to that of Becker (1965). As an illustration, I extend my model to one of household production and shows that it has close ties with that of Aguiar and Hurst (2007). For ease of comparison, I consider the static version of my model.

Suppose that, in addition to spending time shopping, the household can spend time in home production denoted by $h \in \mathbb{R}_{++}$. By using that time input along with market goods, the household can produce some homemade good K by using its (concave) home production function $f(h, \boldsymbol{c})$.[1] The household's problem therefore becomes

$$\max_{(\boldsymbol{c},\boldsymbol{a},K,h)\in C\times A\times\mathbb{R}^2_{++}} u(\boldsymbol{a}, K, h) \quad s.t. \quad \boldsymbol{p}(\boldsymbol{a}, \omega_t)'\mathbf{c} = y_t$$
$$f(\boldsymbol{c}, h) = K.$$

[1]One can think of market goods as comestible such as eggs, sugar and pecans. By spending h unit of time cooking, the household can transform these "raw goods" into a pecan pie, the final good consumed by the household.

One can get rid of the second constraint by substituting it for K in the utility function, yielding

$$\max_{(\boldsymbol{c},\boldsymbol{a},h)\in C\times A\times\mathbb{R}_{++}} u(\boldsymbol{a}, f(\boldsymbol{c}, h), h) \;\; s.t. \;\; \boldsymbol{p}(\boldsymbol{a}, \omega_t)'\mathbf{c} = y_t.$$

Assuming the opportunity cost of time is additively separable, linear, and identical for the shopper and the home producer, the problem boils down to

$$\max_{(\boldsymbol{c},\boldsymbol{a},h)\in C\times A\times\mathbb{R}_{++}} u(f(\boldsymbol{c}, h)) + \boldsymbol{\mu}_t'\boldsymbol{a} + \mu_t h \;\; s.t. \;\; \boldsymbol{p}(\boldsymbol{a}, \omega_t)'\mathbf{c} = y_t,$$

where μ_t denotes the disutility from the time spent on either activity. Since both $u(\cdot)$ and $f(\cdot,\cdot)$ are unobservable concave functions, this maximization problem is observationally equivalent to

$$\max_{(\boldsymbol{c},\boldsymbol{a},h)\in C\times A\times\mathbb{R}_{++}} f(\boldsymbol{c}, h) + \boldsymbol{\mu}_t'\boldsymbol{a} + \mu_t h \;\; s.t. \;\; \boldsymbol{p}(\boldsymbol{a}, \omega_t)'\mathbf{c} = y_t,$$

and we have thereby recovered a model with the same implications to that of Aguiar and Hurst (2007).[2] To see why, assume the solution is interior and take the first-order conditions:

$$\begin{aligned} \frac{\partial f}{\partial c} &= \lambda_t \boldsymbol{p}(\boldsymbol{a}, \omega_t) \\ \boldsymbol{\mu} &= \lambda_t \frac{\partial \boldsymbol{p}(\boldsymbol{a}, \omega_t)}{\partial \boldsymbol{a}} \odot \boldsymbol{c} \\ \mu &= -\frac{\partial f}{\partial h}. \end{aligned}$$

It follows that the marginal rate of transformation (MRT) between time and goods in shopping equals the MRT in home production:

$$\frac{\partial f}{\partial h} / \frac{\partial f}{\partial c_l} = -\frac{\frac{\partial p_l(a_l, \omega_{l,t})}{\partial a_l} \cdot c_l}{p_l(a_l, \omega_{l,t})} \quad \forall l \in \mathcal{L}.$$

This derivation shows that the household production version of my model naturally extends that of Aguiar and Hurst (2007). Conditional on knowing the price function, this last equation

[2] Despite that the two maximization problems are observationally equivalent, eliminating the utility function changes the interpretation of the model.

can be used to identify the home production function, a point that was cleverly exploited by Aguiar and Hurst (2007) in a parametric setting. Finally, note that using exponential discounting as defined in (2.1) would yield $\lambda_t \delta^{-t}$ instead of λ_t in the first-order conditions and leave the MRT unchanged.

A.2 Sample Construction

The Homescan contains information on purchases made by U.S. households in a wide variety of retail outlets. After every trip to a retail outlet, information about the trip is recorded by the panelist via a scanner device. Each trip may have one or many UPC purchases. In total, there are $66,321,848$ purchases in the panel year 2011. Among them, $43,432,246$ pertain to the departments of dry grocery, frozen foods, dairy and packaged meat. Since some purchases in the panel year are outside of the calendar year 2011, I remove them from the sample. This operation drops $751,479$ purchases, leaving a total of $42,680,767$ purchases.

For each household-month, I average UPC prices across trips. Precisely, for any household $i \in \mathcal{N}$ and month $t \in \mathcal{T}$, the weighted average price for a given UPC is given by

$$\bar{p}_{i,UPC,t} = \frac{\sum_{trips_i \in t} p_{i,UPC,trips_i} c_{i,UPC,trips_i}}{\sum_{trips_i \in t} c_{i,UPC,trips_i}},$$

where $trips_i$ denotes a trip of household i. This aggregation is only computed for UPCs that are purchased by a given household in a given month.

The Homescan has a total of $4,510,908$ distinct UPCs, with $1,633,850$ that belong to the four departments considered: dry grocery, frozen foods, dairy, and packaged meat. To keep the analysis tractable and mitigate stockpiling issues, I aggregate UPCs to their department categories. For each household-month, the weighted average price for a given department $l \in \mathcal{L}$ is given by

$$p_{i,l,t} = \frac{\sum_{UPC \in l} \tilde{p}_{i,UPC,t} c_{i,UPC,t}}{\sum_{UPC \in l} c_{i,UPC,t}}.$$

Furthermore, I only keep data from April to September. The main reason for limiting the number of goods and time periods is to control the computational burden. Since the number of parameters to solve for in the model is given by $L \cdot T + T$, the nonlinear optimization problem becomes quickly intractable when either L or T increases.

As the methodology requires the data to be strictly positive, I drop households that do not meet this requirement for any aggregated good and month. These conditions bring down the number of households from 62,092 to 16,025. Further limiting the sample to single households that are at least 50 years old decreases the number of households to 1668. Finally, I drop households that have zero prices paid, thus decreasing the sample size to 1645.[3]

I restrict the sample to single households to avoid the false rejection of the model. As Adams et al. (2014) point out, inconsistencies may arise due to negotiation within a couple household. Jackson and Yariv (2015) further show that time inconsistent behavior will appear if individuals in a non-dictatorial household have different discount factors. By accounting for measurement error in survey data, Aguiar and Kashaev (2021) show that single households behave consistently with exponential discounting while couple households do not.

A.3 Power Analysis

In this section, I show that price search and utility maximization are both refutable under Assumptions 2.1-2.5. I then provide empirical evidence that these additional restrictions are not necessary for the model to be rejected by the data.

A.3.1 Convexity of the Log-linear Shopping Technology

Let the price function for any good $l \in \mathcal{L}$ have the log-linear shopping technology specified by Assumption 2.1:

$$\log(p_{l,t}(a_{l,t}, \omega_{l,t})) = \alpha_l^0 + \alpha_l^1 \log(a_{l,t}) - \omega_{l,t}.$$

It is easy to see that, for any $l \in \mathcal{L}$, the Hessian of the log price function is

$$H(a_{l,t}, \omega_{l,t}) = \begin{bmatrix} -\frac{\alpha_l^1}{a_{l,t}^2} & 0 \\ 0 & 0 \end{bmatrix}.$$

[3]Zero prices may arise because of "free-good" promotions or if the household enters a price equal to zero and no historical information regarding a valid price for the UPC is available.

The principal minors are $D_1 = -\frac{\alpha_l^1}{a_{l,t}^2} \geq 0$, $D_2 = 0$, and $D_3 = 0$. Accordingly, the log price functions are convex and, therefore, the price functions logarithmically convex.[4]

A.3.2 Falsifiability of Price Search

Suppose that Assumptions 2.1-2.5 are satisfied and let $\mathcal{L} = \{1, 2\}$, $\mathcal{T} = \{1, 2\}$. Almost surely, let observed prices be such that $\boldsymbol{p}_1 = [1, 2]'$, $\boldsymbol{p}_2 = [3, 4]'$, shopping intensity be such that $\boldsymbol{a}_1 = [1, 2]'$, $\boldsymbol{a}_2 = [2, 3]'$, and consumption be such that $\boldsymbol{c}_t > 0$ for $t = 1, 2$.

Convexity of the log price functions implies that for all $l \in \mathcal{L}$ and $s, t \in \mathcal{T}$, we have

$$\log\left(\frac{p(a_{l,s}, \omega_{l,s})}{p(a_{l,t}, \omega_{l,t})}\right) \geq \frac{\nabla_a p(a_{l,t}, \omega_{l,t})}{p(a_{l,t}, \omega_{l,t})}(a_{l,s} - a_{l,t}) + \frac{\nabla_\omega p(a_{l,t}, \omega_{l,t})}{p(a_{l,t}, \omega_{l,t})}(\omega_{l,s} - \omega_{l,t}).^5$$

The above expression can be written more concisely as

$$\log\left(\frac{p_{l,s}^*}{p_{l,t}^*}\right) \geq \frac{\rho_{l,t}}{p_{l,t}^* c_{l,t}}(a_{l,s} - a_{l,t}) - (\omega_{l,s} - \omega_{l,t}) \quad \forall s, t \in \mathcal{T}.$$

Summing up these inequalities for each good $l \in \mathcal{L}$ and dividing by L gives

$$\frac{1}{L}\sum_{l=1}^{L} \log\left(\frac{p_{l,s}^*}{p_{l,t}^*}\right) \geq \frac{1}{L}\sum_{l=1}^{L} \frac{\rho_{l,t}}{p_{l,t}^* c_{l,t}}(a_{l,s} - a_{l,t}) - (\bar{\omega}_s - \bar{\omega}_t) \quad \forall s, t \in \mathcal{T},$$

where $\bar{\omega}_t := \frac{1}{L}\sum_{l=1}^{L} \omega_{l,t}$ for all $t \in \mathcal{T}$. Taking the expectation for $s = 1$, $t = 2$ and using the assumptions that $\mathbb{E}[\log(\boldsymbol{p}_t)] = \mathbb{E}[\log(\boldsymbol{p}_t^*)]$ and $\mathbb{E}[\overline{\omega}_t] = 0$ for all $t \in \mathcal{T}$, we get

$$0 > \frac{1}{L}\sum_{l=1}^{L} \left(\mathbb{E}\left[\log(p_{l,1})\right] - \mathbb{E}\left[\log(p_{l,2})\right]\right) \geq -\frac{1}{L}\sum_{l=1}^{L} \mathbb{E}\left[\frac{\rho_{l,2}}{p_{l,2}^* c_{l,2}}\right]. \tag{A.1}$$

Noting that the random variables on the right-hand side are always negative, it follows that the negative of their expectations are positive: $-\mathbb{E}\left[\frac{\rho_{l,2}}{p_{l,2}^* c_{l,2}}\right] \geq 0$ for all $l \in \mathcal{L}$. Clearly, inequality (A.1) yields a contradiction. In other words, the data are inconsistent with the model provided the price functions are log-linear.

[4] A function f is logarithmically convex if the composition of the logarithm with f is itself a convex function.
[5] Note that this expression is well-defined since prices are strictly positive.

A.3.3 Falsifiability of Utility Maximization

Suppose that Assumptions 2.1-2.5 are satisfied and let $\mathcal{L} = \{1,2\}$, $\mathcal{T} = \{1,2\}$. Almost surely, let observed prices be such that $\boldsymbol{p}_1 = [1,2]'$, $\boldsymbol{p}_2 = [3,4]'$, shopping intensity be such that $\boldsymbol{a}_1 = [2,3]'$, $\boldsymbol{a}_2 = [1,2]'$, and consumption be such that $\boldsymbol{c}_1 = [1,1]'$, $\boldsymbol{c}_2 = [2,2]'$. Furthermore, suppose that the discount factor is such that $\delta = 1$ almost surely.

Concavity of the utility function implies that for all $s, t \in \mathcal{T}$

$$u(\boldsymbol{c}_s, \boldsymbol{a}_s) - u(\boldsymbol{c}_t, \boldsymbol{a}_t) \leq \nabla_c u(\boldsymbol{c}_t, \boldsymbol{a}_t)'(\boldsymbol{c}_s - \boldsymbol{c}_t) + \nabla_a u(\boldsymbol{c}_t, \boldsymbol{a}_t)'(\boldsymbol{a}_s - \boldsymbol{a}_t).$$

Summing up these inequalities for $s = 1$, $t = 2$ and $s = 2$, $t = 1$, we can obtain

$$0 \leq \left[(\boldsymbol{p}_2^* - \boldsymbol{p}_1^*)'(\boldsymbol{c}_1 - \boldsymbol{c}_2) + (\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1)'(\boldsymbol{a}_1 - \boldsymbol{a}_2)\right],$$

For concavity to be analytically refuted, it is clear that Assumption 2.2 needs to be changed to $\mathbb{E}[\boldsymbol{p}_t] = \mathbb{E}[\boldsymbol{p}_t^*]$ for all $t \in \mathcal{T}$. Taking the expectation then yields

$$\begin{aligned}
0 &\leq (\mathbb{E}[\boldsymbol{p}_2] - \mathbb{E}[\boldsymbol{p}_1])'(\boldsymbol{c}_1 - \boldsymbol{c}_2) + (\mathbb{E}[\boldsymbol{\rho}_2] - \mathbb{E}[\boldsymbol{\rho}_1])'(\boldsymbol{a}_1 - \boldsymbol{a}_2) \\
&= -4 + \sum_{l=1}^{L}(\mathbb{E}[\rho_{l,2}] - \mathbb{E}[\rho_{l,1}]) \\
&\leq -4 - \sum_{l=1}^{L}\mathbb{E}[\rho_{l,1}] \\
&\leq -4 + \frac{1}{2} + \frac{2}{3} \\
&< 0,
\end{aligned}$$

where the first equality substituted the expected value of true prices for their expected observed values, the second inequality used the assumption that $\boldsymbol{\rho}_t \leq 0$ for all $t \in \mathcal{T}$, and the third inequality exploited the fact that $\alpha_l \in [-1,0]$ for all $l \in \mathcal{L}$. Indeed, the latter allows us to obtain the support of $\boldsymbol{\rho}_1$ since $\rho_{l,1} = \alpha_{l,1} \cdot \frac{p_{l,1}c_{l,1}}{a_{l,1}}$ for all $l \in \mathcal{L}$. Picking $\alpha_{l,1} = -1$ yields the third inequality.

Clearly, the previous inequalities yield a contradiction. As such, utility maximization can be rejected by the data under Assumptions 2.1-2.5 provided $\mathbb{E}[\boldsymbol{p}_t] = \mathbb{E}[\boldsymbol{p}_t^*]$ for all $t \in \mathcal{T}$ (instead of Assumption 2.2) and $\delta = 1$ almost surely.

A.3.4 Falsifiability of the Model: Empirical Evidence

I have shown analytically that the model defined by Assumptions 2.1-2.5 can be rejected by the data with only two time periods if either (1) the price functions are log-linear, or (2) the discount factor equals one almost surely and measurement error satisfies $\mathbb{E}[\boldsymbol{p}_t] = \mathbb{E}[\boldsymbol{p}_t^*]$ for all $t \in \mathcal{T}$.

To complement the above analysis, I now provide empirical evidence that the model can be rejected by the data under Assumptions 2.2-2.5 if the price functions are convex decreasing.[6] This corresponds to the fully nonparametric version of the model. To this end, I consider a data set where $\boldsymbol{p}_1 = [1, 2]'$, $\boldsymbol{p}_2 = [3, 4]$, $\boldsymbol{a}_1 = [1, 2]'$, $\boldsymbol{a}_2 = [2, 3]'$, and $\boldsymbol{c}_1 = [1, 4]'$, $\boldsymbol{c}_2 = [3, 2]'$. I let the sample size be 500 where, for simplicity, every consumer is assumed to have the same data set.

The results derived previously do not allow me to conclude that the model has any empirical content without the log-linearity of the price functions. Nevertheless, an application of the methodology to the constructed data set yields a test statistic of 476.98, well-above the chi-square critical value of 12.59.

A.4 Implementation

In this section, I provide a pseudo-algorithm of the ELVIS approach proposed by Schennach (2014) specialized to my model. Furthermore, I provide pseudo-algorithms for the construction of the conditional distribution $\tilde{\eta}$ and the integration of the latent variables.

A.4.1 Pseudo-Code

Step 1

[6] Formally, a function $f : \mathbb{R}^L \to \mathbb{R}$ is convex if and only if $f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla_y' f(\boldsymbol{y})(\boldsymbol{x} - \boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^L$. It is convex decreasing if it is convex and $\nabla f(\boldsymbol{y}) \leq 0$ for all $\boldsymbol{y}$.

- Fix the number of goods L and the number of time periods T.
- Fix the data set $x = (x_i)_{i=1}^N$, where $x_i = (\boldsymbol{p}_{i,t}, \boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{t \in T}$.
- Fix the moments defining the model: $\boldsymbol{g}_i^u$, $\boldsymbol{g}_i^p$, $\boldsymbol{g}_i^m$, $\boldsymbol{g}^\omega$.
- Fix the support of the structural parameters: $\delta_i \in [\underline{\delta}, 1]$ and $\boldsymbol{\alpha}_i^1 \in [-1, 0]$.
- Fix the conditional distribution of the latent variables $\tilde{\eta}$.

Step 2

for $i = 1 : N$

- Integrate the latent variables under $\tilde{\eta}(\cdot|x_i)$ to obtain $\tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma})$.

end

- Compute $\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma}) = \frac{1}{N} \sum_{i=1}^N \tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma})$.
- Compute $\hat{\tilde{\boldsymbol{\Omega}}}(\boldsymbol{\gamma}) = \frac{1}{N} \sum_{i=1}^N \tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma})\tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma})' - \hat{\tilde{\boldsymbol{h}}}_i(\boldsymbol{\gamma})\hat{\tilde{\boldsymbol{h}}}_i(\boldsymbol{\gamma})'$.
- Compute the objective function: $\text{ObjFct}(\boldsymbol{\gamma}) = N\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma})'\hat{\tilde{\boldsymbol{\Omega}}}(\boldsymbol{\gamma})^{-}\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma})$.

Step 3

- Compute $\text{TS}_N = \min_\gamma \text{ObjFct}(\boldsymbol{\gamma})$.

Step 1 (Construction of $\tilde{\eta}$)

The distribution $\tilde{\eta}$ can be taken to be proportional to a normal distribution:

$$d\tilde{\eta}(\cdot|x_i) \propto \exp(-||\boldsymbol{g}_i^{m,\omega}(x_i, e_i)||^2),$$

where $\boldsymbol{g}_i^{m,\omega}$ is the set of moments on measurement error and search productivity. The following pseudo-code details how to construct the conditional distribution by using rejection sampling and applying Metropolis-Hastings on each passing draw. I draw true prices instead of measurement error as it ensures true prices to be strictly positive. Let $R > 0$.

while $r \leq R$

- Draw candidate latent variables $e_i^c = (\delta_i, \boldsymbol{p}_{i,t}^*, \boldsymbol{\alpha}_i^1, \boldsymbol{\omega}_{i,t})_{t\in T}$ such that their support constraints are satisfied.
- Given x_i and e_i^c, check whether the model is satisfied by using Theorem 2.1. If the model is not satisfied, go a step back.
- Draw ζ from $U[0,1]$
- If $-\left(\|\boldsymbol{g}_i^{m,\omega}(x_i, e_i^c)\|^2 - \|\boldsymbol{g}_i^{m,\omega}(x_i, e_i^{r-1})\|^2\right) > \log(\zeta)$, set e_i^r to e_i^c. Else, set e_i^r to e_i^{r-1}.
- Set $r = r + 1$

end

Step 2 (Latent Variable Integration)

- Fix x_i, $\tilde{\eta}$, and $\boldsymbol{\gamma}$.
- Set $\tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma}) = 0$

while $r \leq R$

- Draw e_i^c proportional to $\tilde{\eta}(\cdot|x_i)$.
- Draw ζ from $U[0,1]$
- If $\left[\boldsymbol{g}_i^{m,\omega}(x_i, e_i^c) - \boldsymbol{g}_i^{m,\omega}(x_i, e_i^{r-1})\right]' \boldsymbol{\gamma} > \log(\zeta)$, set e_i^r to e_i^c. Else, set e_i^r to e_i^{r-1}.
- Compute $\tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma}) = \tilde{\boldsymbol{h}}_i(x_i, \boldsymbol{\gamma}) + \boldsymbol{g}_i^{m,\omega}(x_i, e_i^r)/R$
- Set $r = r + 1$

end

A.5 Instrumental Variable

Following Aguiar and Hurst (2007), I first compute the elasticity of price with respect to shopping intensity using a log-linear regression. That is, for any good $l \in \mathcal{L}$, let the regression equation be given by:

$$\log(p_{i,l,t}) = \alpha_l^0 + \alpha_l^1 \log(a_{i,l,t}) + \alpha_l^2 \log(exp_{i,l,t}) + \epsilon_{i,l,t},$$

where $exp_{i,l,t}$ represents expenditure and $\epsilon_{i,l,t}$ represents the error term.[7] I include expenditure in the regression to control for shopping needs. The elasticities of price with respect to shopping intensity estimated with income as an instrument are presented in Table A.1.

Table A.1: Elasticity of Price With Respect to Shopping Intensity

	Dry Grocery	Frozen foods	Dairy	Packaged Meat
Elasticity (α_l^1)	-1.025	-1.088	-0.847	-1.407
	(0.043)	(0.105)	(0.140)	(0.101)
Instrument	Income	Income	Income	Income
Observations	9870	9870	9870	9870

Notes: The instrument set consists of three income categories. Standard errors are reported in parenthesis.

The relevance of the instrument can be checked by the first stage F-statistics which are 119.19, 15.81, 10.93, and 27.17, respectively. According to the rule of thumb suggested by Staiger and Stock (1997), the instrument is not weak as each F-statistic is above 10.

To get the average elasticity of average price with respect to shopping intensity, Proposition 2.1 states that one has to average the elasticities estimated in Table A.1 and divide them by L, yielding $L^{-1}\overline{\alpha}^1 = -0.273$. This shows that IV gives a downward-biased estimate of the average expected elasticity of average price with respect to shopping intensity in my data.

Although this negative result could be due to a violation of the exogeneity condition, another plausible problem is the presence of nonclassical measurement error in prices. Indeed, Einav et al. (2010) show that a linear regression with the price variable on the left-hand side is not robust to measurement error in prices in the Nielsen Homescan Dataset. This can only be

[7] Any variable other than shopping intensity that could enter in the regression equation is absorbed by search ability in the model.

the case if measurement error is nonclassical. Thus, instrumental variable may be inconsistent even if exogenous.

A.6 Proofs

To reduce the notational burden, I remove the subscript i from the variables in the rest of this section.

A.6.1 Proof of Theorem 2.1

$(i) \implies (ii)$

Suppose the data have been generated by (2.1) where the utility function is continuous, concave, strictly increasing in consumption and decreasing in shopping intensity. Then, the first-order conditions of the consumer's problem are given by

$$\nabla_c u(\boldsymbol{c}_t, \boldsymbol{a}_t) = \lambda_t \delta^{-t} \boldsymbol{p}_t,$$
$$\nabla_a u(\boldsymbol{c}_t, \boldsymbol{a}_t) = \lambda_t \delta^{-t} \frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t.$$

Since the utility function is assumed strictly increasing in consumption, it must be that $\nabla_c u(\boldsymbol{c}, \boldsymbol{a}) = \lambda_t \delta^{-t} \boldsymbol{p}_t > 0$. Accordingly, it follows that $\boldsymbol{p}_t > 0$. Likewise, the assumption that the utility function is decreasing in shopping intensity entails $\nabla_a u(\boldsymbol{c}_t, \boldsymbol{a}_t) = \lambda_t \delta^{-t} \frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t \leq 0$ and, hence, $\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t \leq 0$. Finally, concavity of the utility function implies

$$u(\boldsymbol{c}_s, \boldsymbol{a}_s) - u(\boldsymbol{c}_t, \boldsymbol{a}_t) \leq \nabla_c u(\boldsymbol{c}_t, \boldsymbol{a}_t)'(\boldsymbol{c}_s - \boldsymbol{c}_t) + \nabla_a u(\boldsymbol{c}_t, \boldsymbol{a}_t)'(\boldsymbol{a}_s - \boldsymbol{a}_t) \quad \forall s, t \in \mathcal{T}.$$

Combining the first-order conditions with concavity of the utility function and letting $u_t := u(\boldsymbol{c}_t, \boldsymbol{a}_t)$ for all $t \in \mathcal{T}$ yields

$$u_s - u_t \leq \lambda_t \delta^{-t} \left[\boldsymbol{p}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t \right)' (\boldsymbol{a}_s - \boldsymbol{a}_t) \right] \quad \forall s, t \in \mathcal{T},$$

where $\frac{\partial p_l(a_{l,t},\omega_{l,t})}{\partial a_{l,t}} = \alpha_l^1 a_{l,t}^{\alpha_l^1-1} e^{-(\omega_{l,t}-\alpha_l^0)}$ for all $l \in \mathcal{L}$ due to Assumption 2.1.

$(ii) \implies (i)$

Starting from the first set of inequalities in Theorem 2.1 (ii), note that

$$\boldsymbol{\rho}_t = \frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t,$$

such that

$$u_s \leq u_t + \lambda_t \delta^{-t} \left[\boldsymbol{p}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t \right)' (\boldsymbol{a}_s - \boldsymbol{a}_t) \right] \quad \forall s, t \in \mathcal{T}.$$

Fix some $t \in \mathcal{T}$ and let $t_1 := t$. Consider any sequence of finite indices $\tau = \{t_i\}_{i=1}^m$, $m \geq 2$, $t_i \in \mathcal{T}$. Let $\mathcal{I}$ be the set of all such indices and define

$$\begin{aligned} u(\boldsymbol{c}, \boldsymbol{a}) = \min_{\tau \in \mathcal{I}} \Bigg\{ & \lambda_{t_m} \delta^{-t_m} \left[\boldsymbol{p}_{t_m}'(\boldsymbol{c} - \boldsymbol{c}_{t_m}) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_{t_m}, \omega_{t_m})}{\partial \boldsymbol{a}_{t_m}} \odot \boldsymbol{c}_{t_m} \right)' (\boldsymbol{a} - \boldsymbol{a}_{t_m}) \right] \\ & + \sum_{i=1}^{m-1} \lambda_{t_i} \delta^{-t_i} \left[\boldsymbol{p}_{t_i}'(\boldsymbol{c}_{t_{i+1}} - \boldsymbol{c}_{t_i}) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_{t_i}, \omega_{t_i})}{\partial \boldsymbol{a}_{t_i}} \odot \boldsymbol{c}_{t_i} \right)' (\boldsymbol{a}_{t_{i+1}} - \boldsymbol{a}_{t_i}) \right] \Bigg\}. \end{aligned}$$

This function is the pointwise minimum of a collection of linear functions in $(\boldsymbol{c}, \boldsymbol{a})$. As such, $u(\boldsymbol{c}, \boldsymbol{a})$ is concave and continuous. Moreover, the second set of inequalities in Theorem 2.1 (ii) guarantees that the utility function is strictly increasing in consumption. Likewise, the third set of inequalities implies that it is decreasing in shopping intensity. Finally, note that α_l^0, α_l^1, and $(\omega_{l,t})_{t \in \mathcal{T}}$ directly identify the log-linear price function for good $l \in \mathcal{L}$.

If the budget sets $\{\mathcal{B}_t\}_{t=1}^T$ are convex, then the first-order conditions of the model are necessary and sufficient for a maximum. Therefore, I am left to show that $\lambda_t \delta^{-t} \boldsymbol{p}_t \in \nabla_c u(\boldsymbol{c}_t, \boldsymbol{a}_t)$ and $\lambda_t \delta^{-t} \frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \omega_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t \in \nabla_a u(\boldsymbol{c}_t, \boldsymbol{a}_t)$ for all $t \in \mathcal{T}$.

Let $\epsilon > 0$, $t \in \mathcal{T}$, and note that by definition of $u(\cdot, \cdot)$, there is some sequence of indices

$\tau \in \mathcal{I}$ such that

$$
\begin{aligned}
&u(\boldsymbol{c}_t, \boldsymbol{a}_t) + \epsilon > \lambda_{t_m}\delta^{-t_m}\left[\boldsymbol{p}'_{t_m}(\boldsymbol{c}_t - \boldsymbol{c}_{t_m}) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_{t_m}, \boldsymbol{\omega}_{t_m})}{\partial \boldsymbol{a}_{t_m}} \odot \boldsymbol{c}_{t_m}\right)'(\boldsymbol{a}_t - \boldsymbol{a}_{t_m})\right] \\
&+ \sum_{i=1}^{m-1} \lambda_{t_i}\delta^{-t_i}\left[\boldsymbol{p}'_{t_i}(\boldsymbol{c}_{t_{i+1}} - \boldsymbol{c}_{t_i}) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_{t_i}, \boldsymbol{\omega}_{t_i})}{\partial \boldsymbol{a}_{t_i}} \odot \boldsymbol{c}_{t_i}\right)'(\boldsymbol{a}_{t_{i+1}} - \boldsymbol{a}_{t_i})\right] \\
&\geq u(\boldsymbol{c}_t, \boldsymbol{a}_t).
\end{aligned}
$$

Add any bundle $(\boldsymbol{c}, \boldsymbol{a}) \in C \times A$ to the sequence and use the definition of $u(\cdot, \cdot)$ once again to obtain

$$
\begin{aligned}
&\lambda_{t_m}\delta^{-t_m}\left[\boldsymbol{p}'_{t_m}(\boldsymbol{c}_t - \boldsymbol{c}_{t_m}) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_{t_m}, \boldsymbol{\omega}_{t_m})}{\partial \boldsymbol{a}_{t_m}} \odot \boldsymbol{c}_{t_m}\right)'(\boldsymbol{a}_t - \boldsymbol{a}_{t_m})\right] \\
&+ \sum_{i=1}^{m-1} \lambda_{t_i}\delta^{-t_i}\left[\boldsymbol{p}'_{t_i}(\boldsymbol{c}_{t_{i+1}} - \boldsymbol{c}_{t_i}) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_{t_i}, \boldsymbol{\omega}_{t_i})}{\partial \boldsymbol{a}_{t_i}} \odot \boldsymbol{c}_{t_i}\right)'(\boldsymbol{a}_{t_{i+1}} - \boldsymbol{a}_{t_i})\right] \\
&+ \lambda_t\delta^{-t}\left[\boldsymbol{p}'_t(\boldsymbol{c} - \boldsymbol{c}_t) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \boldsymbol{\omega}_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t\right)'(\boldsymbol{a} - \boldsymbol{a}_t)\right] \\
&\geq u(\boldsymbol{c}, \boldsymbol{a}).
\end{aligned}
$$

Hence,

$$u(\boldsymbol{c}_t, \boldsymbol{a}_t) + \epsilon + \lambda_t\delta^{-t}\left[\boldsymbol{p}'_t(\boldsymbol{c} - \boldsymbol{c}_t) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \boldsymbol{\omega}_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t\right)'(\boldsymbol{a} - \boldsymbol{a}_t)\right] > u(\boldsymbol{c}, \boldsymbol{a}).$$

Since $\epsilon > 0$, $t \in \mathcal{T}$ and $(\boldsymbol{c}, \boldsymbol{a})$ were arbitrary, we get

$$u(\boldsymbol{c}_t, \boldsymbol{a}_t) + \lambda_t\delta^{-t}\left[\boldsymbol{p}'_t(\boldsymbol{c} - \boldsymbol{c}_t) + \left(\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t, \boldsymbol{\omega}_t)}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t\right)'(\boldsymbol{a} - \boldsymbol{a}_t)\right] \geq u(\boldsymbol{c}, \boldsymbol{a}).$$

This corresponds to the definition of concavity and, therefore, it must be that $\lambda_t\delta^{-t}\boldsymbol{p}_t$ and $\lambda_t\delta^{-t}\frac{\partial \boldsymbol{p}(\boldsymbol{a}_t,\boldsymbol{\omega}_t}{\partial \boldsymbol{a}_t} \odot \boldsymbol{c}_t$ are supergradients of $u(\boldsymbol{c}_t, \boldsymbol{a}_t)$. Next, I show that we can construct a utility function that guarantees the solution to exist.

Let $\Gamma := \max_{l\in\mathcal{L},t\in\mathcal{T}}\{a_{l,t}\}$. For every $l \in \mathcal{L}$, let and $h_l(\cdot)$ be a continuously differentiable function satisfying $h_l(0) = 0$, $h'_l(x) > 0$, $h''_l(x) \geq 0$ for $x \in \mathbb{R}_+$ and $\lim_{x\to\infty} h'_l(x) = \infty$.[8] To see that there exists a utility function such that a solution exists, define $\hat{u}(\boldsymbol{c}, \boldsymbol{a}) := u(\boldsymbol{c}, \boldsymbol{a}) -$

[8]This construction is analogous to that of Deb et al. (2018).

$\sum_{l=1}^{L} h_l(\max\{0, a_l - \Gamma\})$. As before, this function is concave, continuous, strictly increasing in consumption, and decreasing in shopping intensity. Furthermore, note that $\hat{u}(\boldsymbol{c}, \boldsymbol{a}) \leq u(\boldsymbol{c}, \boldsymbol{a})$ for all $(\boldsymbol{c}, \boldsymbol{a}) \in C \times A$ and $\hat{u}(\boldsymbol{c}_t, \boldsymbol{a}_t) = u(\boldsymbol{c}_t, \boldsymbol{a}_t)$ for all $t \in \mathcal{T}$. Thus, $(\boldsymbol{c}_t, \boldsymbol{a}_t)_{t\in\mathcal{T}}$ is still a solution to the consumer problem. Finally, note that $\hat{u}(\boldsymbol{c}, \boldsymbol{a}) \to -\infty$ whenever $\boldsymbol{a} \to \infty$ along some dimension. This follows from the piecewise linearity of $u(\cdot, \cdot)$ and the assumption that $\lim_{x\to\infty} h_l'(x) = \infty$.

A.6.2 Proof of Proposition 2.1

Assumption 2.1 states that the price function for any good $l \in \mathcal{L}$ is given by:

$$\log\left(p_{l,t}^*\right) = \alpha_l^0 + \alpha_l^1 \log(a_{l,t}) - \omega_{l,t}.$$

Due to measurement error in prices, we only get to make inference from

$$\log(p_{l,t}) = \alpha_l^0 + \alpha_l^1 \log(a_{l,t}) - \omega_{l,t}.$$

Summing this equation across goods and dividing by L yields

$$\overline{\log(p_{l,t})} = \frac{1}{L}\sum_{l=1}^{L}\left[\alpha_l^0 + \alpha_l^1 \log(a_{l,t})\right] - \bar{\omega}_{l,t},$$

where $\overline{\log(p_{l,t})}$ denotes the average log price paid and $\bar{\omega}_{l,t}$ denotes the average search productivity. Further taking the expectation simplifies the equation to

$$\mathbb{E}\left[\overline{\log(p_{l,t})}\right] = \frac{1}{L}\sum_{l=1}^{L}\left(\mathbb{E}\left[\alpha_l^0\right] + \mathbb{E}\left[\alpha_l^1 \log(a_{l,t})\right]\right),$$

where Assumption 2.4 was used to eliminate the expected average search productivity. By Assumption 2.2, the above can be written as

$$\mathbb{E}\left[\overline{\log\left(p_{l,t}^*\right)}\right] = \frac{1}{L}\sum_{l=1}^{L}\left(\mathbb{E}\left[\alpha_l^0\right] + \mathbb{E}\left[\alpha_l^1 \log(a_{l,t})\right]\right).$$

Taking the derivative with respect to $\log(a_{l,t})$ and invoking Leibniz integration rule, one gets

$$\mathbb{E}\left[\overline{\frac{\partial \log\left(p^*_{l,t}\right)}{\partial \log(a_{l,t})}}\right] = \frac{1}{L}\mathbb{E}\left[\alpha^1_l\right].$$

Finally, summing this equation for each good $l \in \mathcal{L}$ and dividing by L gives

$$\frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\left[\overline{\frac{\partial \log\left(p^*_{l,t}\right)}{\partial \log(a_{l,t})}}\right] = \frac{1}{L}\mathbb{E}\left[\overline{\alpha^1}\right],$$

where $\overline{\alpha^1} := \frac{1}{L}\sum_{l=1}^{L}\alpha^1_l$ is the average shopping technology across goods.

Appendix B

Appendices for Chapter 3

B.1 Elementary Revealed Preference Theory

This section presents the revealed preference terminology and reviews an extension of the static utility maximization model permitting for violations from optimal behavior.[1]

For $e \in (0, 1]$, a consumption bundle $\boldsymbol{c}_t$ is said to be *directly revealed preferred* to a bundle $\boldsymbol{c}_s$ if and only if $\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \leq 0$, where e is designed to remove revealed preference information generating cyclic preferences. Let $R^D(e)$ denote the direct revealed preference relation and let $R(e)$ denote its transitive closure.[2] When the inequality is strict, $\boldsymbol{c}_t$ is said to be *directly revealed strictly preferred* to $\boldsymbol{c}_s$ and is denoted $P^D(e)$. In the case where there is a sequence $\boldsymbol{c}_t R^D(e) \boldsymbol{c}_{t_1}, \boldsymbol{c}_{t_1} R^D(e) \boldsymbol{c}_{t_2}, \ldots, \boldsymbol{c}_{t_m} R^D(e) \boldsymbol{c}_s$ of directly revealed preferences, where $t, t_1, \ldots, t_m, s \in \mathcal{T}$, $\boldsymbol{c}_t$ is said to be *revealed preferred* to $\boldsymbol{c}_s$. Naturally, if any of those preference relations is strict, then $\boldsymbol{c}_t$ is said to be *revealed strictly preferred* to $\boldsymbol{c}_s$. The preceding notation allows me to succinctly define two important concepts.

Definition B.1. *Let $e \in (0, 1]$. A locally nonsatiated utility function $u(\cdot)$ e-rationalizes the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t \in \mathcal{T}}$ if for every observed bundle $\boldsymbol{c}_t \in \mathbb{R}^L_+$, $\boldsymbol{c}_t R^D(e) \boldsymbol{c}$ implies $u(\boldsymbol{c}_t) \geq u(\boldsymbol{c})$ and $\boldsymbol{c}_t P^D(e) \boldsymbol{c}$ implies $u(\boldsymbol{c}_t) > u(\boldsymbol{c})$.*

[1] As noted by Blow et al. (2021), the fact that discounting prices does not change relative prices implies that static rationalizability is the same with either spot prices $(\boldsymbol{p}_t)_{t \in \mathcal{T}}$ or discounted prices $(\boldsymbol{\rho}_t)_{t \in \mathcal{T}}$. I choose to define static utility maximization with discounted prices for notational consistency.

[2] The transitive closure $R(e)$ of a relation $R^D(e)$ is the smallest relation containing $R^D(e)$ satisfying transitivity.

Definition B.2. *Let $e \in (0, 1]$. A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfies the Generalized Axiom of Revealed Preference (GARP(e)) if for all $s, t \in \mathcal{T}$, $\boldsymbol{c}_t R(e) \boldsymbol{c}_s$ implies not $\boldsymbol{c}_s P^D(e) \boldsymbol{c}_t$.*

The generalized axiom gives an intuitive necessary condition for rationalizability by requiring the consumer to have transitive preferences. In particular, note that GARP(e) is a natural generalization of GARP(1) and simply eliminates cycles by reducing the number of revealed preferences. The following result from Halevy et al. (2018) and Heufer and Hjertstrand (2019) extends the influential theorem of Afriat (1967) to consumers violating the model of atemporal utility maximization.[3,4]

e-Afriat's Theorem. *For a given $e \in (0, 1]$, the following statements are equivalent:*

(1) There exists a locally nonsatiated utility function e-rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

(2) The data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfy GARP(e).

(3) There exist numbers u_t, $\lambda_t > 0$, $t = 0, \ldots, T$, such that

$$u_s \leq u_t + \lambda_t \boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

(4) There exists a locally nonsatiated, continuous, monotonic and concave utility function e-rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

For practical purposes, the second condition is the most convenient. As highlighted by Varian (1982), one can use an efficient algorithm called Floyd-Warshall to get the transitive closure $R(e)$ of the direct revealed preference relation $R^D(e)$. For a given $e \in (0, 1]$, one can then directly check for a contradiction of GARP(e) in the data. Alternatively, conditional on $e \in (0, 1]$, one can use linear programming to solve the system of inequalities given in the third condition. The goal is then to verify the existence of a pair $(u_t, \lambda_t)_{t\in\mathcal{T}}$ satisfying it.

From a theoretical standpoint, however, the substance of e-Afriat's Theorem lies in the last condition. It implies that, if the consumer's choices can be thought of as generated by a locally

[3] For the case where $e = 1$, an accessible proof is given by Fostel et al. (2004). An alternative and insightful proof is offered by Geanakoplos (2013).

[4] This theorem, as well as the proposition to follow, could all be written using the efficiency measure of Varian (1990). For my purposes, it is sufficient to consider a common index for all observations.

nonsatiated utility function, then it can further be assumed to be continuous, monotonic and concave.[5,6]

B.2 Efficiency Indices

This section presents the exponential efficiency index, proposes an efficiency index for time-consistency, and investigates which assumption of the exponential discounting model is the most problematic in the data.

B.2.1 Exponential Efficiency Index

For the model of static utility maximization under partial efficiency, it is common to consider the largest efficiency level rationalizing the data. This index is known as the CCEI and was suggested by Afriat (1973).[7] In a similar fashion, one can consider the largest efficiency level rationalizing the data for the exponential discounting model. Formally, I define the exponential efficiency index as

$$\text{EEI} := \sup\{e \in [0,1] : \{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t \in \mathcal{T}} \text{ is } e\text{-ED rationalizable}\}.$$

From the previous analysis, it is clear that the EEI can be interpreted as the smallest proportion of wasted income arising from the selection of a suboptimal consumption stream. Moreover, note that the exponential efficiency index is well-defined as the inequalities in Proposition 3.2 (*ii*) will be trivially satisfied for an e arbitrarily close to zero.

Although the EEI provides a measure of distance between a data set and the exponential discounting model, it does not differentiate between deviations arising from within-period consistency and time consistency. To disentangle their respective contributions to the EEI, an efficiency measure that controls for violations of static utility maximization is needed. For convenience, denote such an index the time consistency efficiency index (TCEI). In what follows, I

[5] A mapping $f : \mathbb{R}^L \to \mathbb{R}^L$ is said to be concave if and only if $f(\boldsymbol{c}_s) \leq f(\boldsymbol{c}_t) + \nabla f(\boldsymbol{c}_t)'(\boldsymbol{c}_s - \boldsymbol{c}_t)$ for all $s, t \in \mathcal{T}$.

[6] Concavity can only be assumed without loss of generality in finite data. In the case of infinite data, it has to be substituted by the weaker assumption of quasiconcavity. I refer the reader to Reny (2015) for a more detailed discussion.

[7] The critical cost efficiency index is defined as $\text{CCEI} := \sup\{e \in [0,1] : \{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t \in \mathcal{T}} \text{ satisfies GARP}(e)\}$.

derive the TCEI based on the 2-step rationalization procedure of Heufer and Hjertstrand (2019) for homothetic rationalizability.

The first stage consists in finding the largest efficiency level rationalizing the data with respect to static utility maximization and yields

$$u_s \leq u_t + \lambda_t \boldsymbol{\rho}_t'(\boldsymbol{c}_s/\text{CCEI} - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T}.$$

Imposing the additional restriction of the exponential discounting model to the CCEI-Afriat inequalities by setting $\lambda_t = \delta^{-t}$ yields

$$u_s \leq u_t + \delta^{-t} \boldsymbol{\rho}_t'(\boldsymbol{c}_s/\text{CCEI} - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T}.$$

The TCEI then corresponds to the largest efficiency level rationalizing the previous system of inequalities with respect to the e-ED model:

$$u_s \leq u_t + \delta^{-t} \boldsymbol{\rho}_t' \left(\frac{\boldsymbol{c}_s}{\text{CCEI} \cdot \text{TCEI}} - \boldsymbol{c}_t \right) \quad \forall s,t \in \mathcal{T}.$$

That is, the TCEI gives the additional adjustment required to the CCEI-adjusted data set to satisfy the e-ED model. Since the largest efficiency level solving the e-ED model is the EEI, it follows that EEI = CCEI · TCEI. One can therefore recover the TCEI by first obtaining the CCEI and the EEI. Moreover, taking the natural logarithm of the previous expression yields the following relationship when EEI < 1:

$$\frac{\log(\text{CCEI})}{\log(\text{EEI})} + \frac{\log(\text{TCEI})}{\log(\text{EEI})} = 1.$$

This identity allows one to obtain the respective contribution of static utility maximization and time consistency to the exponential efficiency index.

Definition B.3. *The contribution of the CCEI to the EEI and of the TCEI to the EEI are respectively given by*

$$C_g := \frac{\log(CCEI)}{\log(EEI)} \quad \text{and} \quad C_t := \frac{\log(TCEI)}{\log(EEI)}.$$

In particular, the contribution of each index is always between zero and one, strictly increases as its efficiency index decreases, and the combined contribution of each index always sum up to one. Finally, note that the sum of C_g and C_t gives the contribution of the EEI to itself.

B.2.2 Empirical Application

In what follows, I am interested in the efficiency indices for static utility maximization, time consistency and exponential discounting, as well as their respective contributions to the EEI. Let μ_i denote the mean and σ_i the standard deviation, where $i \in \{e, C\}$ refers to the object over which the operation is applied. Any object $i \in \{e, C\}$ underlined or overlined represents its smallest and largest value across consumers, respectively.

Using the CCEI for static utility maximization, the TCEI for time consistency and the EEI for exponential discounting, Table B.1 presents summary statistics on the efficiency indices and contributions of each index. These results are obtained with a grid search over $\delta \in (0, 1]$ with a step size of 0.01 and a binary search algorithm for the efficiency indices that guarantees them to be within 2^{-10} of their true values.

Table B.1: Rationalizability Results

Efficiency index	$\underline{e}$	$\overline{e}$	μ_e	σ_e	$\underline{C}$	$\overline{C}$	μ_C	σ_C
CCEI	0.6865	1.0000	0.9551	0.0502	0.0000	1.0000	0.2057	0.2017
TCEI	0.4758	1.0000	0.8365	0.0802	0.0000	1.0000	0.7943	0.2017
EEI	0.3878	0.9561	0.7984	0.0820	1.0000	1.0000	1.0000	0.0000

Notes: The sample size is $N = 494$. $\underline{e}$ denotes the lowest efficiency index, $\overline{e}$ the largest efficiency index, μ_e the average efficiency index, and σ_e the standard deviation of the efficiency index. $\underline{C}$ denotes the lowest contribution of the efficiency index to the EEI, $\overline{C}$ the largest contribution of the efficiency index to the EEI, μ_C the average contribution of the efficiency index to the EEI, and σ_C the standard deviation of the efficiency index's contribution.

Overall, the results in Table B.1 indicate that time consistency is a more stringent assumption than GARP, with an average efficiency level for the TCEI below that of the CCEI by approximately 0.10. The significance of this difference is better grasped by looking at the average contribution of each index to the EEI. Markedly, on average, GARP is responsible for about 20% of a violation from exponential discounting, while 80% of it can be attributed to time consistency.

B.3 Proofs

B.3.1 Proof of Corollary 3.1

The restrictions on the bounds of the discount factor have been derived in the main text. I am thus left to show that CM(e) implies GARP(e), where $e = 1$ corresponds to the special case of Corollary 3.1. I proceed by contraposition. Fix $e \in (0, 1]$ and suppose GARP(e) is violated. Then, for some indices $t_1, t_m \in \mathcal{T}$, $\boldsymbol{c}_{t_1} R(e) \boldsymbol{c}_{t_m}$ and $\boldsymbol{c}_{t_m} P^D(e) \boldsymbol{c}_{t_1}$. Thus, there is a sequence of revealed preferences such that $\boldsymbol{c}_{t_1} R^D(e) \boldsymbol{c}_{t_2}$, $\boldsymbol{c}_{t_2} R^D(e) \boldsymbol{c}_{t_3}$, $\ldots$, $\boldsymbol{c}_{t_{m-1}} R^D(e) \boldsymbol{c}_{t_m}$, where $t_1, t_2, \ldots, t_m \in \mathcal{T}$. By definition, the above implies $\boldsymbol{\rho}'_{t_1}(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) \leq 0$, $\boldsymbol{\rho}'_{t_2}(\boldsymbol{c}_{t_3}/e - \boldsymbol{c}_{t_2}) \leq 0, \ldots,$ $\boldsymbol{\rho}'_{t_{m-1}}(\boldsymbol{c}_{t_m}/e - \boldsymbol{c}_{t_{m-1}}) \leq 0$ and $\boldsymbol{\rho}'_{t_m}(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}) < 0$. Given $\delta \in (0, 1]$, we also have $\delta^{-t_i} \boldsymbol{\rho}'_{t_i}(\boldsymbol{c}_{t_{i+1}}/e - \boldsymbol{c}_{t_i}) \leq 0$ for all $i \in \{1, \ldots, m-1\}$ and $\delta^{-t_m} \boldsymbol{\rho}'_{t_m}(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}) < 0$. Summing up the resulting inequalities yields

$$0 > \delta^{-t_1} \boldsymbol{\rho}'_{t_1}(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) + \delta^{-t_2} \boldsymbol{\rho}'_{t_2}(\boldsymbol{c}_{t_3}/e - \boldsymbol{c}_{t_2}) + \cdots + \delta^{-t_m} \boldsymbol{\rho}'_{t_m}(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}),$$

which violates CM(e).

B.3.2 Proof of Theorem 3.1

$(i) \implies (ii)$

From the first-order condition, we have

$$\nabla u(\boldsymbol{c}_t) \leq \delta^{-t} \boldsymbol{\rho}_t \quad \forall t \in \mathcal{T},$$

where $\nabla u(\boldsymbol{c}_t)$ is some supergradient of $u(\cdot)$ at $\boldsymbol{c}_t$. By continuity and concavity of the instantaneous utility function, we know that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}) \leq u(\boldsymbol{c}_t) + \nabla u(\boldsymbol{c}_t)'(\boldsymbol{c} - \boldsymbol{c}_t).$$

Let N be a set of indices such that $\nabla u(\boldsymbol{c}_t)_j = \delta^{-t} \boldsymbol{\rho}_{t,j}$ for all $j \in N$. It follows that $\nabla u(\boldsymbol{c}_t)_j \leq$

$\delta^{-t}\boldsymbol{\rho}_{t,j}$ for all $j \notin N$. Thus, $\boldsymbol{c}_{t,j} = 0$ is a corner solution for all $j \notin N$. We therefore have

$$\begin{aligned} u(\boldsymbol{c}) - u(\boldsymbol{c}_t) \leq \nabla u(\boldsymbol{c}_t)'(\boldsymbol{c} - \boldsymbol{c}_t) &= \sum_{j \in N} \nabla u(\boldsymbol{c}_t)_j(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) + \sum_{j \notin N} \nabla u(\boldsymbol{c}_t)_j(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) \\ &= \sum_{j \in N} \delta^{-t}\boldsymbol{\rho}_{t,j}(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) + \sum_{j \notin N} \nabla u(\boldsymbol{c}_t)_j(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) \\ &\leq \sum_{j \in N} \delta^{-t}\boldsymbol{\rho}_{t,j}(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) + \sum_{j \notin N} \delta^{-t}\boldsymbol{\rho}_{t,j}(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}), \end{aligned}$$

where the last inequality holds since $\boldsymbol{c}_{t,j} = 0$ and $\boldsymbol{c}_j \geq 0$ for all $j \notin N$. As a result, for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}) \leq u(\boldsymbol{c}_t) + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t).$$

Rearranging gives that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c},$$

where, by assumption, the instantaneous utility function is locally nonsatiated, continuous, monotonic, and concave and $\delta \in (0, 1]$.

$(ii) \implies (i)$

The instantaneous utility function is locally nonsatiated, continuous, monotonic, and concave and the discount factor satisfies $\delta \in (0, 1]$. For all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$, we also have

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}.$$

Rearranging gives that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}) \leq u(\boldsymbol{c}_t) + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t).$$

This inequality corresponds to the definition of concavity and, therefore, it follows that $\delta^{-t}\boldsymbol{\rho}_t$ is a supergradient of $u(\cdot)$ at $\boldsymbol{c}_t$ for all $t \in \mathcal{T}$.

B.3.3 Proof of Proposition 3.2

(i) $\implies$ *(ii)*

Since the data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is e-ED rationalizable, it is the case that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}/e$$

for some $\delta \in (0, 1]$. By the same argument as in the proof of Theorem 3.1, we can obtain

$$u(\boldsymbol{c}_s) \leq u(\mathbf{c_t}) + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T},$$

where one may define $u_t := u(\boldsymbol{c}_t)$ for all $t \in \mathcal{T}$.

(ii) $\implies$ *(iii)*

Starting from the e-ED Afriat inequalities, we have

$$u_s \leq u_t + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

Consider any sequence of indices $\tau = \{t_i\}_{i=1}^m$, $t_i \in \mathcal{T}$, $m \geq 2$, and let $\mathcal{I}$ be the set of all such indices. Summing up the inequalities for the resulting cycle of indices yields

$$0 \leq \delta^{-t_1}\boldsymbol{\rho}_{t_1}'(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) + \ldots + \delta^{-t_m}\boldsymbol{\rho}_{t_m}'(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}),$$

which corresponds to CM(e).

(iii) $\implies$ *(i)*

For some $e \in (0, 1]$, define

$$u(\boldsymbol{c}) = \inf_{\tau\in\mathcal{I}}\Big\{\delta^{-\tau(m)}\boldsymbol{\rho}_{\tau(m)}'(\boldsymbol{c}/e - \boldsymbol{c}_{\tau(m)}) + \sum_{i=1}^{m-1}\delta^{-\tau(i)}\boldsymbol{\rho}_{\tau(i)}'(\boldsymbol{c}_{\tau(i+1)}/e - \boldsymbol{c}_{\tau(i)})\Big\}.$$

This utility function is locally nonsatiated, continuous, monotonic and concave as it is the pointwise minimum of a collection of affine functions. Moreover, the infimum defining $u(\boldsymbol{c})$

has no cycle of indices. To see this, let $s \in \mathcal{T}$ and note that by CM(e) we have

$$0 \leq \delta^{-\tau(1)} \boldsymbol{\rho}'_{\tau(1)}(\boldsymbol{c}_{\tau(2)}/e - \boldsymbol{c}_{\tau(1)}) + \ldots + \delta^{-\tau(m)} \boldsymbol{\rho}'_{\tau(m)}(\boldsymbol{c}_s/e - \boldsymbol{c}_{\tau(m)}) + \delta^{-s} \boldsymbol{\rho}'_s(\boldsymbol{c}_{\tau(1)}/e - \boldsymbol{c}_s)$$

for all $\tau \in \mathcal{T}$. Consider $\boldsymbol{c} \in \mathbb{R}^L_+$ such that $\boldsymbol{c} \neq \boldsymbol{c}_t$ and let $\tau_t \in \mathcal{I}$ be a minimizing sequence for $\boldsymbol{c}_t$. It follows that

$$\begin{aligned} u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}/e &\leq \delta^{-t}\boldsymbol{\rho}'_t(\boldsymbol{c}/e - \boldsymbol{c}_t) + \delta^{-\tau_t(m_t)}\boldsymbol{\rho}'_{\tau_t(m_t)}(\boldsymbol{c}_t/e - \boldsymbol{c}_{\tau_t(m_t)}) \\ &\quad + \sum_{i=1}^{m_t-1} \delta^{-\tau_t(i)}\boldsymbol{\rho}'_{\tau_t(i)}(\boldsymbol{c}_{\tau_t(i+1)}/e - \boldsymbol{c}_{\tau_t(i)}) - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}/e \\ &= \delta^{-\tau_t(m_t)}\boldsymbol{\rho}'_{\tau_t(m_t)}(\boldsymbol{c}_t/e - \boldsymbol{c}_{\tau_t(m_t)}) \\ &\quad + \sum_{i=1}^{m_t-1} \delta^{-\tau_t(i)}\boldsymbol{\rho}'_{\tau_t(i)}(\boldsymbol{c}_{\tau_t(i+1)}/e - \boldsymbol{c}_{\tau_t(i)}) - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}_t \\ &= u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}_t, \end{aligned}$$

where the first inequality holds since $u(\boldsymbol{c})$ uses the sequence achieving the infimum for $\boldsymbol{c}$, the first equality is a mere simplification, and the last equality is a consequence of τ_t being a minimizing sequence for $\boldsymbol{c}_t$. I thus have shown the existence of a locally nonsatiated, continuous, monotonic and concave utility function and a discount factor $\delta \in (0, 1]$ e-ED rationalizing the data.

Appendix C

Appendices for Chapter 4

C.1 Preliminaries

Let $\delta \in [\underline{\delta}, 1]$. For all $t_i, t_j \in \mathcal{T}$, let the entry (t_i, t_j) of a square matrix A be given by $a_{t_i,t_j} := \delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j} - \delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j}$.

Definition C.1. *A square matrix A of dimension T is cyclically consistent if for every chain* $\{t_1, t_2, \ldots, t_m\} \subset \{1, 2, \ldots, T\}$, $a_{t_1,t_2} \leq 0, a_{t_2,t_3} \leq 0, \ldots, a_{t_{m-1},t_m} \leq 0$, $a_{t_m,t_1} \leq 0$ *implies that all terms are zero.*

Lemma C.1. *A data set* $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ *satisfies GAPP(*δ*) if and only if the square matrix A is cyclically consistent.*

Proof

Let $\delta \in [\underline{\delta}, 1]$ and assume that A is cyclically consistent with $\delta^{-t_1}\boldsymbol{\rho}_{t_1}R\delta^{-t_m}\boldsymbol{\rho}_{t_m}$. Thus, there is a sequence of indices $\{t_1, t_2, \ldots, t_m\} \subset \{1, 2, \ldots, T\}$ such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1}R^D\delta^{-t_2}\boldsymbol{\rho}_{t_2}$, $\delta^{-t_2}\boldsymbol{\rho}_{t_2}R^D\delta^{-t_3}\boldsymbol{\rho}_{t_3}$, $\ldots$, $\delta^{-t_{m-1}}\boldsymbol{\rho}_{t_{m-1}}R^D\delta^{-t_m}\boldsymbol{\rho}_{t_m}$. Therefore, this implies $a_{t_1,t_2} \leq 0$, $a_{t_2,t_3} \leq 0$, $\ldots$, $a_{t_{m-1},t_m} \leq 0$. Note that if $\delta^{-t_m}\boldsymbol{\rho}_{t_m}P^D\delta^{-t_1}\boldsymbol{\rho}_{t_1}$, then $a_{t_m,t_1} < 0$. Cyclical consistency then requires that $a_{t_1,t_2} = a_{t_2,t_3} = \cdots = a_{t_m,t_1} = 0$. However, this contradicts the assumption that $a_{t_m,t_1} < 0$. As such, we can't have $\delta^{-t_m}\boldsymbol{\rho}_{t_m}P^D\delta^{-t_1}\boldsymbol{\rho}_{t_1}$, i.e. GAPP($\delta$) holds.

Suppose now that GAPP(δ) is satisfied for some $\delta \in [\underline{\delta}, 1]$. Construct the matrix A of revealed preferences and note that $a_{t,t} = 0$ for all $t \in \mathcal{T}$. Consider any sequence $\{t_1, t_2, \ldots, t_m\} \subset \{1, 2, \ldots, T\}$ such that $a_{t_1,t_2} \leq 0$, $a_{t_2,t_3} \leq 0$, $\ldots$, $a_{t_{m-1},t_m} \leq 0$, $a_{t_m,t_1} \leq 0$. For any element

a_{t_i,t_j} pertaining to that chain, we have $\delta^{-t_i}\boldsymbol{\rho}_{t_i}R^D\delta^{-t_j}\boldsymbol{\rho}_{t_j}$. Moreover, by going along the chain we also obtain $\delta^{-t_j}\boldsymbol{\rho}_t R\delta^{-t_i}\boldsymbol{\rho}_{t_i}$. Since GAPP($\delta$) requires to not have $\delta^{-t_i}\boldsymbol{\rho}_{t_i}P^D\delta^{-t_j}\boldsymbol{\rho}_{t_j}$, it must be that $a_{t_i,t_j} = 0$.

C.2 Proofs

C.2.1 Proof of Theorem 4.1

$(i) \implies (ii)$

Let $\delta \in [\underline{\delta}, 1]$ and note that $\delta^{-t_i}\boldsymbol{\rho}_{t_i}R\delta^{-t_j}\boldsymbol{\rho}_{t_j}$ implies $\delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j} \leq \delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j}$. As such, we have $U(\boldsymbol{c}_{t_i}, -\delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_i}) \geq U(\boldsymbol{c}_{t_j}, -\delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j}) \geq U(\boldsymbol{c}_{t_j}, -\delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j})$, where the first inequality follows from the optimality of $\boldsymbol{c}_{t_i}$ and the second from the revealed preference relation. The same argument can be made for the strict relation P. Suppose now that GAPP(δ) were violated. Then, for any $\delta \in [\underline{\delta}, 1]$ there would be $t_1, t_m \in \mathcal{T}$ such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1}R\delta^{-t_m}\boldsymbol{\rho}_{t_m}$ and $\delta^{-t_m}\boldsymbol{\rho}_{t_m}P^D\delta^{-t_1}\boldsymbol{\rho}_{t_1}$. Thus, there are $t_2, t_3, \ldots, t_{m-1} \in \mathcal{T}$ such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1}R^D\delta^{-t_2}\boldsymbol{\rho}_{t_2}, \delta^{-t_2}\boldsymbol{\rho}_{t_2}R^D\delta^{-t_3}\boldsymbol{\rho}_{t_3}, \ldots, \delta^{-t_{m-1}}\boldsymbol{\rho}_{t_{m-1}}R^D\delta^{-t_m}\boldsymbol{\rho}_{t_m}$. As a result, we would have $U(\boldsymbol{c}_{t_1}, -\delta^{-t_1}\boldsymbol{\rho}'_{t_1}\boldsymbol{c}_{t_1}) \geq U(\boldsymbol{c}_{t_2}, -\delta^{-t_2}\boldsymbol{\rho}'_{t_2}\boldsymbol{c}_{t_2}) \geq \cdots \geq U(\boldsymbol{c}_{t_m}, -\delta^{-t_m}\boldsymbol{\rho}'_{t_m}\boldsymbol{c}_{t_m}) > U(\boldsymbol{c}_{t_1}, -\delta^{-t_1}\boldsymbol{\rho}'_{t_1}\boldsymbol{c}_{t_1})$, an obvious contradiction.

$(ii) \implies (iii)$

In any given time period $t \in \mathcal{T}$, denote lifetime income by $y > 0$ and let $s_t^d := y^d - \delta^{-t}\boldsymbol{\rho}'_t\boldsymbol{c}_t$ denote discounted savings. Extending the data set for savings, one obtains $\{(\boldsymbol{\rho}_t, 1), (\boldsymbol{c}_t, s_t^d)\}_{t\in\mathcal{T}}$ with the value of money priced to 1. Since variables are discounted by interest rates, one can also think of the price for money as evolving according to interest rates but where consumption prices and income are in nominal terms. For all $t_i, t_j \in \mathcal{T}$, note that we have

$$(\delta^{-t_i}\boldsymbol{\rho}_{t_i}, 1)'((\boldsymbol{c}_{t_j}, y - \delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{tj}) - (\boldsymbol{c}_{t_i}, y - \delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_i})) \leq 0 \text{ iff } \delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j} \leq \delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j},$$

where the same equivalence applies with strict inequalities. Now, define $a_{t_i,t_j} := \delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j} - \delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j}$ and let the matrix A be defined by $A_{t_i,t_j} := a_{t_i,t_j}\ \forall t_i, t_j \in \mathcal{T}$. Likewise, define $\tilde{a}_{t_i,t_j} := (\delta^{-t_i}\boldsymbol{\rho}_{t_i}, 1)'((\boldsymbol{c}_{t_j}, y - \delta^{-t_j}\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j}) - (\boldsymbol{c}_{t_i}, y - \delta^{-t_i}\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_i}))$ and let the matrix $\tilde{A}$ be defined by $\tilde{A}_{t_i,t_j} := \tilde{a}_{t_i,t_j}\ \forall t_i, t_j \in \mathcal{T}$. By Lemma C.1, we know that GAPP(δ) holds if and only if A is cyclically consistent and, by the previous equivalence, A is cyclically consistent if and only if $\tilde{A}$ also is. An application of Fostel et al. (2004) (Sections 2 and 3) on the matrix $\tilde{A}$ then guarantees the existence of Afriat inequalities given by

$$u_{t_j} - u_{t_i} \leq \lambda_{t_i}(\delta^{-t_i}\boldsymbol{\rho}_{t_i}, 1)'((\boldsymbol{c}_{t_j}, y - \delta^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}) - (\boldsymbol{c}_{t_i}, y - \delta^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_i})) \quad \forall t_i, t_j \in \mathcal{T},$$

where $u_{t_i} \in \mathbb{R}$ and $\lambda_{t_i} > 0$ for all $t_i \in \mathcal{T}$. We can now construct a well-behaved augmented utility function on $\{(\boldsymbol{\rho}_t, 1), (\boldsymbol{c}_t, s_t^d)\}_{t\in\mathcal{T}}$. To this end, consider sequences of finite indices $\{t_i\}_{i=1}^m$, where $m \geq 2$, $t_i \in \mathcal{T}$, and let $\mathcal{I}$ denote the set of all such sequences. For any $\boldsymbol{c} \in \mathbb{R}^L$ and $t \in \mathcal{T}$, define

$$\begin{aligned}
&\tilde{U}(\boldsymbol{c}, y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}) = \\
&\min_{\tau\in\mathcal{I}}\Big\{\lambda_{t_m}\left[\delta^{-t_m}\boldsymbol{\rho}_{t_m}'(\boldsymbol{c} - \boldsymbol{c}_{t_m}) + (y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}) - (y - \delta^{-t_m}\boldsymbol{\rho}_{t_m}'\boldsymbol{c}_{t_m})\right] \\
&+ \sum_{i=1}^{m-1}\lambda_{t_i}\Big[\delta^{-t_i}\boldsymbol{\rho}_{t_i}'(\boldsymbol{c}_{t_{i+1}} - \boldsymbol{c}_{t_i}) + (y - \delta^{-t_{i+1}}\boldsymbol{\rho}_{t_{i+1}}'\boldsymbol{c}_{t_{i+1}}) - (y - \delta^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_i})\Big]\Big\}.
\end{aligned}$$

Note that $\tilde{U}(\cdot,\cdot)$ defines a continuous, strictly increasing, and concave augmented utility function. Consider $\boldsymbol{c} \in \mathbb{N}^L$ such that $\boldsymbol{c} \neq \boldsymbol{c}_t$ and let $\{t_i^*\}_{i=1}^m \in \mathcal{I}$ be a minimizing sequence for $\boldsymbol{c}_t$. It follows that

$$\begin{aligned}
&\tilde{U}(\boldsymbol{c}, y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}) \leq \lambda_t\left[\delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t) + (y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}) - (y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t)\right] \\
&+ \lambda_{t_m^*}\left[\delta^{-t_m^*}\boldsymbol{\rho}_{t_m^*}'(\boldsymbol{c}_t - \boldsymbol{c}_{t_m^*}) + (y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t) - (y - \delta^{-t_m^*}\boldsymbol{\rho}_{t_m^*}'\boldsymbol{c}_{t_m^*})\right] \\
&+ \sum_{i=1}^{m-1}\lambda_{t_i^*}\Big[\delta^{-t_i^*}\boldsymbol{\rho}_{t_i^*}'(\boldsymbol{c}_{t_{i+1}^*} - \boldsymbol{c}_{t_i^*}) + (y - \delta^{-t_{i+1}^*}\boldsymbol{\rho}_{t_{i+1}^*}'\boldsymbol{c}_{t_{i+1}^*}) - (y - \delta^{-t_i^*}\boldsymbol{\rho}_{t_i^*}'\boldsymbol{c}_{t_i^*})\Big] \\
&= \lambda_{t_m^*}\left[\delta^{-t_m^*}\boldsymbol{\rho}_{t_m^*}'(\boldsymbol{c}_t - \boldsymbol{c}_{t_m^*}) + (y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t) - (y - \delta^{-t_m^*}\boldsymbol{\rho}_{t_m^*}'\boldsymbol{c}_{t_m^*}')\right] \\
&+ \sum_{i=1}^{m-1}\lambda_{t_i^*}\Big[\delta^{-t_i^*}\boldsymbol{\rho}_{t_i^*}'(\boldsymbol{c}_{t_{i+1}^*} - \boldsymbol{c}_{t_i^*}) + (y - \delta^{-t_{i+1}^*}\boldsymbol{\rho}_{t_{i+1}^*}'\boldsymbol{c}_{t_{i+1}^*}) - (y - \delta^{-t_i^*}\boldsymbol{\rho}_{t_i^*}'\boldsymbol{c}_{t_i^*})\Big] \\
&= \tilde{U}(\boldsymbol{c}_t, y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t),
\end{aligned}$$

where the first inequality holds since $\tilde{U}(\boldsymbol{c}, y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c})$ uses the sequence achieving the minimum for $\boldsymbol{c}$, the first equality is a mere simplification, and the last equality is a consequence of $\{t_i^*\}_{i=1}^m$ being a minimizing sequence for $\boldsymbol{c}_t$. Finally, defining $U : \mathbb{R}_+^L \times \mathbb{R}_- \to \mathbb{R}$ by $U(\boldsymbol{c}, -\delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}) := \tilde{U}(\boldsymbol{c}, y - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c})$ yields an augmented utility function rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$. By an identical argument as in Theorem 1 of Deb et al. (2018) it is possible to modify the aforementioned utility function in order to further guarantee the existence of a solution to any set of prices $\boldsymbol{\rho} \in \mathbb{R}_{++}^L$. Intuitively, the idea is to make the utility cost of expenditure prohibitively large outside of the support of the data such that optimal consumption is finite.

C.2.2 Proof of Theorem 4.2

Consider any rational choice type $r \in \mathcal{R}$ and note that any rational choice type $r' \in \mathcal{R}$ with $\boldsymbol{a}^{r'} = \boldsymbol{a}^r$ implies $r' = r$. Let $\alpha \in [0, 1]$, $\delta_a, \delta_b \in (0, 1]$, and $\delta_\alpha := \alpha\delta_a + (1 - \alpha)\delta_b$. For any $t_i, t_j \in \mathcal{T}$, assume that $\delta_a^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_j} \le \delta_a^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}$ if and only if $\delta_b^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_j} \le \delta_b^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}$. It follows that

$$\alpha\delta_a^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_j} + (1-\alpha)\delta_b^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_j} \le \alpha\delta_a^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j} + (1-\alpha)\delta_b^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}$$
$$\iff$$
$$\delta_\alpha^{-t_i}\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_j} \le \delta_\alpha^{-t_j}\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}.$$

In other words, the set of discount factors that is compatible with the same choice profile $\boldsymbol{a}^r$ is convex. Let Δ^r denote such set.

By Theorem 1, the choice profile $\boldsymbol{a}^r$ is associated with decisions that satisfy GAPP(δ) for any $\delta \in \Delta^r$. Thus, any sequence $t_1, t_2, \ldots, t_m$ must be such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1}'\boldsymbol{c}_{t_2} \le \delta^{-t_2}\boldsymbol{\rho}_{t_2}'\boldsymbol{c}_{t_2}$, $\ldots$, $\delta^{-t_{m-1}}\boldsymbol{\rho}_{t_{m-1}}'\boldsymbol{c}_{t_m} \le \delta^{-t_m}\boldsymbol{\rho}_{t_m}'\boldsymbol{c}_{t_m}$ implies not $\delta^{-t_m}\boldsymbol{\rho}_{t_m}'\boldsymbol{c}_{t_1} \le \delta^{-t_1}\boldsymbol{\rho}_{t_1}'\boldsymbol{c}_{t_1}$. Let T^+ be the set of pairs $(t_i, t_j) \in \mathcal{T} \times \mathcal{T}$ such that $\delta^{-t_i}\boldsymbol{\rho}_{t_i}P\delta^{-t_j}\boldsymbol{\rho}_{t_j}$, and T^- be the set of pairs $(t_i, t_j) \in \mathcal{T} \times \mathcal{T}$ such that not $\delta^{-t_i}\boldsymbol{\rho}_{t_i}P\delta^{-t_j}\boldsymbol{\rho}_{t_j}$. Note that $T^- \neq \emptyset$ and $T^+ \cap T^- = \emptyset$. Isolating the discount factor for any $(t_i, t_j) \in T^+$ yields

$$\delta \le \left(\frac{\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}}{\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_i}}\right)^{\frac{1}{t_j - t_i}}.$$

In a similar fashion, isolating the discount factor for any $(t_i, t_j) \in T^-$ yields

$$\delta \ge \left(\frac{\boldsymbol{\rho}_{t_j}'\boldsymbol{c}_{t_j}}{\boldsymbol{\rho}_{t_i}'\boldsymbol{c}_{t_i}}\right)^{\frac{1}{t_j - t_i}}.$$

Compute the previous bounds for every pair (t_i, t_j) and define the least upper bound and the greatest lower bound as

$$lub := \min_{(t_i,t_j)\in T^+} \left(\frac{\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j}}{\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j}}\right)^{\frac{1}{t_j-t_i}},$$

$$glb := \max_{(t_i,t_j)\in T^-} \left(\frac{\boldsymbol{\rho}'_{t_j}\boldsymbol{c}_{t_j}}{\boldsymbol{\rho}'_{t_i}\boldsymbol{c}_{t_j}}\right)^{\frac{1}{t_j-t_i}}.$$

Clearly, it must be that $glb \leq lub$ and $glb \leq 1$ as otherwise no discount factor $\delta \in [\underline{\delta}, 1]$ would be consistent with GAPP(δ). Thus, a rational choice profile $\boldsymbol{a}^r$ satisfies GAPP(δ) for any $\delta \in \Delta^r$ if and only if $glb \leq lub$ and $glb \leq 1$.

C.3 Algorithm to Compute $\mathcal{R}$

The following definitions are used in the pseudo code below.

Definition C.2. *Let $\mathcal{A}_s$ denote the set of matrices of dimension $s \times s$ $(1 < s \leq T)$ as the set containing every matrix A for which an element $a_{t_i,t_j} = 1$ implies $a_{t_j,t_i} = 0$ when $t_i \neq t_j$ and $a_{t_i,t_j} = 0$ when $t_i = t_j$.*

Definition C.3. *Let q_s index matrices in $\mathcal{A}_s$ such that A_{q_s} denote the qth matrix of $\mathcal{A}_s$.*

Definition C.4. *Let $\mathcal{A}_{s|q_{s-1}} \subset \mathcal{A}_s$ denote the set of matrices whose first $s-1$ rows and columns are identical to $A_{q_{s-1}}$. Let $A_{q_s|q_{s-1}}$ be the qth matrix of $\mathcal{A}_{q_s|q_{s-1}}$.*

Definition C.5. *Let m_s index consumption bundles in period $s \in \mathcal{T}$ and denote the m_sth consumption bundle by $\boldsymbol{c}_{s,m_s}$.*

Note that $|\mathcal{A}_s| = 3^{(s^2-s)/2}$, $|\mathcal{A}_{s|q_{s-1}}| = 3^{s-1}$ and $m_s \in \{1, \ldots, I_s\}$. The following pseudo code gives a depth-first search procedure to recover $\mathcal{R}$.

Pseudo Code

1. Initialize $m_1 = \cdots = m_T = 1$.
2. Initialize $q_2 = \cdots = q_T = 1$.
3. Initialize s = 2.
4. If $s = 2$, set the matrix A_{q_2} and continue. Else, go to step 13.

5. Compute the transitive closure of A_{q_2}.
6. If a cycle is detected, go to step 11. Else, continue.
7. Set $\boldsymbol{c}_{1,m_1}, \ldots, \boldsymbol{c}_{s,m_s}$.
8. Compute lub^{q_2} and glb^{q_2}.
9. If ($glb^{q_2} > lub^{q_2}$ or $glb^{q_2} > 1$), go to step 12. Else, continue.
10. Set $s = 3$ and return to step 4.
11. If $q_2 < 3$, set $q_2 = q_2 + 1$ and return to step 4. Else, terminate.

12.a If $m_2 < I_2$, set $m_2 = m_2 + 1$ and return to step 7. Else, continue.

12.b If $m_2 = I_2$ and $m_1 < I_1$, set $m_2 = 1$, $m_1 = m_1 + 1$ and return to step 7. Else, go to step 11.

13. Set the matrix $A_{q_s|q_{s-1}}$.
14. Compute the transitive closure of $A_{q_s|q_{s-1}}$.
15. If a cycle is detected, move to step 21. Else, continue.
16. Set $\boldsymbol{c}_{1,m_1}, \ldots, \boldsymbol{c}_{s,m_s}$.
17. Compute lub^{q_s} and glb^{q_s}.
18. If ($glb^{q_s} > lub^{q_s}$ or $glb^{q_s} > 1$), go to step 22. Else, continue.
19. If $s < T$, set $s = s + 1$, $m_s = 1$, and return to step 13. Else, continue.
20. Extend $\mathcal{R}$ with $(\boldsymbol{c}_{1,m_1}, \ldots, \boldsymbol{c}_{T,m_T})$ and go to step 22.

21.a If $q_s < 3^{(s^2-s)/2}$, set $q_s = q_s + 1$ and return to step 13. Else, continue.

21.b Set $m_s = 1$, $q_s = 1$, $s = s - 1$.

21.c If $q_s < 3^{(s^2-s)/2}$, set $q_s = q_s + 1$ and return to step 13. Else, continue.

21.d If $s > 2$, set $m_s = 1$, $q_s = 1$, $s = s - 1$, and go to step 21.c. Else, terminate.

22.a If $m_s < I_s$, set $m_s = m_s + 1$ and return to step 16. Else, continue.

22.b Set $m_s = 1$ and go to step 21.

Curriculum Vitae

Education

Ph.D. Economics, University of Western Ontario, Canada	2022
M.A. Economics, University of Toronto, Canada	2016
B.A. Economics and Mathematics, Université Laval, Canada	2015

Fellowships and Awards

Social Sciences and Humanities Research Council Doctoral Fellowship	2020–2021
Ontario Graduate Scholarship	2018–2020

Related Experience

Teaching Assistant, University of Western Ontario	2016–2020
Research Assistant for Varouj A. Aivazian, University of Toronto	2016
Teaching Assistant, University of Toronto	2015–2016
Teaching Assistant, Université Laval	2013