

Revealed Preference with Limited Consideration

By THOMAS DEMUYNCK AND CHRISTIAN SEEL*

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.

JEL: D11, D12, D81

Keywords: Revealed preference, bounded rationality, consideration set

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.¹ In this paper, we develop revealed preference tests for different models of bounded rationality in which the choice process is determined by 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 fully rational model.

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 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 behavior. In this canonical model, individuals consider all bundles in their budget

* Demuynck: Maastricht University, School of Business and Economics, Tongersestraat 53, 6211 LM Maastricht, The Netherlands, t.demuynck@maastrichtuniversity.nl. Seel: Maastricht University, School of Business and Economics, Tongersestraat 53, 6211 LM Maastricht, The Netherlands, c.seel@maastrichtuniversity.nl. We would like to thank Ian Crawford, Mark Dean, Arno Riedl, Thomas Kittsteiner and several seminar audiences for useful comments.

¹See Roberts and Lattin (1997) for an overview.

set and choose a bundle that is best according to their utility function.

In many microeconomic consumption datasets, it is found that consumers limit their purchases to a small subset of the set of all available goods. This behavior 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. In this case, the observed choice behavior may not satisfy GARP, although the choices are still optimal given her consideration sets.

Figure 1 illustrates the idea in a two goods setting. There are two choice situations corresponding to the two solid budget lines. In the first situation with the steeper budget line, the decision maker chooses bundle a which contains nothing of good 2. In the second choice situation, the decision maker chooses b which contains positive amounts of both goods. In the first situation, a and b are both affordable and a is chosen. It follows that a is revealed preferred to b . In the second choice situation, again both a and b are available but now b is chosen, so b is also revealed preferred to a . Since a is in the interior of this budget set, the choice behavior violates GARP.²

If the decision maker does not consider good 2 in the first choice situation, the best she can do is to spend her entire disposable income on good 1, i.e., to choose bundle a . However, since bundle b is not ‘feasible’, we cannot conclude that bundle a is preferred over bundle b . If the consideration set in the second choice situation consists of both goods and bundle b is chosen, we can conclude that bundle b is preferred to bundle a . This does not contradict the first observation and hence the decision maker could still be utility maximizing given her consideration sets.

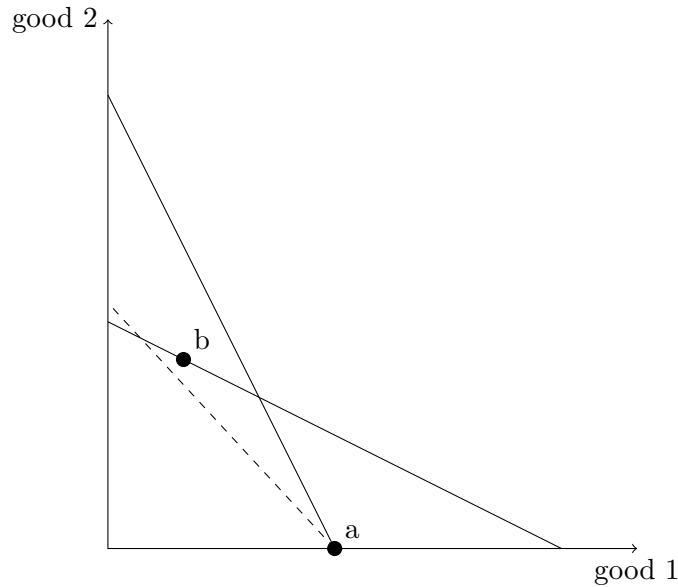
Unfortunately, for real life datasets, consideration sets are not observable. Then, the relevant question is whether it is possible to construct consideration sets such that the observed behavior is rational for these constructed consideration sets. We solve this problem by providing a new revealed preference 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 behavior when taking into account the presence of consideration sets. The axiom remains silent about how the consideration sets are formed. Therefore, we also present an equivalent model, where the construction of the consideration sets is made more explicit. This model considers a utility maximizing decision maker who bases her consumption choice not on the actual prices but on some subjective prices.

Initially, the decision maker forms her consideration set by gathering all goods that she would like to buy, i.e., the goods she would like to consume with strictly positive amounts given her subjective prices. In a next step, the decision maker learns the true prices of the goods in this consideration set. Thus, the subjective

²See the next section for a formal definition of GARP.

FIGURE 1. AN EXAMPLE WITH TWO GOODS



prices for the goods in the consideration set are updated to the true prices. 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. In this case, she learns the true prices of these new goods and she may again choose to update her consideration set by including additional goods that she would like to buy. 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.

For the example in Figure 1, in choice situation 1, the subjective price of good 1 must equal the true price—if good 1 is bought, it must be in the consideration set. The subjective price of good 2, however, might differ from its true price. This subjective price generates a ‘subjective’ budget line, represented by the dashed line, which may differ from the true budget line. Using this new budget line, we observe that the two observations are consistent with GARP.

For real-life datasets, we do not observe subjective prices. Thus, the question is whether there exists a set of subjective prices that can explain the observed choice behavior. In the main text, we show that we can construct such prices if and only if the dataset satisfies LARP. In other words, the model with subjective prices and the model with consideration sets are empirically indistinguishable.

Subsequently, we present three models which restrict the admissible values of the subjective prices. In the first model, subjective prices must not deviate too much from the true prices. In the second model, subjective prices are based on the

average prices over all observations, i.e., subjective prices are on average equal to the true prices, but short term deviations are possible. In the third model, subjective prices are adaptive and determined by previously observed prices. For each of these models, we derive the corresponding revealed preference restrictions.

LITERATURE OVERVIEW

Our paper is related to a growing literature in economics that takes into account choice behavior with consideration sets; see, among others, Eliaz and Spiegler (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 depart from a choice theoretic framework where choice sets are assumed to be discrete, i.e., individuals choose from a finite set of alternatives. In this paper, we extend this literature of boundedly rational choice behavior to models with continuous choice and we look at its testable implications from a revealed preference perspective à la Afriat.

Most of the papers in the literature obtain information on the underlying consideration sets by carefully varying the choice sets and observing where decision makers have revealed preference reversals. In our setting, however, choice sets are exogenously determined by the observed prices and total expenditures in the dataset. Hence, we cannot freely manipulate the choice sets. Instead of retrieving information on consideration sets by revealed preference reversals, we obtain it by the absence of positive consumption.

Independently of our work, Crawford (2014) recently focused on applying revealed preference theory to datasets where prices are not observed for goods that are consumed with zero quantities. He obtains a revealed preference condition which is similar to our LARP condition. Subsequently, he shows how the framework can be adjusted for prediction and welfare analysis. The main difference between his analysis and ours is the focus. While Crawford (2014) focuses on constructing valid revealed preference inference when prices are not fully observable, our focus is on the application of revealed preference theory to consumption models where the decision is influenced by the formation of consideration sets.

OUTLINE

Section I contains a short introduction to revealed preference theory. Readers familiar with this topic can safely skip this section. Section II contains the revealed preference characterization for the limited consideration model. In Section III, we describe the revealed preference model where the decision maker bases her consumption decision on her subjective prices. Section IV contains the empirical illustration. Finally, Section V concludes. Most proofs are in the appendix.

I. 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 behavior \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*

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

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 ϕ_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 λ_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 λ_t at the different observations.

II. 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 $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$.³ 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

³The 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.

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 $I_t = J_t$.

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 . The set \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 $E \in \mathcal{E}_S$. We say that the dataset S satisfies the *Limited Axiom of Revealed Preference* (LARP) if each of these sub-datasets satisfies GARP. We obtain the following characterization:

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 ,

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

- (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 the equivalences in 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 $J_t \neq J_v$. Assume that $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ is rationalizable with limited consideration and assume that, relying on Lemma 1, the consideration sets are $I_t = J_t$. Consider two observations t and v and assume that \mathbf{q}_t is revealed preferred to \mathbf{q}_v , i.e.,

$$\sum_{i=1}^n p_t^i q_t^i \geq \sum_{i=1}^n p_t^i q_v^i.$$

There are two cases to consider. Either $J_v \subseteq J_t$ or $J_v \not\subseteq J_t$. In the first case, every good consumed at observation v was also in the consideration set at observation t . In this case, the bundle \mathbf{q}_v could also have been bought at observation t . By a simple revealed preference argument, the utility from consuming \mathbf{q}_t must be at least as high as the utility from consuming \mathbf{q}_v . For the second case, there are some goods which are consumed at observation v that are not in the consideration set at observation t . In that case, \mathbf{q}_v is not available at observation t , so we cannot say that \mathbf{q}_t is ‘revealed preferred’ to \mathbf{q}_v . In this sense, we can define a limited revealed preference relation \bar{R} such that $\mathbf{q}_t \bar{R} \mathbf{q}_v$ iff $\mathbf{q}_t R \mathbf{q}_v$ and $J_v \subseteq J_t$. Now, for any cycle in this modified revealed preference relation, say

$$\mathbf{q}_t \bar{R} \mathbf{q}_s \bar{R} \mathbf{q}_\ell \dots \mathbf{q}_w R \mathbf{q}_v \bar{R} \mathbf{q}_t,$$

we must have that

$$J_t \subseteq J_v \subseteq J_w \subseteq \dots \subseteq J_s \subseteq J_t,$$

or equivalently

$$J_t = J_v = J_w = \dots = J_t.$$

Thus, in order to find a violation of consistency for this adjusted revealed preference relation, we can restrict ourselves to subsets of observations for which the sets J_t are 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 . To grasp the intuition, assume that the utility function is strictly monotone, differentiable and concave. Relying on Lemma 1, also assume that $I_t = J_t$. Then, from concavity of $u(\cdot)$, we obtain

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

The first order conditions of the utility maximization problem gives

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

Here λ_v is the Lagrange multiplier for the budget constraint and μ_v^i is the Lagrange multiplier for the consideration set constraint, i.e., $q^i = 0, i \notin J_t$. The first condition states that, for goods that are consumed with positive amounts, the marginal utility should be equal to the marginal cost. The second constraint says that for goods that are not consumed but in the consideration set, the marginal utility should be below the marginal cost. Finally, for goods outside the consideration set, marginal utility can be higher or lower than the marginal cost, depending on the value of μ_v^i . Given that J_t only contains goods for which $q_v^i > 0$, we can discard the second first order condition.

If we substitute these conditions in the concavity constraint, we obtain that

$$u(\mathbf{q}_t) - u(\mathbf{q}_v) \leq \lambda_v \left[\sum_{i \in J_v} p_v^i (q_t^i - q_v^i) + \sum_{i \notin J_v} \left(p_v^i + \frac{\mu_v^i}{\lambda_v} \right) (q_t^i - q_v^i) \right].$$

Now, set for $i \notin J_v$, $P_v^i = p_v^i + \mu_v^i/\lambda_v$, $\phi_v = u(\mathbf{q}_v)$ and $\phi_t = u(\mathbf{q}_t)$. Then we obtain

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

This establishes the Afriat-type inequalities. It also demonstrates that the prices P_v^j can be interpreted as the marginal willingnesses 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 I, 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.

III. 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 model where the consideration sets are endogenously determined by the decision maker based on the values of some subjective prices. This provides a conceptually novel, intuitive framework how consideration sets are formed.⁴ Moreover, by imposing more structure on the subjective prices, we provide various alternative models of consumer behavior with additional testable implications.

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 bundles that she would like to buy based on some subjective idea about the prices of the different goods. We denote these subjective prices by $\mathbf{P}_t \in \mathbb{R}_{++}^n$.⁵ They do not necessarily coincide with the true prices of the goods. In an initial phase, the decision maker solves

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

⁴Somewhat similar, Mehta, Rajiv and Srinivasan (2003) construct an econometric model of consideration set formation where consumers are uncertain about the prices on a particular shopping trip.

⁵The case in which a consumer is completely unaware of good i in Period t can be included in this setup by setting P_t^i sufficiently large.

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.,

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

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 :

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

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

The above choice procedure models a situation where the decision maker has some initial idea about the goods she would like to buy (based on the prices \mathbf{P}_t). Given this initial idea, she only gives attention to these goods and, by doing so, she learns their prices. Once she has observed these prices, she may reconsider the set of goods she actually would like to buy. For example, if the price of some good turns out to be much higher than expected, she might consider buying some other goods instead of the expensive one. In this way, the consideration set might expand after learning the prices of some of the goods. This process of updating the consideration set stops at the moment that she decides not to buy any new goods.

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 3: *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:*

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

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

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 II.

THEOREM 3: *Consider a dataset S . The dataset is rationalizable with subjective prices if and only if it satisfies LARP.*

To understand the intuition behind Theorem 3, 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 behavior 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 three 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 $i \in G, t \in T$,*

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

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

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 4: A dataset S is θ -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) $\mathbf{q}_t \in \arg \max_{\mathbf{q}} u(\mathbf{q})$ 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) $P_t^i \leq \theta p_t^i \forall i \notin I_t$.
- (iii) If $i \notin I_t$, then $q_t^i = 0$.

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

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 θ -rationalizable with subjective prices constrained by true prices.
- (ii) The dataset S satisfies θ -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 θ can either be chosen ex ante or determined as the smallest value of θ for which a dataset is still rationalizable (if the dataset satisfies LARP). This critical value of θ 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 this critical value is close to one, then the dataset can be rationalized by a model where the subjective prices are very close to the actual prices. If θ is very large, we need to allow for subjective prices that are far from the actual ones.

SUBJECTIVE PRICES BASED ON AVERAGE PRICES

In our second model with restricted subjective prices, decision makers hold roughly consistent ideas about the (long term) average prices of the goods but do not take into account short term fluctuations from the average price. 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 $\theta \geq 1$.

DEFINITION 5: *A dataset S is θ -rationalizable with subjective prices constrained by average 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 $t \in T$,*

- (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) $P_t^i \leq \theta \bar{p}^i \forall i \notin I_t$.
- (iii) If $i \notin I_t$, then $q_t^i = 0$.

THEOREM 5: *Consider a dataset $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$. The following statements are equivalent:*

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

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

Again, the fact that (i) implies (ii) is a simple application of Lemma 2: if $q_t^j = 0$, then there are two cases, either $j \in I_t$, in which case $P_t^j = p_t^j$ or $j \notin I_t$ in which case $P_t^j \leq \theta \bar{p}^j$. As we do not know which of the two cases hold, we pick the highest price as the relevant one since this provides the weaker revealed preference restrictions.

ADAPTIVE SUBJECTIVE PRICES

In the third model, subjective prices are formed on the basis of past purchases of the goods. 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

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

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

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

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 6: *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 $P_t^j = p_t^j$.

IV. 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).

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 use non-aggregated data implying that most goods are actually purchased in 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.⁶ This leaves us with a total of 814 households.

Table 1 gives some summary statistics on the dataset. On average, over the 27 months, a household chooses among 28 different goods although there is quite some variation among households. Further, more than 85 percent of all quantities (good, observation pairs) are zero which means that the consumption data is quite sparse.

TABLE 1—SUMMARY STATISTICS

Number of households	814
Fraction nonzero consumption observations	0.14912 (0.0026)
Number of distinct goods per household	28.26 (113.31)

Sample variances are between brackets.

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 12% which is quite low. On the other hand, the pass rate for the LARP test (98%) is almost equal to unity implying that nearly all households are consistent with utility maximizing behavior 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 23.73 equivalence classes per household. Given that each dataset has 27 observations, for some households each observation is in a separate element of the partition. In these cases, LARP is trivially satisfied. 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 based on true prices for various values of θ . This pass rate ranges from 41% (for $\theta = 1.05$) to 98% (for $\theta = 1.5$). Given that the revealed preference test becomes weaker as θ increases, the pass rate cannot decrease. We omitted the results for $\theta = 1$ as this test coincides with the usual GARP test. The pass rates for rationalizability with subjective prices constrained by average prices are somewhat higher than the corresponding pass rates for subjective prices constrained by true prices. Finally, the pass rate for rationalizability with adaptive subjective prices is equal to 77%.

⁶In order to obtain reliable price data, we omit goods for which the total quantity bought (over all households) is below 20.

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) behavior based on the model of irrational behavior 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.

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 814 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 behavior is random. As such, the difference determines how much actual behavior agrees better with the revealed preference test compared to what is expected if behavior 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 behavior, since it is at least as good at explaining random behavior. 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 behavior while explaining observed behavior.

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 behavior. LARP seems to perform marginally better than GARP in terms of predictive success. The difference between the two predictive success measures, however, is not statistically significant at the 5% level (see Appendix B for details). The subjective rationalizability model with adaptive subjective prices outperforms both LARP and GARP in terms of predictive success. The refined models θ -LARP-T and θ -LARP-AV with moderate values of θ (equal to 1.1 and 1.05) provide the highest predictive success.

OBSERVABLE HETEROGENEITY

As a final exercise, we investigate whether the predictive success of the different models is related to observable heterogeneity by means of a linear regression of the individual predictive success measures 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* 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 α 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 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 variables *pshopage*, *female college*, *male college*, and *income* have no statistically significant effect for any of the models. The effect of *hhsiz*e 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. One possible explanation for this finding 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 behavior 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). Finally, the variable *conc* is positive and significant for most models with limited consideration. Thus, households where the shopping was mostly conducted by the same person have a higher predictive success. Intuitively, if different shoppers use different subjective prices in their optimization problem, the household need not satisfy the revealed preference conditions anymore even if each individual shopper satisfies the conditions.

V. Conclusion

Models of consumer behavior 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 also derived three 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. All models have positive predictive success, but the refined models with endogenous consideration set formation appear to outperform both GARP and LARP significantly. Thus, our paper provides empirical support for the (refined) consideration set approach.

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.

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 behavior 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 behavior (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.
- Aguiar, Mark, and Erik Hurst.** 2007. "Life-Cycle Prices and Production." *American Economic Review*, 97: 1533–1559.
- 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.
- Bronars, Stephen G.** 1987. "The Power of Nonparametric Tests of Preference Maximization." *Econometrica*, 55: 693–698.
- Browning, Martin, and Pierre-André Chiappori.** 1998. "Efficient Intra-household Allocations: A General Characterization and Empirical Tests." *Econometrica*, 66: 1241–1278.
- 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 Non-parametric 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 Demuyne, and Bram De Rock.** 2011. "Revealed Preference Analysis of Noncooperative Household Consumption." *Economic Journal*, 121: 1073–1096.
- Crawford, Ian.** 2014. "Demand Analysis with Partially Observed Prices." Oxford University.
- Demuyne, Thomas.** 2014. "Statistical Inference for Measures of Predictive Success." METEOR Research Manuscript 14/010.
- Diewert, Walter E.** 1973. "Afriat and Revealed Preference Theory." *Review of Economic Studies*, 40: 419–425.
- Eliasz, Kfir, and Ran Spiegler.** 2011. "Consideration Sets and Competitive Marketing." *Review of Economic Studies*, 78: 235–262.
- 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.
- Houthakker, Hendrik S.** 1950. "Revealed Preference and the Utility Function." *Economica*, 17: 159–174.
- 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.
- 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.

- 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.

MATHEMATICAL APPENDIX

Proof of Lemma 1

(\leftarrow) Straightforward.

(\rightarrow) 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 $t, v \in T$,

$$(A1) \quad \phi_t - \phi_v \leq \lambda_v a_v(\mathbf{q}_t).$$

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 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

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

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., $\mathbf{q}^i = \mathbf{q}^i = \mathbf{0}$ 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 Theorem 3

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

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

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,

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

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.

ADDITIONAL RESULTS

The following table 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 $[-0.0217 \quad 0.0066]$. We refer to Demuyneck (2014) for the construction of these bounds.

TABLE 2—PASS RATES, POWER AND PREDICTIVE SUCCESS

Test	pass rate	power	pred. succ.	95% CI
GARP	0.1179 (0.1040)	0.9312 (0.0191)	0.0491 (0.0776)	[0.0300 0.0682]
LARP	0.9853 (0.0145)	0.0703 (0.0347)	0.0556 (0.0374)	[0.0482 0.0648]
θ -LARP-T				
$\theta = 1.05$	0.4115 (0.2422)	0.6968 (0.0746)	0.1083 (0.1778)	[0.0793 0.1373]
$\theta = 1.10$	0.6671 (0.2221)	0.4440 (0.0966)	0.1111 (0.1577)	[0.0838 0.1384]
$\theta = 1.25$	0.9103 (0.0816)	0.1587 (0.0584)	0.0690 (0.0834)	[0.0492 0.0889]
$\theta = 1.5$	0.9767 (0.0228)	0.0932 (0.0386)	0.0699 (0.0462)	[0.0551 0.0846]
θ -LARP-AV				
$\theta = 1$	0.4570 (0.2482)	0.6418 (0.0849)	0.0988 (0.1629)	[0.0710 0.1265]
$\theta = 1.05$	0.6818 (0.2169)	0.4325 (0.0976)	0.1144 (0.1489)	[0.0878 0.1408]
$\theta = 1.1$	0.8194 (0.1480)	0.2776 (0.0806)	0.0970 (0.1192)	[0.0733 0.1207]
$\theta = 1.25$	0.9386 (0.0576)	0.1301 (0.0498)	0.0687 (0.0691)	[0.0507 0.0868]
$\theta = 1.5$	0.9791 (0.0204)	0.0892 (0.0374)	0.0683 (0.0441)	[0.0539 0.0827]
LARP-AP	0.7706 (0.1768)	0.3125 (0.1055)	0.0830 (0.1158)	[0.0597 0.1065]

Sample variances are between brackets.

TABLE 3—LINEAR REGRESSION OF HOUSEHOLD PREDICTIVE SUCCESS ON OBSERVABLE HOUSEHOLD CHARACTERISTICS

	GARP	LARP	θ -LARP-T $\theta = 1.1$	θ -LARP-AV $\theta = 1.05$	LARP-AP
constant	0.0234 (0.0600)	0.1040*** (0.0351)	0.1290* (0.0869)	0.1272* (0.0859)	0.108* (0.0736)
pshopage	0.0003 (0.0007)	-0.0002 (0.0005)	0.0001 (0.0010)	0.0004 (0.0010)	0.0002 (0.0010)
female college	-0.0266 (0.0213)	-0.0113 (0.0177)	-0.0397 (0.0341)	-0.0149 (0.0336)	-0.0069 (0.0304)
male college	0.0271 (0.0246)	0.0050 (0.0164)	0.0229 (0.0341)	0.0297 (0.0322)	0.0344 (0.0297)
income	-0.0289 (0.0237)	0.0024 (0.0143)	0.0015 (0.0333)	0.0288 (0.0340)	-0.0162 (0.0291)
hhsiz	-0.0013 (0.0077)	-0.0349*** (0.0055)	-0.0270** (0.0122)	-0.0268*** (0.0115)	-0.0302*** (0.0105)
conc	0.0294 (0.0382)	0.0737*** (0.0204)	0.0685* (0.0497)	0.0363 (0.0506)	0.0584* (0.0409)

Heteroscedastic consistent standard errors are between brackets. *** = significant at the 1% level; ** = significant at the 5% level; * = significant at the 10% level.