A Revealed Preference Approach to Identification and Inference in Consumer Models*

Charles Gauthier

September 2022

Abstract

This paper provides a new identification result for a large class of consumer problems using a revealed preference approach. I show that the utility maximization hypothesis nonparametrically identifies production functions via restrictions from the first-order conditions. In addition, I derive a nonparametric characterization of the class of models that operationalizes the identification strategy. Finally, I use a novel and easy-to-apply inference method for the estimation of the production functions. This method can be used to statistically test the model, can deal with any type of latent variables (e.g., measurement error), and can be combined with standard exclusion restrictions. Using data on shopping expenditures from the Nielsen Homescan Dataset, I show that a doubling of shopping intensity decreases prices paid by about 15%. At the same time, I find that search costs more than double within the support of the data, hence largely diminishing net benefits of price search.

1 Introduction

This paper is concerned with what can be learned about production functions that arise in consumer problems. These functions are ubiquitous in economic analyses such as in models of price search, household production, human capital, and general equilibrium. The estimation of these functions often relies on parametric restrictions or exogenous variation whose identifying assumption is untestable. In this paper, I show that the

*Department of Economics, University of Western Ontario, Social Science Centre, London, Ontario, N6A 5C2 (email: cgauth26@uwo.ca). I wish to thank Victor Aguiar, Roy Allen, Elizabeth Caucutt, Rahul Deb, Florian Gunsilius, Per Hjertstrand, Nail Kashaev, Lance Lochner, Nirav Mehta, David Rivers, Baxter Robinson, and Susanne Schennach for useful comments. 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. This project is supported in part by funding from the Social Sciences and Humanities Research Council.

structure of the consumer problem is sufficient to nonparametrically identify the part of the production functions that relates to the consumer preferences. I propose an estimation strategy based on revealed preferences that can be used to test and estimate the model. The approach is exemplified in an application to price search using household expenditures data.

The key insight of my identification strategy is to note that the consumer problem provides a link between preferences and production via the first-order conditions. I show that the consumer preferences are uniquely nonparametrically identified in a wide class of models provided there is sufficient variation in budget sets when the sample size grows. This result uses a revealed preference argument initially proposed by Mas-Colell (1978). My contribution is to show that, once preferences are identified, cross-equation restrictions from the first-order conditions of the consumer problem nonparametrically identify the production functions.

The identification of preferences has first been investigated by Mas-Colell (1977, 1978) in consumer problems with linear budgets when one observes the demand function or an increasingly large set of demands, respectively. The analysis has been extended to nonlinear budgets by Forges and Minelli (2009) when the demand correspondence is fully observed. More recently, there has been an effort to derive identification results when the choice correspondence is not fully observed, a situation referred to as partial observability.[1] In that direction, Gorno (2019) extends results from Mas-Colell (1977) to general choice problems and Chambers, Echenique and Lambert (2021) extends results from Mas-Colell (1978) in the case of binary choice problems.[2] Due to the range of problems considered in this paper, my identification result applies when there is a unique solution to each choice situation faced by the consumer.

The careful reader will note that my identification result shares similarities with Gandhi, Navarro and Rivers (2020) in that I exploit the nonparametric link between production functions and first-order conditions.[3] My approach otherwise differs due to specific challenges arising in consumer problems. First, monotone transformations of the utility function may impact production functions even though preferences remain unchanged. Second, the consumer preferences are unknown such that first-order conditions are not immediately useful to identify the production functions. My main innovation is to unify insights from distinct fields to derive a new identification result that applies to production functions in consumer problems.

The second contribution of this paper is to provide a nonparametric characterization

[1] This terminology follows Chambers, Echenique and Shmaya (2014) and is explained therein.

[2] Related work on the identification of preferences includes Kübler and Polemarchakis (2017) in a dynamic model of consumption with assets and uncertainty and Kübler, Malhotra and Polemarchakis (2020) in finite data with an application to aggregate demand.

[3] Gandhi, Navarro and Rivers (2020) show that using the full structure of the firm problem resolves the lack of identification in the widespread proxy variable approach developed by Olley and Pakes (1996) and refined by Levinsohn and Petrin (2003), Wooldridge (2009), and Ackerberg, Caves and Frazer (2015).

of a large class of models that operationalizes the identification strategy. The characterization gives testable restrictions that are equivalent to the well-known Strong Axiom of Revealed Preference (SARP) when the solution is unique. Since SARP exhausts the empirical content of the model, the restrictions use all of the information entailed by the consumer preferences and the first-order conditions. Thus, inference retains the full identification power underlying the identification strategy.

The previous characterization is rooted in the revealed preference tradition (Afriat, 1967; Diewert, 1973; Varian, 1982; Browning, 1989). The treatment of nonlinear budget sets builds on Matzkin (1991) and, more specifically, Forges and Minelli (2009).[4] I extend their results by allowing for multiple constraints and by relaxing the monotonicity of preferences. This allows me to cover a wide range of models such as models of price search where the utility function is typically decreasing in search intensity. Although the extension of Afriat (1967) to many nonlinear constraints was first analyzed by Chavas and Cox (1993), my results use distinct arguments that may be of interest.[5]

The methodology I employ for the estimation of the production functions is that of Schennach (2014). This is motivated by a result due to Aguiar and Kashaev (2021) that shows how to impose shape constraints without increasing the dimensionality of the problem. I propose a different implementation than Aguiar and Kashaev (2021) due to the wide variety of models encompassed by my results.[6] The method of Schennach (2014) also allows me to statistically test the model and thus check the plausibility of modelling assumptions. Finally, it can be applied in partially identified models without additional complications.[7]

My approach combines revealed preference and nonparametric econometrics to derive identification results and make inference on objects of interest. In a similar fashion, Blundell, Browning and Crawford (2008) use SARP to improve demand responses to price changes, Blundell et al. (2015) use SARP to obtain nonparametric bounds on welfare measures, Cherchye et al. (2015a) use a weaker version of SARP to bound the sharing rule in a collective model, and Deb et al. (2018) bound the welfare implications of a price change via analogous revealed preference restrictions.[8] The estimation strategy put forward in this paper may also be useful for such endeavors.

Another application of my approach is to mitigate certain issues with instrumental variables (IV). The main requirement of IV is the availability of an instrument that only

[4]In a different direction, Nishimura, Ok and Quah (2017) extend the revealed preference analysis to a diverse set of choice environments.

[5]For example, Theorem 3 of this paper is analogous to Proposition 1 of Chavas and Cox (1993) but uses the concavity of the utility function rather than restrictions on saddle points.

[6]Specifically, I use a rejection sampling algorithm that can be applied in models defined by linear or nonlinear constraints indiscriminately, including combinations thereof.

[7]Other methods in the literature includes Chernozhukov, Hong and Tamer (2007) and Andrews and Soares (2010).

[8]Other work include Blundell, Horowitz and Parey (2012) and Blundell, Kristensen and Matzkin (2014), among others.

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).[9] Since instrumental variable is often used within the framework of a model, it is without loss of generality to exploit the structure of the model to obtain (joint) testable restrictions.

The third contribution of this paper is to apply the estimation strategy to a model of price search that allows for unrestricted heterogeneity in preferences. 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.[10] In an influential paper, Aguiar and Hurst (2007) show that price search partially explains the retirement-consumption puzzle.[11] My application provides new insights regarding the validity of price search, recovers the robust impacts of search on prices paid, and quantifies the size of search costs.

My empirical analysis 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. Furthermore, the link between the utility function and the price function given by the first-order conditions allows me to recover search costs.

In a validation study of the Nielsen Homescan Dataset, Einav, Leibtag and Nevo (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.[12] Second, measurement error can complicate empirical analyses by obscuring the true behavior of variables such as expenditure.[13] 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.

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

[10] See, for example, Sorensen (2000), Brown and Goolsbee (2002), and McKenzie, Schargrodsky and Cruces (2011)

[11] The retirement-consumption puzzle was dubbed due to the observed drop in expenditures occurring around retirement that contradicts the life-cycle hypothesis.

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

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

My application 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, Schargrodsky and Cruces, 2011; Nevo and Wong, 2019), and (iii) the wide heterogeneity in prices paid (Kaplan and Menzio, 2015; Kaplan et al., 2019; Hitsch, Hortacsu and Lin, 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, Guler and Taskin, 2021).

While I find support for price search behavior in single households, I reject the model in multi-person households.[14] The latter is likely caused by the implicit assumption that multi-person households behave as a single decision maker. Indeed, there is solid evidence against it in the literature (see e.g., Thomas, 1990, Fortin and Lacroix, 1997, Browning and Chiappori, 1998, and Cherchye and Vermeulen, 2008). As such, this finding provides evidence that the current methodology is successful at detecting erroneous assumptions. Furthermore, it reiterates the importance of recognizing the collective nature of households, including in models of price search.

Using data on single households, the 95% confidence set on the expected elasticity of price with respect to shopping intensity states that a doubling of shopping intensity decreases the price paid by about 15%. My confidence set is consistent with the estimates of Aguiar and Hurst (2007) obtained using an instrumental variable approach.[15] This suggests that the exogeneity requirement of their instruments are fulfilled, and that measurement error does not significantly bias their estimates.

The structure of the model further allows me to recover search costs incurred by the consumer. I find that the 95% confidence set on the expected logarithm of search costs is between 1.75 and 4.75 over the support of the data. That is, the expected search cost at the observed number of shopping trips is between 175 to 475 percent larger than the expected search cost with no search. The expected total search cost represents at most 43% to 120% of the consumer expenditure.

The rest of the paper is organized as follows. Section 2 defines the class of problems covered by the results. Section 3 presents the identification result. Section 4 presents the estimation strategy. Section 5 introduces the model of price search considered in my application. Section 6 describes the data set. Section 7 details the empirical results. Section 8 concludes. The main proofs can be found in Appendix A5.

[14] The term single household refers to households with a single individual in them.

[15] It also rationalizes the calibration of Arslan, Guler and Taskin (2021) used in a different model of price search.

2 Class of Problems

This section defines the notation used throughout the paper, the class of problems under study, and the identified set.

2.1 Environment

Let $\mathcal{N} = \{1, \ldots, N\}$, $\mathcal{L}^k = \{1, \ldots, L^k\}$, $\mathcal{Z} = \{1, \ldots, Z\}$, and $\mathcal{T} = \{1, \ldots, T\}$ denote consumers, goods related to production that enter the utility function, goods related to production that do not enter the utility function, and periods for which data are observable, respectively. The index $k \in \mathcal{K} = \{1, \ldots, K\}$ refers to the constraint in the consumer problem such that L^k gives the number of variables related to production in constraint k. The sets $\mathcal{L}^k$ are mutually exclusive and, for convenience, I let $\mathcal{L} = \bigcup_{k \in \mathcal{K}} \mathcal{L}^k$ and $L = \sum_k L^k$.[16]

An observation for a consumer $i \in \mathcal{N}$ is $(\boldsymbol{F}^k_{i,t}, \boldsymbol{c}^k_{i,t}, \boldsymbol{a}^k_{i,t}, \boldsymbol{z}_{i,t})_{k \in \mathcal{K}} \subseteq \mathbb{R}^L_{++} \times \mathbb{R}^L_{++} \times \mathbb{R}^L_{++} \times \mathbb{R}^Z$, where the variable $\boldsymbol{F}$ is the observed output, the variable $\boldsymbol{c}$ differs from $\boldsymbol{a}$ in that $\boldsymbol{a}$ enters in some production function while $\boldsymbol{c}$ does not, and the variable $\boldsymbol{z}$ differs from $\boldsymbol{a}$ in that $\boldsymbol{a}$ enters in the utility function while $\boldsymbol{z}$ does not.[17] The production functions $F^k_{i,l} : \mathbb{R}^{L^k}_{++} \times \mathbb{R}^Z \times \mathbb{R}^{L^k}_{++} \to \mathbb{R}_{++}$ are absolutely continuous and have no variables in common. A data set for consumer i is denoted $x_i := \{(\boldsymbol{F}^k_{i,t}, \boldsymbol{c}^k_{i,t}, \boldsymbol{a}^k_{i,t}, \boldsymbol{z}_{i,t})\}_{k \in \mathcal{K}, t \in \mathcal{T}}$.

Let $u : \mathcal{C} \times \mathcal{A} \subseteq \mathbb{R}^L_{++} \times \mathbb{R}^L_{++} \to \mathbb{R}$ be a locally nonsatiated, pointwise monotonic, and continuous utility function. A utility function is pointwise monotonic if it is either increasing or decreasing in l for each $l \in \mathcal{L}$.[18] Define the set of all such utility functions by $\mathcal{U}$. Let $\mathcal{I}$ denote the set of characteristics defining the consumer preferences. From the econometrician point of view, a consumer has preferences represented by a random utility function $u_i : \mathcal{I} \to \mathcal{U}$, where consumers are assumed to be i.i.d. draws from $\mathcal{I}$.

2.2 The Consumer Problem

The class of problems considered in this paper is that of a set of a consumer behaving as if maximizing the following problem in every period $t \in \mathcal{T}$:

$$\max_{(\boldsymbol{c},\boldsymbol{a}) \in \mathcal{C} \times \mathcal{A}} u_i(\boldsymbol{c}, \boldsymbol{a}) \;\; s.t. \;\; \tilde{\boldsymbol{F}}^k_i(\boldsymbol{a}^k, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}^k_{i,t})' \boldsymbol{G}^k(\boldsymbol{c}^k) \leq 0 \quad k = 1, \ldots, K, \tag{1}$$

where $F^k_{i,l,t} := \tilde{F}^k_{i,l}(\boldsymbol{a}^k_{i,t}, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}^k_{i,t})$ and $\boldsymbol{G}^k$ is a vector of continuously differentiable functions $G^k_l : \mathcal{C}^k \subseteq \mathbb{R}^{L^k}_{++} \to \mathbb{R}$.[19] It is assumed that constraints hold with equality at the observed data such that $\tilde{\boldsymbol{F}}^k_i(\boldsymbol{a}^k_{i,t}, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}^k_{i,t})' \boldsymbol{G}^k(\boldsymbol{c}^k_{i,t}) = 0$. Moreover, it is assumed that

[16]That is, $\mathcal{L}^k \bigcap \mathcal{L}^{k'} = \emptyset$ for all $k \neq k' \in \mathcal{K}$.

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

[18]This allows the utility function to be increasing in some goods and decreasing in others.

[19]Similarly, we have $\mathcal{A}^k \subseteq \mathbb{R}^{L^k}_{++}$.

the functions $\boldsymbol{G}^k(\boldsymbol{c}^k)$ are known such that the model is specified. For concreteness, note that the standard consumer problem corresponds to $K = 1$, $\boldsymbol{F}_i(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}_{i,t}) = \boldsymbol{F}_t$, and $\boldsymbol{G}(\boldsymbol{c}_{i,t}) = \boldsymbol{c}_{i,t}$, where $\boldsymbol{F}_t$ is interpreted as prices and $\boldsymbol{c}_{i,t}$ is interpreted as consumption.[20]

Example 1. *(Price Search) Consider a model of price search similar to Aguiar and Hurst (2007) with a utility function $u(c, a)$ increasing in consumption (c), decreasing in search intensity (a), and concave. The budget constraint is $F(a, \boldsymbol{z}, \omega)'c = y$, where $F(a, \boldsymbol{z}, \omega)$ is a log-linear price function decreasing in search intensity, $\boldsymbol{z}$ are variables affecting prices paid such as shopping needs, and y is income.*

Example 2. *(Household Production) Consider a model of household production similar to Benhabib, Rogerson and Wright (1991) with a utility function given by $u(c_m, c_h, a_m, a_h)$ $= \log(c_m + c_h) + \alpha \log(1 - a_m + a_h)$, where c_m is the market good, c_h is the homemade good, a_m is the time spent working, a_h is the time spent working home, and $\alpha > 0$ is the value of leisure. Suppose the household can use a_m and a_h to obtain market and homemade goods, i.e. $c_m = wa_m$ and $c_h = F(a_h, \omega_h)$, where w is the wage and ω_h is the household productivity from home working.*

Example 3. *(Limited Attention) Consider a model of bounded rationality similar to Gabaix (2014). The consumer misperceives prices such that $F(a) = ap+(1-a)p^d$, where F is the perceived price, $a \in [0, 1]$ is the degree of attention, p is the actual price and p^d is the default price (e.g., average price). The budget constraint is given by $F(a)c \le y$, where c is consumption and y is income. The cost of attention is captured in reduced form through the utility function such that $u(c, a)$ is increasing in c and decreasing in a.*

Example 4. *(Human Capital) Consider a model of human capital similar to Ben-Porath (1967). The consumer maximizes discounted utility flows $\sum_{t=1}^{T} \delta^{(t-1)} u(c_t, 1-a_{1,t}-a_{2,t})$, where δ is the discount factor, $a_{1,t}$ is hours worked, and $a_{2,t}$ is human capital investment. Human capital evolves according to $z_{t+1} = (1 - d)z_t + F(a_{2,t}, z_t, \omega)$, where z_t is human capital, $d \in (0, 1]$ is depreciation, and ω is learning ability. The consumer budget is $\sum_{t=1}^{T} c_t \le \sum_{t=1}^{T} wz_t a_{1,t}$, where w is the wage.*

Example 5. *(General Equilibrium) Consider a general equilibrium model of consumption and labor choice. The firm maximizes profit $pF(a_1, a_2) - ra_1 - wa_2$, where p is the price of the output, a_1 is capital, a_2 is labor, r is the marginal return of capital, and w is the marginal return of labor. A representative consumer has a utility function $u(c, l - a_2)$, where c is consumption and $1 - a_2$ is leisure. The budget constraint is $c = F(a_1, a_2)$.*

[20] A constant such as income can always be included in the model.

2.3 Identified Set

For the sake of generality, suppose that some variables are mismeasured and let x_i^* denote the true data. Furthermore, let $\boldsymbol{\theta}$ denote the (possibly infinite) parameters of the production functions. Let the identified set be defined by

$$\Theta_I(x_i) := \{\boldsymbol{\theta} : \; x_i^* \text{ solves (1) for some } u_i \in \mathcal{U}\} .$$

In words, the identified set contains every set of parameters for which the true data set can be thought of as solutions of the consumer problem (1) for some well-behaved preferences. The production functions are said to be point identified if the identified set is single-valued such that $\Theta_I(x_i) = \{\boldsymbol{\theta}_0\}$, where $\boldsymbol{\theta}_0$ is the set of parameters of the true production functions. The next section provides general conditions for point identification that may serve as a benchmark, where identification is assumed to arise from $N \to \infty$.

3 Identification

This section introduces concepts and establishes conditions that are necessary for the point identification of the production functions. The estimation strategy proposed in the next section covers both point and set identification.

3.1 Point Identification

The first condition for point identification is that there be no measurement error in the data.

Assumption 1. *For all $i \in \mathcal{N}$, $x_i = x_i^*$.*

This assumption is quite common and necessary in the general setting considered here. That said, it is possible to achieve point identification in some models given a particular structure of measurement error. This is the case, for example, with exogenous additive measurement error in the dependent variable in a log-linear regression.[21]

The next assumption requires that the error term $\boldsymbol{\omega}_{i,t}$ be multiplicatively separable.

Assumption 2. *For all $k \in \mathcal{K}$ and $l \in \mathcal{L}^k$, the following holds:*

$$\tilde{F}_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}_{i,t}^k) = F_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) H_l^k(\boldsymbol{\omega}_{i,t}^k). \tag{2}$$

The separable structure of the error term includes Hicks-neutrality that is widely used in the production function literature. A special case covered by this assumption is

[21] In that case, identification is obtained for the model parameters. More generally, see Wang (2021) for an overview of identification in models with measurement error.

the Cobb-Douglas technology. Therefore, a common empirical specification covered by Assumption 2 is the log-linear regression. Also, note that while $\boldsymbol{F}_i^k$ is time-invariant, it can depend on variables that capture time effects. For example, the production function could have age or experience as variables $\boldsymbol{z}_{i,t}$.

The next restriction is for the functions $\boldsymbol{G}_i^k$ to be nonconstant functions of $\boldsymbol{c}$.

Assumption 3. *For all $k \in \mathcal{K}$ and $l \in \mathcal{L}^k$, $\frac{\partial G_l^k(\boldsymbol{c}^k)}{\partial c_l} \neq 0$.*

Assumption 3 requires that for any good $l \in \mathcal{L}$, there is at least one variable c_l unrelated to production. This assumption guarantees that the production function is invariant to monotone transformations of the utility function. That is, it ensures that identification only depends on preferences, not the arbitrary representation of those preferences. If Assumption 3 only holds for a subset of goods, then only the part of the production function associated with that subset of goods is identified.

Next, I define a notion of revealed preference for the class of models considered.

Definition 1. For any $t \in \mathcal{T}$, a bundle $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is said to be *revealed preferred* to a bundle $(\boldsymbol{c}, \boldsymbol{a})$ if $M_{i,t}^k(\boldsymbol{c}^k, \boldsymbol{a}^k) = \big(\boldsymbol{F}_i^k(\boldsymbol{a}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k)\big)'\boldsymbol{G}^k(\boldsymbol{c}^k) - \big(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k)\big)'\boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k) \leq 0$ for all $k \in \mathcal{K}$. If the inequality is strict for some $k \in \mathcal{K}$, then $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is said to be *strictly revealed preferred* to $(\boldsymbol{c}, \boldsymbol{a})$.

The next notion called the Strong Axiom of Revealed Preference (SARP) requires the revealed preference relation to be acyclic.

Definition 2. (SARP) If $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is revealed preferred to $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$, then $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$ is not revealed preferred to $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ unless $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) = (\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$.

The SARP is a ubiquitous consistency condition in revealed preference analysis that is necessary and sufficient for a data set to be thought of as arising from the maximization of a utility function.[22] This condition is a slight strengthening of the Generalized Axiom of Revealed Preference (GARP) that rules out the consumer problem to have multiple solutions. The next assumption requires the data set $(x_i)_{i \in \mathcal{N}}$ to be rational.

Assumption 4. *The random data set $(x_i)_{i \in \mathcal{N}}$ satisfies SARP.*

A necessary condition for Asssumption 4 to hold is that each individual data set x_i satisfies SARP. In the case where preferences are of the Gorman form (Gorman, 1959), the data $(x_i)_{i \in \mathcal{N}}$ can be treated as that of a representative consumer. Note that point identification is no longer guaranteed if Assumption 4 is relaxed to GARP. This is because indifference between multiple alternatives hinders one's ability to infer a consumer preferences. Indeed, indifference allows one to rationalize any choice pattern as the outcome of coin flips.

[22] In the next section, I show that SARP is indeed a necessary and sufficient condition for the consumer to pertain to the class of models considered in this paper.

Let $B_i^k := \left\{(\boldsymbol{c}^k, \boldsymbol{a}^k) \in \mathcal{C}^k \times \mathcal{A}^k : \left(\boldsymbol{F}_i^k(\boldsymbol{a}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k)\right)' \boldsymbol{G}^k(\boldsymbol{c}^k) \leq 0\right\}$ denote a budget set for constraint $k \in \mathcal{K}$ and $\mathcal{B}^k$ the set of all budget sets B_i^k. The budget sets are assumed to be comprehensive.[23] Let $\beta_1^k = B_1^k$, $\beta_2^k = B_1^k \bigcup B_2^k$, …, $\beta_N^k = \bigcup_{i=1}^N B_i^k$. The next assumption requires variation in budget sets to become rich enough when the sample size grows.

Assumption 5. $\lim\limits_{N\to\infty} \bigcup_{i=1}^N \beta_i^k = \mathcal{B}^k$ *for all* $k \in \mathcal{K}$.

Intuitively, budget variation is necessary as it allows one to learn how the consumer ranks alternatives against each other. Once preferences are identified, the first-order conditions provide nonparametric cross-equation restrictions that can be used to identify the production functions. Hence, the following result obtains.

Theorem 1. *Under Assumptions 1-5,* $\mathbb{E}[\boldsymbol{F}^k(\boldsymbol{a}_t^k, \boldsymbol{z}_t)]$ *is nonparametrically identified up to a function of* $\boldsymbol{z}_t$ *for all* $k \in \mathcal{K}$.

This theorem states that one can point identify the part of the production function that enters preferences in the population in a reasonably large class of problems. The level of variation needed for point identification can always be attained, at least in principle, when the sample becomes large enough.[24] The next assumption allows one to strengthen Theorem 1.

Assumption 6. $\mathbb{E}[\boldsymbol{h}^k(\boldsymbol{\omega}_t^k)|\boldsymbol{z}_t] = 0$, *where* $\boldsymbol{h}^k(\boldsymbol{\omega}_t^k) := \log\big(\boldsymbol{H}^k(\boldsymbol{\omega}_t^k)\big)$.

Assumption 6 amounts to conditional mean independence as assumed in ordinary least squares (OLS). To fix ideas, one can think of $\boldsymbol{z}_{i,t}$ as a discrete variable such as gender $\boldsymbol{z}_{i,t} \in \{\text{Male}, \text{Female}\}$. Assumption 6 requires that the expected error term be the same in either group. As with OLS, this assumption is reasonable if the econometrician is able to include enough variables $\boldsymbol{z}_t$ such as to avoid endogeneity issues.

The equality to zero is a normalization such that the only requirement is for the expectation to be invariant with respect to $\boldsymbol{z}_t$. In particular, note that the error may be correlated with $\boldsymbol{a}_{i,t}$. This is because additional information from the consumer problem is used to identify that part of the production functions. Finally, it is worth noting that Assumption 6 is testable. The following result states that the entire production function is identified when Assumption 6 holds.

Theorem 2. *Under Assumptions 1-6,* $\mathbb{E}[\boldsymbol{F}^k(\boldsymbol{a}_t^k, \boldsymbol{z}_t)]$ *is nonparametrically identified up to scale for all* $k \in \mathcal{K}$.

[23] For concreteness, assume that $\boldsymbol{c}^k$ enters positively in the utility function and that $\boldsymbol{a}^k$ enters negatively. Then, a budget B^k is said to be comprehensive if $\forall(\boldsymbol{c}^k, \boldsymbol{a}^k) \in \mathcal{C}^k \times \mathcal{A}^k$, if $\exists(\tilde{\boldsymbol{c}}^k, \tilde{\boldsymbol{a}}^k) \in B^k$, $\boldsymbol{c}^k \leq \tilde{\boldsymbol{c}}^k$ and $\boldsymbol{a}^k \geq \tilde{\boldsymbol{a}}^k$, then $(\boldsymbol{c}^k, \boldsymbol{a}^k) \in B^k$.

[24] In practice, it may be harder to achieve a dense set of budgets for larger class of budget sets in finite samples.

The first appeal of the identification results presented in this paper is to show that the part of the production function that enters preferences is identified and shielded from endogeneity issues. Thus, if one is only interested in that part of the production function there is no need for exogenous variation. That said, the part of the production function that is unrelated to preferences is subject to the same caveats as OLS. If one is interested in that part of the production function, it may be necessary to use standard exclusion restrictions to achieve point identification.

The second appeal of my results is to show that the structure of the consumer problem contains meaningful identification power. In particular, the restrictions implied by the model may still give useful information about the production functions even when point identification is not achieved. That is, the above assumptions give a benchmark for identification but are not necessary for informative inference. As such, the estimation strategy proposed in the next section provides a practical approach for the estimation of production functions that is applicable in both point and partially identified models.

Remark. *In the case where $T \to \infty$, point identification is obtained under much weaker requirements. Indeed, SARP would only need to hold on individual data sets. Furthermore, one would then be able to point identify individual production functions (up to a function of $\boldsymbol{z}_{i,t}$ without Assumption 6). Thus, while the previous results do not require access to panel data, the panel structure provides additional identification power that should be exploited when available.*

4 Estimation Strategy

In this section, I propose an estimation strategy that operationalizes the identification results. To this end, I derive nonparametric restrictions that exhaust the empirical content entailed by the consumer preferences. I then show how the model can be cast into a set of moment conditions susceptible to statistical inference. The methodology is applicable both in point and partially identified models.

4.1 Shape Restrictions from Preferences

In what follows, it will be useful to be precise about what it means for a data set to be rationalized by a utility function.

Definition 3. A utility function $u_i : \mathcal{C} \times \mathcal{A} \to \mathbb{R}$ rationalizes the data x_i if, for all $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ and $(\boldsymbol{c}, \boldsymbol{a})$, $M_t^k(\boldsymbol{c}^k, \boldsymbol{a}^k) \leq 0$ for all $k \in \mathcal{K}$ implies $u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \geq u_i(\boldsymbol{c}, \boldsymbol{a})$ and, if the inequality is strict for some $k \in \mathcal{K}$, then $u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) > u_i(\boldsymbol{c}, \boldsymbol{a})$.

This definition states that bundles that allow greater production or that are more "expensive" give a higher utility level. The previous section showed that one of the main substantive condition for identification is that the data set $(x_i)_{i\in\mathcal{N}}$ satisfies SARP. The following result shows that rationalizability is equivalent to SARP when the utility maximization problem yields a unique solution.[25]

Proposition 1. *For a given data set x_i, the following are equivalent:*

(i) The data set is rationalized by a locally nonsatiated and continuous utility function that yields a unique maximizer.

(ii) The data set satisfies SARP.

It is often reasonable and desirable to assume that the utility function is concave. The following result relates rationalizability by a concave utility function and conditions that can be checked in the data. Let $\odot$ denote the Hadamard product.

Theorem 3. *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 budgets $(B^k_{i,t})_{k\in\mathcal{K},t\in\mathcal{T}}$ are convex then (ii) implies (i).*

(i) The data set is rationalized by a locally nonsatiated, continuous, pointwise monotonic, and concave utility function.

(ii) There exist numbers $u_{i,t}$, $\lambda^k_{i,t} > 0$, $\dot{F}^k_{i,l,t}$, and $\omega_{i,l,t}$ such that, for all $s, t \in \mathcal{T}$, the following system of inequalities is satisfied

$$u_{i,s} \leq u_{i,t} + \sum_k \lambda^k_{i,t} \Big[\Big(\boldsymbol{F}^k_{i,t} \odot \nabla_c \boldsymbol{G}^k_i(\boldsymbol{c}^k_{i,t}) \Big)' (\boldsymbol{c}^k_{i,s} - \boldsymbol{c}^k_{i,t}) + \Big(\dot{\boldsymbol{F}}^k_{i,t} \odot \boldsymbol{H}^k(\boldsymbol{\omega}^k_{i,t}) \odot \boldsymbol{G}^k(\boldsymbol{c}^k_{i,t}) \Big)' (\boldsymbol{a}^k_{i,s} - \boldsymbol{a}^k_{i,t}) \Big],$$

where $\dot{F}^k_{i,l,t} \geq 0$ if the utility function is increasing in $a_{l,t}$ and $\dot{F}^k_{i,l,t} \leq 0$ otherwise.

The set of inequalities in Theorem 3 (ii) captures the concavity of the utility function, where the numbers $u_{i,t}$ and $\lambda^k_{i,t} > 0$ can be thought of as utility numbers and marginal utilities of "expenditure". The sign of $\dot{F}^k_{i,l,t}$ captures the monotonicity of the utility function with respect to $a^k_{i,l,t}$ and depends on the application.

Theorem 3 states that any data set rationalized by the model must satisfy Theorem 3 (ii) if the utility function is concave. Moreover, any data set that satisfies Theorem 3 (ii) is rationalized by the model if the budget sets are convex. In that case, Proposition 1 guarantees that the data set is consistent with SARP. Hence, these inequalities provide an operational way to use the empirical content lying in the consumer preferences for estimation purposes.

[25] Although the result is stated for an individual data set x_i, it also applies on $(x_i)_{i\in\mathcal{N}}$ since the latter can be viewed as the data set of a representative consumer.

In practice, there are many solutions to the inequalities in Theorem 3 (*ii*). These solutions are observationally equivalent in the sense that the data do not allow one to distinguish one from another. As a consequence, the set of solutions in Theorem 3 (*ii*) directly relates to the identified set.

Corollary 1. *If the inequalities in Theorem 3 (ii) are only necessary for the data to be rationalized by the model, then*

$$\Theta_I(x_i) \subset \{\boldsymbol{\theta} : x_i \text{ satisfies Theorem 3 } (ii)\}.$$

If the budget sets $(B_{i,t}^k)_{k\in\mathcal{K},t\in\mathcal{T}}$ *are convex, then the inequalities in Theorem 3 (ii) are also sufficient for the data to be rationalized by the model and*

$$\Theta_I(x_i) = \{\boldsymbol{\theta} : x_i \text{ satisfies Theorem 3 } (ii)\}.$$

Corollary 1 states that conservative bounds on the identified set can always be recovered via Theorem 3 (*ii*). Moreover, these bounds are sharp whenever the budget sets are convex.

4.2 Characterization via Moment Functions

Let $\mathcal{X} := \mathbb{R}_{++}^L \times \mathcal{C} \times \mathcal{A}$ 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.

For all $l \in \mathcal{L}$, $s, t \in \mathcal{T}$, and $k \in \mathcal{K}$, the class of models defined by Assumption 1-5 is captured by the following moment functions:

$$\begin{aligned} g_{i,s,t,k}^u(x_i, e_i) &:= \mathbb{1}\Bigg(u_{i,s} - u_{i,t} - \sum_k \lambda_{i,t}^k \Big[\Big(\boldsymbol{F}_{i,t}^k \odot \nabla_c \boldsymbol{G}_i^k(\boldsymbol{c}_{i,t}^k)\Big)' (\boldsymbol{c}_{i,s}^k - \boldsymbol{c}_{i,t}^k) \\ &\qquad + \Big(\boldsymbol{\rho}_{i,t}^k \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k)\Big)' (\boldsymbol{a}_{i,s}^k - \boldsymbol{a}_{i,t}^k) \le 0\Big]\Bigg) - 1, \\ g_{i,l,t,k}^\omega(x_i, e_i) &:= h_l^k(\boldsymbol{\omega}_{i,t}^k), \end{aligned}$$

where the first set of functions characterizes the concavity of the utility function and the last set of functions characterizes the normalization of the error term. The latent random variables further satisfy their support constraints: $\lambda_{i,t}^k > 0$ and $\boldsymbol{\rho}_{i,t}^k \le (\ge)\ 0$. Note that a given model may have additional moment functions on the production functions, error term, or measurement error as exemplified in the empirical application.

Let $\boldsymbol{g}_i(x_i, e_i) := (\boldsymbol{g}_i^u(x_i, e_i)', \boldsymbol{g}_i^\omega(x_i, e_i)')'$ denote the set of moments functions that characterize the model. Furthermore, let d_u and d_ω denote their respective number of constraints. Arbitrary combinations of these sets of functions are denoted with their superscripts bundled together. For example, $\boldsymbol{g}_i^{u,\omega}(x_i, e_i)$ is the set of functions on the utility function and the error term. 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.

4.3 Statistical 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.[26]

Definition 4. A random data set $x := \{x_i\}_{i=1}^N$ is statistically 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 statistically 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 4. *The following are equivalent:*

(i) *A random data set x is statistically rationalizable.*

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

where

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

and where $\eta(\cdot|x_i)$ is an arbitrary user-specified distribution function supported on $\mathcal{E}|\mathcal{X}$ such that $\mathbb{E}_{\pi_0}[\log(\mathbb{E}_\eta[\exp(\boldsymbol{\gamma}' \boldsymbol{g}_i^\omega(x_i, e_i))|x_i])]$ exists and is twice continuously differentiable in $\boldsymbol{\gamma}$ for all $\boldsymbol{\gamma} \in \mathbb{R}^{d_\omega}$.

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

[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.

In words, Theorem 4 (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 $\eta(\cdot|x_i)$ can also be achieved under μ. The dimensionality of the problem is then further reduced by noting that the concavity of the utility function only restricts the conditional support of the unobservables. Thus, one can draw from the conditional distribution $\tilde{\eta}(\cdot|x_i) := \mathbb{1}(\boldsymbol{g}_i^u(x_i, \cdot) = 0)\eta(\cdot|x_i)$ rather than leaving the moment functions $\boldsymbol{g}_i^u(x_i, \cdot)$ in the optimization problem.[29]

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

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

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 $\boldsymbol{g}_i^u(x_i, e_i)$ and can be achieved by rejection sampling. Then, a standard Metropolis-Hastings algorithm can be used to draw from the distribution.

4.4 Statistical Inference

The notion of statistical rationalizability together with Theorem 4 provides a feasible way of checking whether the data are consistent with the model. To statistically test the model in 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

$$\mathrm{TS}_N := N \inf_{\boldsymbol{\gamma}\in\mathbb{R}^{d_\omega}} \hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma})'\hat{\tilde{\boldsymbol{\Omega}}}^-(\boldsymbol{\gamma})\hat{\tilde{\boldsymbol{h}}}(\boldsymbol{\gamma})$$

is stochastically bounded by a χ^2 distribution with d_ω degrees of freedom ($\chi^2_{d_\omega}$).[30] As such, the 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_ω degrees of

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

[29] Parametric restrictions on the production functions also only restrict the support of the latent variables as they hold almost surely.

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

freedom. Inference can be made by adding moments on the parameters of interest:

$$r(x_i, e_i; \boldsymbol{\theta}) = \boldsymbol{\theta}_0,$$

where $r(\cdot)$ is some function that may depend on the data x_i and latent variables e_i. As before, this condition can be encapsulated in moment functions:

$$g_i^{\boldsymbol{\theta}}(x_i, e_i) := r(x_i, e_i; \boldsymbol{\theta}) - \boldsymbol{\theta}_0,$$

where $\boldsymbol{\theta}$ may be thought of as parameters of the production functions. A conservative 95% confidence set on $\boldsymbol{\theta}_0$ can be obtained by inverting the test statistic:

$$\{\boldsymbol{\theta}_0 : \mathrm{TS}_N(\boldsymbol{\theta}_0) \leq \chi^2_{d_\omega+1,0.95}\},$$

where $TS_N(\boldsymbol{\theta}_0)$ is the test statistic at a fixed value of $\boldsymbol{\theta}_0$.

5 Application to Price Search

This section considers a flexible model of price search as a means to clarify how the methodology proposed in this paper may be used in practice as well as to show the feasibility and effectiveness of the approach.

5.1 Model

The consumer is assumed to know her realizations of search productivity ($\boldsymbol{\omega}$) and to choose consumption ($\boldsymbol{c}$) and shopping intensity ($\boldsymbol{a}$) accordingly. Formally, a consumer behaves as if maximizing her utility function subject to satisfying her budget constraint:

$$\max_{(\boldsymbol{c}_i, \boldsymbol{a}_i) \in \mathcal{C}^T \times \mathcal{A}^T} u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) + y_i \quad \text{s.t.} \quad \boldsymbol{p}_i(\boldsymbol{a}_{i,t}, \boldsymbol{\omega}_{i,t})' \boldsymbol{c}_{i,t} = y_i, \tag{3}$$

where $u_i : \mathcal{C} \times \mathcal{A} \to \mathbb{R}$ is a utility function that is continuous, concave, strictly increasing in consumption, and decreasing in shopping intensity, $\boldsymbol{p}_i(\boldsymbol{a}, \boldsymbol{\omega}_{i,t})$ is a vector of continuously differentiable good-specific price functions $p_{i,l} : \mathbb{R}^L_{++} \times \mathbb{R}^L \to \mathbb{R}_{++}$ where $\boldsymbol{p}_{i,t} := \boldsymbol{p}_i(\boldsymbol{a}_{i,t}, \boldsymbol{\omega}_{i,t})$, and y_i is income. The econometrician only observes the data set $x_i := \{(\boldsymbol{p}_{i,t}, \boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})\}_{t \in \mathcal{T}}$.

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 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 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 $\mathcal{B}_i := \{(\boldsymbol{c}, \boldsymbol{a}) \in \mathcal{C} \times \mathcal{A} : \boldsymbol{p}_i(\boldsymbol{a}, \boldsymbol{\omega}_i)'\boldsymbol{c} \leq y_i\}$. This set is represented by the shaded area in Figure 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.[31]

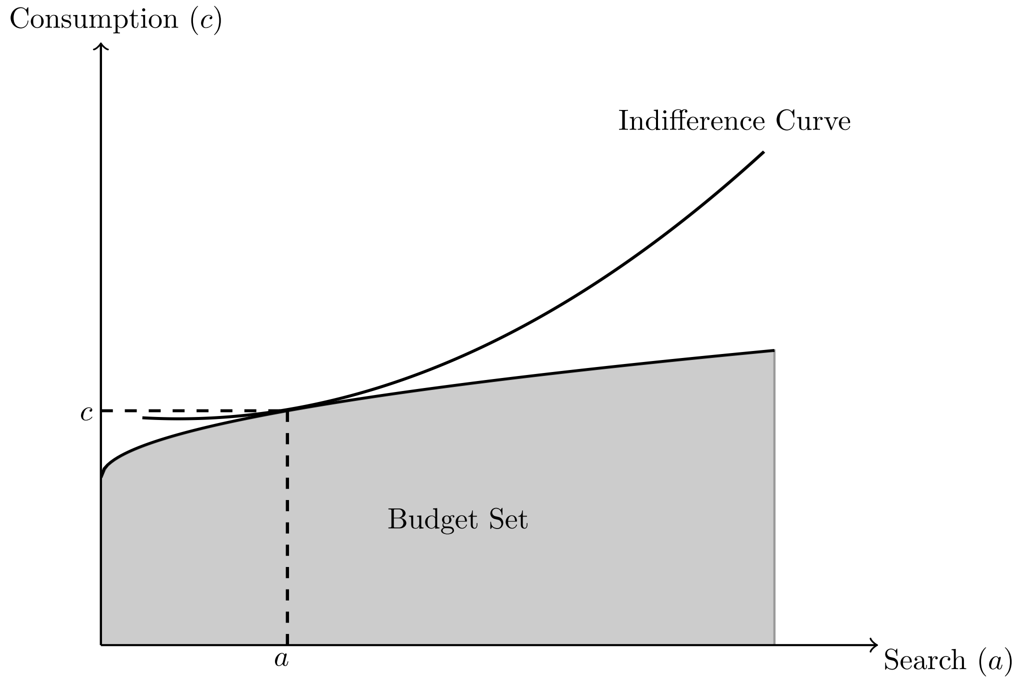

Figure 1: Optimal Choice with Price Search

It is important to note that the quasilinearity assumption could be relaxed in the model.[32] The assumption is motivated by the fact that the data span a period of six months for which changes in income are negligible in my empirical application. Moreover, the data focus on food consumption which tends to be income inelastic.[33] Since quasilinearity is reasonable in this particular data set, it will be useful to impose it in my application. This is because the quasilinear structure provides a useful measure of utility in terms of dollar, hence allowing me to get insights on search costs.

[31]In Appendix A1, I show that my model can be extended to home production and relates it to that of Aguiar and Hurst (2007).

[32]For example, static utility maximization would be more appropriate if the data set spanned a few years such that income varied over time.

[33]Quasilinearity is also used by Echenique, Lee and Shum (2011) and Allen and Rehbeck (2020) in a similar scanner data set on food expenditures.

5.2 Environment

Given the fundamental unobservability of preferences, it is preferable that assumptions on the utility function remain minimal. On the contrary, the production functions are partially observed. Indeed, one has data on prices paid for many values of shopping intensity. As such, one should feel comfortable making more stringent assumptions on the latter.

To gain insights on the behavior of the price functions, Figure 2 displays how log prices averaged across consumers (henceforth, expected prices) vary with log number of shopping trips in the data, where shopping trips capture shopping intensity.

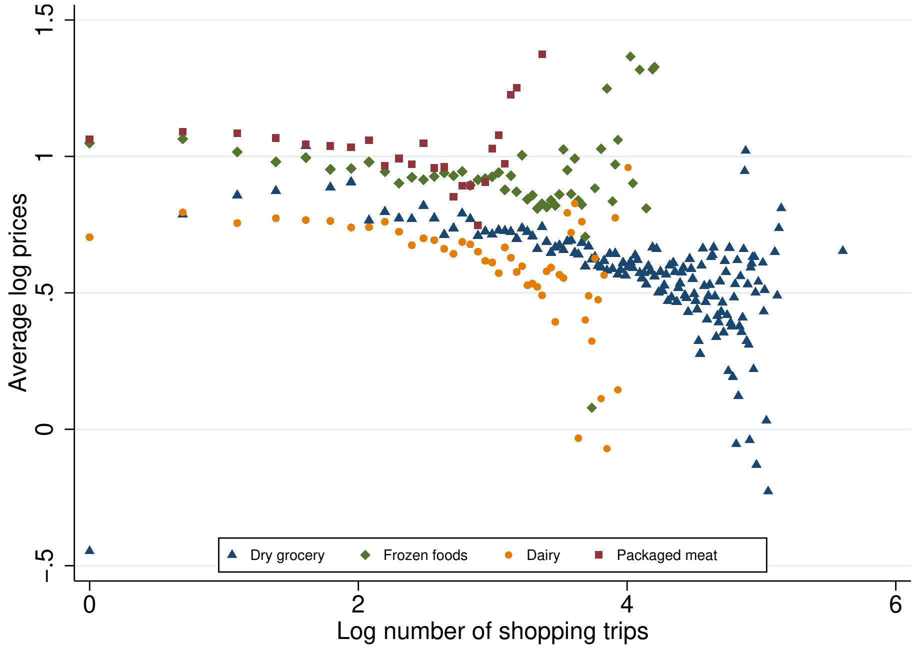

Figure 2: Average Log Price by Log Number of Shopping Trips

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

Consistent with the price search hypothesis, Figure 2 shows a negative relationship between prices paid and shopping intensity. Moreover, we can see that the change in expected log prices as log shopping trips increase can be approximated by a linear function. Accordingly, I follow the price search literature and assume that the price functions are log-linear in shopping intensity.[34]

Assumption 7. *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

[34] See, for example, Aguiar and Hurst (2007) and Arslan, Guler and Taskin (2021).

respect to shopping intensity.

Assumption 7 implies that prices paid decrease at a decreasing rate as shopping intensity or search productivity increases. This requirement captures decreasing marginal returns that arise due to the increasing difficulty of finding discounts surpassing the current best discount.[35] An example of budget set generated by a log-linear price function is illustrated in Figure 1 in the case of a single good. The figure shows that the marginal increase in consumption from lower prices paid decreases as shopping intensity increases.

Conditional on the log-linear specification implied by Assumption 7, 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. 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. This issue is largely mitigated in my application as goods are aggregated to coarse categories.

A look at Figure 2 shows that a log-linear relationship does not hold perfectly for any given good. These deviations are normal in any data set and accounted for by search productivity ($\omega_{i,l}$) in the price functions. It is possible that some consumers that go on many shopping 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. That is, consumers that go on more shopping trips may also purchase more expensive goods.

Although the log-linear relationship is imperfect for any given good, Figure 2 shows that it fits well across goods. That is, one is able to fit a line almost perfectly by averaging expected log prices across goods. Given Assumption 7, this implies that the expected average search productivity is time-invariant. In other words, the unobserved effects of search productivity on prices paid cancel out on average.

Assumption 8. *For all $t \in \mathcal{T}$, $\mathbb{E}[\overline{\omega}_t] = 0$, where $\overline{\omega}_{i,t} := L^{-1} \sum_{l=1}^{L} \omega_{i,l,t}$ denotes the average search productivity across goods.*

Assumption 8 allows search productivity to vary for each individual and each good as long as the overall search productivity remains constant.[36] 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 differ-

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

[36] The equality of the expected average search productivity to zero is a normalization and is thus without loss of generality.

ent baskets of goods in different time periods, prices may vary due to variations in the composition of the baskets of goods.

Other than for this mild centering condition, Assumption 8 is quite general as it does not presume anything about the underlying stochastic process of search productivity. Conditional on the expected 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.[37]

Given the log-linearity of the price functions, Assumption 8 implies that expected log prices should be around the mean, conditional on shopping trips. In accordance with this prediction, Figure 3 shows that the distribution of expected log prices is centered around its mean. I report the unconditional distribution for expositional purposes; similar shapes are obtained for the conditional ones.[38]

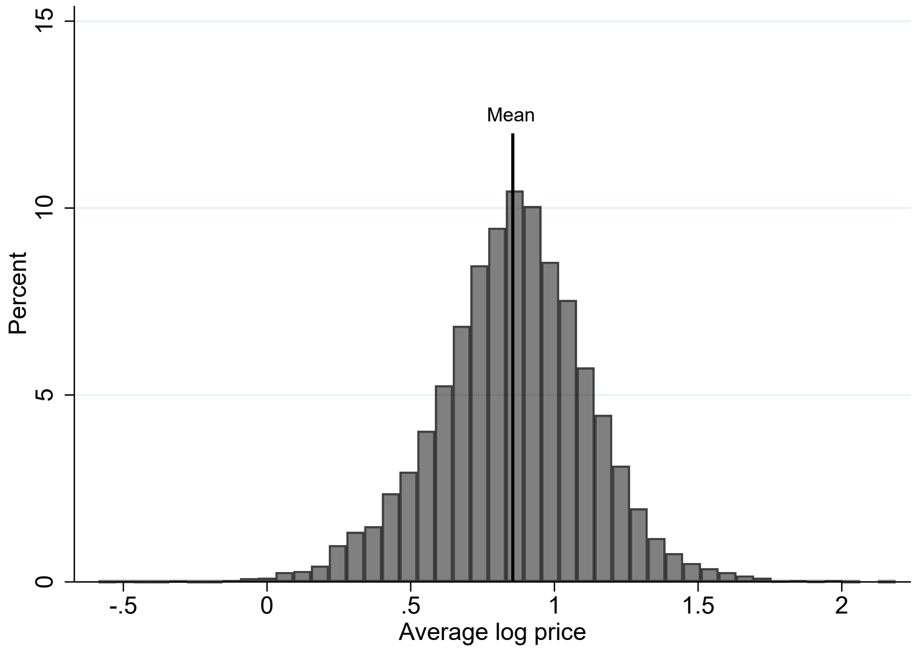

Figure 3: Distribution of Average Log Prices

Another reason why the data may not exhibit a perfect log-linear relationship is the presence of measurement error. Indeed, Einav, Leibtag and Nevo (2010) use transactions from a large retailer in order to document the extent of measurement error in the Nielsen Homescan Dataset on which Figure 2 is derived. They show that measurement error in prices is severe and document that the difference between observed log prices and true

[37] See, for example, Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2015), and Gandhi, Navarro and Rivers (2020).

[38] Consistent with Assumption 7, distributions conditioned on larger values of shopping trips tend to be centered around lower values of log prices.

log prices is zero.[39] This leads to the following assumption.

Assumption 9. *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 9 says that, in expectation, observed log prices and true log prices are the same for each good and time period. Together, they yield a total of $L \cdot T$ moments on measurement error.

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

Assumption 10. *For all $l \in \mathcal{L}$, $\alpha^1_{i,l} \in [-1, 0]$.*

Assumption 10 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 price with respect to shopping intensity using the Homescan 1993-1995.[40] As such, Assumption 10 should give enough flexibility for the needs of the data.

Under the previous assumptions, the concavity of the utility function and the log-linear price functions can be refuted by the data. Intuitively, to see why price search is refutable, note that Assumptions 7-8 imply that the average expected log price paid must decrease whenever shopping intensity increases. Therefore, inconsistencies with price search arise whenever this relationship is violated in the data.[41]

5.3 Price Search Rationalizability

The environment defined in the previous section implies that for all $l \in \mathcal{L}$ and $s, t \in \mathcal{T}$, the model is characterized by the following moment functions:

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

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:

[39] Additional details about the data and measurement error are given in Section 6.

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

[41] See Appendix A2 for analytical power results.

$\boldsymbol{\alpha}_i^1 \in [-1,0]$, $\boldsymbol{\rho}_{i,t} \leq 0$ and $\boldsymbol{p}_{i,t}^* > 0$, where $\rho_{i,l,t}$ further satisfies

$$\rho_{i,l,t} = \alpha_{i,l}^0 \alpha_{i,l}^1 a_{i,l,t}^{\alpha_{i,l}^1 - 1} e^{-\omega_{i,l,t}} c_{i,l,t} \quad \forall l \in \mathcal{L}.$$

This equality constraint implies that $\boldsymbol{\rho}_{i,t}$ is completely determined by the data and latent variables $(u_{i,t}, \boldsymbol{\alpha}_i, \boldsymbol{\omega}_{i,t}, \boldsymbol{m}_{i,t})_{t \in \mathcal{T}}$. Every consumer has a total of $T^2 + L \cdot T + L \cdot T + T$ moment functions, written as $\boldsymbol{g}_i(x_i, e_i) := (\boldsymbol{g}_i^u(x_i, e_i)', \boldsymbol{g}_i^p(x_i, e_i)', \boldsymbol{g}_i^m(x_i, e_i)', \boldsymbol{g}_i^\omega(x_i, e_i)')'$ for short. Arbitrary combinations of these sets of functions are denoted with their superscripts bundled together. For example, $\boldsymbol{g}_i^{m,\omega}(x_i, e_i)$ is the set of functions on measurement error and search productivity.

The moment functions allow me to define the statistical rationalizability of a data set with respect to the model of price search.

Definition 5. Under Assumptions 7-10, 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.

Conditional on the data being consistent with PS-rationalizability, the next step is to make inference on parameters of interest. First, I show that inference on the true expected elasticity of price with respect to shopping intensity is possible in the model.

Proposition 2. *The true expected elasticity of price with respect to shopping intensity is given by*

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

where the line over $\frac{\partial \log(p_t^*)}{\partial \log(a_t)}$ *and* α^1 *denote the average across goods.*

Proposition 2 states that the expected effect of an increase in shopping intensity on the price paid can be recovered from data on prices and search intensity. The reason why it can be achieved in the model is that Assumption 8 restricts the expected (average) search productivity to be time-invariant. Therefore, any variation in expected prices must be caused exclusively by variations in shopping intensity.

Next, I show that the true expected search cost can be recovered.

Proposition 3. *The true expected search cost is given by*

$$\mathbb{E}\left[T^{-1}L^{-1}\sum_{l,t} u(\boldsymbol{c}_t, \boldsymbol{a}_t) - u(\boldsymbol{c}_t, \boldsymbol{a}_{-l,t}, 1_{l,t})\right] = \mathbb{E}\left[\overline{(p_l(a_{l,t}, \omega_{l,t}) - p_l(1, \omega_{l,t}))\, c_{l,t}}\right],$$

where $\boldsymbol{a}_{-l,t}$ denote the shopping intensity on every good but good l, $1_{l,t}$ denote the shopping intensity on good l in period t, and the line over $(p_l(a_{l,t}, \omega_{l,t}) - p_l(1, \omega_{l,t}))\, c_{l,t}$ denote the average across goods and time periods.

In words, Proposition 3 states that the expected disutility of shopping is equal to the expected savings from search. Note that the quasilinear structure of the utility function allows me to interpret the disutility of search in dollars. Additional details about the implementation are given in Appendix A3.

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 9. Likewise, exogenous shocks can be absorbed by search productivity provided they satisfy Assumption 8. Accordingly, the model is robust to a variety of perturbations.*

6 Data

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

6.1 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.[42] This selection leaves over a million distinct UPCs representing

[42] This choice implicitly assumes that food is weakly separable from other categories of goods. This

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". Since the number of moments increases multiplicatively with the number of goods in my application, this level of aggregation will also ensure that the optimization problem remains tractable. The resulting aggregated prices are calculated as the 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.[43]

The data set focuses on households that satisfy the above criteria, participated in the Homescan from April to September of the panel year 2011, and whose head household is at least 50 years old such as to exclude potential online shoppers. The final sample contains X consumers, 4 aggregated goods, and 6 monthly time periods. Additional details about the construction of the data set are provided in Appendix A4.

6.2 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 changes 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, Leibtag and Nevo (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

assumption is empirically plausible (Cherchye et al., 2015b), especially when the presence of measurement error is recognized (Fleissig and Whitney, 2008; Elger and Jones, 2008).

[43] This also avoids imputing prices of zero consumption goods that would overlook the full heterogeneity in prices assumed in the model.

measurement error: $\boldsymbol{m}_{i,t} := \log(\boldsymbol{p}_{i,t}) - \log(\boldsymbol{p}^*_{i,t})$ for all $t \in \mathcal{T}$.[44] Using price data from a large retailer, Einav, Leibtag and Nevo (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.

7 Empirical Results

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

7.1 Price Search Rationalizability

By applying the above methodology to the data, I find that PS-rationalizability is not rejected by the data at the 95% confidence level among single households. More precisely, I obtain a test statistic of 36.38, which is below the chi-square critical value of 43.77. In contrast, I find that PS-rationalizability is rejected by the data in couple households as well as for households of many members.[45]

The rejection of the model in multi-person households is to be expected given the assumption that they behave as a single entity that maximizes a single utility function. Indeed, there is mounting evidence in the literature that multi-person households do not make choices as a single decision maker.[46] That is, the methodology successfully detects modelling assumptions that are inconsistent with the data. This is a desirable outcome as one of the methodology's goals is to alleviate the odds to inadvertently undergo an empirical analysis suffering from misspecification issues.

7.2 Empirical Results

Since the model is not rejected by the data on single households, I can invert the statistical test to obtain a 95% confidence set on the expected elasticity of price with respect to shopping intensity. Doing so, I obtain a confidence set of $[-0.2, -0.1]$. Thus, a doubling of shopping intensity decreases prices paid by about 15% on average.

Despite the important role of price search in reducing expenditures, it only provides one side of the story. Indeed, it does not account for the shadow utility cost of searching.

[44] 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.

[45] The test statistic is 443.38 and X, respectively.

[46] See, for example, Thomas (1990), Fortin and Lacroix (1997), Browning and Chiappori (1998), Cherchye and Vermeulen (2008), Cherchye, Demuynck and De Rock (2011), and Cherchye et al. (2020).

Thus, I next report the log expected search cost such as to recover the percentage increase in search costs over the support of the data.[47] Since log search costs are not defined when equal to zero, which happens when the number of shopping trips equals one, I remove such observations when computing the confidence set.[48] At this point, it is useful to note that consumers going on few shopping trips are likely to have high search costs. Therefore, removing these observations may create a slight downward bias in the expected log search cost.

Using data on single households, I find that the 95% confidence set on the absolute value of expected log search cost is $[1.75, 4.75]$. In other words, consumers search costs increase by about 175% to 475% within the support of the data.

Lastly, the 95% confidence set on the expected marginal search cost (in dollar) is $[-1.25, -3.5]$ at the optimum. Since search costs are weakly increasing in shopping intensity, multiplying the expected marginal search cost at the optimum by the average number of shopping trips gives an upper bound on the expected total search cost. Doing so, I obtain that the expected total search cost is at most 18.5 to 51.8 dollars, or 43 to 120 percent of the observed average expenditure.

7.3 Discussion

In this study, I find that a doubling of shopping trips decreases prices paid by about 15% on average. It is worthwhile to note that my confidence set covers the estimate of Aguiar and Hurst (2007) obtained using an instrumental variable approach.[49] This suggests that the quasilinearity assumption made in the model is reasonable in the data. Similarly, the compatibility of their results with mine provides further evidence that the exogeneity assumption underlying the identification strategy in Aguiar and Hurst (2007) is satisfied, and that measurement error does not create sizable bias in their methodology.[50]

When consumers differ in their price search, it induces heterogeneity in prices paid. In turn, this implies that expenditure gives an erroneous account of consumption under the assumption that prices paid are homogeneous. Building on this insight, Arslan, Guler and Taskin (2021) show that consumption inequality is significantly smaller than expenditure inequality using data from the Homescan 2004. My results show that the monetary savings from price search are partially offset by the utility cost of searching. Thus, one should also recognize the disutility incurred from search to obtain a complete

[47] The log expected search cost is better disciplined than search cost since Assumption 9 imposes restrictions on log prices.

[48] About 6.5% of observations have a number of shopping trips equal to one.

[49] Their estimate of the elasticity of price with respect to shopping intensity is -0.19 using age as an instrument, -0.07 using either income or household size as an instrument, and -0.10 using all three instruments simultaneously.

[50] Recall that measurement error in the dependent variable does not cause bias if measurement error is classical.

picture of inequalities.

8 Conclusion

This paper shows that production functions are nonparametrically identified in a large class of consumer models. This novel result takes advantage of restrictions implied by utility maximization and cross-equation restrictions from the first-order conditions. I provide a simple estimation strategy for the estimation of production functions that is applicable in both point and partially identified models. In my empirical application, I find that search costs more than double within the support of the data, hence showing their importance when evaluating the benefits of price search. The quantitative size of search costs may have implications on within- and between-group inequalities as well as on the level of competition in a market. I leave this to future work.

Appendix

A1: 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).

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 $f(h, \boldsymbol{c})$.[51] The household's problem therefore becomes

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

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

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

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

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

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

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

and we have thereby recovered a model with the same implications to that of Aguiar and Hurst (2007).[52] To see why, assume the solution is interior and take the first-order

[51] 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.

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

conditions:

$$\frac{\partial f}{\partial c} = \lambda_t \boldsymbol{p}(\boldsymbol{a}, \boldsymbol{\omega}_t)$$
$$\boldsymbol{\mu} = \lambda_t \frac{\partial \boldsymbol{p}(\boldsymbol{a}, \boldsymbol{\omega}_t)}{\partial \boldsymbol{a}} \odot \boldsymbol{c}$$
$$\mu = -\frac{\partial f}{\partial h}.$$

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 can be used to identify the home production function, a point that was cleverly exploited by Aguiar and Hurst (2007) in a parametric setting.

A2: Power Analysis

In this section, I show that price search and utility maximization are both refutable under Assumptions 7-10. I then provide empirical evidence that these additional restrictions are not necessary for the model to be rejected by the data.

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 7:

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

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

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

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.[53]

Falsifiability of Price Search

[53] A function f is logarithmically convex if the composition of the logarithm with f is itself a convex function.

Suppose that Assumptions 7-10 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

$$\log\left(\frac{p(a_{l,s},\omega_{l,s})}{p(a_{l,t},\omega_{l,t})}\right) \geq \frac{\nabla_a p(a_{l,t},\omega_{l,t})}{p(a_{l,t},\omega_{l,t})}(a_{l,s}-a_{l,t}) + \frac{\nabla_\omega p(a_{l,t},\omega_{l,t})}{p(a_{l,t},\omega_{l,t})}(\omega_{l,s}-\omega_{l,t}).$$ [54]

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

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

where $\overline{\omega}_t := \frac{1}{L}\sum_{l=1}^{L}\omega_{l,t}$ 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

$$0 > \frac{1}{L}\sum_{l=1}^{L}\left(\mathbb{E}\left[\log(p_{l,1})\right] - \mathbb{E}\left[\log(p_{l,2})\right]\right) \geq -\frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\left[\frac{\rho_{l,2}}{p^*_{l,2}c_{l,2}}\right]. \tag{4}$$

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 (4) yields a contradiction. In other words, the data are inconsistent with the model provided the price functions are log-linear.

Falsifiability of Utility Maximization

Suppose that Assumptions 7-10 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

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

[54]Note that this expression is well-defined since prices are strictly positive.

For concavity to be analytically refuted, it is clear that Assumption 9 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 7-10 provided $\mathbb{E}[\boldsymbol{p}_t] = \mathbb{E}[\boldsymbol{p}_t^*]$ for all $t \in \mathcal{T}$ (instead of Assumption 9) and $\delta = 1$ almost surely.

Falsifiability of the Model: Empirical Evidence

I have shown analytically that the model defined by Assumptions 7-10 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 9-10 if the price functions are convex decreasing.[55] 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.

[55] 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 $\boldsymbol{y}$.

A3: 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.

Pseudo-Code

Step 1

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

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 3. 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

A4: 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

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

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.[56]

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.

[56]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.

A5: Proofs

Proof of Theorem 1 and 2

Lemma 1. *Under Assumptions 1-6, if the utility function is known up to a monotone transformation, then* $\mathbb{E}[\boldsymbol{F}(\boldsymbol{a}_t^k, \boldsymbol{z}_t)]$ *is nonparametrically identified up to scale.*

The first-order conditions of the consumer problem (1) with respect to $\boldsymbol{c}$ and $\boldsymbol{a}$ for any constraint $k \in \mathcal{K}$ is

$$\nabla_{c^k} u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t}^k F_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) H_l^k(\boldsymbol{\omega}_{i,t}^k) \frac{\partial G_l^k(\boldsymbol{c}_{i,t}^k)}{\partial c_{i,l,t}^k} \tag{5}$$

$$\nabla_{a^k} u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t}^k \frac{\partial F_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} H_l^k(\boldsymbol{\omega}_{i,t}^k) G_l^k(\boldsymbol{c}_{i,t}^k) \quad \forall l \in \mathcal{L}^k, \forall k \in \mathcal{K}, \tag{6}$$

where $\nabla u(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ denotes a supergradient of u at the point $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \in \mathcal{C} \times \mathcal{A}$.[57] Dividing (5) by (6) and rearranging yields

$$\frac{\partial f_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} = \frac{\nabla_{a^k} u(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l}{\nabla_{c^k} u(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l} \cdot \frac{\partial g_{i,l}^k(\boldsymbol{c}_{i,t}^k)}{\partial c_{i,l,t}^k} \quad \forall l \in \mathcal{L}^k, \forall k \in \mathcal{K}, \tag{7}$$

where $f_{i,l}^k := \log\Big(F_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t})\Big)$ is absolutely continuous and $g_{i,l}^k := \log\Big(G_{i,l}^k(\boldsymbol{c}_{i,t}^k)\Big)$.[58] Since the marginal rate of substitution (MRS) is invariant to monotone transformations of the utility function, the derivative of the log production function is invariant to such transformations.

Given knowledge of the utility function up to a monotone transformation, equation (7) immediately identifies the derivative of the log production function from the data. Importantly, the left hand side defines a differential equation. Thus, by the fundamental theorem of calculus and since $\boldsymbol{a}$ is a continuous variable, I can integrate the differential equation with respect to a for good l of constraint k, thus yielding

$$\int_{\underline{a}_{i,l}^k}^{a_{i,l,t}^k} \frac{\partial f_{i,l}^k(\boldsymbol{a}_{i,-l,t}^k, a, \boldsymbol{z}_{i,l,t})}{\partial a} \, da = f_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,l,t}) + C_{i,l}^k(\boldsymbol{a}_{i,-l,t}^k, \boldsymbol{z}_{i,l,t}) \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{K}, \tag{8}$$

where $0 < \underline{a}_{i,l}^k < a_{i,l,t}^k$. As Gandhi, Navarro and Rivers (2020) note, the above equations can be combined to obtain the following equality:

$$f_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,l,t}) = \sum_{l=1}^{L} \left(\int_{\underline{a}_{i,l}^k}^{a_{i,l,t}^k} \frac{\partial f_{i,l}^k \left(\boldsymbol{a}_{i,l' \leq l,t}^k, \underline{\boldsymbol{a}}_{i,l'>l}^k, a, \boldsymbol{z}_{i,l,t} \right)}{\partial a} \, da \right) + C_{i,l}^k(\boldsymbol{z}_{i,l,t}). \tag{9}$$

[57] Let $m \in \mathbb{N}$ and $f : \mathbb{R}^m \to \mathbb{R}$ be a concave function. A vector $\mathbf{g}$ is a supergradient of f at $\boldsymbol{y} \in \mathbb{R}^m$ if for every $\boldsymbol{x} \in \mathbb{R}^m$ it satisfies $f(\boldsymbol{x}) \leq f(\boldsymbol{y}) + \mathbf{g}'(\boldsymbol{x} - \boldsymbol{y})$.

[58] This is because the composition of a continuously differentiable function and an absolutely continuous function is absolutely continuous and the fact that $\log(\cdot)$ is continuously differentiable on $\mathbb{R}_{++}$.

Next, subtract equation (9) from $f_{i,l,t}^k$ to get

$$f_{i,l,t}^k - \sum_{l=1}^{L} \left(\int_{\underline{a}_{i,l}^k}^{a_{i,l,t}^k} \frac{\partial f_{i,l}^k \left(\boldsymbol{a}_{i,l' \leq l,t}^k, \underline{\boldsymbol{a}}_{i,l'>l}^k, a, \boldsymbol{z}_{i,l,t} \right)}{\partial a} da \right) = C_{i,l}^k(\boldsymbol{z}_{i,l,t}) - h_l^k(\boldsymbol{\omega}_{i,t}^k), \tag{10}$$

where $h_l^k(\boldsymbol{\omega}_{i,t}^k) = \log\Big(H_l^k(\boldsymbol{\omega}_{i,t}^k)\Big)$. This equation shows that the constant of integration and the log function of the error term are not separately identified.

Without loss of generality, normalize the latter such that $\mathbb{E}[h_l^k(\boldsymbol{\omega}_{l,t}^k)] = 0$. Taking the expectation of equation (10) gives $\mathbb{E}[C_l^k(\boldsymbol{z}_{i,t})]$. Then, differencing equation (10) between $t, t' \in \mathcal{T}$ after expectation identifies the change $\mathbb{E}[h_l^k(\boldsymbol{\omega}_t^k) - h_l^k(\boldsymbol{\omega}_{t'}^k)]$. Therefore, once $\mathbb{E}[h_l^k(\boldsymbol{\omega}_t^k)] = 0$ is normalized for some $t \in \mathcal{T}$, no additional normalization is required. Finally, taking the expectation of equation (9) and then the exponential function yields the expected production function up to a function of $\boldsymbol{z}_t$.

To obtain Theorem 2, simply note that by conditioning on $\boldsymbol{z}_t$ in equation (10) and by assuming $\mathbb{E}[h_l^k(\boldsymbol{\omega}_{l,t}^k)|\boldsymbol{z}_t] = 0$, the constant of integration is identified since the left-hand side is a known quantity. Then, only the constant in the constant of integration changes with the normalization of the error term. Thus, the whole production function is nonparametrically identified up to scale.

Next, I show that the preferences of a representative consumer are identified under Assumptions 1-5. First, note that WARP is implied by GARP and that GARP is in fact SARP once the choice correspondence $h : \mathcal{B} \to \mathcal{C} \times \mathcal{A}$ is assumed single-valued, where $\mathcal{B} := \bigtimes_{k \in \mathcal{K}} \mathcal{B}^k$. Second, note that the utility function is continuous and the budget sets are compact and continuous.[59] Thus, by the maximum theorem the choice function h is continuous. Finally, note that h is trivially closed since it is a function.

Consistent with the assumption that preferences are pointwise monotonic, the choice function is also assumed pointwise monotonic. The choice function h is pointwise monotonic if, for any $k \in \mathcal{K}$, any $B^k \in \mathcal{B}^k$, any $l \in \mathcal{L}^k$, and any $(c_l^k, a_l^k) \in \mathbb{R}_{++}^2$, one of the following is true:

$$[(c_{1,l}^k, a_{1,l}^k) \in h(B^k), c_{2,l}^k > c_{1,l}^k, a_{2,l}^k > a_{1,l}^k] \implies [(c_{2,l}^k, a_{2,l}^k) \notin B^k] \tag{11}$$
$$[(c_{1,l}^k, a_{1,l}^k) \in h(B^k), c_{2,l}^k > c_{1,l}^k, a_{2,l}^k < a_{1,l}^k] \implies [(c_{2,l}^k, a_{2,l}^k) \notin B^k] \tag{12}$$
$$[(c_{1,l}^k, a_{1,l}^k) \in h(B^k), c_{2,l}^k < c_{1,l}^k, a_{2,l}^k > a_{1,l}^k] \implies [(c_{2,l}^k, a_{2,l}^k) \notin B^k] \tag{13}$$
$$[(c_{1,l}^k, a_{1,l}^k) \in h(B^k), c_{2,l}^k < c_{1,l}^k, a_{2,l}^k < a_{1,l}^k] \implies [(c_{2,l}^k, a_{2,l}^k) \notin B^k]. \tag{14}$$

This generalization of monotonicity may appear odd and unnecessary at first but it is in fact quite natural for a choice function to be monotone decreasing for a subset of

[59] For budget sets that are unbounded, it is without loss of generality to take a finite support that includes the unique interior solution.

variables.[60] A preference relation $\succeq$ is pointwise monotonic if, for any $k \in \mathcal{K}$ and any $l \in \mathcal{L}^k$, one of the following is true:

$$[c^k_{1,l} > c^k_{2,l}, a^k_{1,l} > a^k_{2,l}] \implies (c^k_{1,l}, a^k_{1,l}) \succ (c^k_{2,l}, a^k_{2,l}) \tag{15}$$
$$[c^k_{1,l} > c^k_{2,l}, a^k_{1,l} < a^k_{2,l}] \implies (c^k_{1,l}, a^k_{1,l}) \succ (c^k_{2,l}, a^k_{2,l}) \tag{16}$$
$$[c^k_{1,l} < c^k_{2,l}, a^k_{1,l} > a^k_{2,l}] \implies (c^k_{1,l}, a^k_{1,l}) \succ (c^k_{2,l}, a^k_{2,l}) \tag{17}$$
$$[c^k_{1,l} < c^k_{2,l}, a^k_{1,l} < a^k_{2,l}] \implies (c^k_{1,l}, a^k_{1,l}) \succ (c^k_{2,l}, a^k_{2,l}), \tag{18}$$

where $(c^k_{1,l}, a^k_{1,l}) \succ (c^k_{2,l}, a^k_{2,l}) \implies (c^k_{1,l}, a^k_{1,l}) \succeq (c^k_{2,l}, a^k_{2,l})$ and not $(c^k_{2,l}, a^k_{2,l}) \succeq (c^k_{1,l}, a^k_{1,l})$. There are 4^L possible instances of pointwise monotonicity, all of which can be defined in the obvious way. In what follows, I consider pointwise monotonicity as defined in (12) and (16) for each constraint $k \in \mathcal{K}$ and good $l \in \mathcal{L}^k$ as it corresponds to the case in my empirical application, differs from standard monotonicity (more is better), and remains notationally tractable.

For convenience, let $x^k_l = (c^k_{1,l}, a^k_{1,l})$, $y^k_l = (c^k_{2,l}, a^k_{2,l})$, $\boldsymbol{x}^k$, $\boldsymbol{y}^k \in \mathbb{R}^{L^k}_{++} \times \mathbb{R}^{L^k}_{++}$, and $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^L_{++} \times \mathbb{R}^L_{++}$ be their stacked versions. Note that choices between two alternatives $\boldsymbol{x}$ and $\boldsymbol{y}$ can be simulated by the following budget set:

$$B_{\boldsymbol{x},\boldsymbol{y}} = \{(\boldsymbol{c}, \boldsymbol{a}) \in \mathcal{C} \times \mathcal{A} : \boldsymbol{c} \leq \boldsymbol{c}^1, \boldsymbol{a} \geq \boldsymbol{a}^1\} \cup \{(\boldsymbol{c}, \boldsymbol{a}) \in \mathcal{C} \times \mathcal{A} : \boldsymbol{c} \leq \boldsymbol{c}^2, \boldsymbol{a} \geq \boldsymbol{a}^2\}. \tag{19}$$

Let $\mathcal{R}_N$ be the set of continuous and pointwise monotonic preferences that generate h on β_N. Let $\succeq_h$ be the preference relation induced by the choice function h. A preference relation $\succeq$ generates the choice function h on $\mathcal{B}$ if, for all $B \in \mathcal{B}$, $h(B) = \{\boldsymbol{x} \in B : [\boldsymbol{y} \in B] \implies \boldsymbol{x} \succeq \boldsymbol{y}\}$. The following result is due to Forges and Minelli (2009).

Lemma 2. *If the individual choice correspondence $h : \mathcal{B} \to \mathcal{C} \times \mathcal{A}$ has closed values, is pointwise monotonic, upper hemicontinuous, and satisfies WARP, then $\cap_N \mathcal{R}_N = \{\succeq_h\}$.*

The proof is included for completeness.

(1) $\succeq_h$ generates h on $\mathcal{B}$. I have to show that for all $B \in \mathcal{B}$, $h(B) = \{\boldsymbol{x} \in B : [\boldsymbol{y} \in B] \implies \boldsymbol{x} \succeq_h \boldsymbol{y}\}$. Let $\boldsymbol{x}$ be an element of the set on the right-hand side. If $\boldsymbol{x} \notin h(B)$, then there exists $\boldsymbol{y} \in B$, $\boldsymbol{x} \neq \boldsymbol{y}$, $\boldsymbol{y} \in h(B)$ implying $\boldsymbol{y} \succ_h \boldsymbol{x}$. However, this contradicts GARP as $\boldsymbol{x}$ is in the set on the right-hand side and implies $\boldsymbol{x} \succeq_h \boldsymbol{y}$.

(2) $\succeq_h$ is pointwise monotone. Let $\boldsymbol{c}_3 > \boldsymbol{c}_2$ and $\boldsymbol{a}_3 < \boldsymbol{a}_2$. I need to show that $(\boldsymbol{c}_3, \boldsymbol{a}_3) \succeq_h (\boldsymbol{c}_2, \boldsymbol{a}_2)$ and not $(\boldsymbol{c}_2, \boldsymbol{a}_2) \succeq_h (\boldsymbol{c}_3, \boldsymbol{a}_3)$. Take $B_{(\boldsymbol{c}_3,\boldsymbol{a}_3)} = \{(\boldsymbol{c}, \boldsymbol{a}) : \boldsymbol{c} \leq \boldsymbol{c}_3, \boldsymbol{a} \geq \boldsymbol{a}_3\}$. Clearly, $(\boldsymbol{c}_2, \boldsymbol{a}_2)$, $(\boldsymbol{c}_3, \boldsymbol{a}_3) \in B_{(\boldsymbol{c}_3,\boldsymbol{a}_3)}$. If $(\boldsymbol{c}_3, \boldsymbol{a}_3) \notin h(B_{(\boldsymbol{c}_3,\boldsymbol{a}_3)})$, then there is $\boldsymbol{c} \leq \boldsymbol{c}_3$ and $\boldsymbol{a} \geq \boldsymbol{a}_3$ such that $(\boldsymbol{c}, \boldsymbol{a}) \in h(B_{(\boldsymbol{c}_3,\boldsymbol{a}_3)})$, hence contradicting the pointwise monotonicity of h. As such, $(\boldsymbol{c}_3, \boldsymbol{a}_3) \in h(B_{(\boldsymbol{c}_3,\boldsymbol{a}_3)})$ and $(\boldsymbol{c}_3, \boldsymbol{a}_3) \succeq_h (\boldsymbol{c}_2, \boldsymbol{a}_2)$. Since $\succeq_h$ generates h, if

[60] This is the case, for example, in models of price search such as the one considered in my application.

we had $(\boldsymbol{c}_2, \boldsymbol{a}_2) \succeq_h (\boldsymbol{c}_3, \boldsymbol{a}_3)$, then for all B that contains both bundles, if $(\boldsymbol{c}_3, \boldsymbol{a}_3) \in h(B)$ then $(\boldsymbol{c}_2, \boldsymbol{a}_2) \in h(B)$. However, this contradicts the pointwise monotonicity of h.

(3) $\succeq_h$ is continuous. I need to show that the upper and lower contour sets of $\boldsymbol{x}$ are closed for all $\boldsymbol{x} \in \mathcal{C} \times \mathcal{A}$. Let $\succeq_h(\boldsymbol{x}) := \{\boldsymbol{y} \in \mathcal{C} \times \mathcal{A} : \boldsymbol{y} \succeq_h \boldsymbol{x}\}$ denote the upper contour set. Clearly, $\succeq_h(\boldsymbol{x})^c := \{\boldsymbol{y} \in \mathcal{C} \times \mathcal{A} : \boldsymbol{x} \succ_h \boldsymbol{y}\}$ is obviously open (in the usual topology) such that $\succeq_h(\boldsymbol{x})$ is closed. An analogous argument shows that the lower contour set is also closed.

(4) Uniqueness. Suppose there exists $\succeq \in \cap_N \mathcal{R}_N$ such that $\succeq \neq \succeq_h$. Then, we can find $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C} \times \mathcal{A}$ such that $\boldsymbol{x} \succeq \boldsymbol{y}$ and $\boldsymbol{y} \succ_h \boldsymbol{x}$ such that $\boldsymbol{x} \in \succeq_h(\boldsymbol{y})^c$. By pointwise monotonicity of $\succeq$, we can take $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C} \times \mathcal{A}$ such that $\boldsymbol{x} \succ \boldsymbol{y}$ and $\boldsymbol{x} \in \succeq_h(\boldsymbol{y})^c$. Define

$$\begin{aligned} \boldsymbol{c}_{1,\alpha} &= \boldsymbol{c}_1 + (1-\alpha)\eta\mathbf{1} \\ \boldsymbol{a}_{1,\alpha} &= \boldsymbol{a}_1 - (1-\alpha)\eta\mathbf{1}, \end{aligned}$$

where $\boldsymbol{x} = (\boldsymbol{c}_1, \boldsymbol{a}_1)$ and $\alpha \in [0,1]$, and define

$$\begin{aligned} \boldsymbol{c}_{2,\beta} &= \boldsymbol{c}_2 + (1-\beta)\eta\mathbf{1} \\ \boldsymbol{a}_{2,\beta} &= \boldsymbol{a}_2 - (1-\beta)\eta\mathbf{1}, \end{aligned}$$

where $\boldsymbol{y} = (\boldsymbol{c}_2, \boldsymbol{a}_2)$ and $\beta \in [0,1]$. Let $\boldsymbol{x}_\alpha = (\boldsymbol{c}_\alpha, \boldsymbol{a}_\alpha)$ and $\boldsymbol{y}_\beta = (\boldsymbol{c}_\beta, \boldsymbol{a}_\beta)$. By continuity and pointwise monotonicity of $\succeq$ and $\succeq_h$, there exists $\eta > 0$ such that, for all $\alpha \in [0,1]$ and $\beta \in [0,1]$, $\boldsymbol{y}_\beta \notin \succeq(\boldsymbol{x}_\alpha)$ and $\boldsymbol{x}_\alpha \notin \succeq_h(\boldsymbol{y}_\eta)$. Fix $\alpha = \beta = \frac{1}{2}$. For any $\epsilon > 0$, consider the open set around $B_{\boldsymbol{x}_{\frac{1}{2}}, \boldsymbol{y}_{\frac{1}{2}}}$ defined by

$$\mathcal{O}^\epsilon = \left\{ F \in \mathcal{B} : \mathcal{H}(B, B_{\boldsymbol{x}_{\frac{1}{2}}, \boldsymbol{y}_{\frac{1}{2}}}) < \epsilon \right\},$$

where $\mathcal{H}$ is the Hausdorff distance. I claim that, for any $\epsilon < \frac{\eta}{3}$, if $F \in \mathcal{O}^\epsilon$, then $\boldsymbol{x} \in F$ and $\boldsymbol{y} \in F$. Indeed, if $\boldsymbol{x}$ did not belong to F, then, by comprehensiveness of F, none of the points $(\boldsymbol{c}_2, \boldsymbol{a}_2)$ such that $\boldsymbol{c}_2 \geq \boldsymbol{c}_1$ and $\boldsymbol{a}_2 \leq \boldsymbol{a}_1$ would be in F. But the closest point $\boldsymbol{z}$ to $\boldsymbol{x}_{\frac{1}{2}}$ for which it is not case that $\boldsymbol{z} \geq \boldsymbol{x}$ is at distance at least $\frac{\eta}{2}$ from $\boldsymbol{x}_{\frac{1}{2}}$ (remember $\boldsymbol{x}_{\frac{1}{2}}$ is at a distance $\frac{\eta}{2}$ from $\boldsymbol{x}$). Clearly, $\epsilon < \frac{\eta}{3} < \frac{\eta}{2}$ and the argument above contradicts the fact that $F \in \mathcal{O}^\epsilon$. Thus, if $F \in \mathcal{O}^\epsilon$ then $\boldsymbol{x}, \boldsymbol{y} \in F$. The argument is that if $\boldsymbol{x} \notin F$ then the closest point $\boldsymbol{z}$ in F would be outside the open ball, contradicting that F is in the open ball.

Observe that for $\epsilon < \frac{\eta}{3}$, if $F \in \mathcal{O}^\epsilon$, then $F \subset B_{\boldsymbol{x}_0, \boldsymbol{y}_0}$. This is due to comprehensiveness of F and $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ being more "expensive" than $(\boldsymbol{x}, \boldsymbol{y})$. By pointwise monotonicity of $\succeq_h$, if $\boldsymbol{b} \in F$, then either $\boldsymbol{x}_0 \succeq_h \boldsymbol{b}$ or $\boldsymbol{y}_0 \succeq_h \boldsymbol{b}$. If $\boldsymbol{x}_0 \succeq_h \boldsymbol{b}$, then it cannot be that $\boldsymbol{b} \succeq_h \boldsymbol{y}$, because this would imply $\boldsymbol{x}_0 \succeq_h \boldsymbol{y}$, which is false by construction. This is because we took $\eta > 0$ such that, among other things, $\boldsymbol{x}_\alpha \notin \succeq_h(\boldsymbol{y}_\beta)$ for all α and β.

In particular, this implies $\boldsymbol{x}_0 \notin \succeq_h (\boldsymbol{y}_1 = \boldsymbol{y})$. Remember that $\boldsymbol{y} \in F$ such that, if we define $V = [\succeq_h (\boldsymbol{y}) \cap B_{\boldsymbol{y}_0}]$, we must have $h(F) \subset V$. This is because any element of $h(F)$ should be at least as good as $\boldsymbol{y}$.

Using the fact that $\bigcup_N \beta_N$ is dense in $\mathcal{B}$, there exist N and $B \in \beta_N$ such that $B \in \mathcal{O}^\epsilon$. Then, by our argument above we can find $\boldsymbol{b} \in h(B) \subset V$. From the fact that $\boldsymbol{x} \in B$, and that $\succeq$ generates h on β_N, this implies $\boldsymbol{b} \succeq \boldsymbol{x}$. Finally, from $\boldsymbol{b} \in V$, $\boldsymbol{y}_0 \geq b$, and by pointwise monotonicity and continuity of $\succeq$, $\boldsymbol{y}_0 \succeq \boldsymbol{b}$. Therefore, we obtain $\boldsymbol{y}_0 \succeq \boldsymbol{x}$, a contradiction. This is because $\boldsymbol{y}_0 \notin \succeq(x_1 = x)$ by construction.

The first part of the proof (Lemma 1) shows that if the consumer preferences are known, then the production function is nonparametrically identified. The second part of the proof (Lemma 2) then shows that the consumer preferences are nonparametrically identified under mild regularity conditions. Thus, it is possible to nonparametrically identify the production function of the representative consumer $(x_i)_{i\in\mathcal{N}}$ when the number of choice situations goes to infinity ($N \to \infty$). Note that it makes sense to think of the whole data set $(x_i)_{i\in\mathcal{N}}$ as a representative consumer since it satisfies SARP by Assumption 4.

Remark. *In the case of linear budget sets, Mas-Colell (1978) shows that the preference relation is uniquely identified under mild regularity conditions. Interestingly, the result of Mas-Colell (1978) also covers the case of convex budget sets via a clever argument. Namely, Proposition 2 of Forges and Minelli (2009) states that, if the consumer has a concave utility function, then the set of preferences consistent with choices made from convex budget sets is observationally equivalent to the set of preferences consistent with choices from a linearized version of the convex budget sets. Thus, without loss of generality one can treat choices as if stemming from linear budgets.*

Proof of Proposition 1

Definition 6. A square matrix $\overline{M}_i$ of dimension T is strongly cyclically consistent if for every chain $\{t_1, t_2, \ldots, t_m\} \subset \{1, 2, \ldots, T\}$, $\overline{M}_{i,t_1}(\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2}) \leq 0$, $\overline{M}_{i,t_2}(\boldsymbol{c}_{i,t_3}, \boldsymbol{a}_{i,t_3}) \leq 0$, $\ldots$, $\overline{M}_{i,t_{m-1}}(\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m}) \leq 0$ implies $\overline{M}_{i,t_m}(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) > 0$, where $\overline{M}_{i,t} := \max_k \{M^k_{i,t}(\boldsymbol{c}^k_{i,t}, \boldsymbol{a}^k_{i,t})\}$ for all $t \in \mathcal{T}$.

Lemma 3. *SARP holds if and only if the matrix $\overline{M}_i$ of revealed preferences is strongly cyclically consistent.*

For the sake of a contradiction, suppose SARP is violated. Thus, there exists a sequence of indices $\{t_1, t_2, \ldots, t_m\}$ such that $(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1})R(\boldsymbol{c}_{i,t_m}, \boldsymbol{c}_{i,t_m})$ and $(\boldsymbol{c}_{i,t_m}, \boldsymbol{c}_{i,t_m})R^D(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1})$. Construct the matrix of revealed preferences M_i and note that the chain $\{t_1, t_2, \ldots, t_m\}$ violates strong cyclical consistency. Likewise, if strong cyclical consistency is violated, then by extracting a chain causing a violation one obtains a violation of SARP as each element of the matrix represents a revealed preference.

Lemma 4. *If a square matrix $\overline{M}_i$ of dimension T is strongly cyclically consistent, then there exist numbers $(u_{i,t}, \lambda_{i,t})_{t\in\mathcal{T}}$, $\lambda_{i,t} > 0$, such that for all $s, t \in \mathcal{T}$*

$$u_{i,s} - u_{i,t} \leq \lambda_{i,t}\overline{M}_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}).$$

where $\overline{M}_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) = \max_k\{M^k_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})\}$.

The proof is completely analogous to Fostel, Scarf and Todd (2004).

(i) $\implies$ *(ii)*
Suppose that SARP is violated in the data. Thus, there exists $t_1, t_2, \ldots, t_m \in T$ such that $M^k_{i,t_1}(\boldsymbol{c}^k_{i,t_2}, \boldsymbol{a}^k_{i,t_2}) \leq 0$, $M^k_{i,t_2}(\boldsymbol{c}^k_{i,t_3}, \boldsymbol{a}^k_{i,t_3}) \leq 0$, $\ldots$, $M^k_{i,t_m}(\boldsymbol{c}^k_{i,t_1}, \boldsymbol{a}^k_{i,t_1}) \leq 0$ for all $k \in \mathcal{K}$. If there exists a utility function rationalizing the data, then this function must be such that $u_i(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) \geq u_i(\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2})$, $u_i(\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2}) \geq u_i(\boldsymbol{c}_{i,t_3}, \boldsymbol{a}_{i,t_3})$, $\ldots$, $u_i(\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m}) \geq u_i(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1})$. However, this sequence of inequalities is either self-contradictory or violates the assumption that there is a unique maximizer, unless $(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) = \cdots = (\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m})$. Hence, any data set rationalized by a utility function must satisfy SARP.
(ii) $\implies$ *(i)*
Denote the matrix of revealed preferences for constraint k by M^k_i whose element in row s and column t is $M^k_{i,s,t} := M^k_{i,t}(\boldsymbol{c}^k_i, \boldsymbol{a}^k_i)$. Note that the only information that matters is whether $M^k_{i,s,t}$ is positive or negative. Denote the matrix of revealed preferences by $\overline{M}_i$ whose element in row s and column t is $\overline{M}_{i,s,t} := \max_k \{M^k_{i,s,t}\}$. Note that an element $\overline{M}_{i,s,t}$ of $\overline{M}_i$ is negative if and only if $M^k_{i,s,t} \leq 0$ for all $k \in \mathcal{K}$. By Lemma 3, SARP holds if and only if the matrix of revealed preferences $\overline{M}_i$ is strongly cyclically consistent. An application of Lemma 4 thus yields

$$u_{i,s} \leq u_{i,t} + \lambda_{i,t}\overline{M}_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) \quad \forall s, t \in \mathcal{T},$$

where $\overline{M}_{i,t} := \max_k \{M^k_{i,s,t}\}$ and $M^k_{i,s,t} := \big(\boldsymbol{F}^k(\boldsymbol{a}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t})\big)' \boldsymbol{G}^k(\boldsymbol{c}^k) - \big(\boldsymbol{F}^k(\boldsymbol{a}^k_{i,t}, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t})\big)' \boldsymbol{G}^k(\boldsymbol{c}^k_{i,t})$.

Let $u_i(\boldsymbol{c}, \boldsymbol{a}) = \min_{t\in\mathcal{T}} \big\{u_i + \lambda_{i,t}\overline{M}_{i,t}(\boldsymbol{c}, \boldsymbol{a})\big\}$ and note it is a continuous function. To show that it rationalizes the data, first note that for all $s \in \mathcal{T}$, $u_i(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) = \min_{t\in\mathcal{T}} \{u_{i,t} + \lambda_{i,t}\overline{M}_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})\} = u_{i,s}$. This can be seen from the Afriat inequalities and by noting that $M^k_{i,t}(\boldsymbol{c}^k_{i,t}, \boldsymbol{a}^k_{i,t}) = 0$ for all $k \in \mathcal{K}$ such that $\overline{M}_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) = 0$ for all $t \in \mathcal{T}$. Hence, for any $(\boldsymbol{c}, \boldsymbol{a})$ such that $\overline{M}_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq 0$ we have $u_i(\boldsymbol{c}, \boldsymbol{a}) \leq u_{i,t} + \lambda_{i,t}\overline{M}_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq u_{i,t} = u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$. The case where $\overline{M}_{i,t}(\boldsymbol{c}, \boldsymbol{a}) < 0$ is similar. Finally, it is possible to see that the inequalities in Lemma 4 and Proposition 1 are strict whenever $(\boldsymbol{c}_s, \boldsymbol{a}_s) \neq (\boldsymbol{c}_t, \boldsymbol{a}_t)$.

This implies a strict ranking of distinct bundles at each $t \in \mathcal{T}$, i.e. a unique maximizer.

Proof of Theorem 3

$(i) \implies (ii)$

Suppose the data have been generated by (1) where the utility function is locally nonsatiated, continuous, pointwise monotonic, and concave. Then, the first-order conditions of the consumer problem for any constraint $k \in \mathcal{K}$ are given by

$$\nabla_{c^k} u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t}^k F_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) H^k(\boldsymbol{\omega}_{i,t}^k) \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k),$$
$$\nabla_{a^k} u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t}^k \nabla_a F_{i,l}^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) H^k(\boldsymbol{\omega}_{i,t}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k).$$

The concavity of the utility function implies that for all $s, t \in \mathcal{T}$, we have

$$u_i(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) - u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \leq \sum_k \left[\nabla_c^k u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})'(\boldsymbol{c}_{i,s}^k - \boldsymbol{c}_{i,t}^k) + \nabla_a^k u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})'(\boldsymbol{a}_{i,s}^k - \boldsymbol{a}_{i,t}^k) \right].$$

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

$$\begin{aligned} u_{i,s} - u_{i,t} \leq \sum_k \lambda_{i,t}^k \Big[& \Big(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k) \Big)' (\boldsymbol{c}_{i,s}^k - \boldsymbol{c}_{i,t}^k) \\ & + \Big(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k) \Big)' (\boldsymbol{a}_{i,s}^k - \boldsymbol{a}_{i,t}^k) \Big] \quad \forall s, t \in \mathcal{T}. \end{aligned}$$

Note that $\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \equiv \boldsymbol{F}_{i,t}^k$. Also, note that $\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t})$ and $\boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k)$ are unrestricted latent variables. Finally, note that local nonsatiation implies a positive derivative for at least one good $l \in \mathcal{L}^k$ for each constraint k and pointwise monotonicity implies that the sign of each gradient never changes. The inequalities are obtained by letting $\dot{\boldsymbol{F}}_{i,t} := \nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t})$.

$(ii) \implies (i)$

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

$$\begin{aligned} u_i(\boldsymbol{c}, \boldsymbol{a}) = \min_{\tau \in \mathcal{I}} \Bigg\{ & \sum_k \lambda_{i,t_m}^k \Big[\big(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_m}^k, \boldsymbol{z}_{i,t_m}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_m}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t_m}^k) \big)' \big(\boldsymbol{c}^k - \boldsymbol{c}_{t_m}^k \big) \\ & + \big(\nabla_a \boldsymbol{F}_i(\boldsymbol{a}_{i,t_m}^k, \boldsymbol{z}_{i,t_m}) \odot \boldsymbol{H}(\boldsymbol{\omega}_{i,t_m}^k) \odot \boldsymbol{G}(\boldsymbol{c}_{i,t_m}^k) \big)' \big(\boldsymbol{a}^k - \boldsymbol{a}_{t_m}^k \big) \Big] \\ & + \sum_{i=1}^{m-1} \sum_k \lambda_{i,t_i}^k \Big[\Big(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_i}^k, \boldsymbol{z}_{i,t_i}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_i}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t_i}^k) \Big)' \big(\boldsymbol{c}_{t_{i+1}}^k - \boldsymbol{c}_{t_i}^k \big) \\ & + \Big(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_i}^k, \boldsymbol{z}_{i,t_i}) \odot \boldsymbol{H}(\boldsymbol{\omega}_{i,t_i}^k) \odot \boldsymbol{G}(\boldsymbol{c}_{i,t_i}^k) \Big)' \big(\boldsymbol{a}_{t_{i+1}}^k - \boldsymbol{a}_{t_i}^k \big) \Big] \Bigg\}. \end{aligned}$$

This function is the pointwise minimum of a collection of linear functions in $(\boldsymbol{c}, \boldsymbol{a})$. As such, $u_i(\boldsymbol{c}, \boldsymbol{a})$ is continuous and concave. The utility function is also increasing in at least one good in each constraint k and pointwise monotonic from the restrictions on the gradients in (i).

If the budget sets $\{B_{i,t}^k \leq 0\}_{t \in \mathcal{T}}$ are convex, then the first-order conditions of the model are necessary and sufficient for a maximum. Therefore, I am left to show that for all $t \in \mathcal{T}$

$$\begin{aligned}
&\left[\left(\lambda_{i,t}^k \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k)\right), k \in \mathcal{K}\right] \in \nabla_c u_i(\boldsymbol{c}_t, \boldsymbol{a}_t) \\
&\left[\left(\lambda_{i,t}^k \nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k)\right), k \in \mathcal{K}\right] \in \nabla_a u_i(\boldsymbol{c}_t, \boldsymbol{a}_t).
\end{aligned}$$

Let $t \in \mathcal{T}$ and note that by definition of $u_i(\cdot, \cdot)$, there is some sequence of indices $\tau \in \mathcal{I}$ such that

$$\begin{aligned}
u_i(\boldsymbol{c}_t, \boldsymbol{a}_t) \geq \sum_k \lambda_{i,t_m}^k \Big[&\left(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_m}^k, \boldsymbol{z}_{i,t_m}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_m}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t_m}^k)\right)' (\boldsymbol{c}_t^k - \boldsymbol{c}_{t_m}^k) \\
&+ \left(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_m}^k, \boldsymbol{z}_{i,t_m}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_m}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t_m}^k)\right)' (\boldsymbol{a}_t^k - \boldsymbol{a}_{t_m}^k)\Big] \\
+ \sum_{i=1}^{m-1} \sum_k \boldsymbol{\lambda}_{i,t_i}^k \Big[&\left(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_i}^k, \boldsymbol{z}_{i,t_i}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_i}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t_i}^k)\right)' (\boldsymbol{c}_{t_{i+1}}^k - \boldsymbol{c}_{t_i}^k) \\
&+ \left(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_i}^k, \boldsymbol{z}_{i,t_i}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_i}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t_i}^k)\right)' (\boldsymbol{a}_{t_{i+1}}^k - \boldsymbol{a}_{t_i}^k)\Big]
\end{aligned}$$

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

$$\begin{aligned}
&\sum_k \lambda_{i,t_m}^k \Big[\left(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_m}^k, \boldsymbol{z}_{i,t_m}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_m}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t_m}^k)\right)' (\boldsymbol{c}_t^k - \boldsymbol{c}_{t_m}^k) \\
&+ \left(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_m}^k, \boldsymbol{z}_{i,t_m}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_m}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t_m}^k)\right)' (\boldsymbol{a}_t^k - \boldsymbol{a}_{t_m}^k)\Big] \\
&+ \sum_{i=1}^{m-1} \sum_k \lambda_{i,t_i}^k \Big[\left(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_i}^k, \boldsymbol{z}_{i,t_i}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_i}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t_i}^k)\right)' (\boldsymbol{c}_{t_{i+1}}^k - \boldsymbol{c}_{t_i}^k) \\
&+ \left(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t_i}^k, \boldsymbol{z}_{i,t_i}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t_i}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t_i}^k)\right)' (\boldsymbol{a}_{t_{i+1}}^k - \boldsymbol{a}_{t_i}^k)\Big] \\
&+ \sum_k \lambda_{i,t}^k \Big[\left(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k)\right)' (\boldsymbol{c}^k - \boldsymbol{c}_t^k) \\
&+ \left(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k)\right)' (\boldsymbol{a}^k - \boldsymbol{a}_t^k)\Big] \\
&\geq u_i(\boldsymbol{c}, \boldsymbol{a}).
\end{aligned}$$

Hence,

$$u_i(\boldsymbol{c}_t, \boldsymbol{a}_t) + \sum_k \lambda_{i,t}^k \Big[\Big(\boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}^k) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \nabla_c \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k) \Big)' (\boldsymbol{c}^k - \boldsymbol{c}_t^k)$$
$$+ \Big(\nabla_a \boldsymbol{F}_i^k(\boldsymbol{a}_{i,t}^k, \boldsymbol{z}_{i,t}) \odot \boldsymbol{H}^k(\boldsymbol{\omega}_{i,t}^k) \odot \boldsymbol{G}^k(\boldsymbol{c}_{i,t}^k) \Big)' (\boldsymbol{a}^k - \boldsymbol{a}_t^k) \Big] \geq u_i(\boldsymbol{c}, \boldsymbol{a}).$$

Since $t \in \mathcal{T}$ and $(\boldsymbol{c}, \boldsymbol{a})$ were arbitrary, the previous inequality corresponds to the definition of concavity. For the sake of simplicity, say u_i is monotone (more is better) and note that for $(\boldsymbol{c}, \boldsymbol{a}) \leq (\boldsymbol{c}_t, \boldsymbol{a}_t)$, $M_{i,t}^k(\boldsymbol{c}^k, \boldsymbol{a}^k) \leq 0$ for all $k \in \mathcal{K}$ since the budget sets are comprehensive, and $u_i(\boldsymbol{c}_t, \boldsymbol{a}_t) \geq u_i(\boldsymbol{c}, \boldsymbol{a})$, where the inequality is strict if $M_{i,t}^k(\boldsymbol{c}^k, \boldsymbol{a}^k) < 0$ for some $k \in \mathcal{K}$. That is, the utility function rationalizes the data set. Next, I show that we can construct a utility function that guarantees the solution to exist.

For the sake of concreteness, assume that $\boldsymbol{c}$ enters positively in the utility function and that $\boldsymbol{a}$ enters negatively. Let $\Gamma^k := \max_{l \in \mathcal{L}, t \in \mathcal{T}} \{a_{l,t}^k\}$. Let $h_l^k(\cdot)$ be a continuously differentiable function satisfying $h_l^k(0) = 0$, $h_l^{k\prime}(x) > 0$, $h_l^{k\prime\prime}(x) \geq 0$ for $x \in \mathbb{R}_+$ and $\lim_{x \to \infty} h_l^{k\prime}(x) = \infty$.[61] To see that there exists a utility function such that a solution exists, define $\hat{u}(\boldsymbol{c}, \boldsymbol{a}) := u(\boldsymbol{c}, \boldsymbol{a}) - \sum_k \sum_{l=1}^{L^k} h_l^k \left(\max \left\{ 0, a_l^k - \Gamma^k \right\} \right)$. As before, this function is continuous and concave. Furthermore, note that $\hat{u}(\boldsymbol{c}, \boldsymbol{a}) \leq u(\boldsymbol{c}, \boldsymbol{a})$ for all $(\boldsymbol{c}, \boldsymbol{a}) \in \mathcal{C} \times \mathcal{A}$ 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 $\boldsymbol{a} \to \infty$ or $\boldsymbol{c} \to \infty$ along some dimension for those variables that enter negatively. This follows from the piecewise linearity of $u(\cdot, \cdot)$ and the assumption that $\lim_{x \to \infty} h_l^{k\prime}(x) = \infty$.

Proof of Proposition 2

Assumption 7 states that the price function for any good $l \in \mathcal{L}$ is given by:

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

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

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

where $\overline{\log(p_{l,t})}$ denotes the average log price paid and $\overline{\omega}_{l,t}$ denotes the average search

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

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 8 was used to eliminate the expected average search productivity. By Assumption 9, 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 $\log(a_{l,t})$, one gets

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

By Leibniz integration rule, one can insert the derivative inside the expectation to get

$$\mathbb{E}\left[\frac{\partial \log\left(p_{l,t}^*\right)}{\partial \log(a_{l,t})}\right] = \frac{1}{L}\mathbb{E}\left[\alpha_l^1\right].$$ [62]

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

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

where the left-hand side is the expected (average) elasticity of price with respect to shopping intensity and $\overline{\alpha}^1 := \frac{1}{L}\sum_{l=1}^{L}\alpha_l^1$ is the average shopping technology.

Proof of Proposition 3

The first-order condition of the consumer problem with respect to shopping intensity is given by

$$\frac{\partial u(\boldsymbol{c}_t, \boldsymbol{a}_t)}{\partial a_{l,t}} = \frac{\partial p(a_{l,t}, \omega_{l,t})}{\partial a_{l,t}} c_{l,t} \quad \forall l \in \mathcal{L}.$$

Since the utility function is continuous and concave, it follows that it is absolutely continuous in $a_{l,t}$.[63] Hence, by the fundamental theorem of calculus, the search cost on good $l \in \mathcal{L}$ is given by

$$\int_1^{a_{l,t}} \frac{\partial u(\boldsymbol{c}_t, \boldsymbol{a}_t)}{\partial a}\, da = \int_1^{a_{l,t}} \frac{\partial p(a, \omega_{l,t})}{\partial a} c_{l,t}\, da \quad \forall l \in \mathcal{L},$$

[62] This requires the partial derivatives to be continuous.

[63] Note, however, that the utility function is not jointly absolutely continuous (Friedman, 1940).

which gives

$$u(\boldsymbol{c}_t, \boldsymbol{a}_t) - u(\boldsymbol{c}_t, \boldsymbol{a}_{-l,t}, 1_t) = [p(a_{l,t}, \omega_{l,t}) - p(1_t, \omega_{l,t})]c_{l,t},$$

where $\boldsymbol{a}_{-l,t}$ denote shopping intensity on every good except good l and 1_t is the shopping intensity on good l in period t. Using the log-linear structure of the price function due to Assumption 7, the expression simplifies to

$$\frac{u(\boldsymbol{c}_t, \boldsymbol{a}_t) - u(\boldsymbol{c}_t, \boldsymbol{a}_{-l,t}, 1_t)}{p^*_{l,t} c_{l,t}} = 1 - a_l^{-\alpha_l^1},$$

where I divided by true expenditure on both sides. Averaging across goods and time periods and taking the expectation yields the result.

Bibliography

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.

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. 2020. "Satisficing, aggregation, and quasilinear utility." *Available at SSRN 3180302*.

Andrews, Donald WK and Gustavo Soares. 2010. "Inference for parameters defined by moment inequalities using generalized moment selection." *Econometrica*, 78 (1), 119–157.

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.

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

Ben-Porath, Yoram. 1967. “The production of human capital and the life cycle of earnings.” *Journal of political economy*, 75 (4, Part 1), 352–365.

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. “Homework in macroeconomics: Household production and aggregate fluctuations.” *Journal of Political economy*, 99 (6), 1166–1187.

Blundell, Richard, Martin Browning, Laurens Cherchye, Ian Crawford, Bram De Rock, and Frederic Vermeulen. 2015. “Sharp for SARP: nonparametric bounds on counterfactual demands.” *American Economic Journal: Microeconomics*, 7 (1), 43–60.

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, Dennis Kristensen, and Rosa Matzkin. 2014. “Bounding quantile demand functions using revealed preference inequalities.” *Journal of Econometrics*, 179 (2), 112–127.

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.

Browning, Martin and Pierre-André Chiappori. 1998. “Efficient intra-household allocations: A general characterization and empirical tests.” *Econometrica*, 1241–1278.

Chambers, Christopher P, Federico Echenique, and Nicolas S Lambert. 2021. “Recovering preferences from finite data.” *Econometrica*, 89 (4), 1633–1664.

Chambers, Christopher P, Federico Echenique, and Eran Shmaya. 2014. “The axiomatic structure of empirical content.” *American Economic Review*, 104 (8), 2303–19.

Chavas, Jean-Paul and Thomas L Cox. 1993. “On generalized revealed preference analysis.” *The Quarterly Journal of Economics*, 108 (2), 493–506.

Cherchye, Laurens, Bram De Rock, Arthur Lewbel, and Frederic Vermeulen. 2015a. "Sharing rule identification for general collective consumption models." *Econometrica*, 83 (5), 2001–2041.

Cherchye, Laurens, Thomas Demuynck, and Bram De Rock. 2011. "Revealed preference analysis of non-cooperative household consumption." *The Economic Journal*, 121 (555), 1073–1096.

Cherchye, Laurens, Thomas Demuynck, Bram De Rock, and Per Hjertstrand. 2015b. "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*.

Cherchye, Laurens and Frederic Vermeulen. 2008. "Nonparametric analysis of household labor supply: goodness of fit and power of the unitary and the collective model." *The Review of Economics and Statistics*, 90 (2), 267–274.

Chernozhukov, Victor, Han Hong, and Elie Tamer. 2007. "Estimation and confidence regions for parameter sets in econometric models 1." *Econometrica*, 75 (5), 1243–1284.

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

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

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.

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.

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

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

Fortin, Bernard and Guy Lacroix. 1997. "A test of the unitary and collective models of household labour supply." *The economic journal*, 107 (443), 933–955.

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

Friedman, Bernard. 1940. "A note on convex functions." *Bulletin of the American Mathematical Society*, 46 (6), 473–474.

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.

Gorman, William M. 1959. "Separable utility and aggregation." *Econometrica: Journal of the Econometric Society*, 469–481.

Gorno, Leandro. 2019. "Revealed preference and identification." *Journal of Economic Theory*, 183, 698–739.

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

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.

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

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.

Kübler, F, Raghav Malhotra, and Herakles Polemarchakis. 2020. "Identification of preferences, demand and equilibrium with finite data." *University of Warwick Economics Research Paper* (1290).

Kübler, Felix and Herakles Polemarchakis. 2017. "The identification of beliefs from asset demand." *Econometrica*, 85 (4), 1219–1238.

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

Mas-Colell, Andreu. 1977. "The recoverability of consumers' preferences from market demand behavior." *Econometrica: Journal of the Econometric Society*, 1409–1430.

——— 1978. "On revealed preference analysis." *The Review of Economic Studies*, 45 (1), 121–131.

Matzkin, Rosa L. 1991. "Axioms of revealed preference for nonlinear choice sets." *Econometrica: Journal of the Econometric Society*, 1779–1786.

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.

Nishimura, Hiroki, Efe A Ok, and John K-H Quah. 2017. "A comprehensive approach to revealed preference theory." *American Economic Review*, 107 (4), 1239–63.

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."

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

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

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

Thomas, Duncan. 1990. "Intra-household resource allocation: An inferential approach." *Journal of human resources*, 635–664.

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

Wang, Liqun. (2021) "Identifiability in measurement error models." in *Handbook of Measurement Error Models*, 55–70: Chapman and Hall/CRC.

Wooldridge, Jeffrey M. 2009. "On estimating firm-level production functions using proxy variables to control for unobservables." *Economics letters*, 104 (3), 112–114.