

# Revealed Preference with Limited Consideration\*

Thomas Demuynck

Christian Seel<sup>†</sup>

## Abstract

We derive revealed preference tests for models where individuals use consideration sets to simplify their consumption problem. Our basic test provides necessary and sufficient conditions for consistency of observed choices with the existence of consideration set restrictions. The same conditions can also be derived from a model in which the consideration set formation is endogenous and based on subjective prices. By imposing restrictions on these subjective prices, we obtain additional refined revealed preference tests. We illustrate and compare the performance of our tests by means of a dataset on household consumption choices.

*Keywords:* Revealed preference, bounded rationality, consideration set

*JEL code:* D11, D12, D81

## 1 Motivation

Choosing is difficult, especially if the set of available alternatives is large or if the alternatives are difficult to compare. A possible approach to simplify the decision problem is to first narrow down the set of all options into a smaller ‘consideration set’ and then make an optimal choice from this reduced set. Many studies in the marketing and psychology literatures provide strong evidence for such a ‘consider then choose’ decision process.<sup>1</sup> In this paper, we develop revealed preference tests for different models of choice behaviour with consideration sets. We use data from a large homescan consumer dataset on household grocery purchases in order to evaluate the empirical performance of the different models. The models based on consideration sets have positive predictive success and they outperform the model without consideration set restrictions.

Revealed preference analysis started with the seminal contributions of Samuelson (1938) and Houthakker (1950) and was further extended by Afriat (1967), Diewert (1973), and Varian (1982). A key axiom in revealed preference theory is *the Generalized Axiom of Revealed*

---

\*We would like to thank the coeditor John Asker and two anonymous referees for helpful comments on an earlier version of this paper. We also would like to thank Ian Crawford, Mark Dean, Bram Foubert, Thomas Kittsteiner, Arno Riedl and several seminar and conference audiences for useful comments.

<sup>†</sup>Demuynck: Ecares, Université Libre de Bruxelles, Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium and Maastricht University, School of Business and Economics, Tongersestraat 53, 6211 LM Maastricht, The Netherlands, thomas.demuynck@ulb.ac.be. Seel: Maastricht University, School of Business and Economics, Tongersestraat 53, 6211 LM Maastricht, The Netherlands, c.seel@maastrichtuniversity.nl.

<sup>1</sup>See Roberts and Lattin (1997) for an overview.

*Preference*, abbreviated as GARP. GARP provides necessary and sufficient conditions for a finite dataset on prices and quantities to be consistent with the neo-classical model of utility maximizing behaviour. In this canonical model, individuals consider all bundles in their budget set and choose a bundle that is best according to their utility function.

In many microeconomic consumption datasets, e.g. scanner datasets like the one we use in our empirical application, it is found that consumers limit their purchases to a small subset of the set of all available goods. This behaviour satisfies GARP if the decision maker buys none of the other goods because it is utility maximizing for her to do so.

An alternative explanation is that the purchased quantities of some goods are zero because these goods were not considered by the decision maker. Indeed, in a consumer theory setup, it is natural to assume that at a fundamental level what is being considered or not considered are goods and not bundles.<sup>2</sup> Accordingly, a consumer who considers to buy a good chooses the quantity of this good, but she might fail to consider some goods altogether. The consumed quantity of these non-considered goods is zero. If consumers form consideration sets, the observed choice behaviour may actually fail to satisfy GARP, although the choices are still optimal given her consideration sets. The main aim of this paper is to provide revealed preference tests that take into account this explanation.

Unfortunately, for real life datasets, consideration sets are not observable.<sup>3</sup> Thus, given the observed choice data, we ask whether it is possible to construct consideration sets of goods such that observed behavior is optimal for these sets. We solve this problem by providing a new revealed preference axiom, *the Limited Axiom of Revealed Preference* (LARP). Essentially, LARP boils down to verifying GARP on a partition of the original dataset, where the partition is determined by the goods having strictly positive consumption. LARP provides a set of necessary and sufficient conditions for consistency of choice behaviour when taking into account the presence of consideration sets.

Although LARP allows for the existence of consideration sets that determine which goods can be purchased in positive amounts, this basic consideration set model is agnostic about how consideration sets are formed. Therefore, we complement this basic model with a procedural model of consideration set formation in which the consumer also considers goods rather than bundles at the fundamental level.

In an initial stage, the consumer maximizes utility based on some subjective prices rather than the actual prices.<sup>4</sup> She writes down a list of goods that she intends to buy given these subjective prices (i.e. a shopping list). The initial consideration set consists of all goods on the list. Thus, the consumer might not consider to buy some goods because she overestimates their prices.

In the second stage, the consumer goes to the shop and learns the true prices of the goods in this consideration set. Thus, the subjective prices for the goods in the consideration set

---

<sup>2</sup>The same perspective is taken in the marketing literature, e.g., see Roberts and Lattin (1997).

<sup>3</sup>However, see Ben-Akiva and Boccara (1995) who use questionnaire data for the perceived consideration sets in a discrete choice setting with three alternatives; for an econometric approach to brand-choice probabilities based on estimated consideration sets for a small number of alternatives, see also Mehta, Rajiv and Srinivasan (2003).

<sup>4</sup>There is plenty of evidence on the existence and importance of subjective prices in the marketing literature; see, among others, Monroe (1973), Kalyanaram and Winer (1995), Mazumdar, Ray and Sinha (2005).

are updated to the true prices.<sup>5</sup> Given these updated prices, the decision maker may find it optimal to update her consideration set, i.e. she may introduce new goods that she would like to buy. For instance, if a good is much more expensive than expected, the consumer might purchase less of the good and/or look for a substitute good. In this case, she learns the true prices of these new goods and she may again choose to update her consideration set. This process continues until the decision maker no longer desires to update her consideration set. In the end, for all goods in the consideration set, the subjective prices correspond to the actual prices but for goods outside the consideration set, subjective prices may deviate from true prices.

Without any restrictions on the initial subjective prices, the procedural model is shown to be empirically indistinguishable from the basic model. We proceed by imposing two different types of restrictions on the initial subjective prices. The first refined model measures the mistake in the subjective price perception by comparing the minimal distance between subjective and true prices such that the observed behavior is consistent with the procedural model of consideration set formation. Conceptually, this model contributes to the literature on nearly optimal behavior (Afriat, 1973). The second refined model endogenizes the formation of price expectations. To do so, we adapt a simple model from the marketing literature to our setup.<sup>6</sup> In this model, subjective prices are formed based on previously observed prices. We provide revealed preference tests for both refined models.

We provide an empirical application where we implement and test the performance of the different models by means of a large household homescan dataset. Our models are nested in the sense that a household which satisfies GARP also satisfies both refined models and a household which satisfies a refined model also satisfies LARP. Thus, to obtain a reasonable comparison, we compute a measure of predictive success which relates the pass rate and the power (Selten, 1991). All models based on consideration sets have positive predictive success and outperform the standard GARP test. In this sense, limited consideration models appear to provide a better fit of observed consumer behaviour. In addition, we find that the empirical fit is higher for households with fewer members. A possible explanation for these findings could be that multi-member households use a different decision making model instead of the utility maximization model maintained in this paper.

Currently, there is no universally accepted theory of consideration set formation. We think that a revealed preference approach provides a natural environment to study this topic. Our paper provides some evidence in favour of consideration set formation with endogenous subjective prices. In this sense, we hope that it provides inspiration for more empirical research on the topic.

**Related Literature** First, our paper is related to a growing literature in economics that takes into account choice behaviour with consideration sets; see, among others, Eliaz and Spiegler (2011), Caplin, Dean and Martin (2011), Manzini and Mariotti (2012), Masatlioglu, Nakajima and Ozbay (2012), Masatlioglu and Nakajima (2013), Tyson (2013), and Manzini and Mariotti (2014). Most of these papers use a choice theoretic framework where choice

---

<sup>5</sup>The marketing literature also contains models in which the true prices of goods in the consideration set are learned and compared to expected prices; see, for example Bronnenberg and Vonhoner (1996).

<sup>6</sup>The marketing literature contains a myriad of models of price formation; for an overview, see Lowengard (2002).

sets are assumed to be discrete, i.e. individuals choose from a finite set of alternatives. In this paper, we consider consideration set formation for a model with continuous choice and we look at its testable implications from a revealed preference perspective à la Afriat.

Second, this paper is related to several papers on revealed preference theory with non-standard budget sets; see, among others, Forges and Minelli (2009) and Nishimura, Ok and Quah (2013). In our paper, quantities of goods outside the consideration set are by definition equal to zero. Thus, budget sets can be restricted to lie in a certain linear subspace of the full budget set. In addition to some technical differences,<sup>7</sup> Forges and Minelli (2009) and Nishimura, Ok and Quah (2013) assume that the (nonlinear) budget sets for the different choice situations are known. In our setting, however, the budget set is not completely known since the consideration set is not observable. In this sense, our analysis not only deals with having a non-standard constraint set, but moreover this set is not known a priori.

Finally, independently of our work, Crawford and Polisson (2015) apply revealed preference theory to datasets where prices are not observed for goods that are consumed with zero quantities. They obtain a revealed preference condition which has the same testable implications as one part of our characterization for the basic model; for the precise relation, see the discussion at the end of Section 3. The focus and analysis in Crawford and Polisson (2015) is quite different from the present paper. Their main target is on how to deal with missing prices for prediction and welfare analysis. On the other hand, we provide a theory of bounded rationality with a sequential process of consideration set formation.

Section 2 contains a short introduction to revealed preference theory. Section 3 presents the revealed preference characterization for the basic limited consideration model. In Section 4, we describe the procedural revealed preference model where the decision maker bases her consumption decision on her subjective prices. Section 5 contains the empirical application. Finally, Section 6 concludes. Most proofs are in the appendix.

## 2 Revealed Preference

The basic ingredient of revealed preference theory is a finite dataset on prices and quantities. Datasets are denoted by  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ , where  $T$  is a finite set of observations,  $\mathbf{p}_t = [p_t^1, \dots, p_t^n]$  is a row vector of strictly positive prices, and  $\mathbf{q}_t = [q_t^1, \dots, q_t^n]'$  is a column vector of non-negative consumption quantities for the  $n$  different goods. A dataset reflects the observed purchase behaviour  $\mathbf{q}_t$  for a decision maker who faces prices  $\mathbf{p}_t$  at observation  $t$ .

**Definition 1** *A dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  is rationalizable if there exists a (locally non-satiated) utility function  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  and for all observations  $t \in T$ , there exists an expenditure level  $m_t \geq 0$  such that*

$$\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{p}_t \mathbf{q} \leq m_t.$$

---

<sup>7</sup>Forges and Minelli (2009) requires that each (nonlinear) budget set contains at least one strictly positive bundle (see their Lemma 1) which is clearly not the case for settings with consideration sets. In contrast to our Theorem 2, neither Forges and Minelli (2009) nor Nishimura, Ok and Quah (2013) guarantee a rationalization by a concave utility function.

In words, a dataset is rationalizable if the chosen bundles,  $\mathbf{q}_t$ , are consistent with the model of utility maximization subject to a linear budget constraint.

The bundle  $\mathbf{q}_t$  is *directly revealed preferred* to the bundle  $\mathbf{q}_v$  if  $\mathbf{p}_t \mathbf{q}_t \geq \mathbf{p}_t \mathbf{q}_v$ , i.e. the bundle  $\mathbf{q}_t$  was at least as costly as the bundle  $\mathbf{q}_v$  at observation  $t$ . We write this as  $\mathbf{q}_t R \mathbf{q}_v$ . The *indirect revealed preference relation*  $R_T$  is the transitive closure of the relation  $R$ , i.e.  $\mathbf{q}_t R_T \mathbf{q}_v$  if there exists a (possibly empty) sequence  $r, s, \dots, w$  of observations in  $T$  such that  $\mathbf{q}_t R \mathbf{q}_r, \mathbf{q}_r R \mathbf{q}_s, \dots, \mathbf{q}_w R \mathbf{q}_v$ . The dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  satisfies *the Generalized Axiom of Revealed Preference* (GARP) if for all  $\mathbf{q}_t R_T \mathbf{q}_v$  it is not the case that  $\mathbf{p}_v \mathbf{q}_v > \mathbf{p}_v \mathbf{q}_t$ ; if  $\mathbf{q}_t$  is indirectly revealed preferred to  $\mathbf{q}_v$  it is not the case that  $\mathbf{q}_v$  was purchased although  $\mathbf{q}_t$  was less expensive.

In a seminal contribution to the literature, Afriat (1967) showed that GARP is a necessary and sufficient condition for rationalizability.

**Theorem 1 (Afriat's theorem)** *Consider a dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ . Then the following are equivalent:*

- (i)  *$S$  is rationalizable by a locally non-satiated utility function.*
- (ii)  *$S$  satisfies GARP.*
- (iii) *For all observations  $t \in T$ , there exist numbers  $\phi_t$  and  $\lambda_t > 0$  such that for all observations  $t, v \in T$ ,*

$$\phi_t - \phi_v \leq \lambda_v \mathbf{p}_v (\mathbf{q}_t - \mathbf{q}_v).$$

- (iv)  *$S$  is rationalizable by a strictly monotone and concave utility function.*

The linear inequalities in condition (iii) are called the *Afriat inequalities*. In order to grasp the intuition behind these inequalities, let us assume that the dataset is rationalizable by a utility function which is concave, strictly monotone and differentiable. Then, by concavity, we obtain that for all observations  $t, v \in T$ ,

$$u(\mathbf{q}_t) - u(\mathbf{q}_v) \leq \sum_{i=1}^n \frac{\partial u(\mathbf{q}_v)}{\partial q^i} (q_t^i - q_v^i).$$

From the first order conditions of the utility maximization problem, we also have that

$$\begin{aligned} \frac{\partial u(\mathbf{q}_v)}{\partial q^i} &= \lambda_v p_v^i \text{ if } q_v^i > 0, \\ \frac{\partial u(\mathbf{q}_v)}{\partial q^i} &\leq \lambda_v p_v^i \text{ if } q_v^i = 0, \end{aligned}$$

where  $\lambda_t$  is the Lagrange multiplier for the budget constraint. Substituting this first order condition into the concavity restriction and setting  $\phi_t = u(\mathbf{q}_t)$  and  $\phi_v = u(\mathbf{q}_v)$  gives

$$\phi_t - \phi_v \leq \lambda_v \mathbf{p}_v (\mathbf{q}_t - \mathbf{q}_v).$$

Thus, the Afriat inequalities provide us with an estimate of the utility levels  $\phi_t = u(\mathbf{q}_t)$  and marginal utility of income levels  $\lambda_t$  at the different observations.

### 3 Limited Consideration

In this section, we present a model of decision making based on consideration sets. In order to keep the analysis as general as possible, we abstain from placing any restriction on how consideration sets are formed.

Consider a decision maker with a strictly monotone utility function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$ .<sup>8</sup> At each observation  $t$ , the decision maker decides on her consumption bundle  $\mathbf{q}_t$  over the set of goods  $G = \{1, \dots, n\}$ . However, if she goes to the shop to purchase her bundle, she does not always take all goods into consideration.

More precisely, at each observation  $t$ , she only considers a subset  $I_t \subseteq G$  of goods. The set  $I_t$  is called her *consideration set* at observation  $t$ . It contains all goods that the decision maker considers for positive consumption, i.e. the consumption of each good outside this consideration set is equal to zero. For the goods in her consideration set, however, the decision maker chooses the quantities that maximize her utility. Formally, at each observation  $t$ , the decision maker faces the constrained optimization problem

$$\begin{aligned} \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{p}_t \mathbf{q} \leq m_t, \\ q^i = 0 \quad \forall i \notin I_t, \end{aligned}$$

where  $m_t$  is the total disposable income at observation  $t$ . Observe that there might be some goods  $j \in I_t$  for which  $q^j = 0$ . In words, it is possible that the decision maker considers some good  $j$ , but finds it optimal not to buy it.

**Definition 2** *A dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  is rationalizable with limited consideration if there exists a strictly monotone utility function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$  and for all  $t \in T$ , there exist incomes  $m_t$  and consideration sets  $I_t \subseteq G$  such that*

$$\begin{aligned} \mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{p}_t \mathbf{q} \leq m_t, \\ q^i = 0 \quad \forall i \notin I_t. \end{aligned}$$

Definition 2 differs from Definition 1, since we introduce additional constraints based on the consideration sets  $I_t$ , but allow for some freedom in choosing the consideration sets. Note that Definition 2 is weaker, since it boils down to Definition 1 for the special case where  $I_t = G$ .

For a dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  and an observation  $t \in T$ , let us denote by  $J_t$  the subset of goods for which the corresponding elements of  $\mathbf{q}_t$  are non-zero, i.e. the set of goods that have strictly positive components in  $\mathbf{q}_t$ :

$$J_t = \{i \in G \mid q_t^i > 0\}.$$

If the dataset is rationalizable with limited consideration, then we know that  $J_t \subseteq I_t$ : if some good is consumed, then it must be in the consideration set. The following lemma shows that there is no loss in generality by assuming that  $I_t = J_t$ .

---

<sup>8</sup>The strengthening from local non-satiation to strict monotonicity is necessary in the limited consideration framework, because local non-satiation is not sufficient to guarantee that the budget constraint is binding.

**Lemma 1** *A dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  is rationalizable with limited consideration if and only if it is rationalizable with limited consideration using the consideration sets  $I_t = J_t$ .*

Let us partition the set of observations  $T$  according to the equivalence classes generated by the sets  $J_t$ . Two observations  $t$  and  $v$  are in the same element of the partition if and only if  $J_t = J_v$ . Let us denote this partition by  $\mathcal{E}_S$ . Observe that  $\mathcal{E}_S$  is unique. The partition  $\mathcal{E}_S$  partitions the set of observations  $T$  and therefore the dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  into sub-datasets  $S_E = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in E}$ , where  $E \in \mathcal{E}_S$ . This motivates the following definition:

**Definition 3** *The dataset  $S$  satisfies the Limited Axiom of Revealed Preference (LARP) if each sub-dataset  $S_E$  satisfies GARP.*

We obtain the following main result:

**Theorem 2** *Consider a dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ . The following statements are equivalent:*

- (i)  *$S$  is rationalizable with limited consideration by a strictly monotone utility function.*
- (ii)  *$S$  satisfies LARP.*
- (iii) *For all  $t \in T$ , there exist numbers  $\phi_t, \lambda_t > 0$  and vectors  $\mathbf{P}_t \in \mathbb{R}_{++}^n$  such that for all  $t, v$ ,*

$$\phi_t - \phi_v \leq \lambda_v \sum_{i \in J_v} p_v^i (q_t^i - q_v^i) + \lambda_v \sum_{i \notin J_v} P_v^i (q_t^i - q_v^i).$$

- (iv) *There exist vectors  $\mathbf{P}_t \in \mathbb{R}_{++}^n$  such that  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP and for all  $t \in T$  and  $i \in G$ , if  $q_t^i > 0$ , then  $P_t^i = p_t^i$ .*
- (v)  *$S$  is rationalizable with limited consideration by a concave and strictly monotone utility function.*

Let us provide some intuition for Theorem 2.

The equivalence between (i) and (ii) shows that rationalizability with limited consideration is equivalent to LARP. This result might seem a bit surprising, since LARP does not impose any restriction on observations that have different goods with strictly positive consumption, i.e. where  $J_t \neq J_v$ . However, LARP entails a lot of freedom in choosing prices for goods with zero consumption. This allows us to show that in order to find violations of LARP, we can restrict attention to observation for which  $J_t$  is the same.

Condition (iii) gives an equivalence in terms of Afriat-type inequalities. The only difference between these inequalities and the ones in Theorem 1 is that for the goods outside  $J_v$ , the prices  $p_v^j$  are replaced by some (unobserved) prices  $P_v^j$ . These prices correspond to the possible values of the marginal willingness to pay of the decision maker for an increase in the consumption of good  $j$  at observation  $t$ . In the model without consideration sets, utility maximization requires that this marginal willingness to pay for a good is equal to the marginal cost  $p_v^j$  for  $q_v^j > 0$ , i.e.  $j \in J_v$  or is less than or equal to  $p_v^j$  if  $q_v^j = 0$ , i.e.  $j \notin J_v$ .

When consideration sets are taken into account, however, these marginal utilities can also be higher than the prices  $p_v^j$  for goods outside the consideration set.

Condition (iv) is an equivalent statement in terms of GARP for the entire dataset where the price vectors  $\mathbf{p}_t$  are replaced by the vectors  $\mathbf{P}_t$ . These prices are allowed to deviate from the true prices for goods that have zero consumption. The equivalence between (iii) and (iv) follows immediately from the equivalence between (ii) and (iii) in Theorem 1.

An alternative interpretation for condition (iv) is that the decision maker solves the same utility maximization model as in Section 2, but uses the price vector  $\mathbf{P}_t$  instead of the true price vector  $\mathbf{p}_t$ . The next section further explores this idea of  $\mathbf{P}_t$  as subjective prices in order to propose a model of consideration set formation.

Finally, let us compare our Theorem 2 to the result in Crawford and Polisson (2015). They look at the problem of missing prices: if a certain good is not consumed, then expenditures for this good are zero. Thus, the price level (per unit) of this good cannot be identified. Given this missing prices problem, they ask the following question: “what are the necessary and sufficient conditions such that we can construct prices for these non-consumed goods such that the observed purchase behaviour is consistent with utility maximizing behaviour, i.e. for the non-consumed goods, it is utility maximizing to choose a consumption level of zero”. Crawford and Polisson (2015) demonstrate that this problem is equivalent to the rationalizability of a utility maximizing model where goods with zero consumption are rationed to be zero (i.e. they cannot be bought at all). Intuitively, for the rationing model, we can replace actual prices with shadow prices for the rationed goods. Then rationality requires that the consumer satisfies GARP on a dataset of prices and quantities where true prices (of the non-consumed goods) are replaced by these shadow prices.

In our model of consideration sets, we do not look at a problem of missing prices. On the other hand, with consideration sets, goods with zero consumption can either be optimally chosen to be zero, or they can be zero because they are not in the consideration set. If they lie outside the consideration set, this is equivalent to the model where the particular good is rationed at zero. Lemma 1 shows that there is no loss of generality in assuming that all goods with zero consumption are outside the consideration set. Thus, the testable implications boil down to the rationing model where all goods with zero consumption amounts are rationed at zero. As such, both the consideration set model and the missing prices model boil down to a test of a model where non-consumed goods are rationed at zero. Indeed, Crawford and Polisson (2015) show that a dataset is rationalizable with the missing prices model if and only if it satisfies condition (iii) of our Theorem 2 holds, which provides the formal link between the testable implications of the two models.

## 4 A Model of Consideration Set Formation

In the previous section, consideration sets were exogenously determined and therefore beyond the control of the decision maker. In this section, we provide a sequential process which endogeneizes the formation of consideration sets based on the values of some subjective prices. The marketing literature finds a lot of evidence for the importance of such subjective prices (which they call *reference prices*) in consumer decisions; see, e.g. Monroe (1973), Kalyanaram and Winer (1995), and Mazumdar, Ray and Sinha (2005). By com-

binning subjective prices with (constrained) utility maximization, we provide a conceptually novel, intuitive framework how consideration sets are formed.

More specifically, suppose the consumer has some initial idea about the prices. Accordingly, she writes down a shopping list. Then, she goes to the shop, pays attention to the goods on her list and learns their true prices. Up to this point, our model is still in line with the marketing literature (e.g. Bronnenberg and Vonnacker, 1996) which now restricts attention to whether the good is chosen. In contrast, we do not only study whether the consumer still buys the good if the price is higher than expected, but also allow her to buy a lower quantity and/or look for (substitute) goods in this case. In this way, the consideration set might expand after learning the prices of some goods. This process of updating the consideration set stops when she decides not to consider any new goods.

To formalize this process, consider a decision maker that would like to spend her budget  $m_t$  on several goods. Before she goes shopping, she forms an idea about the bundle that she would like to buy based on some subjective idea about the prices of the different goods. We denote these subjective prices by  $\mathbf{P}_t \in \mathbb{R}_{++}^n$ .<sup>9</sup> In an initial phase, the decision maker solves

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{P}_t \mathbf{q} \leq m_t.$$

Let us denote by  $\tilde{\mathbf{q}}_t$  the solution to this problem. Using this bundle, the decision maker forms an initial consideration set by collecting all goods with positive amounts in  $\tilde{\mathbf{q}}_t$ , i.e.

$$\tilde{I}_t = \{i \in G | \tilde{q}_t^i > 0\}.$$

Next, the decision maker goes to the shop and investigates the goods in her consideration set  $\tilde{I}_t$  more thoroughly. Thereby, she learns about the actual prices of these goods. Given this updated information, her subjective prices of the goods in  $\tilde{I}_t$  are replaced by the true prices  $p_t^i$ . Therefore, she now faces the maximization problem

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \sum_{i \in \tilde{I}_t} p_t^i q^i + \sum_{i \notin \tilde{I}_t} P_t^i q^i \leq m_t.$$

Let  $\bar{\mathbf{q}}_t$  be the optimal solution to this problem. If there are new goods  $i \notin \tilde{I}_t$  for which  $\bar{q}_t^i > 0$ , these are added to the updated consideration set  $\bar{I}_t$ :

$$\bar{I}_t = \tilde{I}_t \cup \{i \in G | \bar{q}_t^i > 0\}.$$

In a next step, she learns the true prices of the new goods in the set  $\bar{I}_t$  and the optimization problem becomes

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \sum_{i \in \bar{I}_t} p_t^i q^i + \sum_{i \notin \bar{I}_t} P_t^i q^i \leq m_t.$$

This process of updating the consideration set and solving the associated utility maximization problem is iterated until no additional goods are added to the consideration set. Since the

---

<sup>9</sup>The case in which a consumer is completely unaware of good  $i$  can be included in this setup by setting  $P_t^i$  sufficiently large.

set of goods is finite and the consideration set cannot shrink, this procedure ends after a finite number of rounds.

The final outcome of this decision problem depends on the subjective prices  $\mathbf{P}_t$ . For a good  $i$  that never enters the consideration set, there are two possibilities. First, if  $P_t^i \leq p_t^i$ , the consumer would have chosen  $q^i = 0$  even if she had known the true price. For these goods, the usual revealed preference conditions pose no problem: if consumption of a certain good is zero, we may always replace the price by a higher price without strengthening the revealed preference conditions (see also Lemma 2 below). Second, if  $P_t^i > p_t^i$ , it might have been better for the consumer to include the good in her consideration set.

**Definition 4** *A dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  is rationalizable with subjective prices if there exists a (strictly monotone) utility function  $u(\cdot) : \mathbb{R}_+^n \rightarrow \mathbb{R}$  and for all  $t \in T$ , there exist incomes  $m_t$ , subjective prices  $P_t \in \mathbb{R}_{++}^n$  and consideration sets  $I_t \subseteq G$  such that for all  $t \in T$ , the following two conditions hold:*

- (i)  $\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q})$  s.t.  $\sum_{i \in I_t} p_t^i q^i + \sum_{i \notin I_t} P_t^i q^i \leq m_t$ .
- (ii) If  $i \notin I_t$ , then  $q_t^i = 0$ .

This rationalizability concept coincides with the final stage in the dynamic decision process outlined above. The first condition requires that the chosen consumption bundle  $\mathbf{q}_t$  maximizes the utility given the true prices of the goods in the consideration set and the subjective prices for the goods outside the consideration set. The second restriction ensures that the decision maker does not wish to include any additional goods into her consideration set.

As it turns out, the revealed preference conditions for this model are identical to those for the model with limited consideration presented in Section 3.

**Corollary 1** *Consider a dataset  $S$ . The dataset is rationalizable with subjective prices if and only if it satisfies LARP.*

To understand the intuition behind this Corollary, consider the equivalent formulation (iv) for LARP in Theorem 2. This condition requires the existence of subjective prices  $\mathbf{P}_t$  such that for all  $t \in T$ ,  $j \in G$ , we have  $q_t^j > 0$  implies  $P_t^j = p_t^j$  and  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP. The GARP condition implies consistency with utility maximizing behaviour for the prices  $\mathbf{P}_t$ . The other condition requires that for goods that are consumed, the subjective prices should be equal to the true prices. These correspond to the two conditions required for rationalizability with subjective prices.

**Restrictions on Subjective Prices** So far, we have abstained from imposing any constraints (bounds) on the values of the subjective prices  $\mathbf{P}_t$ . In this section, we consider different types of restrictions. The next lemma will be useful for deriving the results in this section.

**Lemma 2** *Consider two datasets  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  and  $\{\bar{\mathbf{P}}_t, \mathbf{q}_t\}_{t \in T}$  where for all  $i \in G, t \in T$ ,*

- (i) if  $q_t^i > 0$ , then  $\bar{P}_t^i = P_t^i$  and

(ii) if  $q_t^i = 0$ , then  $\bar{P}_t^i \geq P_t^i$ .

If the dataset  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP, then the dataset  $\{\bar{\mathbf{P}}_t, \mathbf{q}_t\}_{t \in T}$  also satisfies GARP.

The lemma states that if a dataset is consistent with GARP, then any dataset that is obtained by increasing the prices for the goods that have zero consumption is also consistent with GARP. Moreover, this lemma shows that imposing lower bounds on the subjective prices does not generate any additional testable restrictions. As such, we focus our analysis on placing upper bounds.

**Subjective Prices Based on True Prices** In the first model with constrained subjective prices, we require that the subjective prices are not too far from the true prices. Given some fixed number  $\theta \geq 1$ , we formalize this by requiring that the subjective prices are such that  $\mathbf{P}_t \leq \theta \mathbf{p}_t$ .

**Definition 5** A dataset  $S$  is  $\theta$ -rationalizable with subjective prices constrained by true prices if there exists a (strictly monotone) utility function  $u(\cdot) : \mathbb{R}_+^n \rightarrow \mathbb{R}$  and for all  $t \in T$ , there exist incomes  $m_t$ , subjective prices  $\mathbf{P}_t \in \mathbb{R}_{++}^n$  and consideration sets  $I_t \subseteq G$ , such that for all  $t \in T$ , the following three conditions hold:

$$(i) \quad \mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \sum_{i \in I_t} p_t^i q^i + \sum_{i \notin I_t} P_t^i q^i \leq m_t.$$

$$(ii) \quad P_t^i \leq \theta p_t^i \quad \forall i \notin I_t.$$

$$(iii) \quad \text{If } i \notin I_t, \text{ then } q_t^i = 0.$$

The following theorem characterizes the datasets that are rationalizable in this sense.

**Theorem 3** Consider a dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ . The following statements are equivalent:

(i) The dataset  $S$  is  $\theta$ -rationalizable with subjective prices constrained by true prices.

(ii) The dataset  $S$  satisfies  $\theta$ -LARP-T, i.e. there exist prices  $P_t \in \mathbb{R}_{++}^n$  such that  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP and for all  $t \in T$  and  $j \in G$ ,

$$P_t^j = \begin{cases} p_t^j & \text{if } q_t^j > 0 \\ \theta p_t^j & \text{if } q_t^j = 0. \end{cases}$$

It is easy to see that (ii) implies (i). The fact that (i) implies (ii) is a consequence of Lemma 2. If  $q_t^i = 0$  there are two cases: either  $i \in I_t$  in which case  $P_t^i = p_t^i$  or  $i \notin I_t$  in which case  $P_t^i \leq \theta p_t^i$ . As we do not know which case is the true one, we may take the highest value since this gives the weaker revealed preference restriction.

The parameter  $\theta$  can either be chosen ex ante or determined as the smallest value of  $\theta$  for which a dataset is still rationalizable (if the dataset satisfies LARP). This critical value of  $\theta$ , which we call the subjective price index (SPI), minimizes the (proportional) difference between the observed true prices  $p_v^i$  and the subjective price  $P_v^i$  that is necessary to rationalize the dataset. It can be found using a simple binary search algorithm. If the SPI is close to one, then the dataset can be rationalized by a model where the subjective prices are very close to the actual prices, i.e. the dataset is close to satisfying GARP. If the SPI is large, we need to allow for subjective prices that are far from the actual ones.

**Adaptive Subjective Prices** The previous model provides a test of how wrong the price expectations have to be such that behavior can be explained as optimal under the consideration set approach. However, it does not provide a procedure how these prices are formed. The marketing literature proposes a myriad of price formation models; see Lowengard (2002) for an overview. We adapt a simple model in which subjective prices are formed on the basis of last observed price of the good to our setting. As Gabor (1977) argues, “The price of the last purchase as remembered represents the price image of the good concerned, and it can be of great importance to the price setter to know how it compares with the actual price, since this will indicate how the market will respond to a price adjustment”.

To formalize this idea, assume that a certain good was in the consideration set at some time in the past. At that time, the subjective price was equal to the true price. Assuming adaptive subjective prices, the decision maker takes the last price she observed as the current subjective price. In order to formalize this setting, consider for any good  $j$  and observation  $t$ , the value  $\ell(j, t)$ , where

$$\ell(j, t) = \begin{cases} \max\{k \in T \mid k \leq t, j \in I_k\} & \text{if } j \in \bigcup_{k \leq t} I_k, \\ 0 & \text{if } j \notin \bigcup_{k \leq t} I_k. \end{cases}$$

The value of  $\ell(j, t)$  gives the last period where good  $j$  was in the consideration set. Observe that if good  $j$  is in  $I_t$ , then  $\ell(j, t) = t$ . Of course, it is possible that  $j$  was in no consideration set before observation  $t$ . In this case  $\ell(j, t)$  is set to zero.

We assume that the subjective prices about good  $j$  in period  $t$  are such that

$$P_t^j = p_{\ell(j,t)}^j \text{ if } \ell(j, t) > 0.$$

On the other hand, if  $j$  was not previously purchased, i.e.  $\ell(j, t) = 0$ , we impose no restrictions on the subjective prices  $P_t^j$ .

**Definition 6** *A dataset  $S$  is rationalizable with adaptive subjective prices if there exists a (strictly monotone) utility function  $u(\cdot) : \mathbb{R}_+^n \rightarrow \mathbb{R}$  and for all  $t \in T$ , there exist incomes  $m_t$ , subjective prices  $P_t \in \mathbb{R}_{++}^n$  and consideration sets  $I_t \subseteq G$ , such that for all  $t \in T$ ,*

- (i)  $\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q})$  s.t.  $\sum_{i \in I_t} p_t^i q^i + \sum_{i \notin I_t} P_t^i q^i \leq m_t$ .
- (ii)  $P_t^i = p_{\ell(i,t)}^i$  if  $\ell(i, t) > 0$ .
- (iii) If  $i \notin I_t$ , then  $q_t^i = 0$ .

Towards the revealed preference characterization, define  $l(j, t)$  as

$$l(j, t) = \begin{cases} \max\{k \in T \mid k \leq t, q_k^j > 0\} & \text{if } j \in \bigcup_{k \leq t} J_k, \\ 0 & \text{if } j \notin \bigcup_{k \leq t} J_k. \end{cases}$$

The function  $l(j, t)$  gives the last observation prior or equal to  $t$  where good  $j$  was consumed with strictly positive amount. If good  $j$  was not consumed prior to observation  $t$ , we set  $l(j, t)$  equal to zero. Observe that if  $q_t^j > 0$ , then  $l(j, t) = t$  and that  $\ell(j, t) \geq l(j, t)$ . If  $l(j, t) > 0$  this follows from the fact that  $j \in I_{l(j,t)}$ . If  $l(j, t) = 0$  this inequality is trivially satisfied. Using the definition of  $l(j, t)$ , we can now state the revealed preference result:

**Theorem 4** Consider a dataset  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ . The following statements are equivalent:

- (i) The dataset  $S$  is rationalizable with adaptive subjective prices.
- (ii) The dataset  $S$  satisfies LARP-AP, i.e. there exist subjective prices  $\mathbf{P}_t \in \mathbb{R}_{++}^n$  such that  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP and for all  $t \in T$ ,  $j \in G$ ,

$$P_t^j = \begin{cases} \max_{v|t \geq v \geq l(j,t)} p_v^j, & \text{if } l(j,t) > 0, \\ \text{unrestricted} & \text{if } l(j,t) = 0. \end{cases}$$

Again, the implication from (i) to (ii) follows from Lemma 2. If  $l(j,t) > 0$ , we know that  $j \in I_{l(j,t)}$ , i.e.  $j$  was in the consideration set at observation  $l(j,t)$ . Also, we cannot exclude the possibility that  $j$  was in any other consideration set after  $l(j,t)$ . Thus, we take the subjective price to be the highest price over all observations  $v$  for which  $l(j,t) \leq v \leq t$ . In particular, if  $l(j,t) = t$ , i.e.  $q_t^j > 0$ , we obtain  $P_t^j = p_t^j$ .

## 5 Application

We illustrate our results using data from ACNielsen's Homescan Panel from the Denver area, a large homescan dataset. The dataset captures a wide variety of grocery packed goods purchased by a large number of households at a large number of retail shops. It covers the period January 1993 through March 1995. The dataset is designed to be representative of the Denver metropolitan area. Households are equipped with an electronic home scanning unit. After every shopping trip, the shopper scans the UPC (Unique Product Code) of all purchased goods and registers the date. Given the UPC, it is possible to match the purchased goods with price information for the stores within the area. The full dataset consists of 2,100 separate households and over 950,000 transactions. A more detailed description of the dataset can be found in Aguiar and Hurst (2007).<sup>10</sup>

We aggregate the purchase data to monthly aggregates. Shorter time frames could lead to situations where purchases of one period may be stored and consumed in another period. Longer periods seem unnecessary since we focus on nondurable grocery purchases. We assume that grocery purchases are separable in the utility function from all other goods and services.

We define a good on the basis of the UPC. As such, each purchase in the dataset with a different UPC is seen as the purchase of a different good. For every shopping trip and each good bought, every household specifies the amount bought and price paid. As such, we define per unit prices for these goods by dividing total expenditure on this UPC (over the month) by the total quantity bought. However, the dataset has a lot of missing prices. If a household does not buy a specific good in a certain month, we do not know the price for this good. In order to minimize the number of missing prices, we look if there was any other household that bought this good in the same period. If so, we take the average price paid by these other households as the relevant price. In this way, we try to minimize the number of missing prices.<sup>11</sup> Given the minor level of aggregation, most goods are actually purchased in

<sup>10</sup>See also Dean and Martin (2016) for another use of the same dataset for revealed preference tests.

<sup>11</sup>In order to obtain reliable price data, we omit goods for which the total 'aggregate' quantity bought (over all households) is below 20 for some month.

Table 1: Summary statistics

Number of households	550
Fraction zero consumption observations	0.84 (0.048)
Number of distinct goods per household	30.57 (10.69)

Sample standard deviations are between brackets.

discrete amounts. This does not immediately fit our framework where quantities are assumed to be continuous. However, Polisson and Quah (2013) recently showed that even in a discrete goods framework, the usual revealed preference conditions hold if utility is separable in the observed goods and an unobserved numeraire good which is infinitely divisible (e.g. money).

Households are observed up to 27 months. This gives us a maximum of 27 observations per household. We use the panel structure in the dataset to verify the revealed preference restrictions for each household separately. We further restrict our panel to households for which we have consumption and relevant price data over all 27 observations. This leaves us with a total of 550 households. We refer to Appendix C for a detailed account of our data cleaning procedure.

Table 1 gives some summary statistics on the dataset. On average, over the 27 months, a household purchases around 30.57 different goods although there is considerable variation among households. Further, on average 84 percent of all quantities (good-observation pairs) are zero which means that the consumption data is quite sparse. As mentioned in the introduction, there might be several reasons why quantities might be zero. In particular, consumers might either find it optimal not to consume certain goods or those the goods were not in their consideration set. Our GARP/LARP tests allow us to separately test these two hypotheses.

**Pass rates** A first metric to evaluate the different revealed preference tests is the pass rate. The pass rate gives the percentage of all households that are consistent with the revealed preference test under consideration. Table 2 provides the results. The pass rate for the GARP test is 5.3% which is quite low. On the other hand, the pass rate for the LARP test (97.8%) is close to unity implying that nearly all households are consistent with utility maximizing behaviour if we allow for the presence of consideration sets. However, one must be cautious. When we look at the size of the partitions that are induced by the LARP test, we see that there are on average 7.74 equivalence classes per household. In other words, the high pass rate of LARP is mainly due to the fact that people tend to buy new goods and stop buying other goods quite frequently. Table 2 also gives the pass rates for rationalizability with subjective prices for various values of  $\theta$ . This pass rate ranges from 66% (for  $\theta = 1.25$ ) to 97% (for  $\theta = 2.25$ ). Given that the revealed preference test becomes weaker as  $\theta$  increases, the pass rate cannot decrease. We omitted the results for  $\theta = 1$  as this test coincides with the usual GARP test. Finally, the pass rate for rationalizability with adaptive subjective prices is about 61%.

Table 2: Pass rates, power and predictive success

Test	pass rate	power	pred. succ.	95% CI
GARP	0.0527 (0.2237)	0.9484 (0.1297)	0.0011 (0.2263)	$[-0.0178 \quad 0.0200]$
LARP	0.9782 (0.1462)	0.0672 (0.1647)	0.0454 (0.1702)	$[0.0311 \quad 0.0596]$
$\theta$ -LARP-T				
$\theta = 1.25$	0.6618 (0.4735)	0.3830 (0.3094)	0.0448 (0.3969)	$[0.0117 \quad 0.0780]$
$\theta = 1.5$	0.8873 (0.3165)	0.1791 (0.2384)	0.0664 (0.2803)	$[0.0430 \quad 0.0898]$
$\theta = 1.75$	0.9473 (0.2237)	0.1212 (0.1981)	0.0684 (0.2336)	$[0.0489 \quad 0.0880]$
$\theta = 2$	0.9655 (0.1828)	0.0992 (0.1812)	0.0646 (0.2020)	$[0.0478 \quad 0.0815]$
$\theta = 2.25$	0.9709 (0.1682)	0.0901 (0.1753)	0.0610 (0.1892)	$[0.0452 \quad 0.0768]$
LARP-AP	0.6091 (0.4884)	0.4512 (0.3320)	0.0603 (0.3928)	$[0.0275 \quad 0.0931]$

Pass rate, power, predictive success and 95% asymptotic confidence intervals for the predictive success for the various revealed preference tests. Sample standard deviations for the different measures are between brackets.

**Power and predictive success** The different revealed preference tests are nested in the sense that, for example, a dataset that passes GARP automatically satisfies LARP or any other revealed preference test. The tests that impose more structure on the subjective prices are stronger than LARP, so their pass rates cannot be higher. In this sense, it is unreasonable to compare the performance of the different models solely on the basis of the pass rates. In order to account for the nestedness of the different models, it is crucial to perform a power analysis. Intuitively, the power of a revealed preference model is given by the probability of rejecting the model when this model is not the true data generating process.

Usually, the power is computed using Bronars' (1987) procedure. This procedure computes the probability that a revealed preference test rejects seemingly irrational (or random) behaviour based on the model of irrational behaviour from Becker (1962). In order to compute the Bronars power of a certain dataset for a specific revealed preference test one first generates a large number of random datasets. Each of these random datasets is obtained by drawing for each observation  $t \in T$  a random consumption bundle uniformly from the budget hyperplane at observation  $t$  (i.e. for prices  $\mathbf{p}_t$  and income  $m_t = \mathbf{p}_t \mathbf{q}_t$ ). The power is then determined by the proportion of these randomly generated datasets that fail the revealed preference test. A higher power therefore means that less randomly generated datasets pass

the revealed preference test.<sup>12</sup>

Computing the Bronars power in our setting is a bit problematic. Given that the random consumption bundles for the Bronars power are drawn from a uniform distribution on the budget hyperplane, the probability of generating zero consumption quantities is zero. For such datasets (where  $q_t^j > 0$  for all  $t$  and  $j$ ), all tests coincide with the usual GARP test. Thus, all models have the same Bronars power.

In order to alleviate this problem, we propose a modified power measure which takes into account the large frequency of zero consumption quantities in the datasets. Instead of drawing uniform bundles from the entire budget hyperplanes, we draw uniform bundles from the sub-hyperplane generated by the non-zero consumption bundles in the vector  $\mathbf{q}_t$ . In other words, all randomly generated bundles are restricted to have zero consumption quantities for the goods in  $G$  that also have zero consumption in the vector  $\mathbf{q}_t$ . For each of the 550 households, we generate 1000 such random datasets. The power for a certain household is then computed as the proportion of these datasets that fail the corresponding revealed preference test. The mean of the power results are given in Table 2. Of course, stronger revealed preference tests have higher power given that they reject a larger fraction of the randomly generated datasets.

As can be seen from Table 2, the pass rates and power of the various models are inversely related, i.e. low power is associated with higher pass rates and vice versa. The next step is to combine the two performance measures into a single index such that they can be used as a reliable criterion for comparing different but possibly nested models. We do this by using the measure of predictive success introduced by Selten (1991) and popularized for use in revealed preference analysis by Beatty and Crawford (2011). Predictive success can be directly calculated as the difference between the pass rate and one minus the power.

$$\text{Predictive success} = \text{Pass rate} - (1 - \text{Power}).$$

The pass rate measures the percentage of households that satisfy a certain revealed preference test. One minus the power measures the percentage of randomly generated datasets that satisfy the same revealed preference test. In other words, it gives the expected pass rate if behaviour is random. As such, the difference determines how much actual behaviour agrees better with the revealed preference test compared to what is expected if behaviour is random. Negative predictive success values—low pass rate combined with low power—suggest that the revealed preference test is inadequate for describing observed consumer behaviour, since it is at least as good at explaining random behaviour. On the other hand, positive predictive success—high pass rate combined with high power—points to a potentially useful model that is able to reject irrational behaviour while explaining observed behaviour.

The predictive success is given in the third column of Table 2. The last column gives the 95% asymptotic confidence intervals. All models have predictive success measures that are larger than zero, which means that they outperform the model which is based on random behaviour. However, the predictive success for the GARP model is not statistically significantly different from zero at the 5% confidence level. The model with subjective prices where  $\theta = 1.75$  has the highest predictive success. Appendix B contains details on the statistical

---

<sup>12</sup>See Andreoni, Gillen and Harbaugh (2013) for various alternative procedures to compute the power of a revealed preference test.

significance of the differences in predictive success between the various models. From this, we see that GARP performs worst in terms of predictive success. The difference between the predictive success of GARP and any other model is negative (at the 5% level). In addition, we see that the predictive success of LARP is smaller than for the models with adaptive subjective prices except for  $\theta = 1.25$ .

**Goodness of Fit** The revealed preference tests tell us whether or not a dataset is consistent with the particular model that is being tested. However, as argued by Varian (1990), in many cases, nearly optimizing behaviour is just as good as optimizing behaviour. In this sense, it is useful to have some indication how close a given dataset is to satisfying GARP when it violates the revealed preference conditions. Usually this is measured by using a goodness-of-fit measure. The most popular goodness-of-fit measure in the revealed preference literature is Afriat’s (1973) critical cost efficiency index. Intuitively, the critical cost efficiency index measures the amount by which each (linear) budget set must be minimally adjusted in order to remove all GARP violations. The left part of Figure 1 shows the kernel estimation of the distribution of the Afriat index for our dataset, i.e. the distribution of the critical cost efficiency index over the households in the sample that violate GARP. We see that the distribution is highly skewed towards one with a peak around 0.95.

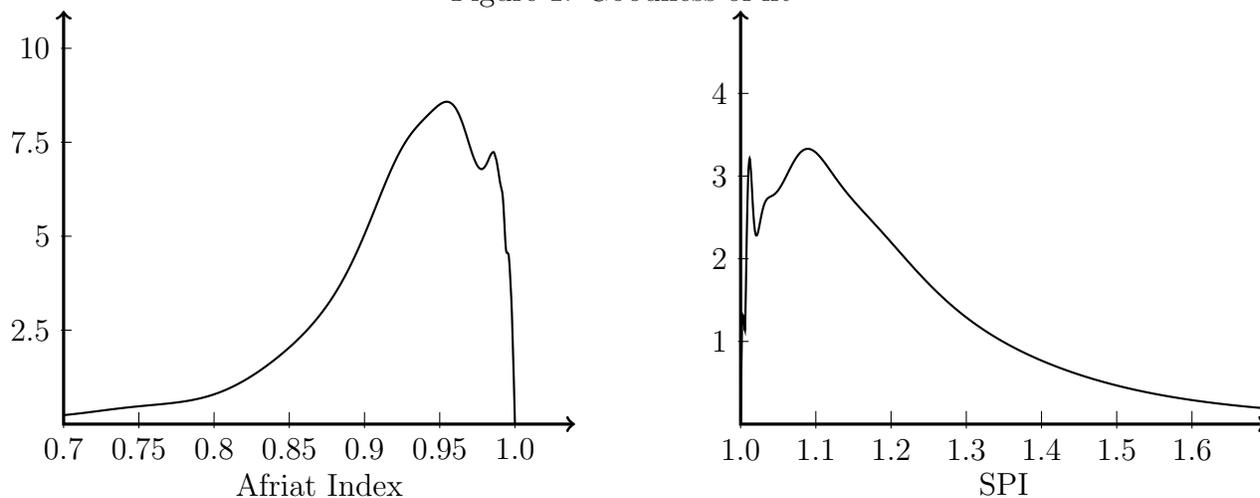
In Section 4, we discussed a similar index, the subjective price index (SPI), as the smallest value of  $\theta$  for which a dataset passes  $\theta$ -LARP-T. The SPI minimizes the (proportional) difference between the observed prices and the subjective prices that are necessary for the dataset to be rationalizable. The closer the SPI to one, the closer the dataset is to satisfying GARP. The distribution of this value is shown on the right part of Figure 1. For this, we restricted our data to those households that satisfy LARP but violate GARP. We see that the distribution is heavily skewed towards one with a peak around 1.1. The SPI and the Afriat critical cost efficiency index turn out to be highly correlated (-0.443),<sup>13</sup> although the underlying motivation behind the two measures is vastly different.

**Observable Heterogeneity** As a final exercise, we investigate whether the predictive success of the different models and the SPI are related to observable heterogeneity by means of a linear regression on several observable household characteristics. Our first characteristic, *pshopage*, is the age of the household member who goes shopping. We choose this variable instead of the variable of the age for the husband and wife separately because this variable is continuously measured. Next, *female and male college* are dummy variables that equal one if the corresponding household members has at least a college degree. *Income* is a dummy variable that equals one if the household (yearly) income is above 50 000 dollars. The variable *hhsiz*e measures the size of the household. Finally, the variable *conc* measures to what extent the shopping was done by the same person. From the dataset, we know for each shopping trip whether the shopper was a male or female. Denote by  $\alpha$  the fraction of times the shopping was done by a male. Next, we define *conc* as  $2 \cdot \max\{\alpha, 1 - \alpha\} - 1$ . As such, *conc* equals 0 if half of the shopping trips were conducted by the household male head and half by the household female head. It equals one if all shopping was done by the

---

<sup>13</sup>This correlation is statistically different from zero at the 1% level.

Figure 1: Goodness of fit



Left figure, kernel density estimation for the Afriat critical cost efficiency index. Right figure, kernel density estimation for the SPI index. The left density is obtained using households that violate GARP. The right densities is obtained from the sample of households that violate GARP but satisfy LARP.

same person. We use this variable to see whether households where the shopping is done by mostly the same person have a higher predictive success.

The results are given in Table 3. The coefficients for the variables *female college*, *pshopage* and *conc* are not statistically significantly different from zero for any of the models (at the 10% level) and their signs vary among the different models of limited consideration. The dummy for *male college* has a somewhat positive effect on the predictive success of the GARP model. On the other hand, the dummy for *income* has a negative effect on the predictive success of the GARP model but has a positive sign for the models with limited consideration although the effect is not statistically significant.

The variable *pshopage* also seems to to have a negative effect on the value of SPI. In other words, the older the age of the household member who goes shopping the better the fit with the standard GARP test.

Importantly, the coefficient for the household size variable, *hhsiz*, is negative and statistically significant at the 5% level for all models with limited consideration. In other words, households with more members appear to have a lower predictive success for limited consideration models. However, there seems to be no statistically significant effect on the predictive success of the GARP test. In addition, a larger household size tends to lower the value of the SPI which means that the datasets of larger households come closer to passing the standard GARP test. A possible explanation for these findings could be that multi-member households use a different decision making model instead of the utility maximization model maintained in this paper. Indeed, parts of the household decision literature argue that multi-person household consumption behaviour differs from a single decision maker that optimizes a household utility function; see, for example, Fortin and Lacroix (1997), Browning and Chiappori (1998) and Cherchye and Vermeulen (2008). Intuitively, if different household members hold different subjective prices, the household need not satisfy the revealed preference conditions even if each individual shopper satisfies the conditions.

Table 3: Linear regression of household predictive success on observable household characteristics

	GARP	LARP	$\theta$ -LARP-T $\theta = 1.75$	LARP-AP	$\theta^*$
constant	0.0549 (0.0627)	0.1206*** (0.0440)	0.1108** (0.0579)	0.1483* (0.1024)	1.3905*** (0.0914)
pshopage	-0.0007 (0.0007)	-0.0003 (0.0006)	0.0005 (0.0008)	0.0004 (0.0013)	-0.0022** (0.0012)
female college	-0.0053 (0.0209)	0.0043 (0.0218)	0.0176 (0.0278)	-0.0170 (0.0412)	-0.0121 (0.0241)
male college	0.0392* (0.0259)	-0.0200 (0.0212)	0.0037 (0.0258)	-0.0314 (0.0429)	-0.0303 (0.0241)
income	-0.0385** (0.0233)	0.0174 (0.0185)	0.0029 (0.0272)	0.0221 (0.0472)	0.0008 (0.0243)
hhsiz	-0.0107 (0.0083)	-0.0277*** (0.0061)	-0.0300*** (0.0075)	-0.0250** (0.0136)	-0.0229*** (0.0075)
conc	0.01160 (0.0390)	0.0193 (0.0305)	0.0152 (0.0375)	-0.0417 (0.0626)	0.0222 (0.0353)

Regression output from household level predictive success on household characteristics. Heteroscedastic consistent standard errors are between brackets. \*\*\* = significant at the 1% level;

\*\* = significant at the 5% level; \* = significant at the 10% level.

## 6 Conclusion

Models of consumer behaviour on the basis of consideration sets impose weaker assumptions than the standard paradigm of utility maximization subject to a budget constraint. Consumers still maximize their consumption bundles given their utility functions and market prices, but they only consider a subset from the set of all available goods—the consideration set. To the best of our knowledge, we have presented and implemented the first revealed preference analysis for this consideration set consumption model.

Without imposing any restrictions on the way consideration sets are formed, the revealed preference restrictions boil down to verifying an axiom which we call the Limited Axiom of Revealed Preference (LARP). In addition, we have derived two refined revealed preference tests that restrict the consideration set formation. Conceptually, the refined models impose stronger assumptions than LARP, but weaker assumptions than the standard GARP test.

We have applied our different revealed preference tests using a large dataset on household grocery consumption. The consideration set models have positive predictive success and the refined models with endogenous consideration set formation appear to outperform GARP (and LARP). Thus, our paper provides empirical support for the consideration set models.

We see several avenues for follow-up research. First, up to now, there seems to be no universally accepted framework to model the formation of consideration sets. We see our paper, which presents models based on subjective prices, as a first step in that direction.

Clearly, more theoretical and empirical work is needed in order to establish a unified theory of consideration set formation. In this respect, the revealed preference approach seems to be a promising way to obtain empirical tests and to compare the empirical performance of different models.

Second, in the empirical part, we have focused mainly on computing and comparing the pass-rate, power and predictive success of the different models. In Sections 4 and 5, we have also introduced the SPI as a new goodness-of-fit measure and compared the correlation with the most well-known measure in the literature, the Afriat Critical Cost Efficiency Index (Afriat, 1973). An interesting extension would be to see whether other measures such as the Varian index (Varian, 1990), the Houtman-Maks index (Houtman and Maks, 1985) or more recent measures suggested by Echenique, Lee and Shum (2011), Apesteguia and Ballester (2015) and Dean and Martin (2016) could also be applied or adapted to the revealed preference tests developed in this paper.

Finally, we have found that the empirical fit of the different consideration set models decreases with the size of the household. One possible explanation for the negative correlation between predictive success and household size is that multi-member households may use a different decision making model. Multi-member households are composed of distinct individuals who are endowed with their own preferences. As a result, household consumption decisions will be determined by the specific intrahousehold decision mechanism which need not coincide with the standard unitary utility maximization model. A fruitful extension of our results would be to combine such collective models of household behaviour with our model of consideration set formation. This extension could build upon our results and other existing revealed preference models for non-unitary models of household consumption behaviour (see, among others, Cherchye, De Rock and Vermeulen, 2007, 2009, 2011 and Cherchye, Demuyne and De Rock, 2011).

## References

- Afriat, Sydney N.** 1967. “The Construction of Utility Functions from Expenditure Data.” *International Economic Review*, 8: 67–77.
- Afriat, Sydney N.** 1973. “On a System of Inequalities in Demand Analysis: An Extension of the Classical Method.” *International Economic Review*, 14: 460-472
- Aguiar, Mark, and Erik Hurst.** 2007. “Life-Cycle Prices and Production.” *American Economic Review*, 97: 1533–1559.
- Andreoni, James, Benjamin J Gillen and William T Harbaugh.** 2013. “The Power of Revealed Preference Tests: Ex-Post Evaluation of Experimental Design.” *Working Paper*
- Apesteguia, José and Miguel A Ballester.** 2015. “A Measure of Rationality and Welfare” *Journal of Political Economy*, 123: 1278-1310.
- Beatty, Timothy K M, and Ian A Crawford.** 2011. “How Demanding is the Revealed Preference Approach to Demand.” *American Economic Review*, 101: 2782–2795.

- Becker, Gary S.** 1962. "Irrational Behavior and Economic Theory." *Journal of Political Economy*, 70: 1–13.
- Ben-Akiva, M, and Bruno Boccara.** 1995. "Discrete choice models with latent choice sets." *International Journal of Research in Marketing*, 12: 9–24.
- Bronars, Stephen G.** 1987. "The Power of Nonparametric Tests of Preference Maximization." *Econometrica*, 55: 693–698.
- Bronnenberg, Bart J., and Wilfried R. Vonhonacker.** 1996. "Limited Choice Sets, Local Price Response, and Implied Measures of Price Competition." *Journal of Marketing Research*, 33: 163-173.
- Browning, Martin, and Pierre-André Chiappori.** 1998. "Efficient Intra-household Allocations: A General Characterization and Empirical Tests." *Econometrica*, 66: 1241–1278.
- Caplin, Andrew, Mark Dean, and Daniel Martin.** 2011. "Search and Satisficing." *American Economic Review*, 101: 2899-2922.
- Cherchye, Laurens, and Frederic Vermeulen.** 2008. "Nonparametric Analysis of Household Labor Supply: Goodness of Fit and Power of the Unitary and the Collective Model." *Review of Economics and Statistics*, 90: 267–274.
- Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen.** 2007. "The Collective Model of Household Consumption: A Nonparametric Characterization." *Econometrica*, 75: 553–574.
- Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen.** 2009. "Opening the Black Box of Intrahousehold Decision Making: Theory and Nonparametric Empirical Tests of General Collective Consumption Models." *Journal of Political Economy*, 117: 1074–1104.
- Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen.** 2011. "The Revealed Preference Approach to Collective Consumption Behaviour: Testing and Sharing Rule Recovery." *Review of Economic Studies*, 78: 176–198.
- Cherchye, Laurens, Thomas Demuynck, and Bram De Rock.** 2011. "Revealed Preference Analysis of Noncooperative Household Consumption." *Economic Journal*, 121: 1073–1096.
- Crawford, Ian. and Matthew Polisson** 2015. "Demand Analysis with Partially Observed Prices." Working Paper.
- Dean, Mark and Daniel Martin.** 2016. "Measuring Rationality with the Minimum Cost of Revealed Preference Violations." *Review of Economics and Statistics*, 98: 524-534.
- Demuynck, Thomas.** 2015. "Statistical Inference for Measures of Predictive Success." *Theory and Decision*, 79: 689-699.

- Diewert, Walter E.** 1973. "Afriat and Revealed Preference Theory." *Review of Economic Studies*, 40: 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: 1201–1223.
- Eliasz, Kfir, and Ran Spiegler.** 2011. "Consideration Sets and Competitive Marketing." *Review of Economic Studies*, 78: 235–262.
- Forges, Françoise, and Enrico Minelli** 2009. "Afriat's Theorem for General Budget Sets." *Journal of Economic Theory*, 144: 135–145.
- Fortin, Bernard, and Guy Lacroix.** 1997. "A Test of the Unitary and Collective Models of Household Labour Supply." *Economic Journal*, 107: 933–955.
- Fostel, Ana, Herbert E Scarf, and Mike J Todd.** 2004. "Two New Proofs of Afriat's theorem." *Economic Theory*, 24: 211–219.
- Gabor, Andre.** 1977. "Pricing, principals and practices", *London: Heinemann Educational*.
- Houthakker, Hendrik S.** 1950. "Revealed Preference and the Utility Function." *Economica*, 17: 159–174.
- Houtman, Martijn, and Johannes A H Maks.** 1985. "Determining all Maximal Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden*, 19: 89–104.
- Kalyanaram, Gurusurthy, and Russell S Winer.** 1995. "Empirical Generalizations from Reference Price Research." *Marketing Science*, 14: G161–G169.
- Lowengart, Oded.** 2002. "Reference Price Conceptualisations: An Integrative Framework of Analysis." *Journal of Marketing Management*, 18: 145–171.
- Manzini, Paola, and Marco Mariotti.** 2012. "Categorize Then Choose: Boundedly Rational Choice and Welfare." *Journal of the European Economic Association*, 10: 1141–1165.
- Manzini, Paola, and Marco Mariotti.** 2014. "Stochastic Choice and Consideration Sets." *Econometrica*, 82: 1153–1176.
- Masatlioglu, Yusufcan, and Daisuke Nakajima.** 2013. "Choice by Iterative Search." *Theoretical Economics*, 8: 701–728.
- Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y Ozbay.** 2012. "Revealed Attention." *American Economic Review*, 102: 2183–2205.
- Mazumdar, Tridib, S.P. Ray, and Indrajit Sinha.** 2005. "Reference Price Research: Review and Propositions." *Marketing Science*, 69: 84–102.

- Mehta, Nitin, Surenda Rajiv, and Kannan Srinivasan.** 2003. “Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation.” *Marketing Science*, 22: 58–84.
- Monroe, Kent B.** 1973. “Buyers’ Subjective Perceptions of Price” *Marketing Science*, 10: 70–80.
- Nishimura, Hiroki, Efe A. Ok, and John Quah.** 2013. “A Unified Approach to Revealed Preference Theory: The Case of Rational Choice.” Working Paper.
- Polisson, Matthew, and John K H Quah.** 2013. “Revealed Preference in a Discrete Consumption Space.” *American Economic Journal: Microeconomics*, 5: 28–34.
- Roberts, John H, and James M Lattin.** 1997. “Consideration: Review of Research and Prospects for Future Insights.” *Journal of Marketing Research*, 34: 406–410.
- Samuelson, Paul A.** 1938. “A Note on the Pure Theory of Consumer’s Behaviour.” *Economica*, 5: 61–71.
- Selten, Reinhard.** 1991. “Properties of a Measure of Predictive Success.” *Mathematical Social Sciences*, 21: 153–167.
- Tyson, Christopher J.** 2013. “Behavioral Implications of Shortlisting Procedures.” *Social Choice and Welfare*, 41: 941–963.
- Varian, Hal R.** 1982. “The Nonparametric Approach to Demand Analysis.” *Econometrica*, 50: 945–973.
- Varian, Hal R.** 1990. “Goodness-of-Fit in Optimizing Models.” *Journal of Econometrics*, 46: 125–140.

## A Mathematical Appendix

### Proof of Lemma 1

(←) Straightforward.

(→) Assume that  $S$  is rationalizable with limited consideration and let  $I_t$  be the consideration sets that provide such a rationalization. Then there exists a utility function  $u$  such that for all  $t \in T$ ,

$$\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{p}_t \mathbf{q} \leq m_t, q^i = 0 \quad \forall i \notin I_t.$$

Since  $J_t \subseteq I_t$  and  $q_t^j = 0$  for all  $j \in I_t \setminus J_t$ , we obtain

$$\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{p}_t \mathbf{q} \leq m_t, q^i = 0 \quad \forall i \notin I_t, q^j = 0 \quad \forall j \in I_t \setminus J_t,$$

given that the last set of restrictions is not binding. The solution of this maximization problem is identical to the solution of the problem

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \mathbf{p}_t \mathbf{q} \leq m_t, q^i = 0 \quad \forall i \notin J_t.$$

Therefore, the sets  $J_t$  provide a rationalization with limited consideration.

## Proof of Theorem 2

(i)  $\rightarrow$  (ii) Assume that the dataset is rationalizable with limited consideration. By Lemma 1, let  $J_t$  and  $m_t$  be the consideration sets and incomes that rationalize the dataset. Fix an equivalence class  $E$ . If  $v$  and  $t$  belong to  $E$ , we have  $J_v = J_t$ . Thus, the inequality  $\mathbf{p}_t \mathbf{q}_t \geq \mathbf{p}_t \mathbf{q}_v$  implies that  $\mathbf{q}_v$  was feasible at observation  $t$ , but not chosen. Therefore, for any two observations  $t, v$  in the same element  $E$  of the partition,

$$\mathbf{p}_t \mathbf{q}_t \geq \mathbf{p}_t \mathbf{q}_v \Rightarrow u(\mathbf{q}_t) \geq u(\mathbf{q}_v).$$

If in addition  $\mathbf{p}_t \mathbf{q}_t > \mathbf{p}_t \mathbf{q}_v$ , there exist a bundle  $\mathbf{q}$  where  $\mathbf{q} > \mathbf{q}_v$ ,  $q^i = 0 \forall i \notin J_t$  and  $\mathbf{p}_t \mathbf{q}_t \geq \mathbf{p}_t \mathbf{q}$ . By strict monotonicity of the utility function, we obtain  $u(\mathbf{q}_t) \geq u(\mathbf{q}) > u(\mathbf{q}_v)$ . Thus, for any two observations  $t$  and  $v$  in the same element of the partition,

$$\mathbf{p}_t \mathbf{q}_t > \mathbf{p}_t \mathbf{q}_v \Rightarrow u(\mathbf{q}_t) > u(\mathbf{q}_v).$$

By contradiction, assume  $S$  violates LARP, i.e. at least one subdataset  $\{\mathbf{p}_t, \mathbf{q}_t\}_{t \in E}$  violates GARP. Then we have a sequence of observations  $t, r, s, \dots, w, v$  such that

$$\mathbf{p}_t \mathbf{q}_t \geq \mathbf{p}_t \mathbf{q}_r, \mathbf{p}_r \mathbf{q}_r \geq \mathbf{p}_r \mathbf{q}_s, \dots, \mathbf{p}_w \mathbf{q}_w \geq \mathbf{p}_w \mathbf{q}_v \text{ and } \mathbf{p}_v \mathbf{q}_v > \mathbf{p}_v \mathbf{q}_t.$$

However, by the reasoning in the first part of this proof, we obtain

$$u(\mathbf{q}_t) \geq u(\mathbf{q}_r) \geq u(\mathbf{q}_s) \geq \dots \geq u(\mathbf{q}_w) \geq u(\mathbf{q}_v) \text{ and } u(\mathbf{q}_v) > u(\mathbf{q}_t),$$

a contradiction.

(ii)  $\rightarrow$  (iii) Assume that  $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$  satisfies LARP. For all  $t \in T$ , define the function  $a_t(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  as

$$a_t(\mathbf{q}) = \sum_{i \in J_t} p_t^i q^i - \mathbf{p}_t \mathbf{q}_t + P \sum_{i \notin J_t} p_t^i q^i.$$

Here,  $P$  is set such that for all  $v \in T$ , if  $J_v \not\subseteq J_t$ , then  $a_t(\mathbf{q}_v) > 0$ , i.e.

$$P > \max_{v, J_v \not\subseteq J_t} \frac{\mathbf{p}_t \mathbf{q}_t - \sum_{i \in J_t} p_t^i q_v^i}{\sum_{i \notin J_t} p_t^i q_v^i}.$$

Such a value can always be found, given that the numerator is bounded and the denominator is strictly positive.

Now, consider a sequence  $t, s, r, \dots, w, v$ . We will first show that if

$$a_t(\mathbf{q}_s) \leq 0, a_s(\mathbf{q}_r) \leq 0, \dots, a_w(\mathbf{q}_v) \leq 0 \text{ and } a_v(\mathbf{q}_t) \leq 0,$$

then all inequalities are in fact equalities, i.e. all these terms are equal to zero. By the definition of the functions  $a_t(\cdot)$  above, this can only happen if

$$J_s \subseteq J_t, J_r \subseteq J_s, \dots, J_v \subseteq J_w \text{ and } J_t \subseteq J_v.$$

If not, by definition of  $P$ , the functions take strictly positive values. Hence, all sets  $J_t, J_s, J_r, \dots, J_v$  are equal, i.e. the observations  $t, r, s, \dots, v$  belong to the same partition. Therefore,

$$\begin{aligned} a_t(\mathbf{q}_s) &= \mathbf{p}_t \mathbf{q}_s - \mathbf{p}_t \mathbf{q}_t \leq 0, \\ a_s(\mathbf{q}_r) &= \mathbf{p}_s \mathbf{q}_r - \mathbf{p}_s \mathbf{q}_s \leq 0, \\ &\dots, \\ a_v(\mathbf{q}_t) &= \mathbf{p}_v \mathbf{q}_t - \mathbf{p}_v \mathbf{q}_v \leq 0. \end{aligned}$$

Now, if one of these inequalities is a strict inequality, this would imply a violation of LARP. Therefore, all inequalities should be equalities, which we needed to show.

Using the above property on the functions  $a_t(\cdot)$ , we can invoke a result of Fostel, Scarf and Todd (2004, Section 2) that shows it is possible to find for all  $t \in T$ , numbers  $\phi_t \in \mathbb{R}$  and  $\lambda_t \in \mathbb{R}_{++}$  such that for all observations  $t, v \in T$ ,

$$\phi_t - \phi_v \leq \lambda_v a_v(\mathbf{q}_t). \tag{1}$$

Setting  $P_t^i = P \cdot p_t^i$  establishes (iii).

(iii)  $\leftrightarrow$  (iv) This is evident from the equivalence between (ii) and (iii) in Theorem 1.

(iii)  $\rightarrow$  (v). Consider the ‘utility’ function

$$u(\mathbf{q}) = \min_{t \in T} \{\phi_t + \lambda_t a_t(\mathbf{q})\}.$$

First of all, this function is strictly monotone and continuous (given that  $a_t(\mathbf{q})$  is continuous and strictly monotone in  $\mathbf{q}$ ). In fact, the function is concave (given that  $a_t(\mathbf{q})$  is linear and and the min operator preserves concavity). Also, we immediately have that

$$u(\mathbf{q}_t) \leq \phi_t + \lambda_t a_t(\mathbf{q}_t) = \phi_t.$$

Let us show that this inequality is in fact an equality. Towards a contradiction, assume that  $u(\mathbf{q}_t) < \phi_t$ . Then there exists an observation  $v$  such that

$$\begin{aligned} u(\mathbf{q}_t) &= \phi_v + \lambda_v a_v(\mathbf{q}_t) < \phi_t, \\ \iff \phi_t - \phi_v &> \lambda_v a_v(\mathbf{q}_t). \end{aligned}$$

However, this contradicts the inequalities in (iii). As such, we see that for all  $t$ ,  $u(\mathbf{q}_t) = \phi_t$ .

Finally, let us show that this utility function provides a rationalization with limited attention. For all  $t$ , set  $I_t = J_t$  and  $m_t = \mathbf{p}_t \mathbf{q}_t$ . Now, take any bundle  $\mathbf{q}$  which is feasible at observation  $t$ , i.e.  $\mathbf{q}^i = \mathbf{q}^j = \mathbf{0}$  for all  $i \notin J_t$  and  $\mathbf{p}_t \mathbf{q} \leq m_t = \mathbf{p}_t \mathbf{q}_t$ . Then, by definition,

$$\begin{aligned} u(\mathbf{q}') &\leq \phi_t + \lambda_t a_t(\mathbf{q}), \\ &= \phi_t + \lambda_t \sum_{i \in J_t} p_t^i q^i - \lambda_t \mathbf{p}_t \mathbf{q}_t + \lambda_t \sum_{i \notin J_t} P_t^i q^i, \\ &= \phi_t + \lambda_t (\mathbf{p}_t \mathbf{q} - \mathbf{p}_t \mathbf{q}_t), \\ &\leq \phi_t = u(\mathbf{q}_t), \end{aligned}$$

which needed to be shown.

(v)  $\rightarrow$  (i) Straightforward.

## Proof of Corollary 1

( $\rightarrow$ ) Assume that  $S$  is rationalizable with subjective prices. Then there is a utility function  $u$  such that

$$\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \sum_{i \in I_t} p_t^i q^i + \sum_{i \notin I_t} P_t^i q^i \leq m_t.$$

From Theorem 1 it follows that  $\{\tilde{\mathbf{P}}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP where  $\tilde{P}_t^j = p_t^j$  if  $j \in I_t$  and  $\tilde{P}_t^j = P_t^j$  else.

Let us show that there exist a vector  $\bar{\mathbf{P}}_t$  such that  $\{\bar{\mathbf{P}}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP and  $\bar{P}_t^j = p_t^j$  if  $j \in J_t$ . If,

$$\begin{aligned} &\sum_{i \in I_t} p_t^i q_t^i + \sum_{i \notin I_t} P_t^i q_t^i \geq \sum_{i \in I_t} p_t^i q_v^i + \sum_{i \notin I_t} P_t^i q_v^i, \\ \iff &\sum_{i \in J_t} p_t^i q_t^i + \sum_{i \in I_t \setminus J_t} p_t^i q_t^i + \sum_{i \notin I_t} P_t^i q_t^i \geq \sum_{i \in I_t} p_t^i q_v^i + \sum_{i \in I_t \setminus J_t} p_t^i q_v^i + \sum_{i \notin J_t} P_t^i q_v^i, \end{aligned}$$

Now, set  $\bar{P}_t^j = p_t^j$  if  $j \in J_t$ ,  $\bar{P}_t^j = p_t^j$  if  $j \in I_t \setminus J_t$  and  $\bar{P}_t^j = P_t^j$  if  $j \notin I_t$ . Then,

$$\sum_{i \in J_t} \bar{P}_t^i q_t^i + \sum_{i \notin J_t} \bar{P}_t^i q_t^i \geq \sum_{i \in J_t} \bar{P}_t^i q_v^i + \sum_{i \notin J_t} \bar{P}_t^i q_v^i,$$

Therefore, we see that  $\{\bar{\mathbf{P}}_t, \mathbf{q}_t\}_{t \in T}$  must also satisfy GARP. Using Theorem 2, we see that the dataset satisfies LARP.

( $\leftarrow$ ) If LARP is satisfied, by Theorem 2, there exist a price vector  $\mathbf{P}_t \in \mathbb{R}_{++}^n$  such that  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  satisfies GARP and  $P_t^j = p_t^j$  if  $q_t^j > 0$ . Then, from Theorem 1, we know that there exist a utility function such that,

$$\mathbf{q}_t = \arg \max_{\mathbf{q}} u(\mathbf{q}) \text{ s.t. } \sum_{i \in J_t} p_t^i q^i + \sum_{i \notin J_t} P_t^i q^i \leq m_t.$$

Setting  $I_t = J_t$  shows that this dataset is also rationalizable with subjective prices.

## Proof of Lemma 2

Assume that,  $\mathbf{q}_t$  is (strictly) revealed preferred to  $\mathbf{q}_v$  according to the prices  $\bar{\mathbf{P}}_t$

$$\begin{aligned} \sum_i \bar{P}_t^i q_t^i &\geq (>) \sum_i \bar{P}_t^i q_v^i, \\ \iff \sum_i \bar{P}_t^i (q_t^i - q_v^i) &\geq (>) 0. \end{aligned}$$

Then, we obtain

$$\begin{aligned} 0 &\leq (<) \sum_i \bar{P}_t^i (q_t^i - q_v^i), \\ &= \sum_{i \in J_t} \bar{P}_t^i (q_t^i - q_v^i) + \sum_{i \notin J_t} \bar{P}_t^i (0 - q_v^i), \\ &\leq \sum_{i \in J_t} P_t^i (q_t^i - q_v^i) + \sum_{i \notin J_t} P_t^i (0 - q_v^i), \\ &= \sum_{i \in J_t} P_t^i (q_t^i - q_v^i) + \sum_{i \notin J_t} P_t^i (q_t^i - q_v^i). \end{aligned}$$

This shows that  $\mathbf{q}_t$  is also (strictly) revealed preferred to  $\mathbf{q}_v$  according to the prices  $\mathbf{P}_t$ . As such, if  $\{\bar{\mathbf{P}}_t, \mathbf{q}_t\}_{t \in T}$  violates GARP, then  $\{\mathbf{P}_t, \mathbf{q}_t\}_{t \in T}$  should also violate GARP which needed to be shown.

Let  $\bar{\mathbf{p}} = \frac{1}{T} \sum_{t \in T} \mathbf{p}_t$  be the vector of average prices. We assume that the subjective prices  $\mathbf{P}_t$  are such that  $\mathbf{P}_t \leq \theta \bar{\mathbf{p}}$ , for some fixed number  $\theta \geq 1$ .

## B Additional Results

Table 4 provides the bounds (ub: upper bound and lb: lower bound) of the 95% confidence intervals for the difference between the predictive success of the different models, i.e. the predictive success of the model in a row minus the predictive success of the corresponding model of the column. For example, the CI for the difference in predictive success of GARP minus the predictive success of LARP is  $[-0.0663 \quad -0.0222]$ . We refer to Demuyneck (2015) for the construction of these bounds.

## C Sample Construction

1. The raw dataset has 957 570 transactions (indexed by  $t$ ) on 2 100 households (indexed by  $j$ ) for 15 890 distinct products (with different UPC, indexed by  $i$ ).
2. First, we drop the observations of April 1995 as these only include a single day. This leaves 956 064 transactions.

Table 4: 95% Confidence intervals for difference between predictive success of the different models

		LARP	$\theta$ -LARP-T $\theta = 1.25$	1.5	1.75	2	2.25	LARP-AP
GARP	lb	-0.0663	-0.0794	-0.0937	-0.0928	-0.0872	-0.0828	-0.0939
	ub	-0.0222	-0.0081	-0.0370	-0.0419	-0.0399	-0.0370	-0.0245
LARP	lb		-0.0318	-0.0407	-0.0371	-0.0285	-0.0224	-0.0477
	ub		0.03280	-0.0014	-0.0091	-0.0101	-0.0089	0.01780
$\theta$ -LARP-T $\theta = 1.25$	lb			-0.0539	-0.0566	-0.0527	-0.0489	-0.0548
	ub			0.0107	0.0094	0.0131	0.0166	0.0238
$\theta = 1.5$	lb				-0.0191	-0.0167	-0.0135	-0.0291
	ub				0.0150	0.02030	0.0243	0.0413
$\theta = 1.75$	lb					-0.0072	-0.0048	-0.0264
	ub					0.0148	0.0197	0.0426
$\theta = 2$	lb						-0.0024	-0.0288
	ub						0.0097	0.0374
$\theta = 2.25$	lb							-0.0322
	ub							0.0336

3. For each transaction  $t$ , household  $j$  and good  $i$ , the raw dataset provides us with the total expenditure  $e_{ijt}$  and total quantity bought  $q_{ijt}$ . We aggregate over months defining expenditures  $e_{ijm}$  and quantities  $q_{ijm}$  for household  $j$  for good  $i$  over month  $m$  by,

$$e_{ijm} = \sum_{t \in m} e_{ijt},$$

$$q_{ijm} = \sum_{t \in m} q_{ijt}.$$

This leaves a total of 743 051 transactions. We also define total quantity per good  $i$  over the month  $m$  as

$$q_{im} = \sum_j q_{ijm}.$$

4. In order to have reliable price information for missing transactions, we remove observations for which  $q_{im}$  is below 20. This leaves 348 022 transactions over 2 026 distinct goods.
5. We define price per good per period per household  $p_{ijm}$  by

$$p_{ijm} = \frac{e_{ijm}}{q_{ijm}}.$$

6. For each UPC  $i$  and month  $m$ , we also define an average price,

$$p_{im} = \frac{1}{T_{im}} \sum_j p_{ijm},$$

where  $T_{im}$  is the total number of households that buy good  $i$  in month  $m$ .

7. We impose a panel structure by creating an observation  $t$  for each  $i, j$  and  $m$  for which  $q_{ijm'} > 0$  for at least one month  $m'$ . This gives us a total of 4 230 954 observations.
8. If we have no transaction for some combination of  $ijm$  in this dataset (i.e. household  $j$  did not consume good  $i$  in month  $m$ ) we set  $q_{ijm} = 0$  and  $p_{ijm} = p_{im}$ , i.e. the missing prices are set to be equal to the average price among all transactions for this UPC during the month  $m$ .
9. We drop observations  $ijm$  where  $p_{ijm}$  is not defined for some time period  $m'$  (i.e. no household bought good  $i$  during some month). This leaves 1 219 887 observations.
10. We drop households  $j$  that have zero total consumption for some month  $m$ , i.e. where  $\sum_i q_{ijm} = 0$  for some  $m$ . This leaves 453 951 observations on 550 households ( $j$ ) and 124 goods ( $i$ ) over 27 months ( $m$ ).
11. The final dataset gives us a balanced panel (over 27 periods) for 550 households. The number of goods per household is on average 30.57 (standard deviation is 10.69) with a median of 30, first quartile of 23 and 3rd quartile of 37.
12. 84% of these observations ( $i, j, m$ ) have zero consumption,  $q_{ijm} = 0$ .
13. On average, 84% of household transactions are zero (standard deviation is 0.0478). The median fraction of zero transactions is 0.84, with first quartile 0.81 and third quartile 0.87.