ECARES

Empirical Revealed Preference

Ian Crawford
Oxford University

Bram De Rock
SBS-EM, ECARES, Université Libre de Bruxelles

ECARES working paper 2013-32

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

Empirical Revealed Preference

Ian Crawford* and Bram De Rock†

July 24, 2013

Abstract

The aim of this article is to provide an introduction to empirical revealed preference and an overview of the current state of the field. We hope to give a sense of how revealed preference methods work, the types of questions which they can address and to assess the strengths and drawbacks of the approach. After briefly recapping the basics of revealed preference theory, we review and critically assess the literature in two main areas which, we suggest, represent the principle fields in which recent research has made significant advances: broadening the scope of revealed preference methods and dealing with empirical issues related to bringing revealed preference to the data. We conclude with a discussion of some future directions.

JEL Classification: D11, D12, D13, C14.

Keywords: consumption behaviour, revealed preferences, nonparametric analysis.

*Institute for Fiscal Studies and University of Oxford, Department of Economics, Manor Road Building, Manor Road, Oxford OX1 3UQ, United Kingdom. E-mail: ian.crawford@economics.ox.ac.uk.

†ECARES and ECORE, Universit Libre de Bruxelles. Avenue F.D. Roosevelt 50, 1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges the European Research Council (ERC) for his Starting Grant.

Contents

1 Introduction

Empirical revealed preference is a "structural" approach to the analysis and interpretation of data by means of economic theory. It is somewhat distinct from structural econometrics because it largely eshews error terms and minimizes the use of untestable assumptions.

The contrast with structural econometrics is a useful starting point for discussing empirical revealed preference. The structural approach to econometrics is very familiar: it proceeds by using economic theory to develop formal mathematical statements concerning causes and effects. The causes (explanatory variables), which may be observed (x) or unobserved (η), and the effects (endogenous variables, y) are linked by these theory-derived statements through structural equations $y = f(x, \eta, \theta)$, where θ represents a set of unknown parameters or functions. Econometricians always then append a statistical structure to the economic model in order to account for the fact that the economic theory as expressed through the structural equations f does not perfectly explain the data. This extra structure entails statistical assumptions regarding the joint distribution of (x, η) and these other unobservables (ϵ) introduced by the econometrician. When combined these economic and statistical assumptions deliver an empirical model that is capable of rationalising any set of observables. The art of structural modelling thus mainly lies in getting this statistical aspect right, because the source and the properties of these econometric errors ϵ can have a critical impact on the estimation results. Unfortunately this is far from straightforward as economic theories, which are by and large completely deterministic, generally have little to say about the statistical model, and the data have generally little to say about unobservables.

Like structural econometrics, empirical revealed preference also begins from economic theory, but the description of the implications of the theory is entirely different to the "$y = f(x, \eta, \theta)$" type of framework. Rather than describing the implications of the theory in terms of parameterised structural equations, empirical revealed preference uses systems of inequalities which depend neither on the form of structural functions nor on unobservables. Statistical error terms and assumptions about the functional structure of the economic model may be added but it is not an essential requirement of the approach. In a sense empirical revealed preference is concerned with what we can learn simply by combining theory with the features of the world that we can observe.

The aim of this article is to provide an introduction to empirical revealed preference and an overview of the current state of the field. We hope to give a sense of how empirical revealed preference methods work, the types of questions which they can address and to assess the strengths and drawbacks of the approach. We begin by briefly recapping the basics of revealed preference theory - namely *Afriat's Theorem* and how it can be used to check data for consistency with the canonical utility maximisation model and, granted this, to make predictions and to allow the recovery of features of the model. We then review and critically assess the literature in two main areas which, we suggest, represent the principal fields in which current and recent research has made significant advances. These relate to efforts which have broadened the scope of revealed preference methods to allow the exploration of a richer variety of economic models and work which has sought to address some of the empirical and statistical challenges involved in applying revealed preference methods to the data. As such we demonstrate the versatility and the attractiveness of empirical revealed preferences as well as areas in which further work is clearly needed. We conclude with a discussion of some future directions.

2 The basic model of rational demand

In this section we set the stage by discussing the basic revealed preference (RP) tools for the most simple case of utility maximization. In the first subsection we present the most fundamental result on this topic, Afriat's Theorem, and show how it can be used to check whether a given dataset with observed consumption choices and prices is consistent with utility maximization. Subsequently, in the second subsection we focus on recovering the underlying preferences and on forecasting behaviour in new situations.

2.1 Afriat's Theorem

We consider a setting with N goods and a finite dataset $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ existing of N-dimensional price vectors $\mathbf{p}_t \in \mathbb{R}^N_{++}$ and N-dimensional quantity vectors $\mathbf{q}_t \in \mathbb{R}^N_+$. The set $T = \{1, \ldots, |T|\}$ corresponds to the set of observations. A utility function $u : \mathbb{R}^N_+ \to \mathbb{R}$ is well-behaved if it is concave, continuous and strict monotone. The following definition is standard.

Definition 1 *A dataset $\{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ is rationalisable by a well-behaved utility function u if for all $t \in T$:*

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

In what follows, we will present several ways to verify if a data set S is rationalisable. The first one is the Generalized Axiom of Revealed Preferences (GARP) introduced in Varian (1982).[1]

[1] In the literature there are several revealed preference axioms. In Samuelson (1938, 1948) the weak axiom of revealed preferences (*WARP*) is introduced. This axiom does not take the transitivity of preferences into account. Therefore Houthakker (1950) introduced the strong axiom of revealed preferences (*SARP*). *SARP* does not allow for indifference curves with flat parts, which is taken into account by GARP. The above axioms all ignore differentiability of the underlying utility function. To take this consideration into account Chiappori and Rochet (1987) introduced strong SARP. In this paper we abstract from all these different revealed preference axioms and we restrict our attention to GARP.

Definition 2 *A dataset* $\{\mathbf{p}_t, \mathbf{q}_t\}_{t\in T}$ *satisfies GARP if and only if we can construct relations* R_0, R *such that*

(i) for all $t, s \in T$*, if* $\mathbf{p}_t\mathbf{q}_t \geq \mathbf{p}_t\mathbf{q}_s$ *then* $\mathbf{q}_t \; R_0 \; \mathbf{q}_s$*;*

(ii) for all $t, s, u, \ldots, r, v \in T$*, if* $\mathbf{q}_t \; R_0 \; \mathbf{q}_s$ *,* $\mathbf{q}_s \; R_0 \; \mathbf{q}_u$*,* $\ldots$*, and* $\mathbf{q}_r \; R_0 \; \mathbf{q}_v$ *then* $\mathbf{q}_t \; R \; \mathbf{q}_v$*;*

(iii) for all $t, s \in T$*, if* $\mathbf{q}_t \; R \; \mathbf{q}_s$*, then* $\mathbf{p}_s\mathbf{q}_s \leq \mathbf{p}_s\mathbf{q}_t$*.*

Condition (i) states that the quantities $\mathbf{q}_t$ are directly revealed preferred over $\mathbf{q}_s$ if $\mathbf{q}_t$ was chosen when $\mathbf{q}_s$ was equally attainable. Next, condition (ii) imposes transitivity on the revealed preference relation R. Finally, condition (iii) states that if a consumption bundle $\mathbf{q}_t$ is revealed preferred to a consumption bundle $\mathbf{q}_s$, then $\mathbf{q}_s$ cannot be more expensive then $\mathbf{q}_t$.

In Cherchye et al. (2011d) it is shown that satisfying GARP is equivalent to having a solution for the Integer Program problem IP-GARP.

Definition 3 *A data* $\{\mathbf{p}_t, \mathbf{q}_t\}_{t\in T}$ *satisfies IP-GARP if and only if there exist for all* $s, t \in T$ *binary variables* $x_{s,t} \in \{0, 1\}$ *such that*

(i) for all $t, s \in T : \mathbf{p}_t\mathbf{q}_t - \mathbf{p}_t\mathbf{q}_s < x_{t,s}\mathbf{p}_t\mathbf{q}_t$*;*

(ii) for all $t, s, v \in T : x_{t,s} + x_{s,v} \leq x_{t,v}$*;*

(iii) for all $t, s \in T : (x_{t,s} - 1)\mathbf{p}_s\mathbf{q}_s \leq \mathbf{p}_s\mathbf{q}_t - \mathbf{p}_s\mathbf{q}_s$*.*

When we interpret $x_{t,s} = 1$ as $\mathbf{q}_t \; R_0 \; \mathbf{q}_s$, we easily observe the similarity between the conditions in Definitions 2 and 3.

The following theorem extends the well-known theorem in introduced in Afriat (1967) and Varian (1982) by adding IP-GARP to it.

Theorem 1 *Let* $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t\in T}$ *be a set of observations. Then the following statements are equivalent:*

(i) There exists a non-satiated utility function that rationalizes S;

(ii) There exists a well-behaved utility function that rationalizes S;

(iii) S satisfies GARP;

(iv) For all $t \in T$, there exist $U_t \in \mathbb{R}_+$ and $\lambda_t \in \mathbb{R}_{++}$ such that for all $t, s \in T$:

$$U_t - U_s \leq \lambda_s(\mathbf{p}_s\mathbf{q}_t - \mathbf{p}_s\mathbf{q}_s);$$

(v) S satisfies IP-GARP.

The equivalence between the first two statements indicates that if the data is rationalisable by any utility function then it also rationalisable by a well-behaved utility function. Inter alia, this implies that concavity does not have testable implications. Statements (iii)-(v) present three alternative ways in which it is possible to verify whether the data are rationalisable.

The first method is a combinatorial one and was originally suggested by Varian (1982). The method consists of three steps, which comply with the three conditions in Definition 2 of GARP. The first step constructs the relation R_0 from the data set $S = \{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$. In particular $\mathbf{q}_t \, R_0 \, \mathbf{q}_s$ if and only if $\mathbf{p}_t\mathbf{q}_t \geq \mathbf{p}_t\mathbf{q}_s$. A second step computes the transitive closure of R_0, i.e. the relation R. Varian (1982) suggests using Warshall's algorithm (Warshall, 1962), which is an efficient algorithm for computing transitive closures. The third step verifies $\mathbf{p}_t\mathbf{q}_t \leq \mathbf{p}_t\mathbf{q}_s$ whenever $\mathbf{q}_s \, R \, \mathbf{q}_t$. If this is the case, the data set satisfies GARP and is, therefore, rationalisable.

The second method verifies the rationalisability conditions by testing feasibility of the corresponding Afriat inequalities. These inequalities are linear in the unknowns U_t and λ_t, which implies that their feasibility can be verified using simple linear programming methods. We refer to Afriat (1967) and Diewert (1973) for discussions of this method. An advantage of this method is that it provides not only an efficient way to verify the rationalisability conditions but also, via the computed values of U_t and λ_t, an estimate for the associated utility levels.

Finally, the third method verifies the rationalisability conditions via the conditions in Definition 3. These conditions are linear in the unknown binary variables $x_{s,t}$. Therefore feasibility can be verified by standard integer programming (IP) methods (branch and bound, cutting plane, etc.). Compared to the other methods, it is very inefficient and should not be recommended for applied work for the basic model developed in this section. However, in contrast to the other two methods, the IP method is be very useful when studying revealed preference characterizations of more complex alternative models; see Section 3.

2.2 Recoverability and forecasting

Recoverability aims at identifying the underlying preferences of the behavioural model under study. In parametric studies this is mostly equivalent to uniquely identifying the structural model parameters of the (in)direct utility function (representing the preferences). However, such an exercise is not feasible on the basis of revealed preference theory since there are usually many types of preferences that rationalize data. So the recoverability question that we have in mind focuses on identifying the set of preferences (or set of utility functions representing the preferences) that are consistent with a given data set.

The recoverability question basically aims at constructing inner and outer bounds for the indifference curves passing through an arbitrary, not necessarily observed, quantity bundle. This construction is primarily based on restrictions upon behaviour imposed by GARP. Let us illustrate the approach by means of Figure 1; the interested reader is referred to Varian (1982) and Varian (2006) for more details.

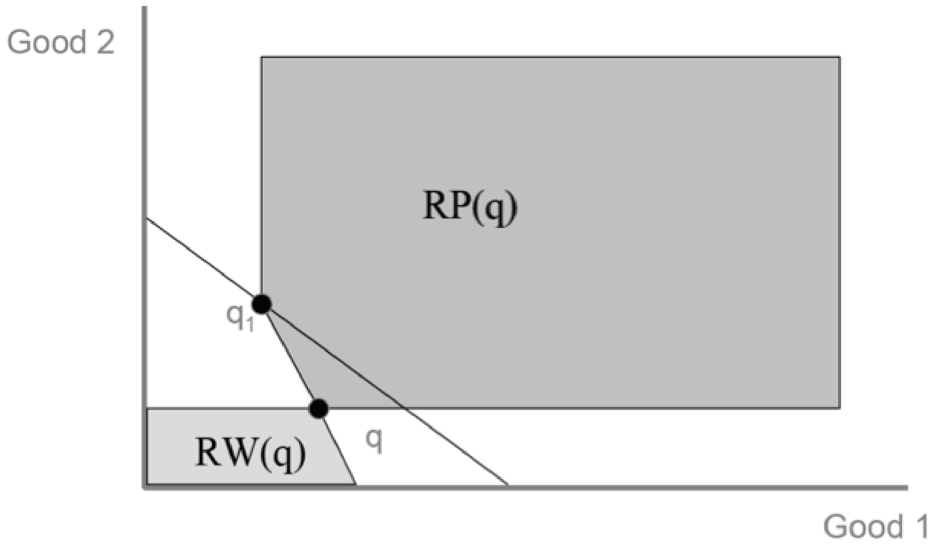

Figure 1: Recovery of the indifference curve

The figure shows a very simple dataset with only 1 observation $\mathbf{q}_1$ and one unobserved bundle $\mathbf{q}$ for which we want to do recovery. The relative prices are represented by the slope of the budget line through $\mathbf{q}_1$. As shown by Varian (1982), the set $RP(\mathbf{q})$ represents the set of all bundles that are revealed preferred to $\mathbf{q}$ and the set $RW(\mathbf{q})$ contains all the bundles that are revealed worse to $\mathbf{q}$. These sets are independent of the prices associated with $\mathbf{q}$. As such the boundaries of these two sets form the inner and outer bound for all indifference curves passing through $\mathbf{q}$ which are consistent with the observed choices and the preferences revealed by those choices.

It is clear from our example that the inner and outer bounds are not necessarily close to each other. This may serve as an illustration of the critique that a revealed preference approach does not have bite: in this particular case indifference curves can be very different from each other and still be consistent with observed behaviour. However, the inner and outer bounds may be much closer together if more observations are available and, indeed, are uniquely determined as the price-quantity data become completely dense. Moreover, recent research by Blundell et al. (2003, 2008) shows that one can dramatically tighten these bounds by combining revealed preference theory and the nonparametric estimation of Engel curves; see also our discussion in Section 4.

Characterizing indifference curves is not the only thing we can do on the basis of revealed preference theory. We can also make predictions of consumer behaviour in new situations. That is, situations were the consumer is faced with a new budget set. Let us illustrate this by means of Figure 2. The figure shows a data set with two observations. Suppose now that the consumer is faced with a new budget line indicated by the dashed line. It is clear that all bundles that exhaust this budget are within the reach of the consumer. However, not all these bundles are consistent with GARP. Actually, only the bundles on the bold line segment are consistent with GARP. The other bundles on the dashed line generate inconsistencies with rationality in the sense that they are not cost minimizing with respect to their revealed preferred set. Once again, it is clear that the set of possible rational outcomes (weakly) shrink if more observations are available.

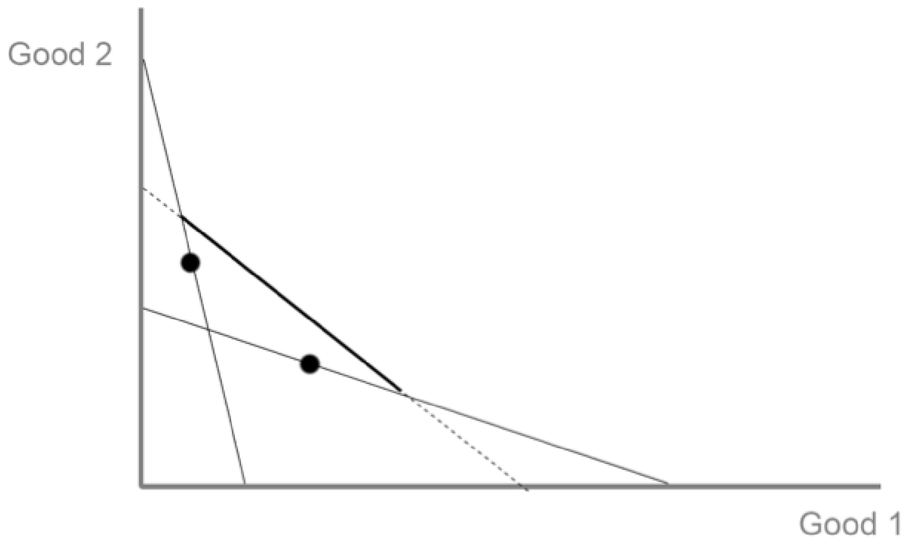

Figure 2: Forecasting new quantity bundles

3 Alternative models of rational demand

One of the main focuses of research in recent years has been the development of revealed preference characterisations of an increasing variety of economic models. Recently Chambers et al. (2013) showed that in principle there exists a set of revealed preference-type conditions for any optimising model that can be expressed as a series of universal statements. Thus many models of interest to economists can be given a revealed preference characterisation. In this section we do not aim at formally stating all the revealed preference results available in the literature. Instead we opt to review some of the more fundamental results in order to provide a good starting point and orientation for the interested reader. This overview is structured around three topics. In the first subsection we focus on special functional form restrictions that are frequently used to add some more structure to the basic model of rational demand. In the second subsection, we discuss extensions of the basic model by relaxing some of the underlying assumptions. Finally, in the last subsection we deal with multi-person behaviour.

3.1 Investigating functional forms

In the basic model of rational demand that we discussed above, we are considering any type of (well-behaved) utility function. In other words, consistency with for instance GARP implies that there exists at least one utility function that allows describing the data in terms of the behavioural model. However, in general the class of utility functions is often restricted in order to simplify the (empirical) analysis.

Revealed preference theory allows researchers to investigate these extra assumptions. That is, if the data satisfies for instance GARP, but not the revealed preference characterization corresponding to the specific class of utility functions, then we can conclude that it is not rationality of preferences per se that is the problem, rather it is the further restriction on the form of preferences. The results discussed below allow for such tests.

Homotheticity. A utility function is homothetic if it is the positive monotonic transformation of a function that is homogeneous of degree 1. This class of functions compromises well-known types of utility functions such as Cobb-Douglas utility functions and Constant Elasticity of Substitution utility functions. Working with this class of utility functions implies for instance that Engel curves are straight lines through the origin, meaning that is straightforward to model income effects for given prices, which in turn is convenient for extrapolating demand behavior.

Based on Afriat (1972) and Diewert (1973), Varian (1983) presents an Afriat-type theorem for characterizing rational demand in terms of homothetic utility functions. Essentially, his characterization combines the Afriat inequalities presented in Theorem 1- we still need that the data is rationalisable- with the extra assumption that for given prices the income effect is constant. This results in a system of linear inequalities that the data need to satisfy in order to be rationalizable by a homothetic utility function. He also provides a combinatorial reformulation of this system, which he labelled HARP (i.e. Homothetic Axiom of Revealed Preference), and he shows that the well-behavedness of the utility function is again not testable.

Cherchye et al. (2013a) extend this discussion to the class of quasi-homothetic utility functions. That is the corresponding Engel curves are straight lines but not through the origin. Remarkably, these authors show that, in the absence of proportional prices, the revealed preference characterization boils down to GARP. In other words, assuming quasi-homotheticity is not restrictive at all; see Cherchye et al. (2011b) for more discussion.

Separability. Separability implies that the goods can be divided in groups and that for each group there is a subutility function capturing the preferences for those goods, which is independent of the consumption of goods outside of that group. On top of that, there is a macro utility function that aggregates the preferences over the groups. If this macro function can be any well-behaved utility function then we are talking about weak separability. While if this macro function is additive in terms of the subutility functions, then we are considering additive separability. Finally latent separability means that goods can be part of several groups.

Separability is a very strong but very useful assumption in applied work. For example, it allows researchers to focus on individual markets for related goods and, combined with (quasi-)homotheticity, it also implies two stage budgeting, which simplifies the analysis of consumer behaviour.

Varian (1983) presents the revealed preference characterizations of both weak and additive separability; see also Afriat (1969) and Diewert and Parkan (1985) for related results. Crawford (2004) contains the revealed preference characterization of latent separability. These characterizations state that we need two types of utility functions. First, for every group we need a subutility function capturing the preferences for these goods. As such, the observed data for each group needs to satisfy the revealed preference conditions discussed in Section 2. Second, we also need a macro utility function aggregating the preferences over the groups. Again this boils down to the usual revealed preference conditions, but this time in terms of unobservable information (i.e. the unobserved utility levels and marginal utilities of income for each group).

All this implies that the data needs to satisfy a system of non-linear inequalities, which is not attractive from an empirical point of view. The only exception are the revealed preference conditions of additive separability, since in that case the marginal utility of income is constant, which allows to rewrite the system as a linear programming problem. This non-testability of the characterization of weak separability has lead to several papers focussing on either necessary or (separate) sufficient conditions for testing weak separability; see, among others, Swofford and Whitney (1987), Swofford and Whitney (1988), Barnett and Choi (1989), Swofford and Whitney (1994), Fleissig and Whitney (2003) Fleissig and Whitney (2007) and Fleissig and Whitney (2008). Finally, in a recent working paper, Cherchye et al. (2013d) presented an integer programming formulation for the setting with two subgroups. Attractively, this makes the revealed preference test easy to apply for this special case of weak separability.

(Generalized-) quasi linear utility functions. A final class of utility functions that are often used are (generalized) quasi linear utility functions. Quasi-linear utility functions are utility functions that are linear in at least one good, usually called the numeraire. This has strong implications (e.g. absence of income effects for all but a single good, risk neutrality, etc.) that simplifies the empirical analysis substantially. Generalized quasi linear utility functions slightly relaxes the linearity assumption by allowing that the numeraire is multiplied by a function defined in terms of a subset of goods. Bergstrom and Cornes (1981, 1983) and Bergstrom (1989) showed that this type of preferences is equivalent to assuming that utility is transferable among consumers as long as the subset of goods are public goods to all these consumers. Transferable utility in turn is a very popular assumption in matching models in order to define stability of the matchings.

Brown and Calsamiglia (2007) present and Afriat-type theorem for quasi-linear utility functions that essentially adds to Theorem 1 that the marginal utility of income should be constant. Cherchye et al. (2011b) extend their results toward generalized quasi-linear utility functions. In both cases the tests are easy to apply, although in the latter case it is via integer programming (which can make it time consuming).

3.2 Investigating richer models

In the previous subsection we focused on extra functional assumptions that restrict the class of utility functions in order to simplify the (empirical) analysis. In this subsection we take a different stance by relaxing the assumptions underlying the basic rationality model. That is, up till now we had a consumer in mind which does not take intertemporal issues into account, who faces linear budget sets and is consuming a set of non-discrete goods. Below we review some seminal contributions that focus on relaxing these assumptions in order to obtain a more realistic model. Importantly, all the results that we present are fairly easy to apply, which again makes revealed preference theory more attractive for using it in applied work.

Intertemporal behaviour. The model studied in Section 2 does not consider the problem of intertemporal allocations. Implicitly, while taking a decision in some observation t, the consumer does not take decisions from the past or for the future into account. There are of course many reasons to argue that this is a naive assumption. But at the same time these dynamic or intertemporal models are also much more complicated. Indeed, since the future is uncertain one should ideally also study risk attitudes and/or work with expected utility. See, Varian (1988) for revealed preference results related to risk aversion and Green and Osband (1991), for a revealed preference analysis focusing on expected utility

The following papers make abstraction from this uncertainty in order to derive some 'benchmark' results. Browning (1989) is the first paper that presents a revealed preference characterization of a life cycle model. In this model, the consumer decides at the beginning his total life consumption plan in order to smooth his/her consumption over all the periods, which obviously takes future decisions into account. This smoothing implies that the marginal utility of income should be constant over the whole time horizon. Moreover, Browning assumes that the decisions for some period are not influenced by consumption in other periods. Given all this, the revealed preference conditions boil down to the linear system Afriat inequalities discussed before, but this time with a constant marginal utility of income. Browning (1989) does not present this set of linear inequalities but instead he presents an equivalent combinatorial condition, which he named cyclical monotonicity. Finally, Crawford (2010) and Demuynck and Verriest (2013) extend Browning (1989) by providing the revealed preference conditions for models that allow habit formation and/or addiction (i.e.consumption in some period depends on consumption in other periods).

Non-linear budget sets. The results stated in Afriat's Theorem crucially depend on the linearity of the budget set. To test GARP, one should check the cost minimization condition and this condition can be easily operationalised due to the linearity of the budget set. This is also clear from the equivalent linear program stated in condition (iv). One important implication of linear budget sets is also that concavity does not have testable implications, essentially since choices in regions of non-convexity could never be observed; see Forges and Minelli (2009) for more discussion.

Besides all this, there is also an empirical motivation to consider non-linear budget sets. Indeed, due to tax systems most labour supply applications of revealed preference theory have to deal with non-linear budget sets. Or, richer intertemporal models that try to incorporate that financial markets are not working perfect, can also lead to non-linear budget sets.

Matzkin (1991) presents the first revealed preference results in the setting of non-linear budget sets and concave utility functions. Her results are extended in Forges and Minelli (2009) and these authors drop the concavity of the utility function. As mentioned above this allows them to show that concavity has testable implications. Finally, Cherchye et al. (2013c) combine the two previous papers by deriving the revealed preference characterizations for very general budget sets and concave utility functions. These authors also provide linear programming formulations of their results, which makes them attractive for applied work. Essentially, all these papers have in common that they present Afriat-type theorems in which the linear budgets are replaced by a convenient representation of the non-linear budget sets.

Discrete goods and characteristics. The results presented in Blow et al. (2008) and Polison and Quah (2013) allow researchers to relax the assumptions related to the consumed goods. The former paper focuses on the setting where consumers are interested in the characteristics of the goods (and not in the goods themselves). While the latter paper deals with a setting in which goods can be discrete in nature. Both assumptions are crucial for the realistic nature of empirical applications, but they also make the (theoretical) analysis more complex.

Indeed models of preferences over characteristics instead of preferences over goods imply that the empirical analyst does no longer directly observes the willingness to pay. That is, the price paid for the good needs to be decomposed into prices that the consumers are willing to pay for the characteristics. Next, if goods are discrete in nature, then this implies that the (non-satiated) consumers cannot any longer exhaust their budget. As such one needs to deal with the fact that there is a remaining budget and that this information is generally not available in the empirical analysis.

3.3 Investigating multi-person behaviour

In the above two subsections we focused on alternative versions of our basic model. However, in the end all these models still have in common that the consumer is maximizing some utility function subject to a budget constraint. In this subsection we want to go one step further by presenting revealed preference results for multi-person behaviour. This is important since empirical applications of revealed preference theory are generally applied to household consumption data. There is a lot of empirical evidence, that such applications should take into account that there are multiple decision makers in multi-member households; see, e.g. Vermeulen (2005) for an overview and Cherchye and Vermeulen (2008) and Cherchye et al. (2009) for evidence based on revealed preference tests.

Models of multi-person behaviour therefore use a different starting point. Instead of assuming that there is a utility function representing the preferences of the group (or household), it takes into account that each individual has its own utility function and that individuals enter into a decision process with the other individuals for deciding how to spend the common budget. The outcome of this decision process should not necessarily lead to a transitive preference ordering, which explains why there should not be a utility function representing the group.

We start by reviewing the classical answer to these type of questions, namely general equilibrium theory and aggregation. These models take a societal viewpoint. Subsequently we discuss the recent revealed preference literature on household models.

Modelling society. The revealed preference analysis of multi-person behaviour started with Brown and Matzkin (1996). This paper focuses on a simple exchange equilibrium in which market prices, individual incomes and aggregate endowments are observed. To obtain testable implications for this set-up, these author derive the conditions which guarantee that the observed data lies on the so-called equilibrium manifold. That is, they showed that individual rationality and market clearing, restrict the response of endogenous aggregate variables to perturbations on individual endowments, which in turn allows them to state their Afriat inequalities for this setting. This surprising result contrasts with the conclusions obtained by Sonnenschein-Mantel-Debreu, which basically states that general equilibrium models do not generate testable implications, and has generated a lot of follow-up research; see Carvajal et al. (2004) for a survey and Cherchye et al. (2011d) for a recent contribution.

Another question related to modelling the society is the aggregation problem. That is, does there exists a social welfare function representing the preferences of the society and if so, what is the relation with the preferences of the individuals in that society. As discussed in Varian (1984) the answer to the first question is equivalent to having that the aggregate data (i.e. the sum of the individual demands and the common price) satisfies GARP. However this social welfare function cannot be used to make normative conclusions, simply because there is no relation at all with the individuals in the society. We refer to Cherchye et al. (2013a) for revealed preference characterizations that do allow for aggregating the preferences of the individuals in the society. Formally this problem is very related to our discussion of separability and its corresponding empirical issues.

Modelling household behaviour. As discussed above, to model household consumption decisions one should take the individual preferences of the household members and the decision process into account. There is a wide variety of possibilities for modelling this decision process of which the so-called collective models is the most popular one; see Chiappori (1988) for a seminal contribution, which also contains some revealed preference theory. Collective models allow for any kind of decision process as long as the outcome is Pareto efficient. The revealed preference characterizations of collective models are presented in Cherchye et al. (2007, 2010, 2011a).

In almost all expenditure surveys, there is only data available at the household level and not at the level of each individual household member. This is a serious data limitation since, as one may expect, the revealed preference tests of a collective model boil down to testing the usual revealed preference conditions at the individual level. To deal with this problem, Cherchye et al. (2008) and Cherchye et al. (2011a) developed integer programming formulations similar to IP-GARP that can easily be applied.

Finally, there are also revealed preference characterizations available of alternative forms of modelling the household decision process. First, one could replace the Pareto efficiency assumption by assuming that the outcome of the decision process should be a Nash equilibrium. Such a model also puts minimal structure on the decision process, but also takes into account that individuals can behave strategically. See Cherchye et al. (2011c) for a more in-depth discussion and for the corresponding revealed preference characterization. Second, one could put more structure on the decision process by assuming that the households take decisions according to some specific bargaining protocol. The revealed preference theory of some of the most popular bargaining models, such as for instance Nash bargaining, is presented in Chambers and Echenique (2011), Cherchye et al. (2013b) and Carvajal and Gonzales (2013),

4 Bringing revealed preference theory to the data

As we discussed above, there has been a significant broadening of the scope of revealed preference methods since the foundational work by Samuelson (1938, 1948), Houthakker (1950) and Afriat (1967). The practical empirical application of revealed preference methods has, arguably, lagged somewhat. This may be because of the relative unfamiliarity of revealed preference methods and the fact that empirical revealed preference often requires researchers to address and find practical solutions to some difficult combinatorial problems. It may also be because empirical revealed preference work with sample data presents a number of important challenges. Consider, for example, the question of a straightforward GARP test. Firstly there is the matter of interpreting the outcome of the test for a single economic agent - what should we make of it if the subject passes/fails? That problem is made more difficult (and indeed the test itself may be hard to conduct) if there are problems with the data such as measurement error or missing data. When we have data on a number of different individuals the question of the pattern and nature of preference heterogeneity arises. Finally there is the problem of going beyond the data at hand and making inferences about some population of interest. In the following subsections, we discuss all these issues in the context of the basic model of rational demand.

4.1 Interpreting RP tests

Consider the $\{\mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ data for a single individual and suppose that everything is measured perfectly. Whether or not this individual's behaviour is rationalisable by the theory is completely deterministic: if the data satisfy GARP then the data are consistent with utility maximisation, otherwise not. Despite this disarming simplicity it can still be hard to know what to make of the result. For example, suppose that utility maximisation was not the data generating process (DGP). Will the RP conditions be sensitive enough to detect it? In statistical hypothesis testing this question concerns to the "power" of the test of a probabilistic model against a probabilistic alternative. In RP tests in this kind of nonstochastic environment the statistical notion of power is not strictly relevant, yet there is clearly a need to consider the same sort of question and many of the same considerations apply. In particular, just as is the case with statistical power calculations, the answer will depend on the alternative DGP considered: the RP test might be quite successful at detecting violations of GARP under some alternative DGPs but less successful given others.

The difficulty is that there are many alternatives to rational choice models but no obvious benchmark. One important, non-rational alternative considered by Becker (1962) was a probabilistic DGP: uniform random choice on the budget constraint. Bronars (1987) applied this in an RP context by calculating the probability of observing a violation of GARP with this DGP operating on the observed constraints. Bronars' approach remains the most popular method but more recent contributions (notably Andreoni et al. (2013)), whilst sticking with the idea of a probabilistic alternative DGP, consider more data-driven alternatives to uniform random choice - they suggest drawing from the empirical distribution of observed choices to allow for a more realistic alternative. Work on this topic is on-going, but the leading approaches which use probability models to frame alternative choice models, are principally variations on Bronars' method.

A different appraoch is to try to avoid the problem of having to specify the alternative DGP, and, instead of asking whether the outcome of an empirical RP test represents a statistically significant departure from a probabilistic DGP, asks whether the results of the test represent an *economically* significant departure from rational choice. The key to this is to see that when a consumer violates RP conditions, that consumer appears to waste money by buying a consumption bundle when a cheaper bundle is available and also revealed preferred to it. The cost-efficiency measure suggested in Afriat (1973) is the smallest amount of this wastage (as a fraction of the overall budget) consistent with the given demand data. This index provides a simple way of measuring the size of a violation of GARP and does so in units which are easy to understand and to interpret economically. The converse of the Afriat cost efficiency index was recently proposed by Andreoni et al. (2013) as a way of interpreting GARP successes: given a dataset in which no revealed preference violations are detected, the Afriat Power Index measures how much the consumer's budget would have to be adjusted in order to *induce* a violation. If the required adjustment is small then the test is considered to be sensitive, if it is high then it is not.

Related to this is the approach that builds on the ideas in de Finetti (1937) concerning "Dutch books" or "money pumps". The idea is that an individual who violates RP conditions has preferences which contain cycles and this means that they are open to being exploited as a "money pump" by an unscrupulous trader who simply buys goods from them at a price they are willing to accept and then sells them back to them again at a (higher) price they are willing to pay. Given a revealed preference cycle of length J with $\mathbf{q}_j R^0 \mathbf{q}_{j-1}$ the intransitivity means that the consumer would also prefer $\mathbf{q}_{j-1}$ to $\mathbf{q}_j$, so $\mathbf{p}_j\mathbf{q}_j - \mathbf{p}_j\mathbf{q}_{j-1}$ can be extracted at each point in the cycle and $\sum_{j=1}^{J} \mathbf{p}_j\mathbf{q}_j - \mathbf{p}_j\mathbf{q}_{j-1}$ in total. Echenique et al. (2011) suggest the money pump (expressed as a proportion of the consumer's total expenditure) as an aid to interpretation when GARP fails; the more money that can be extracted from the individual in this manner, the worse the violation of RP theory. These authors also show how to address the considerable combinatorial/computational challenges involved in calculating the money pump index - as the number of potential cycles which need to be investigated can be huge even when the dataset itself is not large.

A last alternative approach, which has recently been applied to revealed preference tests by Beatty and Crawford (2011), comes from the literature on experimental game theory and is due to Selten and Krischker (1983) and Selten (1991). The key insight is that in their revealed preference guise, shorn of special functional form assumptions, economic models generally generate restrictions in the form of well-defined sets of choices which are consistent with the model of interest. To investigate the performance of models which predict sets, it is useful to consider the feasible outcome space (say) P and the model's prediction as the subset $S \subseteq P$. It is then important to acknowledge the relative size of the predicted/theoretically consistent subset. The essential idea is that if the set of observations explainable by the model (S) is very large relative to the set of behaviours which the consumer could possibly display (P), then simply noting that many of the observed choices lie in S is not a very demanding requirement - they could hardly have done otherwise and the test is therefore not very sensitive. This means that "fit" alone (the proportion of the sample which passes the relevant test) is not a sufficient basis for assessing the outcome of an RP test. A better approach would be to consider the trade off between the pass rate and some sort of measure of how sensitive the test is. Let a denote the size of the theory-consistent subset relative to the outcome space for the model of interest. The relative area of the empty set is zero and the relative area of all outcomes is one so $a \in [0, 1]$. Now suppose that we have some choice/outcome data. Let r denote the pass rate; this is simply the proportion of the data that satisfies the restrictions of the model of interest. Selten (1991) provides an axiomatic argument that the trade-off between the ability to fit the data and the restrictiveness of the theory should be the difference measure: $r - a$. Other axiomatisations would produce different forms for the measure of the outcome of an RP test, but the basic idea that the measure should combine both the pass rate and some measure of sensitivity remains an important and promising area for further work.

4.2 Missing data

Suppose now that the data is less than perfect. In particular consider the case of missing data. It would seem that missing data is fatal to empirical implementation of revealed preference methods. In some cases this is true, but in others progress can still be made. With a "full" set of observations an RP test asks whether there exists a well-behaved utility function which rationalises these data. When some of the data are missing, we can ask a slightly different question: do their exist feasible values for the missing observations such that there exists a well-behaved utility function? In some cases it may be that the answer is always "yes", implying that the utility maximisation hypothesis cannot be falsified and the test collapses. One such situation occurs when all of the price or quantity data for a particular good are missing; Varian (1988) shows that in this situation it is not possible to test RP conditions because it is always possible to find values for the missing price or quantity series such that the data pass the RP conditions. However the situation is not always so bleak. A very common example of missing data in consumer surveys concerns prices which are recorded when the consumer makes a purchase, but which are not recorded when the consumer does not transact. Thus we observe prices when the quantity is positive but not when the quantity is zero. In these situations there often are restrictions on what the missing prices can possibly be and, by the same token, RP conditions can be violated if these conditions are not met. The way to formulate RP tests with this type of missing data is described in Blow et al. (2008) in the context of linear characteristics models, but since these models, in which consumers have preferences for characteristics instead of goods, can be easily transformed into the standard preference-for-goods model, the method they describe also works perfectly for the canonical RP test.

4.3 Statistical errors

As we emphasized in the Introduction, one important difference between structural econometrics and empirical revealed preference lies in the absence of an error term in the latter. Certainly error terms rarely appear in revealed preference theory: there is no mention of an error term in Afriat's Theorem or in any of the other revealed preference characterisations of the various models discussed in Section 3. But as soon as we attempt to take those revealed preference conditions to data, errors can no longer necessarily be ignored. The most obvious situation arises when we consider measurement errors, but identical issues arise when revealed preferences are applied to statistical objects (like estimates of aggregate consumption as in Browning (1989) or nonparametric Engel curves as in Blundell et al. (2003, 2008)). In these cases the price-quantity data we observe is a function of a random variable. This introduces a statistical element to empirical revealed preference and forms an important link between revealed preference with structural econometrics, which, as we discuss in Section 5, appears to be an important future direction for research.

To illustrate the case for classical additive measurement error consider the model

$$\mathbf{q}_t = \mathbf{q}_t^* + \mathbf{e}_t$$

where $\mathbf{q}_t^*$ denote the true values of demands and $\mathbf{e}_t$ is a vector of classical measurement errors. Suppose that we are interested in the null hypothesis that the true data $\{\mathbf{p}_t, \mathbf{q}_t^*\}_{t \in T}$ satisfy GARP. Blundell et al. (2008), building on Varian (1985) construct a statistical test for violations of the revealed preference conditions by supposing that the observed demands are known functions of a finite set of parameters θ_t so that $\mathbf{q}_t = \mathbf{f}(\theta_t)$ for known $\mathbf{f}(.)$. The RP restrictions in the null can be represented by a set of moment inequality restrictions (MIR) involving θ_t. Blundell et al. (2008) then show that it is possible to appeal to results by Manski (2003), Chernozhukov et al. (2007) and Andrews and Guggenberger (2007) for moment inequality estimators of this type. They establish that there always exists values θ_t that satisfy the MIR as long as the support of the θ_t values allow for any positive demands that satisfy adding-up. Generally, there will be a set of values for θ_t that satisfy the RP restrictions and testing consistency with these conditions boils down to verifying if this set includes the observed demands. If the RP conditions fail for the observed demands $\mathbf{q}_t$, it is possible to generate a restricted estimator, $\widehat{\mathbf{q}}_t$ using the following Gaussian quasi-likelihood ratio or minimum distance criterion function:

$$L = \min_{\{\widehat{\mathbf{q}}_t\}_{t \in T}} \sum_{t=1}^{T} (\mathbf{q}_t - \widehat{\mathbf{q}}_t)' \, \Omega_t^{-1} (\mathbf{q}_t - \widehat{\mathbf{q}}_t)$$

subject to the restriction that $\{\mathbf{p_t}, \widehat{\mathbf{q}}_t\}_{t \in T}$ satisfies GARP and where the weight matrix Ω_t^{-1} is the inverse of the covariance matrix of the demands. The solution to this problem leads to demands $\widehat{\mathbf{q}}_t$, which satisfy the RP restrictions and which are unique almost everywhere. Evaluated at the restricted demands, Blundell et al. (2008) show that the above distance function also provides a test statistic for the RP conditions and that this test falls within the general class of misspecification tests investigated in Andrews and Guggenberger (2007, Section 7).

4.4 Heterogeneity

For anyone who has ever looked at consumer microdata, the great variety of behaviour on display amongst consumers and households who are, in most observable respects, very similar is striking. Ideally the researcher would try to model each household individually but most consumer panels are "small T, large N" affairs. This makes it impossible to estimate sufficiently flexible and reliable econometric models at the individual level. The standard structural econometric approach is therefore to pool data across consumers and to model the behaviour of individuals as a combination of a common component and an idiosyncratic component which reflects unobserved heterogeneity. Of course, this immediately requires a combination of often strong assumptions regarding the form of the statistical model and the joint distribution of unobserved heterogeneity with the observables; see for example Brown and Walker (1989) and Lewbel (2001) for more discussion.

Because revealed preference approaches can be applied to very short panels (you only need two observations on an consumer to test GARP, for example) it is generally possible to proceed individual-by-individual even when the T dimension is far too small even to contemplate a statistical approach. This one-at-a-time approach, of course, allows for the maximal amount of heterogeneity - consumers can differ with respect to whether or not they behave in accordance with the theory, and if they are theory-consistent then they can differ with respect to choices and preferences.

However, in some circumstances (a pure cross-section dataset in which individuals are observed only once, for example) heterogeneity cannot be usefully preserved and, indeed, sometimes heterogeneity itself is the object of interest. When this is the case, revealed preference methods can still be used. Instead of applying them to longitudinal data on individual consumers and checking for the existence and stability of well-behaved preferences, they can be applied to cross-sectional data on many different consumers revealed preference restrictions are then interpretable as a check for the commonality of well-behaved preferences.

Gross (1995) applied RP tests to cross sectional consumer data in order to look at the evidence for and against the assumption of homogeneous tastes and concluded that, in a sample drawn from the PSID (Wave IX (1976)), individuals did not share a common utility function. The idea that the choices of all of the consumers in a large microeconomic dataset could be explained perfectly by a single common utility function is probably, as Lewbel (2001) points out, "implausibly restrictive". Recently Dean and Martin (2010) and Crawford and Pendakur (2013) both recognised that tests like this one will reject as soon as one of the consumers has tastes different enough to be detected by the test. The possibility, then, that all of the rest of the data are rationalisable by a single utility function would be masked by the rejection caused by the presence of this single consumer. In order to investigate this further, it would, in principle, require the researcher to look at all possible subsets of the data and to conduct RP tests in all of them to detect the true pattern of preference heterogeneity. This is too computationally demanding as there will be 2^N subsets to check and so this is another example in which researchers have had to take an algorithmic approach.

Dean and Martin (2010) suggest looking for the largest single subset which is consistent with common preferences - this is then a nice summary of an aspect of preference heterogeneity. To do this they develop a new algorithm, which is much more efficient than existing algorithms, that exploits an analogy between the revealed preference problem at hand and the minimum set covering problem, which is a well studied problem in the computer sciences and operations research literature. While the minimum set covering problem is NP hard there are a wide variety of algorithms that are extremely efficient and so, by cleverly translating the revealed preference problem into this form they are able to apply these solution methods. Crawford and Pendakur (2013) take a slightly different approach. Given a result like the one in Gross (1995), the researcher clearly needs more that one utility function to model the data. The question then of course is how many different utility functions we need. Crawford and Pendakur (2013) consider the problem of how to find the minimum number of utility functions necessary to fully explain all observed choices in a dataset. This is a computationally demanding partitioning problem and Crawford and Pendakur (2013) design an algorithm which is able to place tight, two-sided bounds on this minimum number.

4.5 Inference

If the data involved are a random panel sample of households and demands are measured without error, then inference about objects like the proportion of households which satisfy RP restrictions in the population is straightforward. A sample proportion can be viewed as the fraction of "successes" in N independent Bernoulli trials with the same success probability p. The central limit theorem implies that for large N, the sample proportion $\widehat{p} = \sum_{i=1}^{N} I\left(\text{consumer } i \text{ passes RP}\right)$ is normally distributed with mean p and standard deviation $\sqrt{p\left(1-p\right)N}$ so the statistic $z = \left(\widehat{p} - p\right)/\sqrt{p\left(1-p\right)N}$ follows the standard normal distribution. This serves as the basis for statistical inference regarding population proportions.

Inference with repeated cross-sections from a heterogeneous population is more difficult. The issue here is that we do not see the same consumer twice, so we cannot proceed on an consumer-by-consumer basis, checking the RP conditions for each one as before. The object of interest remains the population proportion of consumers who satisfy the RP conditions. However, this parameter depends on the joint distribution of choices over different budget sets and repeated cross-sectional data do not reveal this: only its marginal distributions can be observed. Thus, the population parameter of interest is not point identified. Hoderlein and Stoye (2013) show that in the context of the Weak Axiom of Revealed Preference, that it can be partially identified: i.e. bounded. They describe the problem as a copula problem and use copula techniques to analyse it. They also show that inference on the bounds is an application of partial identification through moment inequalities. This approach is somewhat in the tradition of the literature on the partial identification of treatment effects and it emphasizes the conceptual value of understanding clearly how much might be learned from the data without identifying assumptions. Consequently the approach is careful to impose no or very weak homogeneity assumptions and as a result it seems that WARP may be hard to reject. However, it is important to note that WARP does not exploit transitivity of preferences - a much stronger assumption - so it remains to be seen how this approach might be fruitfully extended to RP conditions which are more demanding.

5 Conclusion

In this review we have focussed on the present state of empirical revealed preference in two general respects: work which extends the scope of these methods to a variety of richer models of behaviour, and work which seeks to apply these methods to data. To conclude we briefly consider how each of these areas might develop in the future. In terms of scope we note that all of the revealed preference characterisations of models which we have discussed have remained firmly embedded within the neoclassical tradition in which the whole literature began. It therefore seems to be an interesting and open question to ask whether these methods might be applied to non-standard "behavioural" models.

As far as empirical applications are concerned we have described how empirical revealed preference differs from structural econometrics and the benefits and drawbacks of the approach. One of the principal drawbacks of RP methods, which readers will have picked up on, is that, compared to structural econometric methods, they are relatively ungainly. That is to say they produce characterisations of preferences, for example, which are difficult to represent succinctly and awkward to interpret (e.g. piecewise linear bounds computed on individual indifference curves). This compares unfavourably with the traditional econometric approach, which focusses on simple functional forms with parameters that have useful and straightforward economic interpretations. On the other hand revealed preference methods seem to make fewer maintained assumptions than standard methods. An important area for research may therefore be to investigate whether is may be possible to blend empirical RP and econometric methods and preserve the strengths of both approaches.

5.1 Behavioural models

Recently there has been renewed academic interest to economic models that moves somewhat away from the neoclassical tradition of treating people as always-rational decision makers. This "behavioural economics" approach focuses on models which combine conventional neoclassical microeconomic methods with behavioural and modelling assumptions that have more plausible sociological and psychological foundations. For instance, they allow for situations in which people are influenced by others, may make mistakes or may come to regret their choices. Behavioural economics promises much in terms of its potential to help us understand choices which could otherwise prove resistant to straightforward explanation by standard rational choice models. At present,it remains something of an open question as to whether these models might be amenable to a revealed preference characterisation.

Neoclassical models in economics, for all of the (often justifiable) criticisms which they attract, are at least falsifiable in a revealed preference sense - it is possible in principle to detect when the data and the model are not rationalisable. There is, as far as we know, nothing like an Afriat's Theorem for behavioural economic models. This is of interest because it is important to know whether, without auxiliary hypotheses, observational data is able to tell us when behavioural models are unable to rationalise behaviour. In this respect we want to make some concluding remarks on two particularly interesting classes of behavioural models: reference-dependent preferences and time-inconsistent choices. As we will discuss, these classes are sufficiently close to existing neoclassical models which do have revealed preference characterisations. So it might be possible to investigate whether or not they are characterisable by an Afriat-type theorem.

Reference-dependent preferences. Reference-dependent preferences incorporate ideas from prospect theory. Tversky and Kahneman (1991) posit that individuals understand their options in decision problems as gains or losses relative to a reference point. The reference point is not generally observable (to the researcher): sometimes it is modelled as the current position (i.e. the status quo) of the individual, but it might also depend on past consumption, expectations, social comparisons, social norms, etc. A feature of prospect theory, which reference dependence inherits, is that the value function exhibits loss aversion, so that negative departures from one's reference consumption level decrease utility by a greater amount than positive departures increase it. Another feature of prospect theory is that the value function exhibits diminishing sensitivity for both gains and losses, which means that the value function is concave over gains and convex over losses. Taken together this implies that changes in an unobservable reference point is capable to altering the individual's preferences. It would therefore seem that giving this model a revealed preference characterisation with empirical content is going to be far from straightforward yet the model is deterministic and indeed rational in the sense that, conditional on the reference point, preferences are well-behaved. Although this appealing set-up is deterministic and indeed rational in the sense that, conditional on the reference point, preferences are well-behaved, it seems that giving this model a revealed preference characterization with empirical content is going to be far from straightforward due to the unobserved reference point.

Time-inconsistent choices. Models of time inconsistent choice relax the standard assumption that all of the disparate motives underlying intertemporal allocations can be condensed into a single parameter - the discount rate - which is constant. Constant discounting entails an even-handedness in the way a person evaluates time. It implies that a person's intertemporal preferences are time-consistent, which means that later preferences "confirm" earlier preferences. This consistency was exploited by Browning (1989) in his derivation of a revealed preference characterisation of the strong rational expectations hypothesis. However, whilst the standard model assumes constant discounting, the leading behavioural alternative, suggests that discounting is hyperbolic - that a person has a declining rate of time preference. This implies that when subjects are asked to compare a smaller-sooner reward to a larger-later reward, the implicit discount rate over longer time horizons is lower than the implicit discount rate over shorter time horizons (see, e.g., (Thaler and Shefrin, 1981)). Once again the model is perfectly rational and deterministic, but whether or not it has a revealed preference characterisation akin to Browning (1989) is the subject of on-going work.

5.2 Empirical Revealed Preference and Structural Econometrics

At the start of this article we emphasised a key difference between empirical revealed preference and structural econometrics - whilst empirical revealed preference focusses on observables, the introduction of unobservables (error terms) is an essential aspect of structural econometrics. These error terms are there in part because the structural functions alone generally do not rationalise the data. The empirical revealed preference approach, being based on the weaker requirements of inequality restrictions, generally has no need to resort to error terms. The great advantage, however, of structural econometrics is that it generally seeks to recover the structural functions of interest uniquely. Empirical revealed preference, by contrast, can typically only place bounds on these. Moreover if the bounds are wide then the revealed preference approach is, arguably, of little utility.

An important area of future research, therefore, lies at the boundary between econometrics and revealed preference. In particular there is the question of whether the inequality restrictions from revealed preference arguments can be used to help guide the estimation of structural econometric models. In a sense this may be simply a question of augmenting traditional econometric loss functions (sum of squared residuals, least absolute deviations, etc) with loss functions motivated by revealed preference theory. Blundell et al. (2008) represents an initial step in this direction. There the authors estimate a system of Engel curves and impose revealed preference restrictions with the result that the resulting estimated Engel curves minimise least-squares losses subject to the weak axiom of revealed preference. However the revealed preference restrictions are only applied locally - which is to say at a particular point in the income distribution - and this means that the entire Engel curve is not necessarily constrained to be consistent with a single set of well-behaved preferences. Imposing global consistency in an easily interpretable way is more challenging. A recent development in this direction is by Halevy et al. (2012). They aim to fit a single, simple parametric utility function to choice data subject to revealed preference conditions. The loss function in this case is based on the Afriat Efficiency index applied at each observation. The objective of the approach is to select a simple, tractable representation of preferences which minimises the inconsistency between the empirical revealed preference information contained in the choices and the ranking information contained in the recovered preferences. Of course one could just compute a piecewise linear, perfectly rationalising utility function as in Afriat (1967), but that method requires recovering twice the number of parameters as there are observations and the behavioural content of the utility function is almost impossible to interpret. The tradition in econometrics is to work with the simplest model that allow the researcher to adequately fit it to the data and also to interpret it - e.g. to have simple characterisation of concepts like elasticity of demand, risk aversion or time-preference. Drawing on this econometric approach the authors' method rather neatly trades off the inevitable misspecification of the necessarily overly simple rationalising utility function against parsimony/interpretability and represents what might turn out to be an important first step in this broad research program.

References

Afriat, S. N., 1967. The construction of utility functions from expenditure data. International Economic Review 8, 67–77.

Afriat, S. N., 1969. The construction of separable utility functions from expenditure data. Tech. rep., Mimeo, University of North Carolina.

Afriat, S. N., 1972. Efficiency estimation of production functions. International

Economic Review 13, 568–598.

Afriat, S. N., 1973. On a system of inequalities in demand analysis: An extension of the classical method. International Economic Review 14, 460–472.

Andreoni, J., Gillen, B. J., Harbaugh, W. T., 2013. The power of revealed preference tests: Ex-post evaluation of experimental design. Tech. rep.

Andrews, D. W. K., Guggenberger, P., 2007. Validity of subsampling and "plug-in" asymptotic inference for parameters defined by moment inequalities. Tech. rep.

Barnett, W. A., Choi, S., 1989. A Monte Carlo study of tests of blockwise weak separability. Journal of Business and Economic Statistics 7, 363–377.

Beatty, T. K. M., Crawford, I., 2011. How demanding is the revealed preference approach to demand. American Economic Review 101, 2782–2795.

Becker, G. S., 1962. Irrational behavior and economic theory. Journal of Political Economy 70, 1–13.

Bergstrom, T., 1989. A fresh look at the rotten-kid theorem and other household mysteries. Journal of Political Economy 97, 1138–1159.

Bergstrom, T. C., Cornes, R. C., 1981. Gorman and Musgrave are dual: an antipodean theorem on public goods. Economics Letters 7, 371–378.

Bergstrom, T. C., Cornes, R. C., 1983. Independence of allocative efficiency from distribution in the theory of public goods. Econometrica 51, 1753–1765.

Blow, L., Browning, M., Crawford, I., 2008. Revealed prefrence analysis of characteristics models. Review of Economic Studies 75, 371–389.

Blundell, R., Browning, M., Crawford, I., 2008. Best nonparametric bounds on demand responses. Econometrica 76, 1227–1262.

Blundell, R. W., Browning, M., Crawford, I. A., 2003. Nonparametric engel curves and revealed preference. Econometrica 71, 205–240.

Bronars, S. G., 1987. The power of nonparametric tests of preference maximization. Econometrica 55, 693–698.

Brown, B., Walker, M. B., 1989. The random utiity hypothesis and inference in demand systems. Econometrica 57, 815–829.

Brown, D. J., Calsamiglia, C., 2007. The nonparametric approach to applied welfare analysis. Economic Theory 31, 183–188.

Brown, D. J., Matzkin, R. L., 1996. Testable restrictions on the equilibrium manifold. Econometrica 64, 1249–1262.

Browning, M., 1989. A nonparametric test of the life-cycle rational expectations hypothesis. International Economic Review 30, 979–992.

Carvajal, A., Gonzales, N., 2013. On refutability of the Nash bargaining solution , forthcoming.

Carvajal, A., Ray, I., Snyder, S., 2004. Equilibrium behavior in markets and games: Testable restricions and identification. Journal of Mathematical Economics 40, 1–40.

Chambers, C., Echenique, F., 2011. Testable implications of bargaining theories. Tech. Rep. 1348, Caltech SS.

Chambers, C. P., Echenique, F., Shamaya, S., 2013. The axiomatic structure of empirical content. Tech. rep.

Cherchye, L., Crawford, I., De Rock, B., Vermeulen, F., 2013a. Gorman revisited: Nonparametric conditions for exact linear aggregation. Tech. rep., KULeuven.

Cherchye, L., De Rock, B., Sabbe, J., Vermeulen, F., 2008. Nonparametric tests of collective rational consumption behavior: An integer programming procedure. Journal of Econometrics 147, 258–265.

Cherchye, L., De Rock, B., Vermeulen, F., 2007. The collective model of household consumption: a nonparametric characterization. Econometrica 75, 553–574.

Cherchye, L., De Rock, B., Vermeulen, F., 2009. Opening the black box of intra-household decision-making. Journal of Political Economy 117, 1074–1104.

Cherchye, L., De Rock, B., Vermeulen, F., 2010. An Afriat theorem for the collective model of household consumption. Journal of Economic Theory 145, 1142–1163.

Cherchye, L., De Rock, B., Vermeulen, F., 2011a. The revealed preference approach to collective consumption behavior: Testing and sharing rule recovery. Review of Economic Studies 78, 176–198.

Cherchye, L., Demuynck, T., De Rock, B., 2011b. Is utility transferable? a revealed preference analysis. Ces dp.11.02.

Cherchye, L., Demuynck, T., De Rock, B., 2011c. Revealed preference analysis of noncooperative household consumption. The Economic Journal 121, 1073–1096.

Cherchye, L., Demuynck, T., De Rock, B., 2011d. Testable implications of general equilibrium models: an integer programming approach. Journal of Mathematical Economics 47, 564–575.

Cherchye, L., Demuynck, T., De Rock, B., 2013b. Nash bargained consumption decisions: a revealed preference analysis. Economic Journal 123, 195–235.

Cherchye, L., Demuynck, T., De Rock, B., 2013c. Revealed preference analysis for convex rationalizations on nonlinear budget sets. KULeuven working paper.

Cherchye, L., Demuynck, T., De Rock, B., Hjertstrand, P., 2013d. Revealed preference tests for weak separability: an integer programming approach. CES DP 11.25.

Cherchye, L., Vermeulen, F., 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.

Chernozhukov, V., Hong, H., Tamer, E., 2007. Estimation and confidence regions for parameter sets in econometric models. Econometrica 75, 1243–1283.

Chiappori, P., 1988. Rational household labor supply. Econometrica 56, 63–89.

Chiappori, P., Rochet, J., 1987. Revealed preferences and differentiable demand. Econometrica 55, 687–691.

Crawford, I., 2004. Necessary and sufficient conditions for latent separability. Tech. Rep. 02/04, cemmap.

Crawford, I., 2010. Habits revealed. Review of Economic Studies 77, 1382–1402.

Crawford, I., Pendakur, K., 2013. How many types are there? Economic Journal 123, 77–95.

de Finetti, B., 1937. La prevision ses lois logiques, ses sources subjectives. Ann. Inst. H. Poincaré 7, 1–68, (Translated in Kyburg, H. E., JR. and Smokler, H. E., (eds.) (1964). Studies in Subjective Probability. Wiley, New York).

Dean, M., Martin, D., 2010. How rational are your choice data? Tech. rep.

Demuynck, T., Verriest, E., 2013. I'll never forget my first cigarette: A revealed preference analysis of the habits as durables model. International Economic Review , forthcoming.

Diewert, E. W., Parkan, C., 1985. Tests for the consistency of consumer data. Journal of Econometrics 30, 127–147.

Diewert, W. E., 1973. Afriat and revealed preference theory. The Review of Economic Studies 40, 419–425.

Echenique, F., Lee, S., Shum, M., 2011. The money pump as a measure of revealed preference violations. Journal of Political Economy 119, 1201–1223.

Fleissig, A., Whitney, G. A., 2003. A new PC-based test for Varian's weak separability condition. Journal of Business and Economic Statistics 21, 133–145.

Fleissig, A., Whitney, G. A., 2007. Testing additive separability. Economics Letters 96, 215–220.

Fleissig, A., Whitney, G. A., 2008. A nonparametric test of weak separability and consumer preferences. Journal of Econometrics 147, 275–281.

Forges, F., Minelli, E., 2009. Afriat's theorem for general budget sets. Journal of Economic Theory 144, 135–145.

Green, E. J., Osband, K., 1991. A revealed preference theory for expected utility. Review of Economic Studies 58, 677–695.

Gross, J., 1995. Testing data for consistency with revealed preference. Review of Economics and Statistics 77, 701–710.

Halevy, Y., Persitz, D., Zrill, L., 2012. Parametric recoverability of preferences. Tech. rep., Working paper.

Hoderlein, S., Stoye, J., 2013. Revealed preference in a heterogeneous population. Review of Economics and Statistics , forthcoming.

Houthakker, H. S., 1950. Revealed preference and the utility function. Economica 17, 159–174.

Lewbel, A., 2001. Demand systems with and without errors. American Economic Review 91, 611–618.

Manski, C., 2003. Partial identiOfication of probability distributions. Springer-Verlag, New York.

Matzkin, R. L., 1991. Axioms of revealed preference for nonlinear choice sets. Econometrica 59, 1779–1786.

Polison, M., Quah, J., 2013. Revealed preference in a discrete consumption space. American Economic Journal: Microeconomics 5, 28–34.

Samuelson, P. A., 1938. A note on the pure theory of consumer's behavior. Economica 5, 61–71.

Samuelson, P. A., 1948. Consumption theory in terms of revealed preference. Economica 15, 243–253.

Selten, R., 1991. Properties of a measure of predictive success. Mathematical Social Sciences 21, 153–167.

Selten, R., Krischker, S., 1983. Aspiration Levels in Bargaining and Economic Decision Making. Springer, Ch. Comparison of two theories for characteristic function experiments.

Swofford, J. L., Whitney, G. A., 1987. Nonparametric tests of utility maximization and weak separability for consumption, leisure and money. The Review of Economics and Statistics 69, 458–464.

Swofford, J. L., Whitney, G. A., 1988. A comparison of nonparametric tests of weak separability for annual and quarterly data on consumption, leisure, and money. Journal of Business and Economic Statistics 6, 241–246.

Swofford, J. L., Whitney, G. A., 1994. A revealed preference test for weakly separable utility maximization with incomplete adjustment. Journal of Econometrics 60, 235–249.

Thaler, R., Shefrin, H. M., 1981. An economic theory of self-control. Journal of Political Economy 89, 392–406.

Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: A reference-dependent model. The Quarterly Journal of Economics 106, 1039–1061.

Varian, H. R., 1982. The nonparametric approach to demand analysis. Econometrica 50, 945–974.

Varian, H. R., 1983. Non-parametric tests of consumer behavior. The Review of Economic Studies 50, 99–110.

Varian, H. R., 1984. The nonparametric approach to production analysis. Econometrica 52, 579–597.

Varian, H. R., 1985. Non-parametric analysis of optimizing behavior with measurement error. Journal of Econometrics 30, 445–458.

Varian, H. R., 1988. Estimating risk aversion from Arrow-Debreu portfolio choice. Econometrica 56, 973–979.

Varian, H. R., 2006. Samuelson economics and the 21st century. Oxford University Press, Ch. Revealed Preference.

Vermeulen, F., 2005. And the winner is... an empirical evaluation of unitary and collective labour supply models. Empirical Economics 30, 711–734.

Warshall, S., 1962. A theorem of boolean matrices. Journal of the American Association of Computing Machinery 9, 11–12.