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

Charles Gauthier

Abstract

This paper provides a new identification result for a large class of models in which consumers participate in production. I show that consumer preferences are necessary and sufficient to identify production functions through cross-equation restrictions implied by first-order conditions. In addition, I derive a nonparametric revealed preference characterization of the class of models that exhausts its empirical implications. Finally, I use a novel and easy-to-apply inference method that is valid under partial identification. 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

*This paper previously circulated as “Price Search and Consumption Inequality: Robust, Credible, and Valid Inference” (first version, March 2021). Department of Economics, KU Leuven. Naamsestraat 69 Leuven, Belgium. (email: charles.gauthier@kuleuven.be). 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 data are those of the researcher and do not reflect the views of NielsenIQ. NielsenIQ 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.

data on shopping expenditures and shopping intensity from the NielsenIQ 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 are significant, hence largely diminishing benefits of price search.

Keywords: Production function, price search, demand analysis

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. This paper provides a new identification result that uses the structure of the consumer problem to show that preferences are necessary and sufficient to identify production functions. 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. If preferences are identified, cross-equation restrictions from the first-order conditions nonparametrically identify the part of the production functions that relates to consumer preferences. Likewise, if production functions are identified, the cross-equation restrictions can be used to identify preferences.

The careful reader will note that the identification strategy shares similarities with Gandhi, Navarro and Rivers (2020) in that I exploit restrictions implied by optimizing behavior to identify production functions.[1] My approach otherwise differs due to specific challenges arising in consumer problems. First, mono-

[1] Gandhi, Navarro and Rivers (2020) show that using the structure of the firm problem resolves the lack of identification of gross production functions in the proxy variable approach developed by Olley and Pakes (1996).

tone 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 derive a new identification result that exploits the structure of the consumer problem to show that one can learn about production functions from preferences, and learn about preferences from production functions.

The second contribution of this paper is to provide a nonparametric characterization of a large class of models via shape restrictions. These restrictions can be shown to imply the well-known Generalized Axiom of Revealed Preference (GARP). Although GARP captures the full empirical content of the model, it does not guarantee the utility function to be concave when budget sets are nonlinear. Hence, shape restrictions provide a preferable alternative to use the identification power that stems from the consumer model. Importantly, they can be used as the basis for the estimation strategy.

The previous characterization is rooted in the revealed preference tradition (Afriat, 1967; Diewert, 1973; Varian, 1982; Browning, 1989). As such, it gives minimal testable conditions for the data to be rationalized by the model. Due to the nonlinearity of the production functions, the treatment of nonlinear budget sets builds on Matzkin (1991) and, more specifically, Forges and Minelli (2009).[2] A difference is that I allow for the utility function to be decreasing in some of its arguments to cover additional models of interest such as those of price search.

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

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

(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.[3] 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 via a chi-square approximation.[4] This feature is useful as identification may not always be achieved such as in the presence of measurement error.

My approach combines revealed preference and econometrics to make inference on objects of interest.[5] In a similar fashion, Blundell, Browning and Crawford (2008) use the Strong Axiom of Revealed Preference (SARP) to improve predictions on 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. (2023) bound the welfare implications of a price change via analogous revealed preference restrictions.[6] The estimation strategy put forward in this paper may also be useful for such endeavors.

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

[3]Specifically, I use a rejection sampling algorithm for the integration of the latent variables that can be applied in models defined by linear or nonlinear constraints indiscriminately, including combinations thereof.

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

[5]The shape restrictions imply GARP but further ensure that the consumer problem is well-behaved.

[6]Other work include Blundell, Browning and Crawford (2003, 2007), Blundell, Horowitz and Parey (2012) and Blundell, Kristensen and Matzkin (2014), among others.

paper of Stigler (1961) and has gained strong empirical support over the years.[7] In an influential paper, Aguiar and Hurst (2007) show that price search partially explains the retirement-consumption puzzle.[8] My application provides new insights regarding the validity of price search, the robust impacts of search on prices paid, and the size of search costs.

My empirical analysis uses the NielsenIQ 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 NielsenIQ 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.[9] Second, measurement error can complicate empirical analyses by obscuring the true behavior of variables such as expenditure. 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

[7] For a general survey, see Baye et al. (2006).

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

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

in a standard regression setting.

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).

Using my methodology in the NielsenIQ Homescan Dataset, I find support for price search behavior in single households. However, the model is rejected in multi-person households.[10] This outcome can be rationalized by the implicit assumption that multi-person households behave as a single decision maker and the 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 negative 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.

Restricting the empirical analysis to single households, the conservative 95% confidence set on the expected elasticity of price with respect to shopping intensity is $[-0.2, -0.1]$. In other words, 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.[11] This finding suggests that the exogeneity requirement of their instruments are fulfilled,

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

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

and that measurement error does not significantly bias their estimates.

Since shopping intensity enters preferences, the structure of the model further allows me to assess the utility cost associated with price search. Once again restricting the analysis to single consumers, I get that the conservative 95% confidence set on a lower bound of the expected search cost is $[10, 50]$. That is, the expected search cost is about 10 to 50 dollars, representing at least 20 to 116 percent of consumers observed expenditures. This finding shows that search costs are significant and should be accounted for when measuring the consumer welfare.

The rest of the paper is organized as follows. Section 2 defines the class of problems covered. Section 3 presents the identification result. Section 4 presents the estimation strategy. Section 5 details the application. Section 6 concludes. The main proofs can be found in Appendix A5.

2 Class of Models

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

2.1 Environment

Let $\mathcal{N} = \{1, \ldots, N\}$, $\mathcal{L} = \{1, \ldots, L\}$, and $\mathcal{T} = \{1, \ldots, T\}$ denote the set of consumers, goods, and periods for which data are observable, respectively. I denote consumption by $\boldsymbol{c} \in \mathcal{C} = \mathbb{R}^L_{++}$.[12] For each good $l \in \mathcal{L}$, there is an output F_l produced using a set of variable inputs $\boldsymbol{a}_l \in \mathcal{A}_l = \mathbb{R}^{A_l}_{++}$ and a set of fixed inputs $\boldsymbol{z} \in \mathcal{Z} \subseteq \mathbb{R}^Z_{++}$, where $A_l, Z \in \mathbb{N}$ are natural numbers. The set of all variable inputs is denoted by $\mathcal{A} = \mathbb{R}^A_{++}$, where $A = A_1 + \cdots + A_L$. A data set for consumer i is

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

denoted by $x_i := \{(\boldsymbol{F}_{i,t}, \boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t})\}_{t\in\mathcal{T}}$.

Let $u : \mathcal{C}\times\mathcal{A} \to \mathbb{R}$ be a utility function that is twice continuously differentiable, increasing in $\boldsymbol{c}$, and decreasing in $\boldsymbol{a}$. 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, consumers have preferences given by a random utility function $u : \mathcal{I} \to \mathcal{U}$. A consumer is a draw $i \in \mathcal{I}$ with preferences u_i.

2.2 The Consumer Problem

The class of problems considered in this paper is that of a set of consumers behaving as if maximizing their utility function:

$$\max_{(\boldsymbol{c},\boldsymbol{a})\in\mathcal{C}\times\mathcal{A}} u_i(\boldsymbol{c}, \boldsymbol{a}) \tag{1}$$

subject to a constraint of the form

$$\tilde{\boldsymbol{F}}_i(\boldsymbol{a}, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}_{i,t})'\boldsymbol{c} \leq y_{i,t}, \tag{A}$$

or

$$c \leq \tilde{F}_i(\boldsymbol{a}, \boldsymbol{z}_{i,t}, \omega_{i,t}), \tag{B}$$

where the production function $\tilde{F}_{i,l} : \mathcal{A}_l \times \mathcal{Z} \times \mathbb{R} \to \mathbb{R}_{++}$ is continuously differentiable and monotone[13], $\omega_{i,l,t} \in \mathbb{R}$ is the productivity shock of the production function, and $y_{i,t} > 0$ is observed expenditure in period $t \in \mathcal{T}$. I assume that the constraints of the problem hold with equality at the observed data and that the

[13] The production function is assumed decreasing in $\boldsymbol{a}$ for budgets of type A and increasing in $\boldsymbol{a}$ for budgets of type B. For ease of exposition, I assume that the production function shares the same type of monotonicity in $\boldsymbol{z}$ as it does in $\boldsymbol{a}$.

production function $\tilde{F}_l$ has no variables in common with $\tilde{F}_{l'}$ for any $l' \neq l$.[14]

Example 1. *(Price Search) Consider a model of price search similar to Aguiar and Hurst (2007) with a concave utility function $u(c, a)$ that is increasing in consumption (c) and decreasing in search intensity (a). The log price function is $\log(F(a, z, \omega)) = \alpha_0 + \alpha_1 \log(a) + \alpha_2 z + \omega$, where α_0, $\alpha_1 < 0$, and α_2 are parameters, z is a variable affecting prices paid such as shopping needs, and ω is a productivity shock. The budget is $F(a, z, \omega)c = y$, where 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. 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. *(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 $c = F(a_1, a_2)$.*

[14] The production functions have no variables in common in the sense that $\mathcal{A}_l \bigcap \mathcal{A}_{l'} = \emptyset$ for all $l \neq l'$.

3 Identification

This section defines the identified set and investigates identification in the class of models.

3.1 Identified Set

Let Θ_u and Θ_F be compact subsets of the Euclidean space. For each consumer $i \in \mathcal{N}$, let $\boldsymbol{\theta}^0_{u_i} \in \Theta_u$ be a finite-dimensional parameter of the true utility function and $\boldsymbol{\theta}^0_{F_i} \in \Theta_F$ be a possibly infinite-dimensional parameter of the true production functions. In what follows, I may write the dependence of the utility function and the production functions on those parameters explicitly by $u_i(\boldsymbol{c}, \boldsymbol{a}; \boldsymbol{\theta}_{u_i})$ and $\tilde{F}_i(\boldsymbol{a}, \boldsymbol{z}, \omega; \boldsymbol{\theta}_{F_i})$ when relevant. The identified set for a consumer $i \in \mathcal{N}$ is defined as

$$\Theta_I(x_i) := \left\{ (\boldsymbol{\theta}_{u_i}, \boldsymbol{\theta}_{F_i}) \in \Theta_u \times \Theta_F : \ x_i \text{ solves } (1) \right\}.$$

In words, the identified set contains every combination of parameters that are consistent with the utility maximization problem at the observed data. The model is said to be point identified if the identified set is single-valued such that $\Theta_I(x_i) = \{(\boldsymbol{\theta}^0_{\boldsymbol{u_i}}, \boldsymbol{\theta}^0_{F_i})\}$, the utility function is said to be point identified if $\Theta_I(x_i) = \{(\boldsymbol{\theta}^0_{\boldsymbol{u_i}}, \boldsymbol{\theta}_{F_i}), \ \boldsymbol{\theta}_{F_i} \in \Theta_F\}$, and the production functions are said to be point identified if $\Theta_I(x_i) = \{(\boldsymbol{\theta}_{u_i}, \boldsymbol{\theta}^0_{F_i}), \ \boldsymbol{\theta}_{u_i} \in \Theta_u\}$. The next section investigates identification by letting $T \to \infty$.

3.2 Identification

The first condition for identification is that the consumer problem has a unique solution.

Assumption 1. *The following statements hold:*

(i) The utility function is concave and parameterized by a finite-dimensional parameter $\boldsymbol{\theta}_{u_i}$*.*

(ii) The production functions are convex (concave) for budgets of type A (B).

Assumption 1 imposes mild shape restrictions that guarantee a unique solution to the consumer problem. This is because a downward sloping convex production function leads to a convex budget set for budgets of type A, whereas an upward sloping concave production function leads to a convex budget set for budgets of type B. The requirement that the utility function be parameterized is necessary to learn about preference parameters from the marginal rate of substitution (MRS). Importantly, this assumption ensures that there is a unique MRS tangent to the budget at the chosen allocations.[15] Next, I assume that productivity shocks are Hicks-neutral.

Assumption 2. *For all* $l \in \mathcal{L}$, $\tilde{F}_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}, \boldsymbol{\omega}_{i,t}; \boldsymbol{\theta}_{F_i}) = F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}; \boldsymbol{\theta}_{F_i}) e^{-\omega_{i,l,t}}$.

Assumption 2 implies that productivity shocks do not change marginal rates of substitution between variables $(c_{i,l}, \boldsymbol{a}_{i,l,t})$ that pertain to the same group. It is necessary for identification as its multiplicatively separable structure allows to obtain restrictions that are independent from productivity shocks. Hicks-neutrality is a widely used specification in the production function literature and includes as a common empirical specification the log-linear regression. Lastly, I require that productivity shocks induce changes in optimal input allocations.

Assumption 3. *Income expansion paths for variable inputs are nonzero such that* $\frac{\partial \boldsymbol{a}_{i,l,t}}{\partial \omega_{i,l,t}} \neq 0$ *for all* $l \in \mathcal{L}$.

[15] Note that goods cannot be perfect complements as the utility function would not be differentiable. In that case, the MRS may not be unique.

This assumption ensures that the model can generate sufficient variation in the optimal allocations $\boldsymbol{a}$ to identify production functions. Indeed, Hicks-neutral productivity shocks induce no substitution effect when there is a single good. Thus, Assumption 3 guarantees that productivity shocks induce variation from their income effects. The following result characterizes identification in the class of models.

Theorem 1. *Suppose Assumptions 1-3 hold. The utility function is identified up to a monotone transformation if and only if the log production functions are identified up to an additively separable function of* $\boldsymbol{z}$*.*

The idea behind Theorem 1 is to use the consumer problem to derive cross-equation restrictions between preferences and production functions. Precisely, these restrictions state that the rate of change of the log production functions is proportional to the MRS. Since the MRS is invariant to monotone transformations of the utility function, the production functions are also invariant to such transformations. If preferences are known, the cross-equation restrictions identify partial derivatives of the log production functions. The key is to note that these partial derivatives can be integrated to identify the log production functions up to a constant of integration that is a function of $\boldsymbol{z}$. Intuitively, since the utility function does not include fixed inputs $\boldsymbol{z}$, there is no structure on $\boldsymbol{z}$ that can propagate from preferences to production functions. Conversely, if log production functions are known up to an additively separable function of $\boldsymbol{z}$, the cross-equation restrictions identify preference parameters that enter the MRS.

Theorem 1 implies that it may be possible to identify preferences through the production function under weaker assumptions. This is because the utility derived from a set of goods is unknown, but the output produced from a set of inputs is observed. The additional information available on the production side can facilitate (partial) identification, an observation exploited in the application. Conversely, preferences may be known such as in a macroeconomic model. There, Theorem 1 implies that a set of production parameters can be obtained for free.

A consequence of Theorem 1 is that cross-sectional heterogeneity in preferences and production functions must be the same.

Corollary 1. *Suppose Assumption 1-3 hold. Two consumers have identical utility functions up to a monotone transformation if and only if their log production functions are identical up to an additively separable function of* $\boldsymbol{z}$*.*

This result implies that a model where consumers have heterogeneous preferences but identical production functions would be inconsistent.

4 Estimation Strategy

In this section, I derive nonparametric restrictions that exhaust the empirical content of the class of models and propose an estimation strategy that can be used to make inference even if the model is partially identified.

4.1 Shape Restrictions

In what follows, it will be useful to define what it means for a bundle to be revealed preferred to another directly from the data. Let $\odot$ define the Hadamard product such that for any two vectors $\boldsymbol{v}$ and $\tilde{\boldsymbol{v}}$ of equal dimension, $(\boldsymbol{v} \odot \tilde{\boldsymbol{v}})_l = v_l \tilde{v}_l$.

Definition 1. For any $s, t \in \mathcal{T}$, a bundle $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is said to be directly revealed preferred to a bundle $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$ if $M_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) = \left(\boldsymbol{F}_i(\boldsymbol{a}_{i,s}, \boldsymbol{z}_{i,t}) \odot e^{-\boldsymbol{\omega}_{i,t}}\right)' \boldsymbol{c}_{i,s} -$

$\left(\boldsymbol{F}_i(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t}) \odot e^{-\boldsymbol{\omega}_{i,t}}\right)' \boldsymbol{c}_{i,t} \leq 0.$[16]

Next, I define what it means for a utility function to rationalize a data set.

Definition 2. 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 $(\boldsymbol{c}, \boldsymbol{a})$, $M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq 0$ implies $u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \geq u_i(\boldsymbol{c}, \boldsymbol{a})$.

This definition states that within the set of affordable bundles, those that are more expensive give a higher utility level. A data set should be rationalizable if observed bundles can be thought of as arising from the maximization of a utility function.

For convenience, let $e^{-\boldsymbol{\omega}_{i,l,t}}$ denote $e^{-\omega_{i,l,t}}$ stacked A_l times and $\boldsymbol{c}_{i,l,t}$ denote $c_{i,l,t}$ stacked A_l times, where A_l is the dimension of $\boldsymbol{a}_l$. The following result relates rationalizability by a utility function with conditions that can easily be checked in the data.

Theorem 2. *The following statements are equivalent:*

(i) The data set x_i is rationalized by a utility function that is continuous, increasing in $\boldsymbol{c}$, decreasing in $\boldsymbol{a}$, and concave, and where the production functions are convex.

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

$$u_{i,s} \leq u_{i,t} + \lambda_{i,t}\Big[\boldsymbol{F}'_{i,t}(\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) + \sum_{l \in \mathcal{L}} \left(\dot{\boldsymbol{F}}_{i,l,t} \odot e^{-\boldsymbol{\omega}_{i,l,t}} \odot \boldsymbol{c}_{i,l,t}\right)' (\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t})\Big]$$

$$\phi_{i,l,s} \geq \phi_{i,l,t} + \dot{\boldsymbol{F}}'_{i,l,t}(\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) + \ddot{\boldsymbol{F}}'_{i,l,t}(\boldsymbol{z}_{i,s} - \boldsymbol{z}_{i,t}) \quad \forall l \in \mathcal{L},$$

[16] The definition for budgets of type B is the same with $M_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) = (c_{i,s} - F(\boldsymbol{a}_{i,s}, \boldsymbol{z}_{i,t})e^{-\omega_{i,t}}) - (c_{i,t} - F(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t})e^{-\omega_{i,t}})$. The results below also hold for this type of budgets.

and $\phi_{i,l,t}e^{-\omega_{i,l,t}} = F_{i,l,t}$ *for all* $l \in \mathcal{L}$ *and* $t \in \mathcal{T}$.

In words, this result states that any data set rationalized by the model must satisfy the inequalities in Theorem 2 (*ii*) provided the utility function is concave and the production functions are convex. Conversely, any data set that satisfies the inequalities in Theorem 2 (*ii*) is rationalized by a concave utility function where the production functions are convex.[17]

The first set of inequalities in Theorem 2 (*ii*) captures the concavity of the utility function, where the numbers $u_{i,t}$ and $\lambda_{i,t} > 0$ can be thought of as utility numbers and marginal utilities of expenditure. The sign of $\dot{F}^k_{i,l,t}$ captures the monotonicity of the production function with respect to $a^k_{i,l,t}$ and depends on the application. In a model of price search similar to the one of Example 1, $\dot{\boldsymbol{F}}_{i,t} < 0$ captures the fact that an increase in search intensity decreases prices paid.

The second set of inequalities in Theorem 2 (*ii*) captures the convexity of the production functions, where the numbers $\phi_{i,l,t}$ can be thought of as output numbers.[18] The sign of $\dot{F}^k_{i,l,t}$ captures the monotonicity of the production function with respect to $a^k_{i,l,t}$ and are the same numbers as those appearing in the concavity of the utility function. Likewise, the sign of $\ddot{F}^k_{i,l,t}$ captures the monotonicity of the production function with respect to $z^k_{i,t}$. Lastly, the restriction that $\phi_{i,l,t}e^{-\omega_{i,l,t}} = F_{i,l,t}$ captures the fact that the candidate output should match the actual output at the observed data.

[17]The Appendix shows that the Generalized Axiom of Revealed Preference (GARP) is equivalent to rationalizability without the nice structure on the utility and production functions. A similar proof is also in Forges and Minelli (2009).

[18]In models with budgets of type B, it is more natural to assume that the production function is concave. The statement of the theorem can be modified accordingly.

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 variables conditional on $\mathcal{X}$. Moreover, let $x_i \in \mathcal{X}$ denote the observed data and $e_i \in \mathcal{E}|\mathcal{X}$ denote the latent variables. The restrictions of the model derived in Theorem 2 (ii) imply the following moment functions for all $l \in \mathcal{L}$ and all $s, t \in \mathcal{T}$:

$$
\begin{aligned}
g^u_{i,s,t}(x_i, e_i) := \mathbb{1}\Bigg(& u_{i,s} - u_{i,t} - \lambda_{i,t}\Big[\boldsymbol{F}'_{i,t}(\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) \\
& + \sum_{l \in \mathcal{L}} \left(\dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,l,t}} \odot \boldsymbol{c}_{i,l,t}\right)' (\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) \leq 0\Big]\Bigg) - 1 \\
g^F_{i,l,s,t}(x_i, e_i) := \mathbb{1}\Bigg(& \phi_{i,l,s} - \phi_{i,l,t} - \Big[\dot{\boldsymbol{F}}'_{i,l,t}(\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) + \ddot{\boldsymbol{F}}'_{i,l,t}(\boldsymbol{z}_{i,s} - \boldsymbol{z}_{i,t}) \leq 0\Big]\Bigg) - 1,
\end{aligned}
$$

where the latent variables further satisfy their support constraints: $\lambda_{i,t} > 0$, $\dot{\boldsymbol{F}}_{i,t} < 0$, $\ddot{\boldsymbol{F}}_{i,t} < 0$, and $\phi_{i,l,t} e^{-\omega_{i,l,t}} = F_{i,l,t}$. In addition, note that a model may have additional restrictions on the productivity shock, written as $g^\omega_{i,l,t}(x_i, e_i)$. These restrictions could represent orthogonality conditions and depend on the application. In general, these additional restrictions may be necessary to refute the model.

Let $\boldsymbol{g}_i(x_i, e_i) := (\boldsymbol{g}^u_i(x_i, e_i)'$, $\boldsymbol{g}^F_i(x_i, e_i)$', $\boldsymbol{g}^\omega_i(x_i, e_i)')'$ denote the set of moment functions that characterize the model. Furthermore, let d_u, d_F 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, $\boldsymbol{g}^{u,\omega}_i(x_i, e_i)$ is the set of moment functions on the utility function and the productivity shock. 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 $\mathcal{M}_{\mathcal{X}}$ and $\mathcal{M}_{\mathcal{E}|\mathcal{X}}$ denote the set of all probability measures defined over $\mathcal{X}$ and $\mathcal{E}|\mathcal{X}$, respectively. Moreover, let $\mathbb{E}_{\mu\times\pi}[\boldsymbol{g}(x,e)] := \int_{\mathcal{X}}\int_{\mathcal{E}|\mathcal{X}} \boldsymbol{g}(x,e)\, d\mu\, d\pi$, 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.[19]

Definition 3. A data set $x := \{x_i\}_{i\in\mathcal{N}}$ is statistically rationalizable if

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

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 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.[20]

Theorem 3. *The following are equivalent:*

(i) A data set x is statistically rationalizable.

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

where

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

[19] This definition follows the notion of identified set in Schennach (2014).

[20] See Aguiar and Kashaev (2021) for the weak technical assumptions required for this result to hold.

and where $\eta(\cdot|x_i)$ is an arbitrary user-specified distribution function supported on $\mathcal{E}|\mathcal{X}$ such that $\mathbb{E}_{\pi_0}[\log(\mathbb{E}_\eta[\exp(\boldsymbol{\gamma}'\boldsymbol{g}^\omega(x,e))|x])]$ 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). □

In words, Theorem 3 (*ii*) averages out the unobservables in $\boldsymbol{g}_i(x_i, e_i)$ according to some conditional distribution.[21] 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 further reduced by noting that the concavity of the utility function and the convexity of the production functions only restrict the conditional support of the unobservables. This can be seen from the fact that if the data are exactly statistically rationalizable, then the moment functions $\boldsymbol{g}_i^{u,F}(x_i, e_i)$ should hold exactly for each consumer. Thus, one can draw from the conditional distribution $\tilde{\eta}(\cdot|x_i) := \mathbb{1}(\boldsymbol{g}_i^{u,F}(x_i, \cdot) = 0)\eta(\cdot|x_i)$ rather than leaving the moment functions $\boldsymbol{g}_i^{u,F}(x_i, \cdot)$ in the optimization problem.

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

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

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

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

that satisfy the moment functions $\boldsymbol{g}_i^{u,F}(x_i, e_i)$ and can be achieved by rejection sampling. Then, a standard Metropolis-Hastings algorithm can be used to draw from the distribution.[22]

4.4 Statistical Inference

The notion of statistical rationalizability together with Theorem 3 provides a feasible way of checking whether the data are consistent with the model. Indeed, 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 ($\chi^2_{d_\omega}$). 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 freedom. Note that the panel structure contributes to the test statistic via the shape constraints as they give cross-equation restrictions that limit the support of the latent variables across time. Inference on an expected parameter

[22] Further details about the implementation are given in Appendix A3.

of interest θ_0 can be made by adding a moment function of the form:

$$g_i^\theta(x_i, e_i) - \theta_0,$$

where θ_0 may be thought of as a parameter of the production function. A conservative 95% confidence set on θ_0 is obtained by inverting the test statistic:

$$\{\theta_0 : \mathrm{TS}_N(\theta_0) \le \chi^2_{d_\omega + d_\theta, 0.95}\},$$

where $TS_N(\theta_0)$ is the test statistic at a fixed value of θ_0 and d_θ is the number of moments on parameters of interest.

4.5 Related Literature

The previous methodology provides an approach to conduct specification testing and inference with unrestricted heterogeneity in preferences. In this respect, it is similar to Kitamura and Stoye (2018) and Deb et al. (2023). However, their framework only assumes access to a repeated cross-section of consumers rather than a panel of consumers. As such, they analyze the data through a random utility framework where preferences are unrestricted in the cross-section and in time.[23] In contrast, the current approach analyzes the data through a random utility framework where preferences are unrestricted in the cross-section but fixed in time. Furthermore, the current framework assumes that individual choices are observed rather than only the distribution of choices.[24] More importantly, the current statistical framework can analyze models that include moment conditions.

[23]Strictly speaking, they assume that the distribution of preferences is time invariant.

[24]This allows me to impose individual rationality. In contrast, a data set can be stochastically rationalizable even if it contains individuals that are not rational (Im and Rehbeck, 2022).

This feature is crucial for many structural models including the class of models considered in this paper.

5 Application to Price Search

This section considers an application to price search as a means to show the feasibility and effectiveness of the approach. I propose a flexible model of price search to test some of the key assumptions in the literature, to make inference on the impacts of search intensity on prices paid that is robust to measurement error, and to quantify the size of search costs.

5.1 Model

Consider a model of price search where consumers have a price search technology that can be used to search for lower prices. The consumer is assumed to know her realization of search productivity ($\boldsymbol{\omega}$) and to choose consumption ($\boldsymbol{c}$) and shopping intensity ($\boldsymbol{a}$) accordingly. Formally, in each period the consumer behaves as if maximizing her quasilinear utility function subject to satisfying her budget constraint:

$$\max_{(\boldsymbol{c},\boldsymbol{a},r)\in\mathcal{C}\times\mathcal{A}\times\mathbb{R}} u_i(\boldsymbol{c},\boldsymbol{a}) + r_{i,t} \quad \text{s.t.} \quad \boldsymbol{p}_i(\boldsymbol{a},\boldsymbol{\omega}_{i,t})'\boldsymbol{c} + r_{i,t} = y_{i,t}, \tag{2}$$

where $u_i : \mathcal{C}\times\mathcal{A}\to\mathbb{R}$ is a utility function that is twice continuously differentiable, increasing in consumption, decreasing in shopping intensity, and concave, $r_{i,t}$ is the numeraire good, $\boldsymbol{p}_i(\boldsymbol{a},\boldsymbol{\omega}_{i,t})$ is a vector of twice continuously differentiable good-specific price functions $p_{i,l} : \mathbb{R}_{++}\times\mathbb{R}\to\mathbb{R}_{++}$ where $\boldsymbol{p}_{i,t} := \boldsymbol{p}_i(\boldsymbol{a}_{i,t},\boldsymbol{\omega}_{i,t})$, and $y_{i,t}$ 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. Thus, 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 of a single good. The consumer has to choose a bundle that lies within her budget set represented by the shaded area. The bundle that maximizes the consumer utility is the point where the indifference curve is tangent to the budget line, denoted (c^*, a^*).[25]

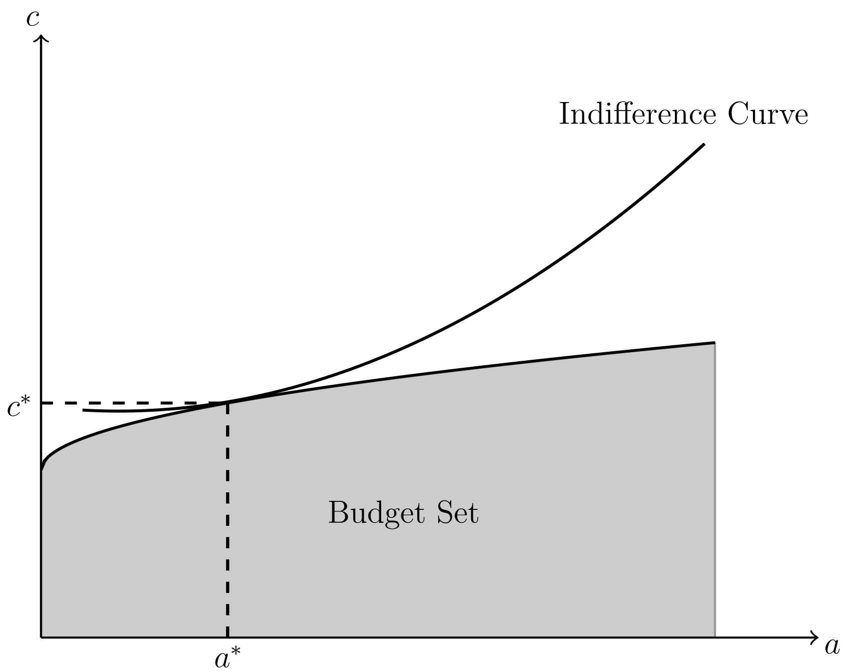

Figure 1: Optimal Choice with Price Search

It is important to note that the quasilinearity assumption could be relaxed in

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

the model. The assumption is motivated by the fact that the data in my application span a period of six months for which changes in income are negligible. Moreover, the data focus on food consumption which tends to be income inelastic.[26] Since the quasilinear structure provides a useful measure of utility in terms of dollar, it will allow me to get a straightforward interpretation of search costs. A description of the data set is provided in Appendix A4.

5.2 Environment

Given the fundamental unobservability of preferences, it is prudent to keep assumptions on the utility function minimal. On the contrary, the price functions are partially observed since one has data on prices paid at 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 log prices) vary with log number of shopping trips in the data, where shopping trips capture shopping intensity.

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 the log number of shopping trips increases can be approximated by a linear function. Thus, I follow the price search literature and assume that the price functions are log-linear in shopping intensity.[27]

Assumption 4. *For all $l \in \mathcal{L}$, the log price function is given by*

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

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

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

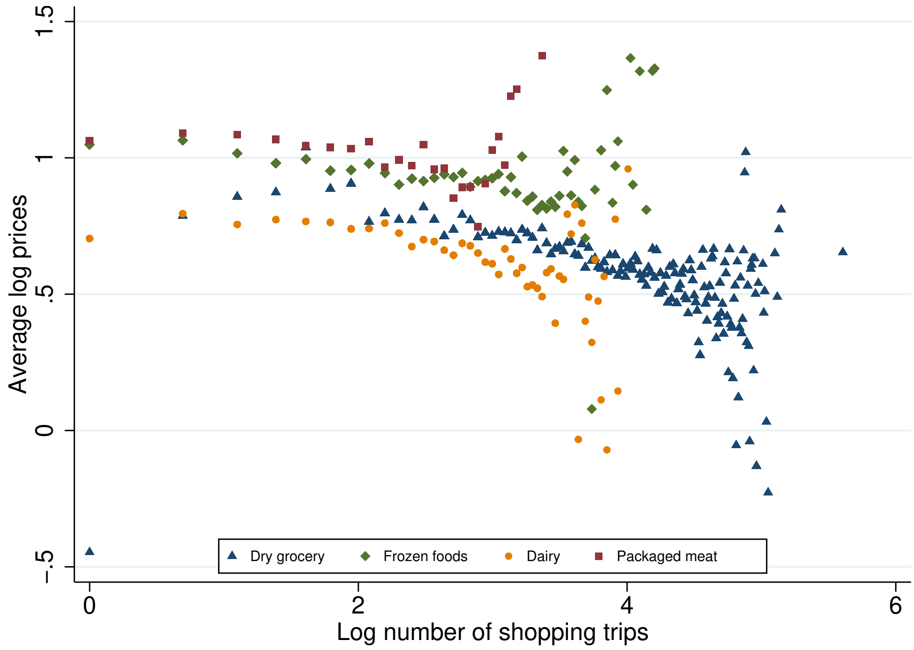

Figure 2: Average Log Price by Log Number of Shopping Trips

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

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

Assumption 4 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.[28]

Conditional on the log-linear specification assumed in Assumption 4, 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

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

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 glance at Figure 2 shows that the log-linear relationship does not hold perfectly for any given good. These deviations are normal in any data set and are accounted for by search productivity ($\boldsymbol{\omega}$) 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. For example, 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 4, this implies that the expected average search productivity is zero. In other words, the unobserved effects of search productivity on prices paid cancel out on average.

Assumption 5. *For all $t \in \mathcal{T}$, $\mathbb{E}\left[\overline{\omega}_t\right] = 0$, where $\overline{\omega}_{i,t} := L^{-1} \sum_{l=1}^{L} \omega_{i,l,t}$ denotes the average search productivity across goods.*

Assumption 5 allows search productivity to vary for each individual and each good as long as the overall search productivity remains constant.[29] Permitting search productivity for a particular good to change over time is important in my application because of the coarse aggregation of the data. Indeed, since a consumer may purchase different baskets of goods in different time periods, prices may vary

[29]The expected search productivity and the expected intercept in the price function are not separately identified. Thus, the empirical bite of Assumption 5 is the time invariance of the expected search productivity.

due to variations in the composition of the baskets of goods.

Other than for this mild centering condition, Assumption 5 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.[30]

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 NielsenIQ Homescan Dataset on which Figure 2 is derived. They show that measurement error in prices is severe and document that the expected difference between observed log prices and true log prices is zero.[31] Let $(\boldsymbol{p}^*_{i,t})_{t\in\mathcal{T}}$ denote true prices paid by the consumer. The previous information motivates the following assumption on measurement error.

Assumption 6. *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 6 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, where measurement error is defined as $\boldsymbol{m}_{i,t} := \log(\boldsymbol{p}_{i,t}) - \log(\boldsymbol{p}^*_{i,t})$.[32] The structure of measurement error implies that for

[30] See, for example, Gandhi, Navarro and Rivers (2020).

[31] Additional details about the data and measurement error are given in Appendix A4.

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

each category of goods in Figure 2, the true value of the data points above (below) the log-linear relationship could be lower (higher) as long as measurement error averages out.

Lastly, I bound the support of the elasticity of price with respect to shopping intensity to strengthen the power of the statistical test.

Assumption 7. *For all* $l \in \mathcal{L}$, $\alpha^1_{i,l} \in [-1, 0]$.

This assumption constrains the elasticity of price with respect to shopping intensity to be in $[-1, 0]$ for every good $l \in \mathcal{L}$ and is motivated by the estimate of Aguiar and Hurst (2007). Indeed, they obtain a point estimate of -0.074 for the elasticity of price with respect to shopping intensity using the Homescan 1993-1995.[33] As such, Assumption 7 should give enough flexibility for the needs of the data.

Under the previous assumptions, the concavity of the utility function and the log-linearity of the price functions can be refuted by the data. Intuitively, to see why price search is refutable, note that Assumptions 4-5 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.[34] Hence, if any of the assumptions is inconsistent with the data, this will turn up in the statistical test and the model may be rejected as a consequence.

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

[34] See Appendix A2 for analytical power results.

5.3 Price Search Rationalizability

The environment defined in the previous section allows me to define a notion of statistical rationalizability as defined in Section 4 with the moment functions specialized to the model.[35] A data set consistent with this notion is said to be price search rationalizable (PS-rationalizable).

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 1. *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 1 states that the true expected effect of an increase in shopping intensity on the price paid can be recovered. The reason why it can be achieved in the model is that Assumption 5 restricts the expected (average) search productivity to be time-invariant. Therefore, any variation in expected log prices must be caused exclusively by variations in shopping intensity. As such, I can make inference on the true expected elasticity of price with respect to shopping intensity (θ_0) by adding the moment function:

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

Next, I am interested in learning about search costs. It is possible to see that

[35] The definition of statistical price search rationalizability is derived in Appendix A6.

inference on a lower bound of the expected search cost (θ_0) can be made by adding the moment function:

$$g_i^{sc}(x_i, e_i) := -\frac{1}{L \cdot T} \sum_{l \in \mathcal{L}, t \in \mathcal{T}} \dot{p}_{i,l,t} a_{i,l,t} c_{i,l,t} - \theta_0.$$[36]

For either of these objects, inference is made by test inversion.

5.4 Empirical Results

The application aims to check whether a panel data set can be rationalized by a heterogeneous model of price search under mild centering conditions on productivity shocks and measurement error in prices. If the data are compatible with price search behavior, then inference on the impacts of price search can be computed. Summary statistics on the sample are displayed in Table 1.

Table 1: Summary Statistics

Variable	Household Size			Annual Income			Education[a]		
	1	2	3+	$< 40k$	$[40k, 70k]$	$> 70k$	$\leq$ High school	Some college	$\geq$ College
Observations	1645	6364	3539	4133	3849	3566	4070	3532	3946
Total	11548			11548			11548		

[a] If both spouses are present in a household, the education of the male member is reported.

By applying the methodology of Section 4, 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 and households of many members with a test statistic of 443.38 and 230.01, respectively.

[36] The derivation is available in Appendix A6.

Since the model is not rejected by the data on single households, I can invert the statistical test to obtain a conservative 95% confidence set on the expected elasticity of price with respect to shopping intensity. Doing so, I obtain a confidence set of $[-0.2, -0.1]$. That is, a doubling of shopping intensity decreases prices paid by about 15% on average. Likewise, I find that the conservative 95% confidence set on the lower bound of the expected search cost is $[10, 50]$. In comparison, the average observed expenditure on any good is 43.14 dollars.

6 Conclusion

This paper shows that the identification of preferences is necessary and sufficient to identify production functions in a large class of models where consumers are involved in production. This observation results from cross-equation restrictions entailed by the consumer problem. I provide a simple estimation strategy that exploits natural shape restrictions to make inference on both production functions and preferences. In my empirical application, I recover the impacts of shopping intensity on prices paid and quantify the size of search costs. I find that price search is effective in reducing prices paid but comes at a significant utility cost. As such, my results highlight the importance of accounting for search costs to evaluate the consumer welfare. In particular, recognizing the utility cost from price search may provide new insights on within- and between-group inequalities.

Conflict of Interest

The author reports there are no competing interests to declare.

Additional Information

Funding

This project is supported in part by funding from the Social Sciences and Humanities Research Council.

Bibliography

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.

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

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

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. 2007. "Improving revealed preference bounds on demand responses." *International Economic Review*, 48 (4), 1227–1244.

——— 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.

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

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.

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, 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 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 KH Quah, and Jörg Stoye. 2023. "Revealed price preference: theory and empirical analysis." *The Review of Economic Studies*, 90 (2), 707–743.

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.

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.

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

Im, Changkuk and John Rehbeck. 2022. "Non-rationalizable individuals and stochastic rationalizability." *Economics Letters*, 219, 110786.

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.

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

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.

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

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

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.

Appendix

A1: Relationship with Models of Household Production

Although the focus of my application 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 $f(h, \boldsymbol{c})$.[37] The household 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

[37] 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.

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 $u(\cdot)$ and $f(\cdot,\cdot)$ are both unobservable concave functions and $u(\cdot)$ is increasing, 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).[38] To see why, assume the solution is interior and take the first-order conditions:

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

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

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

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

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

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 the model is refutable under Assumption 4-7. 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 $l \in \mathcal{L}$ be log-linear as specified in Assumption 4 such that:

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

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 the price functions logarithmically convex.[39]

Falsifiability of Price Search

Suppose that Assumption 4-7 are satisfied and let $\mathcal{L} = \{1, 2\}$, $\mathcal{T} = \{1, 2\}$. Almost

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

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 $t = 1, 2$.

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

$$\log\left(\frac{p(a_{l,s}, \omega_{l,s})}{p(a_{l,t}, \omega_{l,t})}\right) \geq \frac{\nabla_a p(a_{l,t}, \omega_{l,t})}{p(a_{l,t}, \omega_{l,t})}(a_{l,s} - a_{l,t}) + \frac{\nabla_\omega p(a_{l,t}, \omega_{l,t})}{p(a_{l,t}, \omega_{l,t})}(\omega_{l,s} - \omega_{l,t}).$$ [40]

The above expression can be written more concisely as

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

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

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

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

$$0 > \frac{1}{L}\sum_{l=1}^{L} \left(\mathbb{E}\left[\log(p_{l,1})\right] - \mathbb{E}\left[\log(p_{l,2})\right]\right) \geq -\frac{1}{L}\sum_{l=1}^{L} \mathbb{E}\left[\frac{\nabla_a p(a_{l,2}, \omega_{l,2})}{p^*_{l,2} c_{l,2}}\right]. \tag{3}$$

Noting that the random variable on the right-hand side is always negative, it follows that the negative of its expectation is positive: $-\mathbb{E}\left[\frac{\nabla_a p(a_{l,2}, \omega_{l,2})}{p^*_{l,2} c_{l,2}}\right] \geq 0$ for all $l \in \mathcal{L}$. Clearly, inequality (3) yields a contradiction. In other words, the data are inconsistent with the model provided the price functions are logarithmically

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

convex, and that is the case for log-linear price functions.

A3: Implementation

This section provides a pseudo-algorithm of the ELVIS method developed by Schennach (2014) specialized to my model of price search.

Pseudo-Code

Step 1

- Fix the number of goods L and the number of time periods T.
- Fix the data set $x = (x_i)_{i \in \mathcal{N}}$, 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: $\boldsymbol{\alpha}_i^1 \in [-1, 0]$.
- 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 pseudo-code below details how to construct the conditional distribution by using rejection sampling and by applying the Metropolis-Hastings algorithm. Note that tuning parameters from the Metropolis-Hastings algorithm (proposal distribution and number of iterations) appear in the implementation of ELVIS. In general, one should aim for an acceptance rate of about 25%. This requires the user to calibrate the variance of the proposal distribution. In practice, the acceptance rate of the Metropolis-Hastings algorithm can be computed such that the user is guided through the process. In what follows, I draw true prices instead of measurement error as it ensures true prices to be strictly positive. Let $R > 0$ denote the number of iterations.

while $r \leq R$

- Draw candidate latent variables $e_i^c = (\boldsymbol{p}_{i,t}^*, \boldsymbol{\alpha}_i^0, \boldsymbol{\alpha}_i^1, \boldsymbol{\omega}_{i,t})_{t \in T}$ such that their support constraints are satisfied. Recover $\dot{p}_{i,l,t} = \alpha_{i,l}^0 \alpha_{i,l}^1 a_{i,l,t}^{\alpha_{i,l}^1 - 1} e^{-\omega_{i,l,t}}$ for all $l \in \mathcal{L}$ and $t \in \mathcal{T}$. Add it to e_i^c.

- Given x_i and e_i^c, check whether the model is satisfied using Theorem 2. 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: Data Set

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

Sample Construction

For my empirical application, I use the NielsenIQ 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 NielsenIQ 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.[41] This selection leaves over a

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

million distinct UPCs representing about 40% of all products in the Homescan. For each household, I also aggregate the data to monthly observations to further reduce stockpiling issues. The resulting UPC prices are calculated as the average UPC prices weighted by quantities purchased.

To obtain regular observations on each good, I aggregate UPCs to their department categories, yielding a total of four "goods". 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.[42]

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 11548 households, 4 aggregated goods, and 6 monthly time periods.

Measurement error

The data collection process employed by NielsenIQ 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 NielsenIQ with point-of-sale data, the price reported

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

(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 $(\boldsymbol{p}^*_{i,t})_{t\in\mathcal{T}}$. 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 NielsenIQ'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.

Details about 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 restriction to this time window achieves three purposes. First, it increases the plausibility of the stability of preferences. Second, it reduces the likelihood of income shocks, hence increasing the plausibility of the quasilinearity of preferences. Third, it reduces the computational burden. Indeed, since the number of parameters to solve for in the model is given by $L \cdot T + T$, the nonlinear optimization problem can become prohibitive if either L or T is too large.

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 number of single households to 1645.[43] Starting from the sample of $16,025$ households and making the same operations bring down the number of couples to 6364 and the number of multi-person households with more than two members to 3539.

A5: Proofs

Proof of Theorem 1

Budget type A

The first-order conditions of the consumer problem (1) with respect to $\boldsymbol{c}$ and $\boldsymbol{a}$ are

$$\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t} F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}) e^{-\omega_{i,l,t}} \tag{4}$$

$$\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k} = \lambda_{i,t} \frac{\partial F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a^k_{i,l,t}} e^{-\omega_{i,l,t}} c_{i,l,t} \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{A}_l, \tag{5}$$

where $\nabla u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ denotes the gradient of u_i at the point $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ and $\mathcal{A}_l$ denotes

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

the set of variable inputs for good $l \in \mathcal{L}$. Dividing (5) by (4) and rearranging yields

$$\frac{\partial f_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} = \frac{\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k}}{\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l} \cdot \frac{1}{c_{i,l,t}} \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{A}_l, \tag{6}$$

where $f_{i,l} := \log(F_{i,l})$. Since the MRS is invariant to monotone transformations of the utility function, the derivative of the log production function is invariant to such transformations.

Budget type B

The first-order conditions of the consumer problem (1) with respect to $\boldsymbol{c}$ and $\boldsymbol{a}$ are

$$\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t} \tag{7}$$

$$\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k} = -\lambda_{i,t} \frac{\partial F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} e^{-\omega_{i,l,t}} \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{A}_l, \tag{8}$$

where $\nabla u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ denotes a gradient of u_i at the point $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ and $\mathcal{A}_l$ denotes the set of variable inputs for good $l \in \mathcal{L}$. Dividing (8) by (7) and rearranging yields

$$\frac{\partial F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} = -\frac{\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k}}{\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l} \cdot \frac{1}{e^{-\omega_{i,l,t}}} \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{A}_l. \tag{9}$$

Dividing both sides by $F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})$, one obtains

$$\frac{\partial f_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} = -\frac{\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k}}{\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l} \cdot \frac{1}{F_{i,l,t}} \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{A}_l, \tag{10}$$

where $F_{i,l,t} = F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})e^{-\omega_{i,l,t}}$ is the observed output. Since the MRS is invariant to monotone transformations of the utility function, the derivative of the log production function is invariant to such transformations.

Suppose the MRS is known, then equations (6) and (10) immediately identify 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_l^k to obtain

$$\int_{\underline{a}_{i,l}^k}^{a_{i,l,t}^k} \frac{\partial f_{i,l}(\boldsymbol{a}_{i,l,t}^{-k}, a, \boldsymbol{z}_{i,t})}{\partial a} \, da = f_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}) + C_{i,l}(\boldsymbol{a}_{i,l,t}^{-k}, \boldsymbol{z}_{i,t}) \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{A}_l, \tag{11}$$

where $0 < \underline{a}_{i,l}^k < a_{i,l,t}^k$ and $\boldsymbol{a}_{i,l,t}^{-k}$ represents the vector $\boldsymbol{a}_l \in \mathcal{A}_l$ stripped of its kth element. As Gandhi, Navarro and Rivers (2020) note, the above equations can be combined to obtain the following equality:

$$f_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}) = \sum_{k=1}^{A_l} \left(\int_{\underline{a}_{i,l}^k}^{a_{i,l,t}^k} \frac{\partial f_{i,l}\left(\boldsymbol{a}_{i,l,t}^{k'<k}, \underline{\boldsymbol{a}}_{i,l}^{k'>k}, a, \boldsymbol{z}_{i,t}\right)}{\partial a} \, da \right) - C_{i,l}(\boldsymbol{z}_{i,t}). \tag{12}$$

This equation shows that log production functions are identified up to an additive function of $\boldsymbol{z}$. Also, note that the constant in the constant of integration and the intercept of the log production function are not separately identified. As such, one can normalize one or the other. Suppose now that log production functions are known up to an additive function of $\boldsymbol{z}$, $C_{i,l}(\boldsymbol{z}_{i,t})$. It follows that partial derivatives with respect to $\boldsymbol{a}$ are known as

$$\frac{\partial C_{i,l}(\boldsymbol{z}_{i,t})}{\partial a_{i,l,t}^k} = 0.$$

Thus, equations (6) and (10) directly identify preference parameters as there are $|\mathcal{A}_l| \cdot T \to \infty$ equations and the marginal rate of substitution is a function of a finite number of preference parameters.

Proof of Theorem 2

$(i) \implies (ii)$

Budget type A

The first-order conditions of the consumer problem for any good $l \in \mathcal{L}$ is given by

$$\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t} F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}) e^{-\omega_{i,l,t}}$$
$$\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k} = \lambda_{i,t} \frac{\partial F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a^k_{i,l,t}} e^{-\omega_{i,l,t}} c_{i,l,t} \quad \forall k \in \mathcal{A}_l,$$

where the equalities hold for some supergradient of $u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$.[44]

Budget type B

The first-order conditions of the consumer problem for any good $l \in \mathcal{L}$ is given by

$$\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_l = \lambda_{i,t} \tag{13}$$
$$\nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})_{l,k} = -\lambda_{i,t} \frac{\partial F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})}{\partial a^k_{i,l,t}} e^{-\omega_{i,l,t}} \quad \forall k \in \mathcal{A}_l, \tag{14}$$

where the equalities hold for some supergradient of $u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$.

In what follows, I only include the derivation for budgets of type A; an analogous argument can be made for budgets of type B. The concavity of the utility function implies that for all $s, t \in \mathcal{T}$, we have

$$u_i(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) - u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \leq \left[\nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})'(\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) + \sum_{l \in \mathcal{L}} \nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})'_l (\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) \right].$$

Combining the first-order conditions with the concavity of the utility function and

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

letting $u_{i,t} := u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ for all $t \in \mathcal{T}$ yields

$$
\begin{aligned}
u_{i,s} - u_{i,t} \leq \lambda_{i,t} \Big[&\big(\boldsymbol{F}_i(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t}) \odot e^{-\boldsymbol{\omega}_{i,t}}\big)' (\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) \\
&+ \sum_{l \in \mathcal{L}} \big(\nabla_a \boldsymbol{F}_i(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t})_l \odot e^{-\boldsymbol{\omega}_{i,l,t}} \odot \boldsymbol{c}_{i,l,t}\big)' (\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) \Big] \quad \forall s, t \in \mathcal{T},
\end{aligned}
$$

where $\boldsymbol{F}_i(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t}) \odot e^{-\boldsymbol{\omega}_{i,t}} \equiv \boldsymbol{F}_{i,t}$ is the observed output, $e^{-\boldsymbol{\omega}_{i,l,t}}$ is $e^{-\omega_{i,l,t}}$ stacked A_l times, $\boldsymbol{c}_{i,l,t}$ is $c_{i,l,t}$ stacked A_l times, and A_l is the dimension of $\boldsymbol{a}_l$.

In what follows, $(\dot{F}^k_{i,l,t})_{l \in \mathcal{L}, k \in \mathcal{A}_l, t \in \mathcal{T}}$ are latent variables that replace $\nabla_a \boldsymbol{F}_i(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t})_{l,k}$ and I write the stacked vectors $\dot{\boldsymbol{F}}_{i,l,t} \odot e^{-\boldsymbol{\omega}_{i,l,t}} \odot \boldsymbol{c}_{i,l,t}$, $l = 1, \ldots, L$, as $\dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}} \odot \boldsymbol{c}_{i,t}$ to simplify notation. The previous inequalities imply the existence of numbers $(u_{i,t})_{t \in \mathcal{T}}$, positive $(\lambda_{i,t})_{t \in \mathcal{T}}$, negative $(\dot{F}^k_{i,l,t})_{l \in \mathcal{L}, k \in \mathcal{A}_l, t \in \mathcal{T}}$, and $(\omega_{i,l,t})_{l \in \mathcal{L}, t \in \mathcal{T}}$ such that for all $s, t \in \mathcal{T}$

$$
u_{i,s} - u_{i,t} \leq \lambda_{i,t} \Big[\boldsymbol{F}'_{i,t} (\boldsymbol{c}_{i,s} - \boldsymbol{c}_{i,t}) + \Big(\dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}} \odot \boldsymbol{c}_{i,t}\Big)' (\boldsymbol{a}_{i,s} - \boldsymbol{a}_{i,t}) \Big]. \tag{15}
$$

Next, using the convexity of the production function for each $l \in \mathcal{L}$, we have

$$
\begin{aligned}
F_{i,l}(\boldsymbol{a}_{i,l,s}, \boldsymbol{z}_{i,s}) - F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}) &\geq \nabla_a F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})'_l (\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) \\
&\quad + \nabla_z F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})' (\boldsymbol{z}_{i,s} - \boldsymbol{z}_{i,t}) \quad \forall s, t \in \mathcal{T}.
\end{aligned}
$$

These inequalities imply the existence of positive numbers $(\phi_{i,l,t})_{l \in \mathcal{L}, t \in \mathcal{T}}$, negative numbers $(\dot{F}^k_{i,l,t})_{l \in \mathcal{L}, k \in \mathcal{A}_l, t \in \mathcal{T}}$ that are the same as those in equation (15), and negative $(\ddot{F}_{i,l,t})_{l \in \mathcal{L}, t \in \mathcal{T}}$ such that for each $l \in \mathcal{L}$

$$
\phi_{i,l,s} - \phi_{i,l,t} \geq \dot{\boldsymbol{F}}'_{i,l,t} (\boldsymbol{a}_{i,l,s} - \boldsymbol{a}_{i,l,t}) + \ddot{\boldsymbol{F}}'_{i,l,t} (\boldsymbol{z}_{i,s} - \boldsymbol{z}_{i,t}) \quad \forall s, t \in \mathcal{T}.
$$

Combining the restrictions from the concavity of the utility function and the convexity of the production functions yields the inequalities of Theorem 2 (ii). Finally, it must be the case that the numbers $\phi_{i,l,t}$ and $\omega_{i,l,t}$ are such that $\phi_{i,l,t}e^{-\omega_{i,l,t}} = F_{i,l,t}$ since $F_{i,l,t} = F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})e^{-\omega_{i,l,t}}$ for all $t \in \mathcal{T}$.

$(ii) \implies (i)$

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

$$
\begin{aligned}
u_i(\boldsymbol{c}, \boldsymbol{a}) = \min_{\tau \in \mathcal{I}} \Bigg\{ & \lambda_{i,t_m} \Big[\boldsymbol{F}'_{i,t_m} \big(\boldsymbol{c} - \boldsymbol{c}_{i,t_m} \big) + \big(\dot{\boldsymbol{F}}_{i,t_m} \odot e^{-\boldsymbol{\omega}_{i,t_m}} \odot \boldsymbol{c}_{i,t_m} \big)' \big(\boldsymbol{a} - \boldsymbol{a}_{i,t_m} \big) \Big] \\
& + \sum_{j=1}^{m-1} \lambda_{i,t_j} \Big[\boldsymbol{F}'_{i,t_j} \big(\boldsymbol{c}_{t_{i,j+1}} - \boldsymbol{c}_{i,t_j} \big) + \Big(\dot{\boldsymbol{F}}_{i,t_j} \odot e^{-\boldsymbol{\omega}_{i,t_j}} \odot \boldsymbol{c}_{i,t_j} \Big)' \big(\boldsymbol{a}_{t_{i,j+1}} - \boldsymbol{a}_{i,t_j} \big) \Big] \Bigg\}.
\end{aligned}
$$

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. Moreover, note that the utility function is increasing in $\boldsymbol{c}$ and decreasing in $\boldsymbol{a}$.

If the budget sets $(B_{i,t})_{t \in \mathcal{T}}$ are convex, then the first-order conditions of the model are necessary and sufficient for a maximum.[45] Therefore, I have to show that for all $t \in \mathcal{T}$,

$$
\begin{aligned}
\lambda_{i,t} \boldsymbol{F}_{i,t} &\in \nabla_c u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \\
\lambda_{i,t} \dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}} \odot \boldsymbol{c}_{i,t} &\in \nabla_a u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}),
\end{aligned}
$$

or that these vectors are supergradients of the constructed utility function. Let $t \in \mathcal{T}$ and note that by definition of $u_i(\cdot, \cdot)$, there is some sequence of indices $\tau \in \mathcal{I}$

[45] I show that the convexity of the production functions implies the convexity of the budget sets. The observed data can thus be thought of as optimal choices and, as I show later, rationalizability obtains.

such that

$$u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \geq \lambda_{i,t_m}\Big[\boldsymbol{F}'_{i,t_m}\big(\boldsymbol{c}_{i,t} - \boldsymbol{c}_{i,t_m}\big) + \Big(\dot{\boldsymbol{F}}_{i,t_m} \odot e^{-\boldsymbol{\omega}_{i,t_m}} \odot \boldsymbol{c}_{i,t_m}\Big)'\big(\boldsymbol{a}_{i,t} - \boldsymbol{a}_{i,t_m}\big)\Big]$$
$$+ \sum_{j=1}^{m-1} \lambda_{i,t_j}\Big[\boldsymbol{F}'_{i,t_j}\big(\boldsymbol{c}_{t_{i,j+1}} - \boldsymbol{c}_{i,t_j}\big) + \Big(\dot{\boldsymbol{F}}_{i,t_j} \odot e^{-\boldsymbol{\omega}_{i,t_j}} \odot \boldsymbol{c}_{i,t_j}\Big)'\big(\boldsymbol{a}_{i,t_{j+1}} - \boldsymbol{a}_{i,t_j}\big)\Big].$$

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

$$\lambda_{i,t}\Big[\boldsymbol{F}'_{i,t}\big(\boldsymbol{c} - \boldsymbol{c}_{i,t}\big) + \Big(\dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}}\Big)'\big(\boldsymbol{a} - \boldsymbol{a}_{i,t}\big)\Big]$$
$$+ \lambda_{i,t_m}\Big[\boldsymbol{F}'_{i,t_m}\big(\boldsymbol{c}_{i,t} - \boldsymbol{c}_{i,t_m}\big) + \Big(\dot{\boldsymbol{F}}_{i,t_m} \odot e^{-\boldsymbol{\omega}_{i,t_m}} \odot \boldsymbol{c}_{i,t_m}\Big)'\big(\boldsymbol{a}_{i,t} - \boldsymbol{a}_{i,t_m}\big)\Big]$$
$$+ \sum_{j=1}^{m-1} \lambda_{i,t_j}\Big[\boldsymbol{F}'_{i,t_j}\big(\boldsymbol{c}_{i,t_{j+1}} - \boldsymbol{c}_{i,t_j}\big) + \Big(\dot{\boldsymbol{F}}_{i,t_j} \odot e^{-\boldsymbol{\omega}_{i,t_j}} \odot \boldsymbol{c}_{i,t_j}\Big)'\big(\boldsymbol{a}_{i,t_{j+1}} - \boldsymbol{a}_{i,t_j}\big)\Big]$$
$$\geq u_i(\boldsymbol{c}, \boldsymbol{a}).$$

Hence,

$$u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) + \lambda_{i,t}\Big[\boldsymbol{F}'_{i,t}(\boldsymbol{c} - \boldsymbol{c}_{i,t}) + \Big(\dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}} \odot \boldsymbol{c}_{i,t}\Big)'\big(\boldsymbol{a} - \boldsymbol{a}_{i,t}\big)\Big] \geq u_i(\boldsymbol{c}, \boldsymbol{a}). \tag{16}$$

Since $t \in \mathcal{T}$ and $(\boldsymbol{c}, \boldsymbol{a})$ were arbitrary, the previous inequality corresponds to the definition of concavity. Thus, the said vectors are indeed supergradients of u_i.

I am left to show that the constructed utility function rationalizes the data. For every $l \in \mathcal{L}$, define the function

$$F_{i,l}(\boldsymbol{a}_l, \boldsymbol{z}) = \max_{t \in \mathcal{T}} \left\{ \phi_{i,l,t} + \sum_k \Big(\dot{F}^k_{i,l,t}(a^k_l - a^k_{i,l,t}) + \ddot{F}^k_{i,l,t}(z^k - z^k_{i,t})\Big) \right\}.$$

This production function is the pointwise maximum of a collection of downward sloping linear functions in $\boldsymbol{a}$ and $\boldsymbol{z}$. As such, $F_{i,l}(\boldsymbol{a}_l, \boldsymbol{z})$ is continuous, decreasing in $\boldsymbol{a}$ and $\boldsymbol{z}$, and convex, where the monotonicity property is a direct consequence of $\dot{F}^k_{i,l,t} < 0$ and $\ddot{F}^k_{i,l,t} < 0$. Next, define the function

$$M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) = \left(\boldsymbol{F}_i(\boldsymbol{a}, \boldsymbol{z}_{i,t}) \odot e^{-\boldsymbol{\omega}_{i,t}}\right)' \boldsymbol{c} - \boldsymbol{F}'_{i,t}\boldsymbol{c}_{i,t}.$$

Since $\boldsymbol{F}_i(\cdot,\cdot)$ and $\boldsymbol{c}$ are positive and convex functions, it follows that $M_{i,t}(\boldsymbol{c}, \boldsymbol{a})$ is convex for all $t \in \mathcal{T}$.[46] The convexity of $M_{i,t}(\cdot,\cdot)$ implies that for any $(\boldsymbol{c}, \boldsymbol{a})$ and $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$

$$M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) - M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \geq \nabla_c M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{c} - \boldsymbol{c}_{i,t}) + \nabla_a M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{a} - \boldsymbol{a}_{i,t}),$$

where $\nabla_c M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ and $\nabla_a M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ denote subgradients of $M_{i,t}(\cdot,\cdot)$ at $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$. From the second set of inequalities in Theorem 2 (ii), we have

$$\phi_{i,l,t} \geq \phi_{i,l,s} + \dot{\boldsymbol{F}}'_{i,l,s}(\boldsymbol{a}_{i,l,t} - \boldsymbol{a}_{i,l,s}) + \ddot{\boldsymbol{F}}'_{i,l,s}(\boldsymbol{z}_{i,t} - \boldsymbol{z}_{i,s}) \quad \forall s \in \mathcal{T},$$

which, combined with the definition of $F_{i,l}(\boldsymbol{a}_l, \boldsymbol{z})$, implies that $F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t}) = \phi_{i,l,t}$ for all $t \in \mathcal{T}$.[47] Using the restriction that $\phi_{i,l,t}e^{-\omega_{i,l,t}} = F_{i,l,t}$, it follows that $M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) = 0$ for all $t \in \mathcal{T}$. As such, the inequalities for the convexity of the function $M_{i,t}(\cdot,\cdot)$ simplify to

$$M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \geq \nabla_c M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{c} - \boldsymbol{c}_{i,t}) + \nabla_a M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{a} - \boldsymbol{a}_{i,t}).$$

[46]Note that the convexity of $M_{i,t}(\cdot,\cdot)$ implies the convexity of the budget sets. This is because the budget sets $B_{i,t} := \{(\boldsymbol{c}, \boldsymbol{a}) | M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq 0\}$ are the lower contour sets of $M_{i,t}(\cdot,\cdot)$.

[47]It also follows that $\dot{\boldsymbol{F}}_{i,t}$ and $\ddot{\boldsymbol{F}}_{i,t}$ are subgradients of $F_{i,l}(\boldsymbol{a}_{i,l,t}, \boldsymbol{z}_{i,t})$.

Consider $(\boldsymbol{c}, \boldsymbol{a})$ such that $M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq 0$. Then, we have

$$0 \geq M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \geq \nabla_c M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{c} - \boldsymbol{c}_{i,t}) + \nabla_a M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{a} - \boldsymbol{a}_{i,t}). \tag{17}$$

Furthermore, from the definition of $M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$, we have

$$\begin{aligned}
\nabla_c M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) &= \boldsymbol{F}_{i,t} \\
\nabla_a M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) &= \dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}} \odot \boldsymbol{c}_{i,t}.
\end{aligned}$$

Plugging those into equation (17), one obtains that

$$0 \geq \boldsymbol{F}'_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})(\boldsymbol{c} - \boldsymbol{c}_{i,t}) + \left(\dot{\boldsymbol{F}}_{i,t} \odot e^{-\boldsymbol{\omega}_{i,t}} \odot \boldsymbol{c}_{i,t}\right)' (\boldsymbol{a} - \boldsymbol{a}_{i,t}). \tag{18}$$

Using this last inequality in equation (16), one concludes that $u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) \geq u_i(\boldsymbol{c}, \boldsymbol{a})$. In other words, the data are rationalized.

Proof of Proposition 1

Assumption 4 states that the price function for any good $l \in \mathcal{L}$ is given by:

$$\log\left(p^*_{l,t}\right) = \alpha^0_l + \alpha^1_l \log(a_{l,t}) - \omega_{l,t}.$$

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

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

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 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 5 was used to eliminate the expected average search productivity. By Assumption 6, 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

$$\mathbb{E}\left[\frac{\partial \log\left(p_{l,t}^*\right)}{\partial \log(a_{l,t})}\right] = \frac{1}{L}\mathbb{E}\left[\alpha_l^1\right].$$ [48]

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 (average) expected elasticity of price with respect

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

to shopping intensity and $\overline{\alpha}^1 := \frac{1}{L}\sum_{l=1}^{L}\alpha_l^1$ is the average shopping technology.

Remark. *Measurement error implicitly accommodates various shocks that may occur outside the model. For example, it could capture changes in prices that arise due to fluctuations in transportation costs, or changes in prices that arise due to fluctuations in crop yields. Likewise, exogenous shocks may be absorbed by search productivity. For example, it could capture random fluctuations in the consumer degree of attention to prices. Accordingly, the model is robust to a variety of perturbations.*

A6: Price Search Rationalizability

The environment defined in Section 5.2 implies that for all $l \in \mathcal{L}$ and $s, t \in \mathcal{T}$, the model is characterized by the following moment functions:

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

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 variables satisfy their support constraints: $\boldsymbol{\alpha}^1_i \in [-1, 0]$, $\dot{\boldsymbol{p}}_{i,t} < 0$ and $\boldsymbol{p}^*_{i,t} > 0$, where $\dot{p}_{i,l,t}$ further satisfies

$$
\dot{p}_{i,l,t} = \alpha^0_{i,l}\alpha^1_{i,l}a_{i,l,t}^{\alpha^1_{i,l}-1}e^{-\omega_{i,l,t}} \quad \forall l \in \mathcal{L}.
$$

This equality constraint implies that $\dot{\boldsymbol{p}}_{i,t}$ 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 4. Under Assumptions 4-7, a data set $x := \{x_i\}_{i \in \mathcal{N}}$ is price search rationalizable (PS-rationalizable) if

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

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

This definition is exactly the same as that of statistical rationalizability with the moment functions specialized to the model of price search.

Lower Bound on Search Costs

Note that the first-order conditions of the consumer problem (2) give rise to the following relationship:

$$\frac{\partial u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})}{\partial a_{i,l,t}} = \frac{\partial p_{i,l}(a_{i,l,t}, \omega_{i,l,t})}{\partial a_{i,l,t}} c_{i,l,t} \quad \forall l \in \mathcal{L}.$$

This equation states that the disutility from price search is equal to the savings from price search at the optimum. Importantly, it allows one to obtain a simple lower bound on the search cost of any good $l \in \mathcal{L}$. Let $\boldsymbol{a}_{i,-l,t}$ denote $\boldsymbol{a}_{i,t}$ stripped of its lth element, then

$$\int_{a_{i,l,t}}^{0} \frac{\partial u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,-l,t}, a)}{\partial a}\, da = \int_{a_{i,l,t}}^{0} \frac{\partial p_{i,l}(a, \omega_{i,l,t})}{\partial a} c_{i,l,t}\, da \geq -\frac{\partial p_{i,l}(a_{i,l,t}, \omega_{i,l,t})}{\partial a_{i,l,t}} a_{i,l,t} c_{i,l,t},$$

where the inequality is obtained from the fact that the derivative of the price function at $a_{i,l,t}$ is smaller (in absolute) than at any $a < a_{i,l,t}$. From Theorem 2, we know that price search rationalizability is equivalent to the rationalizability of each individual data set by a well-behaved utility function and the moment conditions on search productivity. As such, the expected search cost is well-defined.

Relationship Between Rationalizability and GARP

A bundle $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is said to be directly revealed preferred to a bundle $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$ if $M_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) \leq 0$. Let R^D denote the direct revealed preference relation and let R denote its transitive closure.[49] When the inequality is strict, $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is said to be directly revealed strictly preferred to $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$ and is denoted P^D. In the case where there is a sequence $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) R^D (\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}), (\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) R^D (\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2}), \ldots,$ $(\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m}) R^D (\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$ of directly revealed preferences, where $t, t_1, \ldots, t_m, s \in \mathcal{T}$, $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is said to be revealed preferred to $(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$. Denote the direct revealed preference relation by R^D and the indirect revealed preference relation by R.

Definition 5. (GARP) If $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$ is revealed preferred to $(\boldsymbol{c}, \boldsymbol{a})$, then $(\boldsymbol{c}, \boldsymbol{a})$ is not strictly directly revealed preferred to $(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$.

The following result shows that GARP is not sufficient to obtain a concave utility function nor a convex production function. Thus, rationalizability as stated in Theorem 2 implies GARP but is not implied by it.

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

Proposition 2. *The following are equivalent:*

(i) The data set x_i is rationalized by a utility function that is continuous, increasing in $\boldsymbol{c}$, and decreasing in $\boldsymbol{a}$.

(ii) The data set x_i satisfies GARP.

Proof of Proposition 2

Definition 6. A square matrix M_i of dimension T is cyclically consistent if for every chain $\{t_1, t_2, \ldots, t_m\} \subset \{1, 2, \ldots, T\}$, $M_{i,t_1,t_2} \leq 0$, $M_{i,t_2,t_3} \leq 0$, $\ldots$, $M_{i,t_{m-1},t_m} \leq 0$ implies $M_{i,t_m,t_1} > 0$, where M_{i,t_i,t_j} represents the element in row t_i and column t_j of M_i.

Lemma 1. *GARP holds if and only if the matrix of revealed preferences M_i is cyclically consistent.*

For the sake of a contradiction, suppose GARP is violated. Thus, there is a sequence of indices $\{t_1, t_2, \ldots, t_m\}$ such that $(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) R (\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m})$ and $(\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m}) R^D (\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1})$. By definition, this implies that $M_{i,t_1}(\boldsymbol{c}_{t_2}, \boldsymbol{a}_{t_2}) \leq 0$, $M_{i,t_2}(\boldsymbol{c}_{t_3}, \boldsymbol{a}_{t_3}) \leq 0, \ldots, M_{i,t_m}(\boldsymbol{c}_{t_1}, \boldsymbol{a}_{t_1}) < 0$. Construct the matrix of revealed preferences M_i and note that the chain $\{t_1, t_2, \ldots, t_m\}$ violates cyclical consistency. Likewise, if cyclical consistency is violated, then by extracting a chain causing a violation one obtains a violation of GARP as each element of the matrix represents a revealed preference.

Lemma 2. *If a square matrix M_i of dimension T is cyclically consistent, then there exist numbers $u_{i,t}$, $\lambda_{i,t} > 0$, $t = 1, \ldots T$, such that for all $s, t \in \mathcal{T}$*

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

The proof is a complete analogue of Section 2 and 3 in Fostel, Scarf and Todd (2004).

I am now ready to prove Proposition 2.

$(i) \implies (ii)$

For the sake of a contradiction, suppose that GARP is violated in the data. Thus, there exists $t_1, t_2, \ldots, t_m \in \mathcal{T}$ such that $M_{i,t_1}(\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2}) \leq 0$, $M_{i,t_2}(\boldsymbol{c}_{i,t_3}, \boldsymbol{a}_{i,t_3}) \leq 0$, $\ldots$, $M_{i,t_m}(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) < 0$. If there exists a utility function that rationalizes the data, then this function must be such that $u_i(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1}) \geq u_i(\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2})$, $u_i(\boldsymbol{c}_{i,t_2}, \boldsymbol{a}_{i,t_2}) \geq u_i(\boldsymbol{c}_{i,t_3}, \boldsymbol{a}_{i,t_3})$, $\ldots$, $u_i(\boldsymbol{c}_{i,t_m}, \boldsymbol{a}_{i,t_m}) > u_i(\boldsymbol{c}_{i,t_1}, \boldsymbol{a}_{i,t_1})$. However, this sequence of inequalities is self-contradictory.

$(ii) \implies (i)$

Let M_i be a square matrix of dimension T whose element in row s and column t is $M_{i,s,t} := M_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})$. By Lemma 1, GARP holds if and only if the matrix of revealed preferences M_i is cyclically consistent. An application of Lemma 2 thus implies the existence of numbers $u_{i,t}$, $\lambda_{i,t} > 0$, $t = 1, \ldots T$, such that

$$u_{i,s} \leq u_{i,t} + \lambda_{i,t} M_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) \quad \forall s, t \in \mathcal{T}. \tag{19}$$

Let $u_i(\boldsymbol{c}, \boldsymbol{a}) = \min_{t \in \mathcal{T}} \{u_i + \lambda_{i,t} M_{i,t}(\boldsymbol{c}, \boldsymbol{a})\}$ and note that the utility function is locally nonsatiated, continuous, increasing in $\boldsymbol{c}$, and decreasing in $\boldsymbol{a}$. For budgets of type A, the latter is a consequence of the assumption that the production function is decreasing in $\boldsymbol{a}$. For budgets of type B, it is a consequence of the assumption that the production function is increasing in $\boldsymbol{a}$ and the distinct definition of $M_{i,t}(\boldsymbol{c}, \boldsymbol{a})$, i.e. $M_{i,t}(c, \boldsymbol{a}) = (c - F(\boldsymbol{a}, \boldsymbol{z}_{i,t})e^{-\omega_{i,t}}) - (c_{i,t} - F(\boldsymbol{a}_{i,t}, \boldsymbol{z}_{i,t})e^{-\omega_{i,t}})$.

To show that the utility function rationalizes the data, first note that for all $s \in \mathcal{T}$, $u_i(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s}) = \min_{t \in \mathcal{T}} \{u_{i,t} + \lambda_{i,t} M_{i,t}(\boldsymbol{c}_{i,s}, \boldsymbol{a}_{i,s})\} = u_{i,s}$. This can be seen from the inequalities in (19) and by noting that $M_{i,t}(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t}) = 0$ for all $t \in \mathcal{T}$. Hence, for any $(\boldsymbol{c}, \boldsymbol{a})$ such that $M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq 0$ we have $u_i(\boldsymbol{c}, \boldsymbol{a}) \leq u_{i,t} + \lambda_{i,t} M_{i,t}(\boldsymbol{c}, \boldsymbol{a}) \leq u_{i,t} = u_i(\boldsymbol{c}_{i,t}, \boldsymbol{a}_{i,t})$.