# Revealed Preference Tests for Weak Separability: an integer programming approach[*]

Laurens Cherchye[†] Thomas Demuynck[‡] Bram De Rock[§] Per Hjertstrand[¶]

**Abstract**

We focus on the revealed preference conditions that characterize the collection of finite data sets that are consistent with the maximization of a weakly separable utility function. From a theoretical perspective, we show that verifying these revealed preference conditions is a difficult problem, i.e. it is NP–complete. From a practical perspective, we present an integer programming approach that can verify the revealed preference conditions in a straightforward way, which is particularly attractive in view of empirical analysis. We demonstrate the versatility of this integer programming approach by showing that it also allows for testing homothetic separability and weak separability of the indirect utility function. We illustrate the practical usefulness of the approach by an empirical application to Spanish household consumption data. In this application we also include two statistical tests in which we account for measurement error.

*JEL Classification:* C14, C60, D01, D10.
*Keywords:* weak separability, revealed preference, integer programming

[*]We thank Ian Crawford for generously providing us the data of the Encuesta Continua de Presupestos Familiares (ECPF), which we use in our empirical study. We are also grateful to Adrian Fleissig, James Swofford and Gerald Whitney for helpful comments. This paper is a merger of two working papers, Cherchye, Demuynck, and De Rock (2011c) and Hjertstrand (2011).

[†]Center for Economic Studies, University of Leuven. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven-kortrijk.be. Laurens Cherchye gratefully acknowledges financial support from the Research Fund K.U.Leuven through the grant STRT1/08/004.

[‡]Center for Economic Studies, University of Leuven, E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: thomas.demuynck@kuleuven-kortrijk.be. Thomas Demuynck gratefully acknowledges the Fund for Scientific Research - Flanders (FWO-Vlaanderen) for his postdoctoral fellowship.

[§]ECARES-ECORE, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges the European Research Council (ERC) for his Starting Grant.

[¶]Research Institute of Industrial Economics (IFN), P.O. Box 55665, SE–102 15 Stockholm, Sweden. E-mail: per.hjertstrand@ifn.se. Per Hjertstrand gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation under research grant number W2009-0079:1

*The universe cannot be dealt with in one stroke and so a bit has to be broken off and treated as if the rest did not matter.*

Afriat, 1969

## 1 Introduction

We focus on the revealed preference conditions for consistency of a finite data set with the maximization of a weakly separable utility function. Our main contribution is twofold. First, we show that verification of these revealed preference conditions is a difficult problem. In particular, the problem is NP–complete, which essentially means that it cannot be solved in polynomial time. As we will discuss below, this actually motivates our second contribution. Specifically, we show that the revealed preference conditions can be verified by means of elementary integer programming procedures, which are easily implemented in practice. We demonstrate the versatility of this integer programming approach by showing that it can also assess homothetic separability and weak separability of the indirect utility function. Finally, we illustrate our approach by applying it to a Spanish panel data set. Here, we also consider extending our integer programming approach to account for measurement error in the data.

Weak separability of the utility function is a frequently used assumption in theoretical and applied demand analysis. A group of goods is said to be weakly separable if the marginal rate of substitution between any two goods in the group is independent from the quantities consumed of any good outside this group (Leontief, 1947; Sono, 1961). Weak separability has several convenient implications.[1] First of all, it allows for representing consumption in terms of two stage budgeting. This means that, in order to determine the demanded quantities of the goods in the separable group, it suffices to know the prices of the goods in this group and the total within–group expenditure. Further, weak separability is a crucial condition for the construction of group price and quantity indices. Such aggregates can be useful, for example, to compute group cost of living indices to be used in welfare analysis. Finally, from an empirical point of view, weak separability significantly reduces the number of parameters of the demand system to be estimated in practical applications.[2]

Considering these advantages for both theoretical and empirical work, an important issue concerns empirically testing the validity of the separability assumption (prior to effectively imposing it). In the literature, there are two approaches to test for weak separability. One approach uses econometric techniques to verify certain parameter restrictions given a specific demand model. Although this approach is fairly flexible in terms of the demand model that is used, it also poses a number of problems.

First, the separability restriction is often tested using Wald or likelihood ratio test procedures which require estimation of the full (unrestricted) demand model. Consequently, these tests may suffer from a degrees of freedom problem in the sense that too many parameters must be estimated given the amount of data.[3] Next, if the hypothesis of weak separability is rejected, it is impossible to verify whether this

[1] See also Deaton and Muellbauer (1980) for a more thorough discussion.

[2] In this respect, it is important to point out the importance of employing a correct separability structure in empirical demand modeling. On the one hand, using a too narrow structure (i.e. omitting goods that should be included in the separable grouping) leads to an omitted variables problem, which consequently produces inconsistent parameter estimates in the estimated demand model. On the other hand, including redundant goods in the separability structure may inflate the variances of the parameters, which may cause inefficient parameter estimates.

[3] The degrees of freedom problem could in theory be circumvented by instead using Lagrange multiplier tests. However, similar to Wald tests, Lagrange multiplier tests require a consistent estimate of the covariance matrix. Although it is relatively easy to obtain such estimates, these are often biased in small samples, implying that the Lagrange multiplier test may suffer from a small sample bias.

implies a rejection of weak separability as such or, instead, a rejection of the specific functional form imposed on demand a priori. In other words, if the null hypothesis of weak separability is rejected, this may well be due to the use of a wrong functional form rather than a non–separable utility structure per se.[4] Finally, most econometric tests for separability are based on separability of the indirect utility function (i.e. separability in prices), which does not imply separability of the direct utility function (i.e. separability in quantities).[5]

An alternative approach to test for weak separability is based on revealed preference theory. In several seminal contributions to the literature, Afriat (1969), Varian (1983) and Diewert and Parkan (1985) developed revealed preference conditions that characterize the collection of data sets that are rationalizable by a (weakly) separable utility function.[6] The revealed preference approach remedies different problems associated with the econometric approach. First, the revealed preference conditions can meaningfully be applied to data sets with as few as two observations, which avoids the degrees of freedom problem discussed above. Further, the revealed preference approach abstains from imposing a specific functional form on the utility functions. As such, the tests are insensitive to model misspecification. Finally, the revealed preference approach does not require additional assumptions like homotheticity of the subutility function or separability of the indirect utility function (although such additional assumptions can be imposed and tested; see below).

Unfortunately, the revealed preference conditions have the drawback that they take the form of a set of nonlinear, quadratic inequalities, which are very hard to verify. In order to avoid this problem, several heuristics have been proposed that provide separate sufficient and necessary conditions for data consistency with weak separability (see Section 2 for an overview). The lack of an efficient algorithm to verify the revealed preference conditions raises the question whether such an algorithm exists at all. In this study, we show that the answer is no. In particular, we prove that the verification of the revealed preference conditions for weak separability is an NP–complete problem.[7] This NP–completeness result implies that it is impossible to find a polynomial time algorithm that verifies whether a data set is consistent with the maximization of a weakly separable utility function (unless one can prove P = NP). This indicates that we should better look for a widely applied and (for moderately sized problems) reasonably quick non–polynomial time algorithm to verify the revealed preference conditions. Given this, we present an easy–to–implement (non–polynomial time) integer programming procedure to verify the revealed preference conditions. Our approach exploits the equivalence of the generalized axiom of revealed preference (GARP) and a set of mixed integer inequalities. Such an integer programming approach has proven very useful in the literature that applies revealed preference theory to collective consumption models, which studies the behavior of multi-person households, and in the literature that investigates the testable implications of general equilibrium models.[8] We extend the insights from this literature to

[4] Imposing separability conditions on a particular functional form might lead to additional difficulties. In particular, Blackorby, Primont, and Russell (1978) showed that testing for separability using several econometric specifications based on local approximations of the true model (i.e. flexible functional forms) is actually equivalent to testing a much stronger condition. For example, it turns out impossible to test separability for the translog model without imposing the much more stringent assumption of additive separability. Barnett and Choi (1989) confirmed this result by means of Monte Carlo simulations.

[5] A well known sufficient condition to obtain that direct and indirect separability coincide is that the subutility function is homothetic. We refer to Blackorby and Russel (1994) for more discussion.

[6] The revealed preference conditions for weak separability have been used in many different types of applications. See, for example, Swofford and Whitney (1987, 1988), Barnhart and Whitney (1988), Patterson (1991), Belongia and Chrystal (1991), Choi and Sosin (1992), Swofford and Whitney (1994), Jones and Mazzi (1996), Cox (1997), Fisher and Fleissig (1997), Rickertsen (1998), Spencer (2002), Fleissig and Whitney (2003, 2008), Swofford (2005), Serletis and Rangel-Ruiz (2005), Jones, Dutkowsky, and Elger (2005), Jha and Longjam (2006), Blundell, Browning, and Crawford (2007), Hjertstrand (2007, 2009), Elger, Jones, Edgerton, and Binner (2008), Elger and Jones (2008), and Drake and Fleissig (2008).

[7] We refer to Garey and Johnson (1979) for an introduction into the theory of NP–completeness.

[8] See Cherchye, De Rock, and Vermeulen (2007, 2009, 2011), Cherchye, De Rock, Sabbe, and Vermeulen (2008), and Cher-

the model of utility maximization with a weakly separable utility function.

From a theoretical point of view, the core motivation for adopting an integer programming approach is that this is a widely accepted and a well known approach to handle NP–complete problems. Besides this, we also have a number of other motivations. First of all, our approach can be applied to data sets with any number of observations. Second, any mixed integer program can be solved in finite time. Hence, our approach implies the possibility to verify in finite time the necessary and sufficient conditions for a given data set to be consistent with maximization of a weakly separable utility function. A third important argument pro our integer programming approach is that it provides a versatile framework for analyzing testable implications of different model specifications: we will show that our approach can easily accommodate for homotheticity of the subutility functions, and that we can readily adjust our integer programming procedure to test for separability of the indirect utility function. Finally, we show how our approach can be used to design simple statistical tests for weak separability that account for measurement error in the data.

We demonstrate the practical usefulness of our approach by applying it to data drawn from the Encuesta Continua de Presupestos Familiares (ECPF), a Spanish household survey. In this application we first investigate the performance of our integer programming formulation. We do this by comparing it to Varian's three step procedure, which provides separate necessary and sufficient conditions for weak separability (see below for more details). We also study the computational performance of the integer programming formulation. Secondly, we compare the empirical fit of the four alternative model specifications mentioned above: the standard utility maximization model, the model that additionally imposes weak separability, the homothetic separability model, and the model that assumes a weakly separable indirect utility function. Specifically, following a recent proposal of Beatty and Crawford (2011), we evaluate these different model specifications in terms of their 'predictive success'. In our final exercise, we introduce two statistical tests that allow us to take the possibility of measurement error into account.

Section 2 introduces the revealed preference conditions for rationalizability under a weakly separable utility function and presents our NP–completeness result. Section 3 presents our integer programming approach. Section 4 discusses our empirical application. Section 5 concludes.

## 2 Revealed preferences conditions

To set the stage, we briefly recapture the known revealed preference conditions for the standard utility maximization model and for the model that additionally imposes weak separability on the utility function. These results will be useful for our discussion in the following sections. In this section, we also state our NP–completeness result.

**Standard utility maximization.** Consider a finite data set $D = \{\mathbf{p}_t; \mathbf{x}_t\}_{t\in T}$, which consists of strictly positive price vectors $\mathbf{p}_t \in \mathbb{R}^n_{++}$ and nonnegative consumption bundles $\mathbf{x}_t \in \mathbb{R}^n_+$ for consumption observations $t$ in a (finite) set $T$. This data set $D$ is said to be *rationalizable* if there exists a well–behaved (i.e. increasing, continuous and concave) utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ such that, for all observations $t \in T$,

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

In other words, for each observation $t$ it must be the case that the consumed bundle $\mathbf{x}_t$ maximizes the utility function $u$ over the set of all affordable consumption bundles.

chye, Demuynck, and De Rock (2011b) for integer programming characterizations of household consumption models and Cherchye, Demuynck, and De Rock (2011d) for integer programming characterizations of general equilibrium models.

Next, consider the following concepts. The direct revealed preference relation $R^D$ over the set $\{\mathbf{x}_t\}_{t\in T}$ is defined by $\mathbf{x}_t R^D \mathbf{x}_v$ if $\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}_v$. In words, we have that $\mathbf{x}_t R^D \mathbf{x}_v$ if $\mathbf{x}_t$ was chosen while $\mathbf{x}_v$ was also affordable. The indirect revealed preference relation $R$ is the transitive closure of the relation $R^D$; $\mathbf{x}_t R \mathbf{x}_v$ if there exist bundles $\mathbf{x}_w, \mathbf{x}_r, \ldots, \mathbf{x}_m$ such that $\mathbf{x}_t R^D \mathbf{x}_w$, $\mathbf{x}_w R^D \mathbf{x}_r, \ldots, \mathbf{x}_m R^D \mathbf{x}_v$. Finally, we say that $\{\mathbf{p}_t, \mathbf{x}_t\}_{t\in T}$ satisfies the Generalized Axiom of Revealed Preferences (GARP) if for all $\mathbf{x}_t R \mathbf{x}_v$ it is not the case that $\mathbf{p}_v\mathbf{x}_v > \mathbf{p}_v\mathbf{x}_t$. In words, if $\mathbf{x}_t$ is indirectly revealed preferred to $\mathbf{x}_v$, then it is not the case that $\mathbf{x}_v$ was more expensive than $\mathbf{x}_t$ when $\mathbf{x}_v$ was bought.

Using these concepts, we can state the following result, which is probably the single most important theorem in revealed preference theory.

**Theorem 1.** *[Varian (1982), based on Afriat (1967)]*
*The following statements are equivalent:*

*(i) The data set $D = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in T}$ is rationalizable,*

*(ii) The data set $D = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in T}$ satisfies* GARP,

*(iii) There exist strictly positive numbers $\lambda_t$ and numbers $U_t$ such that, for all $t, v \in T$,*

$$U_t - U_v \leq \lambda_v \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v).$$

*(iv) There exist numbers $u_t$ such that for all $t, v \in T$*

$$\begin{aligned} &\text{if } u_v \geq u_t, \quad \text{then } \mathbf{p}_t\mathbf{x}_t \leq \mathbf{p}_t\mathbf{x}_v, \\ &\text{if } u_v > u_t, \quad \text{then } \mathbf{p}_t\mathbf{x}_t < \mathbf{p}_t\mathbf{x}_v. \end{aligned}$$

Condition (ii) states that GARP is necessary and sufficient for rationalizability. Condition (iii) provides an equivalent characterization of utility maximization in terms of so–called Afriat inequalities. Intuitively, these Afriat inequalities allow us to obtain an explicit construction of the utility levels and the marginal utility of income associated with each observation $t$: they define a utility level $U_t$ and a marginal utility of income $\lambda_t$ (associated with the observed income $\mathbf{p}_t\mathbf{x}_t$) for each observed $\mathbf{x}_t$. Condition (iv) is a reformulation of GARP in the way it is usually presented in the closely related nonparametric production literature (see Varian (1984)); in this setting this formulation is known as the 'strong axiom of cost minimization'.[9] The basic idea behind this condition is very simple: if the utility at observation $v$ exceeds the utility at observation $t$ (i.e. $u_v \geq (>) u_t$), then it is not the case that $\mathbf{x}_t$ was more expensive than $\mathbf{x}_v$ when $\mathbf{x}_t$ was bought (i.e. $\mathbf{p}_t\mathbf{x}_t \leq (<)\mathbf{p}_t\mathbf{x}_v$). Otherwise, the rational individual would not be utility maximizing at $t$, because (s)he could also afford the preferred bundle $\mathbf{x}_v$.

Theorem 1 provides three methods to verify whether a data set is rationalizable. The first method was originally suggested by Varian (1982) and focuses on verifying the GARP condition. The method consists of three steps, which comply with the three steps in the definition of GARP. The first step constructs the relation $R^D$ from the data set $D = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in T}$. A second step computes the transitive closure of $R$. Here, Varian suggests using Warshall (1962)'s algorithm, which provides an efficient procedure for computing transitive closures. The third step verifies if $\mathbf{p}_v\mathbf{x}_v \leq \mathbf{p}_v\mathbf{x}_t$ whenever $\mathbf{x}_t R \mathbf{x}_v$. If this is the case, the data set satisfies GARP and is, therefore, rationalizable. Due to its efficiency, this procedure is very popular in applied work. The second and third method verifies the rationalizability conditions by testing feasibility of either the Afriat inequalities in condition (iii) or the inequalities in condition

[9] This condition is related to the notion of semi–strict quasi–concavity, see Hjertstrand (2008).

(iv). The Afriat inequalities are linear in the unknowns $U_t$ and $\lambda_t$, which implies that their feasibility can be verified using elementary linear programming methods (see Afriat (1967) and Diewert (1973) for discussions of this method). In a similar vein, feasibility of the inequalities in condition (iv) can be checked by solving a linear programming problem (in the unknowns $u_t$) applied to the contrapositive statement of this condition.

**Weak separability.** To introduce the notion of weak separability, we first partition the set of goods $N = \{1, \ldots, n\}$ in two groups. Accordingly, we can split any given consumption bundle into two separate bundles. The first bundle $\mathbf{x}$ contains all consumption quantities of the goods from the first group and the second bundle $\mathbf{y}$ captures the remaining goods. We denote the full consumption bundle as $(\mathbf{x}, \mathbf{y})$. Likewise, we can split any price vector into a price vector of all goods in the first group $\mathbf{p}$ and a vector of prices for the goods in the second group $\mathbf{q}$. Now, consider a data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$. We say that this data set is *rationalizable by a weakly separable utility function* if there exists a well–behaved utility function $u$ and a well–behaved subutility function $s$ such that, for all observations $t \in T$,

$$(\mathbf{x}_t, \mathbf{y}_t) \in \arg\max_{\mathbf{x},\mathbf{y}} u(\mathbf{x}, s(\mathbf{y})) \qquad \text{s.t. } \mathbf{p}_t\mathbf{x} + \mathbf{q}_t\mathbf{y} \leq \mathbf{p}_t\mathbf{x}_t + \mathbf{q}_t\mathbf{y}_t.$$

Varian (1983) provides the following characterization of behavior that is rationalizable by a weakly separable utility function.

**Theorem 2.** *[Varian (1983)]*
*The following statements are equivalent:*

*(i) The data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ is rationalizable by a weakly separable utility function.*

*(ii) For all $t \in T$ there exist nonnegative numbers $S_t$ and strictly positive numbers $\delta_t$ such that, for all $t$, $v \in T$,*

$$S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v), \tag{ii.1}$$
$$\{\mathbf{p}_t, 1/\delta_t; \mathbf{x}_t, S_t\}_{t \in T} \textit{satisfies } \text{GARP}. \tag{ii.2}$$

*(iii) For all $t \in T$, there exist nonnegative numbers $S_t$ and $U_t$ and strictly positive numbers $\delta_t$ and $\lambda_t$ such that, for all $t, v \in T$,*

$$S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v), \tag{iii.1}$$
$$U_t - U_v \leq \lambda_v \left[ \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) + \frac{1}{\delta_v}(S_t - S_v) \right]. \tag{iii.2}$$

*(iv) For all $t \in T$, there exist numbers $S_t$ and $u_t$ and strictly positive numbers $\delta_t$ such that, for all $t, v \in T$,*

$$S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v), \tag{iv.1}$$
$$\textit{if } u_v \geq u_t, \textit{ then } \mathbf{p}_t\mathbf{x}_t + \frac{1}{\delta_t}S_t \leq \mathbf{p}_t\mathbf{x}_v + \frac{1}{\delta_t}S_v, \tag{iv.2}$$
$$\textit{if } u_v > u_t, \textit{ then } \mathbf{p}_t\mathbf{x}_t + \frac{1}{\delta_t}S_t < \mathbf{p}_t\mathbf{x}_v + \frac{1}{\delta_t}S_v. \tag{iv.3}$$

In contrast to the conditions in Theorem 1, the conditions in this theorem are not easily verified. The main problem is that, when checking (ii.2), the 'prices' $1/\delta_t$ and the corresponding 'quantities' $S_t$, which must satisfy condition (ii.1), are unobserved. This is also reflected in condition (iii.2), which is a set of quadratic inequalities.

The literature brings forward several methods to test the weak separability conditions. Probably the best known alternative is Varian (1983)'s three step procedure. In the first step, this method tests GARP consistency of the data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$. If the data fail GARP, they are not rationalizable and, hence, we can reject weak separability.[10] By contrast, if the data set satisfies GARP, the second step tests whether the data set $\{\mathbf{q}_t, \mathbf{y}_t\}_{t \in T}$ satisfies GARP. This GARP condition is equivalent to condition (ii.1). If GARP consistency is rejected in this second step then, again, the data set is not rationalizable by weak separability. Finally, the third step verifies GARP of a data set $\{\mathbf{p}_t, 1/\delta_t^*; \mathbf{x}_t, V_t^*\}_{t \in T}$ for some specific values $\delta_t^*$ and $S_t^*$ that satisfy condition (ii.1). If for this last step GARP is not rejected, then we conclude that the data are consistent with weak separability.

Unfortunately, Varian (1983)'s test is not an exact one. In particular, it is possible that a data set is rationalizable by a weakly separable utility function while the algorithm does not reach this conclusion. Simulation results indicate that this may actually occur quite frequently; see, for example, Barnett and Choi (1989), Fleissig and Whitney (2003), Hjertstrand (2009) and our empirical application in Section 4. The problem is that the third step of the procedure fixes the values of both $\delta_t^*$ and $V_t^*$ in an arbitrary way. In this respect, however, certain values may be more probable than others. This idea provides the intuition behind the linear program developed by Fleissig and Whitney (2003). In particular, these authors determine the values of $1/\delta_t^*$ and $V_t^*$ based on the theory of superlative index numbers (see Diewert (1976, 1978)). A superlative index number provides an exact index number for some order approximation of the underlying (in casu homogeneous) utility function $s$. However, this test is again only sufficient but not necessary for weak separability to hold.

An alternative testing strategy is explored by Swofford and Whitney (1994), Elger and Jones (2008) and Fleissig and Whitney (2008), who use nonlinear programming methods to solve (iii.1) and (iii.2) simultaneously. This is done by reformulating the problem as a nonlinear minimization problem subject to a number of linear and nonlinear restrictions. The data set is then rationalizable if the global optimal solution of this problem is equal to zero. Alas, nonlinear programming problems (with nonlinear restrictions) become computationally burdensome even for moderate sized problems.[11] A second problem with such programming problems is that they do not always yield an optimal solution: most algorithms search for local optima, which need not be globally optimal (unless some additional concavity assumptions are true). Generally, finding a global optimum requires a fine grid search over the set of initial values. But even a very fine grid search cannot exclude that weak separability is rejected while the assumption effectively holds. We refer to Hjertstrand (2009) for a Monte Carlo comparison of the different test procedures cited in this paragraph.

**An NP–completeness result.** Consider a data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$. For any such data set, we can ask the question whether this data set is rationalizable by a weakly separable utility function (i.e. whether it satisfies the conditions of Theorem 2). Basically, this decision problem asks for testing rationalizability by a weakly separable utility function for an arbitrary data set. The following theorem shows that this problem is NP–complete. The proof is given in Appendix A.

[10] If the data set fails to satisfy GARP, then it is not rationalizable by any utility function, whether it is weakly separable or not.

[11] See Jones, McCloud, and Edgerton (2007) and Hjertstrand (2009) for more discussion on this. As an example, an empirical analyst handling a data set of 60 observations would have to solve an optimization problem with at least 3540 linear and 3540 nonlinear constraints to verify the conditions (iii.1) and (iii.2) in Theorem 2.

**Theorem 3.** *The question whether a given data set is rationalizable by a weakly separable utility function is an* NP*–complete problem.*

This result considers the general case without any restriction on the number of goods or observations. Of course, it does not rule out specific instances for which verification of the rationalizability conditions might be performed efficiently. Nevertheless, our result does indicate that it is highly improbable that the problem of rationalizability by a weakly separable utility function can be solved by means of an efficient algorithm (like, for example, linear programming).

Essentially, Theorem 3 implies that one should not waste time trying to construct a polynomial time algorithm that verifies the conditions in Theorem 2 (unless one has taken up the ambitious task of showing that P = NP). In turn, this suggests considering easy–to–implement non–polynomial time algorithms for tackling the testing problem. Therefore, we next propose a widely used method called Mixed Integer Programming (MIP).

# 3 The mixed integer program

MIP problems look like standard linear programming problems except that certain variables are restricted to be integer valued (in our case either 0 or 1). The MIP formulation has a number of advantages. First of all, a MIP always gives a result in finite time and, moreover, every local solution of a MIP is in fact a global solution. This last property directly addresses the issue of finding a global optimum which we argued is a problem for any of the nonlinear approaches discussed above. Second, as we will show below, our MIP formulation is a joint test of the necessary and sufficient conditions. This means that such a MIP test is theoretically unbiased, and therefore, will by definition always outperform any sequential procedure for implementing the weak rationalizability conditions from Theorem 2. Third, MIP problems are a frequently used and widely accepted approach to handle NP–complete problems. As such, there exist well performing software programs to solve such problems. Finally, we demonstrate the flexibility of our approach by deriving MIP conditions for two related rationalizability problems. First, we consider the specific case where the subutility function is homothetic. Next, we focus on the case where separability is imposed on the indirect utility function, i.e. the case of weak separability in prices. These two cases are particularly interesting because they are widely used in econometric analyses involving separability concepts (see also our discussion in the Introduction).

We proceed by translating conditions (iv.1)–(iv.3) to an integer programming setting. The basic idea is to notice that conditions (iv.2)–(iv.3) are equivalent to the following set of conditions:

$$\begin{aligned} &\text{if } u_v \geq u_t, \text{ then } \delta_t \mathbf{p}_t \mathbf{x}_t + S_t \leq \delta_t \mathbf{p}_t \mathbf{x}_v + S_v, \\ &\text{if } u_v > u_t, \text{ then } \delta_t \mathbf{p}_t \mathbf{x}_t + S_t < \delta_t \mathbf{p}_t \mathbf{x}_v + S_v. \end{aligned}$$

This equivalence follows from multiplying both sides of the right hands side inequalities by $\delta_t$ $(> 0)$. As such, we see that the inequalities on both the right and left hand side become linear. We make use of binary variables to capture the logical relation between the different inequalities. This leads to the following mixed integer linear program.

**CS.WS** For all $t, v \in T$, there exist numbers $S_t, u_t \in [0, 1[$, $\delta_t \in ]0, 1]$ and binary variables $X_{t,v} \in \{0, 1\}$

such that, for all observations $t$ and $v \in T$, [12]

$$S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v), \tag{cs.1}$$

$$u_t - u_v < X_{t,v}, \tag{cs.2}$$

$$(X_{t,v} - 1) \leq u_t - u_v, \tag{cs.3}$$

$$\delta_t \mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) < X_{t,v} A_t, \tag{cs.4}$$

$$(X_{t,v} - 1)A_v \leq \delta_v \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v). \tag{cs.5}$$

Here, we let $A_t$ be some fixed and large number (larger than $\mathbf{p}_t\mathbf{x}_t + 1$). First of all, observe that the restriction of $S_t$, $u_t$ and $\delta_t$ to the unit interval is harmless as it is possible to rescale these variables without changing the revealed preference conditions (iv.1)–(iv.3). Condition (cs.1) reproduces condition (iv.1). The interpretation behind the binary variables is that $X_{t,v}$ should be equal to one if and only if $u_t \geq u_v$. This requirement is formalized by conditions (cs.2) and (cs.3). Finally, conditions (cs.4) and (cs.5) reformulate conditions (iv.2) and (iv.3) by making use of these binary variables. The following theorem formalizes the equivalence between the above MIP conditions and the rationalizability conditions for weak separability in Theorem 2. The proof is given in Appendix B.

**Theorem 4.** *The data set $D = \{\delta_t \mathbf{p}_t, 1; \mathbf{x}_t, S_t\}_{t \in T}$ satisfies* (iv.2)–(iv.3) *if and only if conditions* (cs.2)–(cs.5) *have a solution.*

**Homothetic and indirect weak separability.** The above MIP formulation is very flexible in terms of incorporating additional (separable) preference structure. We illustrate this by considering two special cases. The first case requires that the subutility function $s$ is homothetic. The second case requires separability of the indirect utility function.

A data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ is *rationalizable by homothetic separability* if there exist a well–behaved utility function $u$ and a well–behaved and homothetic subutility function $s$ such that, for all observations $t \in T$,

$$(\mathbf{x}_t, \mathbf{y}_t) \in \arg\max_{\mathbf{x},\mathbf{y}} u(\mathbf{x}, s(\mathbf{y})) \qquad \text{s.t. } \mathbf{p}_t\mathbf{x} + \mathbf{q}_t\mathbf{y} \leq \mathbf{p}_t\mathbf{x}_t + \mathbf{q}_t\mathbf{y}_t.$$

The following theorem characterizes data sets that are consistent with homothetic separability. The result directly follows from combining Varian (1983)'s rationalizability conditions for a homothetic utility function with Theorem 2.

**Theorem 5.** *The following statements are equivalent:*

*(i) The data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ is rationalizable by homothetic separability.*

*(ii) For all $t \in T$ there exist nonnegative numbers $U_t$ and strictly positive numbers $S_t$ such that, for all $t, v \in T$,*

$$S_t - S_v \leq \frac{S_v}{\mathbf{q}_v\mathbf{y}_v}\mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v),$$

$$\left\{\mathbf{p}_t, \frac{\mathbf{q}_t\mathbf{y}_t}{S_t}; \mathbf{x}_t, \mathbf{y}_t\right\}_{t \in T} \textit{ satisfies } \textsc{garp}.$$

[12] The strict inequalities in cs.2 and cs.4 are difficult to handle. Therefore, in practice, we use a weak inequality and subtract a very small but fixed number from the right hand side.

*(iii) For all $t \in T$, there exist nonnegative $U_t$ and strictly positive numbers $S_t$ and $\lambda_t$ such that, for all $t, v \in T$,*

$$S_t - S_v \leq \frac{S_v}{\mathbf{q}_v \mathbf{y}_v} \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v),$$
$$U_t - U_v \leq \lambda_v \left[ \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) + \frac{\mathbf{q}_v \mathbf{y}_v}{S_v}(S_t - S_v) \right].$$

*(iv) For all $t \in T$, there exist numbers $u_t$ and strictly positive numbers $S_t$ such that, for all $t, v \in T$,*

$$S_t - S_v \leq \frac{S_v}{\mathbf{q}_v \mathbf{y}_v} \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v), \tag{iv.1}$$
$$\textit{if } u_v \geq u_t, \textit{ then } \mathbf{p}_t \mathbf{x}_t + \frac{\mathbf{q}_t \mathbf{y}_t}{S_t} S_t \leq \mathbf{p}_t \mathbf{x}_v + \frac{\mathbf{q}_t \mathbf{y}_t}{S_t} S_v, \tag{iv.2}$$
$$\textit{if } u_v > u_t, \textit{ then } \mathbf{p}_t \mathbf{x}_t + \frac{\mathbf{q}_t \mathbf{y}_t}{S_t} S_t < \mathbf{p}_t \mathbf{x}_v + \frac{\mathbf{q}_t \mathbf{y}_t}{S_t} S_v. \tag{iv.3}$$

In other words, to impose homotheticity of the subutility function, we only need to add the additional (linear) restriction that $\delta_t = S_t/\mathbf{q}_t\mathbf{y}_t$ to the earlier weak separability conditions. As such, by substituting in the MIP problem **CS.WS** each occurrence of $\delta_t$ by $S_t/\mathbf{q}_t\mathbf{y}_t$ (or by imposing the additional restriction that $\delta_t = S_t/\mathbf{q}_t\mathbf{y}_t$), we obtain a MIP formulation of the necessary and sufficient conditions for homothetic separability. In view of our following empirical application, it is also worth noting that Theorem 5 implies two necessary conditions for the data to be rationalized by homothetic separability. More precisely, the whole dataset $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ needs to satisfy GARP and, secondly, the data $\{\mathbf{q}_t, \mathbf{y}_t\}_{t \in T}$ also needs to satisfy the homothetic axiom of revealed preference (HARP); see Varian (1983) for a detailed discussion of HARP.

As a final result we state the revealed preference conditions for indirect weak separability. First of all, let us normalize the prices $\mathbf{p}_t$ and $\mathbf{q}_t$ such that, for all $t$, $\mathbf{p}_t\mathbf{x}_t + \mathbf{q}_t\mathbf{y}_t = 1$. Then, we say that the data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ is *rationalizable by indirect weak separability* if there exist a well-behaved (i.e. decreasing, convex and continuous) indirect utility function $v$ and a well–behaved indirect subutility function $w$ such that, for all observations $t \in T$,

$$\{\mathbf{p}_t, \mathbf{q}_t\} \in \arg\min v(\mathbf{p}, w(\mathbf{q})) \qquad \text{s.t. } \mathbf{p}\mathbf{x}_t + \mathbf{q}\mathbf{y}_t \leq 1. \tag{1}$$

The next theorem gives a characterization of data sets that are rationalizable in terms of an indirect weakly separable utility function. The result is obtained by combining the result in Theorem 2 with Brown and Shannon (2000) 's rationalizability conditions for an indirect utility function. See also Hjertstrand and Swofford (2012) for a similar result.

**Theorem 6.** *The following statements are equivalent:*

*(i) The data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ is rationalizable by indirect weak separability.*

*(ii) For all $t \in T$ there exist numbers $V_t$ and $W_t$ and strict positive numbers $\lambda_t$ and $\delta_t$ such that, for all $t, v \in T$,*

$$W_t - W_v \geq -\delta_v \mathbf{y}_v(\mathbf{q}_t - \mathbf{q}_v), \tag{v.1}$$
$$V_t - V_v \geq -\lambda_v \left( \mathbf{x}_v(\mathbf{p}_t - \mathbf{p}_v) + \frac{1}{-\delta_v}(W_t - W_v) \right). \tag{v.2}$$

If we introduce the variables $S_t = -W_t$ and $U_t = -V_t$, we can reformulate (v.1)-(v.2) as:

$$S_t - S_v \leq \delta_v \mathbf{y}_v(\mathbf{q}_t - \mathbf{q}_v), \tag{v.1}$$

$$U_t - U_v \leq \lambda_v \left( \mathbf{x}_v(\mathbf{p}_t - \mathbf{p}_v) + \frac{1}{\delta_v}(S_t - S_v) \right). \tag{v.2}$$

Observe that the conditions in this theorem are formally equivalent to the conditions (iii.1)–(iii.2) in Theorem 2 with prices and quantities interchanged. Thus two necessary conditions for the data to be rationalized by indirect separability are that the data sets $\{\mathbf{x}_t, \mathbf{y}_t; \mathbf{p}_t, \mathbf{q}_t\}_{t \in T}$ and $\{\mathbf{y}_t, \mathbf{q}_t\}_{t \in T}$ both satisfy GARP. Finally, from (v.1)–(v.2) and by a direct application of Theorem 4, we can show that the rationalizability conditions in Theorem 6 are equivalent to the following set of MIP constraints:

**CS.WSI** There exist numbers $S_t, u_v \in [0, 1[$, $\delta_t \in ]0, 1]$ and binary variables $X_{t,v} \in \{0, 1\}$ such that, for all $t, v \in T$,

$$S_t - S_v \leq \delta_v \mathbf{y}_v(\mathbf{q}_t - \mathbf{q}_v), \tag{csi.1}$$

$$u_t - u_v < X_{t,v}, \tag{csi.2}$$

$$(X_{t,v} - 1) \leq u_t - u_v, \tag{csi.3}$$

$$\delta_t \mathbf{x}_t(\mathbf{p}_t - \mathbf{p}_v) + (S_t - S_v) < X_{t,v} A_t, \tag{csi.4}$$

$$(X_{t,v} - 1)A_v \leq \delta_v \mathbf{x}_v(\mathbf{p}_t - \mathbf{p}_v) + (S_t - S_v). \tag{csi.5}$$

Again, $A_t$ is a fixed number larger than $\mathbf{p}_t \mathbf{x}_t + 1$.

# 4 Empirical Application

We apply our integer programming tests to data drawn from the Encuesta Contunua de Presopuestos Familieares (ECPF) Survey. The ECPF is a quarterly budget survey (1985–1997) that interviews about 3200 Spanish households on their consumption expenditures. For each household, the survey provides consumption observations for a maximum of eight consecutive quarters. See Browning and Collado (2001) and Crawford (2010) for a more detailed explanation of this data set. We exclude all households with less than eight observations. In the end, this obtains a panel with 1585 households. The data set covers consumption decisions for 15 (nondurable) goods: (i) food and non-alcoholic drinks at home, (ii) alcohol, (iii) tobacco, (iv) energy at home, (v) services at home, (vi) nondurables at home, (vii) nondurable medicines, (viii) medical services, (ix) transportation, (x) petrol, (xi) leisure, (xii) personal services, (xiii) personal nondurables, (xiv) restaurant and bars and (xv) traveling holiday. We follow Blundell, Browning, and Crawford (2007) and define the separable group to include all goods except food (i.e. the separable group contains all goods except (i), (ii) and (xiv)). This separability assumption is frequently used in empirical analysis of consumption behavior.

## 4.1 Evaluating the integer programming method

To assess the performance of our integer programming method, we will compare it with Varian's three step procedure. Next, another main focus is on evaluating the computational speed of our integer programming approach for substantially large data sets. To this end, we will consider a preference homogeneity assumption that parallels an assumption often used in econometric demand analysis. This will allow us to conduct our separability tests on data sets that bring together information on multiple (similar) households.

**Comparison with Varian's three step procedure.** Our MIP formulation provides exact conditions for rationalizability by a weakly separable utility function. To evaluate the practical usefulness of our MIP procedure, it is interesting the associated test results with the ones generated by the frequently used three step procedure of Varian (1983). As discussed in Section 2, the first two steps of Varian's procedure imply necessary conditions for rationalizability by a weakly separable utility function. The first step verifies GARP consistency of the data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ and the second step verifies GARP consistency of the data set $\{\mathbf{q}_t, \mathbf{y}_t\}_{t \in T}$. Both steps are very easy to implement. Actually, if this two step procedure can identify almost all non–rationalizable data sets (i.e. the test has a low type II error), then this may plead for using this (efficiently implementable) procedure instead of our computationally more demanding (necessary and sufficient) MIP procedure.

We find that 83% of our Spanish households (1323 out of 1585) meet the the two GARP conditions of Varian's two step procedure. By contrast, only 54% of the households (853 out of 1585) satisfy our MIP conditions. In other words, 64% of the households that are not rationalizable by a weakly separable utility function still satisfy Varian's necessary conditions. In our opinion, this difference between the two test procedures is rather significant.

Let us now turn to Varian's sufficient conditions for weak separability. These conditions add one step to the above two step procedure. Like before, it first verifies GARP consistency of the data sets $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$ and $\{\mathbf{q}_t, \mathbf{y}_t\}_{t \in T}$. Subsequently, it verifies GARP consistency of a data set $\{\mathbf{p}_t, 1/\delta_t^*; \mathbf{x}_t, S_t^*\}_{t \in T}$ for some specific values $\delta_t^*$ and $S_t^*$ that satisfy condition (ii.1). If in this last step GARP is not rejected, then we conclude that the data set is rationalizable by a weakly separable utility function. We find that 40% of all households (636 out of 1585) pass this three step test. Thus, when comparing to our test results for the MIP conditions, 25% of all households that are effectively consistent with weak separability do not satisfy Varian's sufficient conditions. Once more, we find this difference quite big.

As a final exercise, we consider the 'adjusted' version of Varian's sufficiency test that was introduced by Fleissig and Whitney (2003). As discussed in Section 2, these authors use the theory of superlative index numbers to define the Afriat numbers in the last step of Varian's three step procedure. We find that 50% of the households (794 out of 1584) satisfy the resulting sufficient conditions for weak separability.[13] Thus, the difference with our MIP test results decreases quite substantially. However, we still have that about 6% of the households that are consistent with weak separability do not pass this adjusted three step test.

At a more general level, we believe that these exercises also demonstrate that our exact MIP conditions for weak separability can be fruitfully applied to assess (and compare) the empirical performance of tests for weak separability that are not exact but very efficiently implementable. For example, for our data set we conclude that Varian's procedures generate test results that are considerably different from our MIP results, while Fleissig and Withney's procedure delivers much more similar (and thus 'better') results.

**Computational speed.** The above empirical application considers data sets with (only) 8 observations. Not very surprisingly, for such small data sets the MIP method we propose comes to a test result very rapidly. Here, it seems interesting to assess whether this 'speediness' also holds if we increase the size of the data sets. As is well known, MIP problems might become increasingly hard to solve as the size of the problem gets larger. To assess whether our MIP method also works well for substantially large data sets, we assume identical preferences for all households with the same age of the male and female household members. In practice, this means that we perform our separability tests on pooled data sets

[13] We applied the test using the chain–linked Fisher ideal quantity index. We performed robustness exercises using other popular superlative index numbers but the results did not vary very much.

Table 1: Computational speed

| size of data set | number of datasets | pass rate | (mean) time (in seconds) | var time |
|---|---|---|---|---|
| 8 | 152 | 84 | 0.023 | 0.00001 |
| 16 | 87 | 5 | 0.099 | 0.0002 |
| 24 | 49 | 1 | 0.256 | 0.001 |
| 32 | 38 | 0 | 0.559 | 0.009 |
| 40 | 48 | 0 | 1.125 | 0.021 |
| 48 | 22 | 0 | 1.911 | 0.067 |
| 56 | 22 | 0 | 3.223 | 0.098 |
| 64 | 14 | 0 | 5.206 | 0.144 |
| 72 | 7 | 0 | 7.74 | 1.256 |
| 80 | 9 | 0 | 12.042 | 1.835 |
| 88 | 6 | 0 | 15.319 | 1.223 |
| 96 | 5 | 0 | 21.644 | 5.367 |
| 104 | 1 | 0 | 39.267 | N.A. |
| 120 | 2 | 0 | 50.543 | 13.256 |

containing all households with equally aged household members. A similar homogeneity assumption is frequently used in econometric demand analysis, i.e. demand estimation is often conditioned on ages of the household members as demographic factors.

As can be seen from Table 1, the size of our newly constructed data sets varies from 8 to 120 observations, with the average number of observations equal to 27.44. Clearly, this implies relatively big data sets as compared to other data sets that have been considered in empirical revealed preference analysis.[14]

The third column of Table 1 reports the pass rates for the data sets of different sizes. However, our main interest here is in the fourth column of the table, which gives the average computation time of our algorithm for the different data set sizes that we consider.[15] Generally, these results provide a fairly strong case in favor of our MIP approach. For example, checking the revealed preference conditions for weak separability takes (on average) less than a second for data sets with up to 32 observations. However, if we keep increasing the number of observations, the computational time increases rapidly. Nevertheless, even for the largest data sets with 120 observations we obtain an outcome in less than a minute (on average), which —in our opinion— is still reasonably fast.

### 4.2 Comparing alternative behavioral models

In this section we consider the four models introduced in Sections 2 and 3: the standard utility maximization model, the model that additionally imposes weak separability, the homothetic separability model, and the model that assumes a weakly separable indirect utility function. We will evaluate the different models in terms of their pass rates, discriminatory power and predictive success.

**Test results: pass rates and power.** Table 2 reports the pass rates of the four revealed preference tests. About 91% of the households (1445 out of 1585) satisfy the revealed preference conditions for the stan-

[14]See, for example, Cherchye, De Rock, Sabbe, and Vermeulen (2008) for a discussion on the typical size of data sets considered in empirical revealed preference analysis.

[15]We performed all our computations on a laptop computer with 2.4GHz clock speed and 4GB RAM with a standard configuration. For solving the integer programming problem, we used the commercial solver CPLEX©.

dard utility maximization model (i.e. the conditions in Theorem 1). By contrast, only 853 households (or approximately 54%) satisfy the revealed preference conditions for rationalizability by weak separability (as given by Theorem 2). Remarkably, none (!) of the households satisfy the conditions for rationalizability by homothetic separability (see Theorem 5).[16] This already indicates that weak separability and, to a much greater extent, homothetic separability are rather stringent assumptions. Finally, 1264 households (or approximately 80%) pass the rationalizability conditions for indirect weak separability (see Theorem 6), which is substantially more than for the other separability assumptions.

Our diverging results for weak separability and indirect weak separability can seem surprising to some, as one may have expected these two assumptions to be about equally stringent. Still, our pass rate results suggest that the latter assumption has a better empirical fit than the former one for our sample of households. In a sense, this may be a useful result from the perspective of econometric applications, which often invoke indirect weak separability (see our discussion in the Introduction). Our results reveal that observed behavior is largely consistent with such indirect separability.

Table 2: Pass rate and Power (in percentages)

| model | pass rate | power | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | mean | min | 1st quartile | median | 3rd quartile | max |
| general ut. max. | 91.17 | 11.13 | 0.0 | 0.2 | 6.5 | 19.9 | 68.4 |
| weak separability | 53.82 | 47.82 | 4.3 | 31.3 | 48.5 | 64.1 | 98.9 |
| homothetic separability | 0 | 99.99 | 99.9 | 100 | 100 | 100 | 100 |
| indirect separability | 79.75 | 15.99 | 0.0 | 0.9 | 12.8 | 27.3 | 80.9 |

Importantly, to meaningfully compare the different models, one should not merely consider the corresponding pass rates. For example, as the weak separability model is nested within the standard utility maximization model, the former model will have a lower pass rate than the latter model by construction. Indeed, Bronars (1987) and, more recently, Andreoni and Harbaugh (2008) and Beatty and Crawford (2011) —rather convincingly— argue that revealed preference test results (indicating pass or fail of the data for some behavioral condition) should be complemented with power measures to obtain a fair empirical assessment of the rationalizability conditions under evaluation. Here, power is measured as the probability of rejecting the revealed preference test given that the model does not hold. Favorable test results (i.e. a high pass rate for some given data), which prima facie suggest a good empirical fit, have little value if the test has little discriminatory power (i.e. the conditions are hard to reject for the data at hand).

For all revealed preference tests under evaluation, we compute a power measure for every individual household. This measure quantifies discriminatory power in terms of the probability to detect random behavior, and is based on Bronars (1987). More precisely, we simulated 1000 random series of eight consumption choices by drawing, for each of the eight observed household budgets, a random quantity bundle from a uniform distribution on the given budget hyperplane for the corresponding prices and total expenditure. The power measure is then calculated as one minus the proportion of these randomly generated consumption series that are consistent with the rationalizability conditions under evaluation. The distribution of this power measure for the different models is given in Table 2. We see that the

[16]To investigate the source of these violations we checked whether the households satisfy the necessary conditions for homothetic separability (see Section 3). When checking whether the second of these conditions hold (i.e. homotheticity of the subutility function) we found that, indeed, none of the households passes the corresponding test. This means that the data $\{\mathbf{q}_t, \mathbf{y}_t\}_{t \in T}$ cannot be rationalized by a homothetic utility function for any of the households, as required by homothetic separability.

standard utility maximization model has a rather low power. On average only about 11% of all random data sets violate the revealed preference conditions of Theorem 1. By contrast, the power distribution for the homothetic separability test is entirely centered around 1, with almost no spread. In other words, nearly all random data sets fail this test, which confirms its stringency. Finally, the weak separability test has reasonably high power while the power of the indirect weak separability test is fairly low.

This last finding suggests that, from an empirical point of view, indirect separability is much less stringent than weak separability. That is, while the indirect weak separability model was associated with a higher pass rate for the sample at hand, it seems that this better fit may simply be due to a lower discriminatory power rather than a better model per se. Our following exercise accounts for the possible trade–off between pass rate and power.

**Predictive success.** The above analysis compares the four behavioral models in terms of their pass rates and discriminatory power. Beatty and Crawford (2011) recently suggested to combine these two (often inversely related) performance measures into a single metric. More specifically, building further on an original idea of Selten (1991), they suggest to assess the empirical performance of a model by a so–called predictive success measure which, for a given household, is computed as the difference between the pass rate (either 1 or 0) and 1 minus the power. By construction, this measure takes values between -1 and 1. Negative values then suggest that the model under study is rather inadequate to describe the household data at hand: the model provides a poor fit of the household behavior (pass rate is zero) even though the model's power is low (i.e. the model is difficult to reject empirically). Conversely, a high and positive predictive success value points to a potentially useful model: it is able to explain the observed consumption behavior (i.e. pass rate equals 1) while its power is high (i.e. the model would rapidly be rejected in case of random behavior).

Table 3 presents some statistics of the predictive success measures associated with the four models under study. We observe that the standard utility maximization model achieves the highest mean predictive success. However, the value of 0.023 is still very low. In general, the mean predictive success values do not provide a strong empirical case in favor of one or the other model.

We obtain a more balanced picture when considering the quartile values. For the homothetic separability model, the predictive success measure is entirely centered around zero with (practically) no spread. This result directly follows from the fact that this model has, for each household, a zero pass rate combined with power (close to) unity. Next, the distributions of the predictive success measures are almost identical for the standard utility maximization model and the indirect weak separability model. In other words, indirect separability seems to add little value (if any) over and above basic utility maximization in terms of predictive success. Finally, the predictive success distribution of the weak separability model seems to be bimodal: on the one hand, there are a lot of households with very negative predictive success values for weak separability but, on the other hand, there are also a lot of households with large and positive predictive success values. One interpretation is that the weak separability model performs rather well empirically for one subgroup of households while it does a fairly poor job for other households. Given this, it can be useful to investigate which household characteristics determine the good fit of the weak separability model. Because our empirical application is mainly meant to be illustrative, we will not explore this route here, but we do see this as a potentially interesting avenue for follow–up research.

A related point concerns the observation that the weak separability model dominates the indirect weak separability model in terms of predictive success for the median, third quartile and maximum values. This suggests that weak separability may effectively constitute an appropriate model to describe the consumption behavior of most households in our sample. While it provides a worse fit than indirect separability at the overall sample level, for those households that do pass the weak separability test the higher discriminatory power effectively makes this model more useful from an empirical point of view. That is, for many households we obtain a predictive success value that is substantially above zero.

Table 3: Predictive success

| model | mean | min | 1st quartile | median | 3rd quartile | max |
|---|---|---|---|---|---|---|
| general ut. max. | 0.023 | $-1$ | 0 | 0.035 | 0.177 | 0.649 |
| weak separability | 0.016 | $-0.957$ | $-0.414$ | 0.139 | 0.444 | 0.964 |
| homothetic separability | 0 | $-0.001$ | 0 | 0 | 0 | 0 |
| indirect separability | $-0.042$ | $-1$ | 0 | 0.024 | 0.207 | 0.757 |

### 4.3 Accounting for measurement error

Until now, we considered the basic revealed preference test of weak separability. This is a 'sharp' test in the sense that it does not take possible measurement error into account. It only tells us, for the data at hand, whether the households are exact optimizers in terms of, for example, a weakly separable utility function. However, since consumption data are often measured with error, this exactness is not innocuous. There are two possible cases where the sharp test can produce the wrong answer. In the first case, the true data are rationalizable but, due to the measurement error, the observed data are not, i.e. we have a so called 'false negative'. In the second case, the true data are not rationalizable although the observed data are, i.e. we have a 'false positive'.

We complement our application with an analysis that develops a statistical test procedure to account for these two situations. At this point, two remarks are in order. Firstly, although our focus here will be on measurement error in the quantities, our analysis can easily be adapted to take into account measurement error in the prices.[17] Secondly, for brevity we will only consider tests for weak separability in what follows. However, our following reasoning is also directly applicable to the models of homothetic separability and indirect separability.

We consider the following optimization problem.

**OP.WS**

$$\min_{S_t,\delta_t,X_{t,v},F} F$$

$$\begin{aligned}
&\text{s.t.}\\
&S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v) + \delta_v F,\\
&u_t - u_v < X_{t,v},\\
&(X_{t,v} - 1) \leq u_t - u_v,\\
&\delta_t \mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) < X_{t,v} A_t + \delta_t F,\\
&(X_{t,v} - 1) A_v \leq \delta_v \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) + \delta_v F,\\
&S_t \geq 0,\\
&\delta_t > 0,\\
&X_{t,v} \in \{0,1\}.
\end{aligned}$$

[17] See for instance Crawford (2010) and Cherchye, Demuynck, and De Rock (2011a) for examples of revealed preference tests that take into account measurement error in the prices.

When comparing this problem with the earlier program **CS.WS**, we observe that the optimal solution of **OP.WS**, say $F^*$, must be smaller than or equal to zero if and only if the data set is rationalizable by a weakly separable utility function. In other words, the data set $\{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t\in T}$ satisfies **CS.WS** if and only if $F^* \leq 0$.

**A false negative.** Let $(\mathbf{x}_t, \mathbf{y}_t)$ represent the observed quantities at observation $t$ and assume that the true quantities are given by $(\mathbf{x}_t^*, \mathbf{y}_t^*)$, where

$$\mathbf{x}_t^* = \mathbf{x}_t + \varepsilon_t \text{ and,}$$
$$\mathbf{y}_t^* = \mathbf{y}_t + v_t,$$

with $\varepsilon_t$ and $v_t$ defining the unobserved measurement error.

If we have a false negative, then the true data set is rationalizable by a weakly separable utility function (i.e. it satisfies the conditions of **CS.WS**), while the observed data set is not because the quantities are measured with error. This means that the optimal solution of **OP.WS** is larger than zero for the actual data but, if we had used the true quantities, then this solution value would not have exceeded zero. As such, if the true data set is rationalizable, we should have that $F^*$ is not too large. The following theorem formalizes this intuition by giving an upper bound on the optimal value of **OP.WS**. The proof is given in Appendix C.

**Theorem 7.** *Assume that $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{x}_t^*, \mathbf{y}_t^*\}_{t\in T}$ satisfies the constraints of CS.WS and let $F^*$ be the optimal value of OP.WS for the observed data set $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{x}_t, \mathbf{y}_t\}_{t\in T}$. Then,*

$$F^* \leq \max\left\{\max_{t,v} \mathbf{p}_t(\varepsilon_t - \varepsilon_v); \max_{t,v} \mathbf{q}_t(v_t - v_v)\right\}.$$

This theorem motivates the following formulation of the null hypothesis that $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{x}_t^*, \mathbf{y}_t^*\}_{t\in T}$ is rationalizable by a weakly separable utility function:

$$H_0 : F^* \leq \max\left\{\max_{t,v} \mathbf{p}_t(\varepsilon_v - \varepsilon_v); \max_{t,v} \mathbf{q}_t(v_v - v_v)\right\};$$
$$H_1 : F^* > \max\left\{\max_{t,v} \mathbf{p}_t(\varepsilon_v - \varepsilon_v); \max_{t,v} \mathbf{q}_t(v_v - v_v)\right\}.$$

It is directly apparent that a test of this null hypothesis for false negatives will be a conservative one. A similar qualification applies to the test for false positives which will be developed below. The distribution of the errors $\varepsilon_t$ and $v_t$ is unknown, so we resort to a simulation procedure in order to implement the hypothesis test. This procedure takes the following steps:

1. Compute the optimal value of **OP.WS**.
2. Simulate errors $\varepsilon_t$ and $v_t$ drawn from some predefined distribution and calculate the value $\max\{\max_{t,v} \mathbf{p}_t(\varepsilon_v - \varepsilon_t); \max_{t,v} \mathbf{q}_t(v_v - v_t)\}$. We have 5000 draws (per household).
3. Compute the percentage of these values that exceeds the optimal value of **OP.WS** computed in the first step.
4. If this percentage is smaller than $\alpha$, then we reject the hypothesis that the true data set is rationalizable by a weakly separable utility function for a significance level of $\alpha$.

This test procedure is a variation of the one originally developed by Fleissig and Whitney (2008) and Jones and Edgerton (2009).[18] In order to apply it, two issues must be resolved. First of all, the optimal value of **OP.WS** must be computed. This problem is nonlinear in the variable $F$ and might therefore be considered difficult. However, notice that if **OP.WS** has a feasible solution for a particular value of $F$, then it also has a feasible solution for all values of $F' \geq F$. From this monotonicity condition, it follows that we can solve the problem quite efficiently using a binary search algorithm.[19] The second issue concerns the distribution and structure of the errors $\varepsilon_t$ and $v_t$. We assume a multiplicative error structure: $\varepsilon_t = \eta_t \mathbf{x}_t$ and $v_t = \zeta_t \mathbf{y}_t$ where $\eta_t$ and $\zeta_t$ are diagonal matrices with the diagonals being i.i.d. mean zero normally distributed variables with standard deviation $\sigma$. Although other error structures are of course amenable, we choose the multiplicative one here because it more efficiently accounts for differences in the scale of expenditure across goods and observations. In our application, we performed the computations for different values of the standard deviation.

Table 4 presents the result of our procedure for the 732 households that fail the 'sharp' weak separability test. The entries in the table give the percentage of households for which we reject the null hypothesis of rationalizability by a weakly separable utility function for various levels of $\alpha$ and $\sigma$. As an example on how to interpret these numbers, take a standard deviation $\sigma$ of 0.3% and a significance level $\alpha$ of 0.05. Then, we have that the null of rationalizability by weak separability is rejected for 3.97% of all households (that violated the 'sharp' conditions **CS.WS**). In other words, almost 96% of the households are labelled as false negatives. Not surprisingly, Table 4 also reveals that the number of households rejecting the null is decreasing drastically in the level of $\sigma$.

Table 4: percentage of households for which $H_0$ is rejected at the given significance level

| | Significance level | | |
|---|---|---|---|
| Standard deviation $\sigma$ | $\alpha = 0.01$ | $\alpha = 0.05$ | $\alpha = 0.10$ |
| 0.0025/100 | 95.50 | 96.73 | 97.14 |
| 0.005/100 | 91.25 | 93.72 | 94.39 |
| 0.01/100 | 81.14 | 84.84 | 86.75 |
| 0.05/100 | 42.08 | 49.44 | 53.01 |
| 0.10/100 | 23.63 | 30.06 | 32.51 |
| 0.15/100 | 13.39 | 19.54 | 21.44 |
| 0.20/100 | 6.83 | 11.07 | 13.80 |
| 0.25/100 | 3.82 | 6.83 | 8.06 |
| 0.30/100 | 2.32 | 3.97 | 5.46 |
| 0.35/100 | 1.63 | 2.60 | 3.69 |
| 0.40/100 | 0.83 | 2.32 | 2.60 |
| 0.45/100 | 0.83 | 1.50 | 2.32 |
| 0.50/100 | 0.55 | 0.96 | 1.50 |
| 0.55/100 | 0.42 | 0.83 | 0.96 |
| 0.60/100 | 0.27 | 0.42 | 0.96 |

[18] See also Varian (1985) and Epstein and Yatchew (1985) for procedures to account for measurement error in revealed preference tests.

[19] A binary search algorithm departs with an infeasible lower bound, $F_\ell$, and a feasible upper bound, $F_u$, for the objective function. For each step of the algorithm, the procedure evaluates whether the midpoint $(F_u + F_\ell)/2$ is feasible. If it is, then in the next iteration the upper bound $F_u$ is replaced by this midpoint. If the midpoint is not feasible, then the midpoint replaces the lower bound $F_\ell$. At each iteration of the algorithm, the range $[F_\ell, F_u]$, which contains the solution of the problem, is halved. As such, the width of the interval decreases exponential in the number of iterations.

**A false positive.** A false positive means that the true data set is actually not rationalizable by a weakly separable utility function but, due to measurement error, we conclude that the observed data set is rationalizable. In this case, the optimal solution of **OP.WS** should be less than or equal to zero, while this value would be larger than zero if the true data set were used. As such, if the true data set is not rationalizable, we would expect that $F^*$ is not too far below zero. The following theorem formalizes this intuition. Like before, the proof is given in Appendix C.

**Theorem 8.** *Assume that* $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{x}_t^*, \mathbf{y}_t^*\}_{t \in T}$ *does not satisfy the constraints of CS.WS and let* $F^*$ *be the optimal value of OP.WS for the observed data set* $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$*. Then,*

$$F^* > \min \left\{ \min_{t,v} \mathbf{p}_t(\varepsilon_v - \varepsilon_t); \min_{t,v} \mathbf{q}_t(v_v - v_t) \right\}.$$

This result allows us to formulate the following null hypothesis based on the null that the true data $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{x}_t^*, \mathbf{y}_t^*\}_{t \in T}$ is not rationalizable by a weakly separable utility function.

$$H_0 : F^* > \min \left\{ \min_{t,v} \mathbf{p}_t(\varepsilon_t - \varepsilon_v); \min_{t,v} \mathbf{q}_t(v_v - v_t) \right\};$$
$$H_1 : F^* \leq \min \left\{ \min_{t,v} \mathbf{p}_t(\varepsilon_t - \varepsilon_v); \min_{t,v} \mathbf{q}_t(v_v - v_t) \right\}.$$

Thus, one rejects the null of no rationalizability for large negative values of $F^*$. Analogous to above, we use a simulation based procedure to implement the corresponding hypothesis test.

1. Compute the optimal value of **OP.WS**.
2. Simulate errors $\varepsilon_t$ and $v_t$ drawn from some predefined distribution and calculate the value $\min\{\min_{t,v} \mathbf{p}_t(\varepsilon_t - \varepsilon_v); \min_{t,v} \mathbf{q}_t(v_v - v_t)\}$. We have 5000 draws (per household).
3. Compute the percentage of these values that are below the optimal value of **OP.WS** computed in the first step.
4. If this percentage is smaller than $\alpha$, then we reject the hypothesis that the true data set is not rationalizable for a significance level of $\alpha$.

Table 5 presents the results of the procedure for the 853 households that satisfy the sharp rationalizability test. Its interpretation is similar to the results from Table 4. As an example, let us again consider a standard deviation $\sigma$ of 0.3%, and a significance level $\alpha$ of 0.05. We then reject the null hypothesis of non–rationalizability for 63.1% of the households for which the observed data did satisfy the sharp test. In other words, about 37% of the households can be labeled as false positives. Again as one can expect, we conclude that the number of household rejecting the null decreases if the standard deviation increases.

Finally, comparing the test results in Tables 4 and 5 may suggest that we should care much more about false negatives than about false positives, if we believe measurement error is an issue for the data at hand. However, drawing such a conclusion is misleading for several reasons. First of all, in order to compare the two tests, we should control for the power of the different tests. This power will generally depend on the nature of the data that are involved. For compactness, we choose not to explore this power issue further in this paper. Second, the test results in the two tables pertain to distinct subsets of observations, with possibly different characteristics. Finally, we observe that both tests are conservative by construction. As such, for a given significance level, they only identify lower bounds on the numbers of false positives and negatives. This also means that the true probability values (under the null) of the test statistics are unknown, which makes the test results not comparable.

Table 5: percentage of households for which $H_0$ is rejected at the given significance level

| | Significance level | | |
|---|---|---|---|
| Standard deviation $\sigma$ | $\alpha = 0.01$ | $\alpha = 0.05$ | $\alpha = 0.10$ |
| 0.0025/100 | 98.21 | 98.94 | 99.07 |
| 0.005/100 | 97.54 | 98.23 | 98.23 |
| 0.01/100 | 94.36 | 96.00 | 96.83 |
| 0.05/100 | 80.61 | 83.42 | 84.62 |
| 0.10/100 | 72.73 | 75.79 | 76.61 |
| 0.15/100 | 67.33 | 70.03 | 71.91 |
| 0.20/100 | 64.28 | 66.40 | 68.39 |
| 0.25/100 | 62.99 | 64.52 | 65.56 |
| 0.30/100 | 61.82 | 63.10 | 64.03 |
| 0.35/100 | 60.29 | 62.28 | 63.10 |
| 0.40/100 | 58.88 | 61.46 | 62.40 |
| 0.45/100 | 57.81 | 60.29 | 61.69 |
| 0.50/100 | 56.64 | 59.34 | 60.63 |
| 0.55/100 | 54.76 | 58.06 | 59.58 |
| 0.60/100 | 54.16 | 57.35 | 58.75 |

## 5 Conclusion

We considered the revealed preference conditions for weak separability. From a theoretical perspective, we found that verifying these conditions is a difficult (= NP–complete) problem. Given this, we introduced an integer programming approach to test data consistency with the conditions. We illustrated the versatility of this approach by deriving formally similar integer programming tests for the cases of homothetic separability and indirect weak separability.

Further, we showed the empirical viability of our integer programming approach by providing an application to Spanish household consumption data. In this application, we focused on separability between food expenditures and other expenditures (on nondurables). An interesting observation was that indirect weak separability was associated with a higher pass rate than weak separability for the sample of households at hand. However, we also found that the weak separability test had substantially more discriminatory power than the indirect separability test. As a result, the weak separability model was associated with a rather favorable predictive success measure (indicating a high degree of empirical usefulness) for most households considered. Finally, we presented two statistical tests that account for measurement error in the data.

We see multiple avenues for further research. First of all, at the theoretical level, we have concentrated on the three most commonly used types of separability, which have been established and implemented in the literature for a long time: weak separability, homothetic separability and indirect weak separability. More recently, Blundell and Robin (2000) introduced the notion of latent separability, a generalization of weak separability that provides an attractive empirical and theoretical framework for investigating the grouping of goods and prices. Crawford (2004) has derived the revealed preference conditions for latent separability. As in the weak separability case, the latent separability conditions are nonlinear (quadratic) and thus hard to verify. We believe it would be interesting to explore whether and to what extent the integer programming approach set out in the current paper may help to derive necessary and/or sufficient testable (integer programming) formulations of Crawford's conditions for latent separability.

Next, at the methodological level, we focused our discussion by only considering revealed preference tests for alternative separability specifications. If observed behavior is consistent with a particular specification (i.e. can be rationalized), then a natural next question pertains to recovering/identifying the structural features of the model under consideration. For example, in the present context such recovery can focus on identifying group (price/quantity) indices that are consistent with a separable representation of the utility structure. Because the revealed preference approach does not require a prior specification for the utility functions, it addresses recovery questions by 'letting the data speak for themselves' (i.e. it only uses the information that is directly revealed by the data). See, for example, Afriat (1967) and Varian (1982) for detailed discussions of revealed preference recoverability. These authors consider the standard utility maximization model. By using the integer programming formulations developed in the current paper, one can address similar recovery questions under alternative separability assumptions.[20]

## References


Afriat, S. N. (1967). "The Construction of Utility Functions from Expenditure Data". *International Economic Review* 8, pp. 67–77.

Afriat, S. N. (1969). "The Construction of Separable Utility Functions from Expenditure Data". Tech. rep. Mimeo, University of North Carolina.

Andreoni, J and W Harbaugh (2008). "Power indices for revealed preference tests". Tech. rep. 2005-10. University of Wisconsin-Madison Department of Economics.

Barnett, W. A. and S. Choi (1989). "A Monte Carlo Study of Tests of Blockwise Weak Separability". *Journal of Business and Economic Statistics* 7, pp. 363–377.

Barnhart, S. W. and G. A. Whitney (1988). "Nonparametric Analysis in Parametric Estimation: An Application ot Translog Demand Systems". *The Review of Economics and Statistics* 70, pp. 149–153.

Beatty, T. K. M. and I. A. Crawford (2011). "How Demanding is the Revealed Preference Approach to Demand". *American Economic Review* 101, pp. 2782–2795.

Belongia, M. T. and K. A. Chrystal (1991). "An Admissible Monetary Aggregate for the United Kingdom". *Review of Economics and Statistics* 73, pp. 497–503.

Blackorby, C., D. Primont, and R. R. Russell (1978). *Duality, Separability and Functional Structure: Theory and Economic Applications*. New York: Elsevier.

Blackorby, C. and R. R. Russel (1994). "The Conjunction of Direct and Indirect Separability". *Journal of Economic Theory* 62, pp. 480–498.

Blundell, R., M. Browning, and I. Crawford (2007). "Improving Revealed Preference Bounds on Demand Responses". *International Economic Review* 48, pp. 1227–1244.

Blundell, R. and J. Robin (2000). "Latent Separability: Grouping Goods Without Weak Separability". *Econometrica* 68, pp. 53–84.

Bronars, S. G. (1987). "The Power of Nonparametric Tests of Preference Maximization". *Econometrica* 55, pp. 693–698.

Brown, D. J. and C. Shannon (2000). "Uniqueness, Stability, and Comparative Statics in Rationalizable Walrasian Markets". *Econometrica* 68, pp. 1529–1540.

Browning, M. and M. D. Collado (2001). "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates". *American Economic Review* 91, pp. 681–692.


[20] Compare with Cherchye, De Rock, and Vermeulen (2011), who address recovery questions for collective consumption models by using a closely similar integer programming approach.

Cherchye, L., B. De Rock, and F. Vermeulen (2007). "The Collective Model of Household Consumption: a Nonparametric Characterization". *Econometrica* 75, pp. 553–574.
Cherchye, L., B. De Rock, and F. Vermeulen (2009). "Opening the Black Box of Intra-Household Decision-Making". *Journal of Political Economy* 117, pp. 1074–1104.
Cherchye, L., B. De Rock, and F. Vermeulen (2011). "The Revealed Preference Approach to Collective Consumption Behavior: Testing and Sharing Rule Recovery". *Review of Economic Studies* 78, pp. 176–198.
Cherchye, L., T. Demuynck, and B. De Rock (2011a). "Is utility transferable? A revealed preference analysis". CES DP. 11.02,
Cherchye, L., T. Demuynck, and B. De Rock (2011b). "Revealed preference analysis of noncooperative household consumption". *The Economic Journal* 121, pp. 1073–1096.
Cherchye, L., T. Demuynck, and B. De Rock (2011c). "Revealed Preference tests for Weak Separability: an integer programming approach". CES DP 11.25.
Cherchye, L., T. Demuynck, and B. De Rock (2011d). "Testable implications of general equilibrium models: an integer programming approach". *Journal of Mathematical Economics* 47, pp. 564–575.
Cherchye, L. et al. (2008). "Nonparametric tests of collective rational consumption behavior: An integer programming procedure". *Journal of Econometrics* 147, pp. 258–265.
Choi, S. and K. Sosin (1992). "Structural Change in the Demand for Money". *Journal of Money, Credit and Banking* 24, pp. 226–238.
Cox, J. C. (1997). "On Testing the Utility Hypothesis". *The Economic Journal* 107, pp. 1054–1078.
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, pp. 1382–1402.
Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behavior*. Cambridge University Press.
Diewert, E. W. (1976). "Exact and Superlative Index Numbers". *Journal of Econometrics* 4, pp. 115–145.
Diewert, E. W. (1978). "Superlative Index Numbers and Consistency in Aggregation". *Econometrica* 46, pp. 883–900.
Diewert, E. W. and C. Parkan (1985). "Tests for the Consistency of Consumer Data". *Journal of Econometrics* 30, pp. 127–147.
Diewert, W. E. (1973). "Afriat and revealed preference theory". *The Review of Economic Studies* 40, pp. 419–425.
Drake, L. and A. R. Fleissig (2008). "A Note on the Policy Implications of Using Divisia Consumption and Monetary Aggregates". *Macroeconomic Dynamics* 12, pp. 132–149.
Elger, C. T. et al. (2008). "A Note on the Optimal Level of Monetary Aggregation in the United Kingdom". *Macroeconomic Dynamics* 12, pp. 117–131.
Elger, T. and B. E. Jones (2008). "Can rejections of non-separability be attributed to random measurement errors in the data?" *Economics Letter* 99, pp. 44–47.
Epstein, L. G. and A. J. Yatchew (1985). "Non-Parametric Hypothesis Testing Procedures and Applications to Demand Analysis". *Journal of Econometrics* 30, pp. 149–169.
Fisher, D. and A. R. Fleissig (1997). "Monetary Aggregation and the Demand for Assets". *Journal of Money, Credit and Banking* 29, pp. 458–475.
Fleissig, A. and G. A. Whitney (2003). "A New PC-Based Test for Varian's Weak Separability Condition". *Journal of Business and Economic Statistics* 21, pp. 133–145.
Fleissig, A. and G. A. Whitney (2008). "A Nonparametric test of Weak Separability and Consumer Preferences". *Journal of Econometrics* 147, pp. 275–281.
Garey, M. R. and D. S. Johnson (1979). *Computers and Intractability*. Bell Telephone Laboratories, Inc.

Hjertstrand, P and J. L. Swofford (2012). “Revealed Preference Tests for Consitency with Weakly Separable Indirect Utility”. *Theory and Decision* 72, pp. 245–256.
Hjertstrand, P. (2007). “Functional Structure Inference”. Ed. by A. Serletis. Emerald Group Publishing Limited. Chap. Food Demand in Sweden: a Nonparametric Approach, pp. 157–182.
Hjertstrand, P. (2008). “Testing for rationality, separability and efficiency”. Phd. Thesis. Lund University.
Hjertstrand, P. (2009). “Measurment Error: Consequences, Applications and Solutions”. Ed. by T. B. Fomby and R. C. Hill. Emerald Group Publishing Ltd. Chap. A Monte Carlo Study of the Necessary and Sufficient Conditions for Weak Separabillity, pp. 151–182.
Hjertstrand, P. (2011). “Linear and Joint Revealed Preference Tests of the Necessary and Sufficient Conditions for Weak Separability”. mimeo.
Jha, R. and I. S. Longjam (2006). “Structure of Financial Savings During Indian Economic Reforms”. *Empirical Economics* 31, pp. 861–869.
Jones, A. and M. G. Mazzi (1996). “Tobacco Consumption and Taxation in Italy: an Application of the QUAIDS model”. *Applied Economics* 28, pp. 595–603.
Jones, B. E., D. H. Dutkowsky, and T. Elger (2005). “Sweep Programs and Optimal Monetary Aggregation”. *Journal of Banking and Finance* 29, pp. 483–508.
Jones, B. E. and D. L. Edgerton (2009). “Advances in Econometrics, Volume 24”. Ed. by J. M. Binner, D. L. Edgerton, and T. Elger. Emerald Group Publishing Limited. Chap. Testing Utility Maximization with Measurement Errors in the Data, pp. 199–236.
Jones, B. E., N. McCloud, and D. L. Edgerton (2007). “Functional Structure Inference”. Ed. by A. Serletis. Emerald Group Publishing Limited. Chap. Nonparametric Tests of the Necessary and Sufficient Conditions for Separability, pp. 33–56.
Leontief, W. (1947). “Introduction to a Theory of the Internal Structure of Functional Relationships”. *Econometrica* 15, pp. 361–373.
Patterson, K. D. (1991). “A Non-Parametric Analysis of Personal Sector Decisions on Consumption, Liquid Assets and Leisure”. *The Economic Journal* 101, pp. 1103–1116.
Rickertsen, K. (1998). “The Demand for Food and Beverages in Norway”. *Agricultural Economics* 18, pp. 89–100.
Selten, R. (1991). “Properties of a Measure of Predictive Success”. *Mathematical Social Sciences* 21, pp. 153–167.
Serletis, A. and R. Rangel-Ruiz (2005). “Microeconometrics and measurement Matters: Some Results from Monetary Economics in Canada”. *Journal of Macroeconomics* 27, pp. 307–330.
Sono, M. (1961). “The Effect of Price Changes on the Demand and Supply of Separable Goods”. *International Economic Review* 2, pp. 239–271.
Spencer, P. (2002). “The Inpact of Information and Communications Technology Investment on UK Productive Potential 1986-2000: New Statistical Methods and Tests”. *The Manchester School* 70, pp. 107–126.
Swofford, J. L. (2005). “Tests of Microeconomic Foundations of a North American Common Currency Area”. *The Canadian Journal of Economics* 38, pp. 420–429.
Swofford, J. L. and G. A. Whitney (1987). “Nonparametric Tests of Utility Maximization and Weak Separability for Consumption, Leisure and Money”. *The Review of Economics and Statistics* 69, pp. 458–464.
Swofford, J. L. and G. A. Whitney (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, pp. 241–246.
Swofford, J. L. and G. A. Whitney (1994). “A revealed preference test for weakly separable utility maximization with incomplete adjustment”. *Journal of Econometrics* 60, pp. 235–249.

Varian, H. (1984). “The Nonparametric Approach to Production Analysis”. *Econometrica* 52, pp. 579–597.
Varian, H. (1982). “The Nonparametric Approach to Demand Analysis”. *Econometrica* 50, pp. 945–974.
Varian, H. (1983). “Non-Parametric Tests of Consumer Behavior”. *The Review of Economic Studies* 50, pp. 99–110.
Varian, H. (1985). “Non-Parametric Analysis of Optimizing Behavior with Measurement Error”. *Journal of Econometrics* 30, pp. 445–458.
Warshall, S (1962). “A Theorem of Boolean Matrices”. *Journal of the American Association of Computing Machinery* 9, pp. 11–12.

## Appendix A: proof of Theorem 3

*Proof.* In order to show that the problem of rationalizability by a weakly separable utility function is in the class NP, we need to reduce a known NP-complete problem to this decision problem. For this we use the problem of Monotone 3SAT (M3SAT).
**M3SAT**
**INSTANCE:** A set of binary variables $b_1, \dots, b_t$ and a set of clauses $C_1, \dots, C_r$. Each clause $C_\ell$, $\ell = 1, \dots, r$, contains three literals $l_{1,\ell}, l_{2,\ell}$ and $l_{3,\ell}$ and each literal either equals a variable or its negation. The condition monotone refers to the fact that for every clause all literals within this clause are either negated or unnegated.
**QUESTION:** Does there exist an assignment to the variables $b_1, \dots, b_t$ (either 1 or 0) such that each clause contains at least one literal with the values equal to 1?

Now, consider an instance of M3SAT. We first construct the set of observations $T$ and the sets of goods $T$ and $S$.:

- For every literal $l_{k,\ell}$ ($\ell = 1, \dots, r$ and $k = 1, 2, 3$), we construct two observations $t(k, \ell)$ and $v(k, \ell)$. These observations are gathered in the set $T'$.
- For every literal $l_{k,\ell}$ ($\ell = 1, \dots, r$ and $k = 1, 2, 3$), we create two goods $g(t, k, \ell)$ and $g(v, k, \ell)$.
- For every literal $l_{k,\ell}$ ($\ell = 1, \dots, r$ and $k = 1, 2, 3$), we create two goods $h(t, k, \ell)$ and $h(v, k, \ell)$.

For two literals $l$ and $l'$, we say that they are opposites if $l$ corresponds to a variable $b_i$ and $l'$ corresponds to $(1 - b_i)$ or $l$ corresponds to $(1 - b_i)$ and $l'$ corresponds to $b_i$ (i.e. $l \equiv (1 - l')$). We consider some special subsets of the set of goods.

- $\mathcal{G}_t = \{g(t, k, \ell) | k = 1, 2, 3; \ell = 1, \dots, r\}$.
- $\mathcal{G}_v = \{g(v, k, \ell) | k = 1, 2, 3; \ell = 1, \dots, r\}$.
- $\mathcal{O}(t, k, \ell) = \left\{ g(v, k', \ell') \,\middle|\, \begin{array}{l} \text{the k-th literal in clause } \ell \text{ and the } k'\text{th literal} \\ \text{in clause } \ell' \text{ are opposites} \end{array} \right\}$.
- $\mathcal{H}_t = \{h(t, k, \ell) | k = 1, 2, 3; \ell = 1, \dots, r\}$.
- $\mathcal{H}_v = \{h(v, k, \ell) | k = 1, 2, 3; \ell = 1, \dots, r\}$.

Table 6: Prices and quantities for instance of weak separability

| observation | $g(t,k,\ell)$ | $\mathcal{G}_t - \{g(t,k,\ell)\}$ | $\mathcal{G}_v - \mathcal{O}(t,k,\ell)$ | $O(t,k,\ell)$ |
|---|---|---|---|---|
| $t(k,\ell)$ | $\mathfrak{p}\vert 1$ | $1\vert 2$ | $1\vert 2$ | $1\vert 1-\frac{1}{\mathfrak{p}}$ |

| observation | $h(t,k,\ell)$ | $\mathcal{H}_t - \{h(t,k,\ell)\}$ | $h(v,k,\ell)$ | $\mathcal{H}_v - \{h(v,k,\ell)\}$ |
|---|---|---|---|---|
| $t(k,\ell)$ | $\mathfrak{z}\vert 2$ | $1\vert 3$ | $1\vert 1-\frac{1}{\mathfrak{y}}$ | $1\vert 2$ |

| observation | $g(v,k,\ell)$ | $\mathcal{G}_t$ | $\mathcal{G}_v - \{g(v,k,\ell)\}$ | |
|---|---|---|---|---|
| $v(k,\ell)$ | $\mathfrak{p}\vert 1$ | $1\vert 2$ | $1\vert 2$ | |

| observation | $h(t,k\oplus 2,\ell)$ | $\mathcal{H}_t - \{h(t,k\oplus 2,\ell)\}$ | $h(v,k,\ell)$ | $\mathcal{H}_v - \{h(v,k,\ell)\}$ |
|---|---|---|---|---|
| $v(k,\ell)$ | $1\vert 1$ | $1\vert 3$ | $\mathfrak{y}\vert 1$ | $1\vert 2$ |

The goods in the separable group (bundle $\mathbf{y}$) are the goods $g(t,k,\ell)$ and $g(v,k,\ell)$. For $k$ and $l \in \mathbb{N}$ denote by $k \oplus l$ the number $(k+l) \mod 3$. The remaining goods are the goods for the non–separable group (bundle $\mathbf{x}$). The prices and quantities for each observation and good are summarized in the following tables for all $k = 1,2,3$ and $\ell = 1,\dots,r$ (prices are before the separator '|', quantities after).

Here, the numbers $\mathfrak{p}$, $\mathfrak{z}$ and $\mathfrak{y}$ are given by:

$$\mathfrak{p} = 14 + 35r, \qquad \mathfrak{z} = 16 + 42r, \qquad \mathfrak{y} = 11 + 29r,$$

with $r$ the number of clauses.

We have to show that M3SAT has a solution if and only if the data set constructed above is weakly separable rationalizable. First let us assume that the data set is weakly separable rationalizable. Let $S_{t(k,\ell)}$ and $S_{v(k,\ell)}$ and $U_{t(k,\ell)}, U_{v(k,\ell)}$ be the Afriat numbers for the observations $t(k,\ell)$ and $v(k,\ell)$ that correspond to this rationalization. The idea is to set the value of the variables in such a way as to guarantee that the $k$th literal in the $\ell$th clause is equal to one whenever $S_{t(k,\ell)} \geq S_{v(k,\ell)}$. We need to verify that this is possible and that this leads to a solution of M3SAT. The following facts will be helpful.

**Fact 1.** *For all $k,k' = 1,2,3$ and $\ell,\ell' = 1,\dots,r$, if the $k$th literal in the $\ell$th clause and the $k'$th literal in the $\ell'$th are opposites, then $S_{v(k,\ell)} > S_{t(k',\ell')}$.*

*Proof.* We have that:

$$\begin{aligned} S_{t(k',\ell')} - S_{v(k,\ell)} &\leq \delta_{v(k,\ell)} \mathbf{q}_{v(k,l)} \left[\mathbf{y}_{t(k',\ell')} - \mathbf{y}_{v(k,\ell)}\right] \\ &= \delta_{v(k,\ell)} \left[-1 - 1 + (-2 + 1 - 1/\mathfrak{p}) \left|O(t,k',\ell') \cap (\mathcal{G}_v - \{g(v,k,\ell)\})\right|\right] \\ &< 0 \end{aligned}$$

□

**Fact 2.** *For all $\ell,\ell' = 1,\dots,r$ and $k,k' = 1,2,3$ if the $k$th literal in the $\ell$th clause and the $k'$th literal in the $\ell'$th clause are opposites then it is not the case that both $S_{t(k,\ell)} \geq S_{v(k,\ell)}$ and $S_{t(k',\ell')} \geq S_{v(k',\ell')}$.*

*Proof.* If, on the contrary, $S_{t(k,\ell)} \geq S_{v(k,\ell)}$ and $S_{t(k',\ell')} \geq S_{v(k',\ell')}$, we would have that (by fact 1):

$$S_{t(k,\ell)} \geq S_{v(k,\ell)} > S_{t(k',\ell')} \geq S_{v(k',\ell')} > S_{t(k,\ell)},$$

a contradiction. □

Facts 1 and 2 show that above construction above can be performed (i.e. it is never the case that two opposite literals have the value of one). The following fact demonstrates that it provides a solution to M3SAT.

**Fact 3.** *For all $\ell = 1, \dots, r$, there is at least one value $k = 1, 2, 3$ such that $S_{t(k,\ell)} \geq S_{v(k,\ell)}$.*

*Proof.* Let us first show that for all $k = 1, 2, 3$ and $\ell = 1, \dots, r$, $U_{t(k,\ell)} > U_{v(k\oplus 1,\ell)}$. Indeed,

$$\begin{aligned}
U_{v(k\oplus 1,\ell)} - U_{t(k,\ell)} &\leq \lambda_{t(k,\ell)} \mathbf{p}_{t(k,\ell)} \left[\mathbf{x}_{v(k\oplus 1,\ell)} - \mathbf{x}_{t(k,\ell)}\right] + \lambda_{t(k,\ell)} \mathbf{q}_{t(k,\ell)} \left[\mathbf{y}_{v(k\oplus 1,\ell)} - \mathbf{y}_{t(k,\ell)}\right] \\
&= \lambda_{t(k,\ell)} \begin{bmatrix} \mathfrak{p} - 1 + (2 - 1 + 1/\mathfrak{p}) \left|O(t,k,\ell) \cap (\mathcal{G}_v - \{g(v,k,\ell\})\right| \\ -\mathfrak{z} + (2 - 1 + 1/\mathfrak{y}) + (1 - 2) \end{bmatrix} \\
&\leq \lambda_{t(k,\ell)} \left[\mathfrak{p} - 1 + 6r - \mathfrak{z} + 2 - 1\right] \\
&= \lambda_{t(k,\ell)} \left[(14 + 35r) + 6r - (16 + 42r)\right] < 0
\end{aligned}$$

Now, consider the identity

$$\begin{aligned}
0 = &\left[U_{v(k\oplus 1,\ell)} - U_{t(k,\ell)}\right] + \left[U_{v(k\oplus 2,\ell)} - U_{t(k\oplus 1,\ell)}\right] + \left[U_{v(k\oplus 3,\ell)} - U_{t(k\oplus 2,\ell)}\right] \\
&+ \left[U_{t(k,\ell)} - U_{v(k,\ell)}\right] + \left[U_{t(k\oplus 1,\ell)} - U_{v(k\oplus 1,\ell)}\right] + \left[U_{t(k\oplus 2,\ell)} - U_{v(k\oplus 2,\ell)}\right]
\end{aligned}$$

The first three terms on the right hand side are negative, hence,

$$\begin{aligned}
0 &< \left[U_{t(k,\ell)} - U_{v(k,\ell)}\right] + \left[U_{t(k\oplus 1,\ell)} - U_{v(k\oplus 1,\ell)}\right] + \left[U_{t(k\oplus 2,\ell)} - U_{v(k\oplus 2,\ell)}\right] \\
&\leq \frac{\lambda_{v(1,\ell)}}{\delta_{v(1,\ell)}} \left[S_{t(1,\ell)} - S_{v(1,\ell)}\right] + \frac{\lambda_{v(2,\ell)}}{\delta_{v(k\oplus 1,\ell)}} \left[S_{t(2,\ell)} - S_{v(2,\ell)}\right] + \frac{\lambda_{v(3,\ell)}}{\delta_{v(3,\ell)}} \left[S_{t(3,\ell)} - S_{v(3,\ell)}\right]
\end{aligned}$$

As such at least for one $k = 1, 2, 3$ it must be that $S_{t(k,\ell)} > S_{v(k,\ell)}$. □

Now, consider a 'yes' instance of M3SAT. We need to construct Afriat numbers $S$ and $\delta$ for each observation that satisfy the the conditions for rationalizability by weak separability (see Theorem 2). Let us start by constructing a binary relation $\succ$. For $k, k' = 1, 2, 3$ and $\ell, \ell' = 1, \dots, r$ if the $k$-th literal in the $\ell$th clause and the $k'$th literal in the $\ell'$th clause are opposites, we set $v(k,\ell) \succ t(k',\ell')$. Further, for all $k = 1, 2, 3$ and $\ell = 1, \dots, r$ if the $k$th literal in the $\ell$th clause has the value 1, we set $t(k,\ell) \succ v(k,\ell)$. These are the only comparisons in $\succ$. Observe that $\succ$ has no cycles and any path in $\succ$ contains no more than 4 observations.

Let $M_1$ be the set of $\succ$-maximal elements of $T'$:

$$M_1 = \{a \in T |\ \nexists b \in T', b \succ a\}.$$

For all observations $a$ in $M_1$, we set $S_a = 4$. Let $M_2$ be the set of $\succ$-maximal elements in $T' - M_1$. For all $a \in M_2$, set $S_a = 3$. Next, let $M_3$ be the set of $\succ$-maximal elements in $T' - (M_1 \cup M_2)$ and set $S_a = 2$ for all $a \in M_3$. Finally let $M_4$ be the set of $\succ$-maximal element in $T' - (M_1 \cup M_2 \cup M_3)$ and set for all $a \in M_4$, $S_a = 1$. It is easy to see that $M_1 \cup M_2 \cup M_3 \cup M_4 = T$, hence all observations are allocated a value. Observe that when the $k$th literal in the $\ell$th clause equals one, then $S_{t(k,\ell)} > S_{v(k,\ell)}$. Finally, for all $k = 1, 2, 3$ and $\ell = 1, \ldots, r$, set $\delta_{t(k,\ell)} = 1$ and set $\delta_{v(k,\ell)} = \dfrac{1}{3 + 7r}$, where $r$ is the number of clauses.

We need to proof two things. First we need to verify that all Afriat inequalities hold for every two observations in the set $\left\{t(k,\ell), v(k,\ell), t(k',\ell'), v(k',\ell')\right\}_{k,k'=1,2,3;\ell,\ell'=1,\ldots,r}$ (i.e. condition (ii.1) of Theorem 2). Second, we need to show that the data set $\{\mathbf{p}_w, 1/\delta_w, \mathbf{x}_w, S_w\}_{w\in T'}$ satisfies GARP (condition (ii.2)). For the first, it is a straightforward but cumbersome exercise to verify every possible combination of states. As such we refer to the appendix C. Now, let us verify the second claim. Consider the direct revealed preference relation $R^D$ for the data set $\{\mathbf{p}_w, 1/\delta_w; \mathbf{x}_w, S_w\}_{w\in T'}$. We have following results.

**Fact 4.** *For all $k = 1, 2, 3$ and $\ell = 1, \ldots, r$, we have that the observation $t(k,\ell)$ is directly revealed preferred to the observation $v(k \oplus 1, \ell)$ (i.e. $(t(k,\ell), v(k \oplus 1, \ell)) \in R^D$).*

*Proof.* We have that:

$$\begin{aligned}
&\mathbf{p}_{t(k,\ell)}\left[\mathbf{x}_{t(k,\ell)} - \mathbf{x}_{v(k\oplus 1,\ell)}\right] + \frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)} - S_{v(k\oplus 1,\ell)}\right]\\
&= \mathfrak{z} + (1 - 1/\mathfrak{y} - 2) + (2 - 1) + \left[S_{t(k,\ell)} - S_{v(k\oplus 1,\ell)}\right]\\
&\geq \mathfrak{z} - 2 - 3 = 16 + 42r - 5 > 0
\end{aligned}$$

□

**Fact 5.** *For all $k = 1, 2, 3$ and $\ell = 1, \ldots, r$, $(v(k,\ell), t(k,\ell)) \in R^D$ if and only if $S_{v(k,\ell)} \geq S_{t(k,\ell)}$ (which implies that the kth literal in the $\ell$th clause is equal to zero).*

*Proof.* We have that,

$$\mathbf{p}_{v(k,\ell)}(\mathbf{x}_{v(k,\ell)} - \mathbf{x}_{t(k,\ell)}) + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k,\ell)}\right] = \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k,\ell)}\right].$$

This is positive or negative depending on the sign of $S_{v(k,\ell)} - S_{t(k,\ell)}$. □

**Fact 6.** *The relation $R^D$ contains no comparisons except for the cases mentioned by Facts 4 and 5.*

*Proof.* See appendix E. □

Now, assume a violation of GARP. Above Facts show that this implies the following cycle for some $\ell = 1, \ldots, r$:

$$\begin{aligned}
&(t(1,\ell), v(2,\ell)),\ (v(2,\ell), t(2,\ell)),\ (t(2,\ell), v(3,\ell))\\
&\qquad\qquad (v(3,\ell), t(3,\ell)),\ (t(3,\ell), v(1,\ell)),\ (v(1,\ell), t(1,\ell)).
\end{aligned}$$

Fact 5 shows that in this case $S_{v(1,\ell)} \geq S_{t(1,\ell)}$, $S_{v(2,\ell)} \geq S_{t(2,\ell)}$ and $S_{v(3,\ell)} \geq S_{t(3,\ell)}$. This can only be the case if all literals in the clause $\ell$ are zero, a contradiction. □

## Appendix B: proof of Theorem 4

*Proof.* Assume that the data set $D = \{\delta_t \mathbf{p}_t, 1; \mathbf{x}_t, S_t\}_{t\in T}$ satisfies (iv.1)–(iv.3). It is always possible to rescale the values $u_t$ such that, for all $t \in T$, $u_t < 1$. For all $t, v \in T$, define $X_{t,v} = 1$ if and only if $u_t \geq u_v$ (i.e. $X_{t,v} = 0$ if $u_t < u_v$) . We must show that conditions (cs.2)-(cs.5) hold. By definition, conditions (cs.2) and (cs.3) are always satisfied. Let $\delta_t \mathbf{p}_t \mathbf{x}_t + S_t \geq \delta_t \mathbf{p}_t \mathbf{x}_v + S_v$. Then form the contraposition of (iv.3)we obtain that $u_v > u_t$ can not occur. As such, $u_t \geq u_v$ and thus $X_{t,v} = 1$. This demonstrates that condition (cs.4) holds. Next, assume that $X_{t,v} = 1$, which is equivalent to $u_t \geq u_v$. From condition (iv.2), we then obtain that $\delta_v \mathbf{p}_v \mathbf{x}_v + S_v \leq \delta_v \mathbf{p}_v \mathbf{x}_t + S_t$. This shows that (cs.5) is also satisfied.

For the reverse, assume that (cs.2)-(cs.5) has a solution. We need to show that $D$ satisfies (iv.2)-(iv.3). If $\delta_t \mathbf{p}_t \mathbf{x}_t + S_t \geq \delta_t \mathbf{p}_t \mathbf{x}_v + S_v$, we have, from (cs.4), that $X_{t,v} = 1$. Condition (cs.3) then requires that $u_t \geq u_v$. This demonstrates condition (iv.3). Next, if $u_t \geq u_v$, then $X_{t,v} = 1$ (from (cs.2)) and using condition (cs.5) we see that $\delta_v \mathbf{p}_v \mathbf{x}_v + S_v \leq \delta_v \mathbf{p}_v \mathbf{x}_t + S_t$. As such, condition (iv.2) is also satisfied.

$\square$

## Appendix C: proof of Theorems 7 and 8

**Proof of Theorem 7.** Assume that $\{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t^*, \mathbf{y}_t^*\}_{t\in T}$ satisfies the conditions in **CS.WS**. Then there exist numbers $S_t$ and $u_t$, strict positive numbers $\delta_t$ and binary numbers $X_{t,v}$ such that:

$$\begin{aligned}
&S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t^* - \mathbf{y}_v^*),\\
&u_t - u_v < X_{t,v},\\
&(X_{t,v} - 1) \leq u_t - u_v,\\
&\delta_t \mathbf{p}_t(\mathbf{x}_t^* - \mathbf{x}_v^*) + (S_t - S_v) < X_{t,v} A_t,\\
&(X_{t,v} - 1) A_v \leq \delta_v \mathbf{p}_v(\mathbf{x}_t^* - \mathbf{x}_v^*) + (S_t - S_v).
\end{aligned}$$

Then, given that $\mathbf{x}_t^* = \mathbf{x}_t + \varepsilon_t$ and $\mathbf{y}_t^* = \mathbf{y}_t + \upsilon_t$, we obtain that:

$$\begin{aligned}
&S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v) + \delta_v \mathbf{q}_v(\upsilon_t - \upsilon_v),\\
&u_t - u_v < X_{t,v},\\
&(X_{t,v} - 1) \leq u_t - u_v,\\
&\delta_t \mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) < X_{t,v} A_t + \delta_t \mathbf{p}_t(\varepsilon_v - \varepsilon_t),\\
&(X_{t,v} - 1) A_v \leq \delta_v \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) + \delta_v \mathbf{p}_v(\varepsilon_t - \varepsilon_v).
\end{aligned}$$

From this, we see that $\max\{\max_{t,v} \mathbf{p}_v(\varepsilon_t - \varepsilon_v); \max_{t,v} \mathbf{q}_v(\upsilon_t - \upsilon_v)\}$ is a feasible solution for **OP.WS**, from which the theorem follows.

**Proof of Theorem 8.** Assume that $\{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t\in T}$ satisfies the conditions in **CS.WS**. Then there exist numbers $S_t$ and $u_t$, strict positive numbers $\delta_t$ and binary numbers $X_{t,v}$ such that the optimal solution $F$

of **OP.WS** satisfies:

$$
\begin{aligned}
&S_t - S_v \leq \delta_v \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v) + \delta_v F,\\
&u_t - u_v < X_{t,v},\\
&(X_{t,v} - 1) \leq u_t - u_v,\\
&\delta_t \mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) < X_{t,v}A_t + \delta_t F,\\
&(X_{t,v} - 1)A_v \leq \delta_v \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) + (S_t - S_v) + \delta_v F.
\end{aligned}
$$

This implies that for all $t$ and $v$:

$$
\begin{aligned}
&F \geq \frac{S_t - S_v}{\delta_v} - \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v),\\
&F \geq \mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_v) + \frac{S_t - S_v}{\delta_t} - X_{t,v}A_t,\\
&F \geq (X_{t,v} - 1)A_v - \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) - \frac{S_t - S_v}{\delta_v}.
\end{aligned}
$$

As the data set $\{\mathbf{p}_t \mathbf{q}_t, \mathbf{x}_t^*, \mathbf{y}_t^*\}_{t \in T}$ is not rationalizable, a similar reasoning as above shows that there must be a $t$ and $v$ such that at least one of the following inequalities holds

$$
\begin{aligned}
&\frac{S_t - S_v}{\delta_v} - \mathbf{q}_v(\mathbf{y}_t - \mathbf{y}_v) > \mathbf{q}_v(\upsilon_t - \upsilon_v),\\
&\mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_v) + \frac{S_t - S_v}{\delta_t} - X_{t,v}A_t > \mathbf{p}_t(\varepsilon_v - \varepsilon_v),\\
&(X_{t,v} - 1)A_v - \mathbf{p}_v(\mathbf{x}_t - \mathbf{x}_v) - \frac{S_t - S_v}{\delta_v} > \mathbf{p}_v(\varepsilon_t - \varepsilon_v).
\end{aligned}
$$

Therefore, we can conclude that at least

$$
F > \min \left\{ \min_{t,v} \mathbf{q}_v(\upsilon_t - \upsilon_v); \min_{t,v} \mathbf{p}_v(\varepsilon_t - \varepsilon_v)) \right\}.
$$

## Appendix D: supplement to Appendix A

**Case 1:** $(t(k,\ell), t(k',\ell'))$

$$
\begin{aligned}
\delta_{t(k,\ell)} \mathbf{q}_{t(k,\ell)} \left[\mathbf{y}_{t(k',\ell')} - \mathbf{y}_{t(k,\ell)}\right] = \;& \mathfrak{p} + (1-2) + (1 - 2 - 1/\mathfrak{p}) \left|\mathcal{G} - \mathcal{O}(t,k,\ell) \cap O(t,k',\ell')\right| \\
& +(2 - 1 + 1/\mathfrak{p}) \left|\mathcal{O}(t,k,\ell) \cap \left(\mathcal{G}_v - \mathcal{O}(t,k',\ell')\right)\right| \\
\geq \;& \mathfrak{p} - 1 - 6r = 14 + 35r - 6r > 3 \geq S_{t(k',\ell')} - S_{t(k,\ell)}
\end{aligned}
$$

**Case 2:** $(v(k,\ell), v(k',\ell'))$

$$
\begin{aligned}
\delta_{v(k,\ell)} \mathbf{q}_{v(k,\ell)} \left[\mathbf{y}_{v(k',\ell')} - \mathbf{y}_{v(k,\ell)}\right] &= \delta_{v(k,\ell)} \left[\mathfrak{p} + (1-2)\right] \\
&= \frac{14 + 35r - 1}{3 + 7r} > 3 \geq S_{v(k',\ell')} - S_{v(k,\ell)}
\end{aligned}
$$

**Case 3:** $(t(k,\ell), v(k',\ell'))$

$$\begin{aligned}\delta_{t(k,\ell)}\mathbf{q}_{t(k,\ell)}\left[\mathbf{y}_{v(k',\ell')}-\mathbf{y}_{t(k,\ell)}\right] = \;& \begin{array}{l}\mathfrak{p}+(1/\mathfrak{p})\left|\mathcal{O}(t,k,l)\cap\{g(v,k',\ell')\}\right| \\ +(1-2)\left|(\mathcal{G}_v-\mathcal{O}(t,k,\ell))\cap\{g(v,k',\ell'\}\right| \\ +(2-1+1/\mathfrak{p})\left|\mathcal{O}(t,k,l)\cap\left(\mathcal{G}_v-\{g(v,k',\ell')\}\right)\right|\end{array} \\ & \geq \mathfrak{p}-1 = 14+35r-1 > 3 \geq S_{v(k',\ell')}-S_{t(k,\ell)}\end{aligned}$$

**Case 4:** $(v(k,\ell), t(k',\ell'))$ and the $k$th literal in the $\ell$th clause and the $k'$th literal in the $\ell'$th clause have opposite signature. First of all, observe that $-1 \geq S_{t(k',\ell')} - S_{v(k,\ell)}$.

Then:

$$\begin{aligned}\delta_{v(k,\ell)}\mathbf{q}_{v(k,\ell)}\left[\mathbf{y}_{t(k',\ell')}-\mathbf{y}_{v(k',\ell')}\right] &= \delta_{v(k,\ell)}\left[\begin{array}{l}\mathfrak{p}(1-1/\mathfrak{p}-1)+(1-2) \\ +(1-2-1/\mathfrak{p})\left|(\mathcal{G}_v-\{g(v,k,\ell)\})\cap\mathcal{O}(t,k',\ell')\right|\end{array}\right] \\ &\geq \frac{-1-1-6r}{3+7r} > -1 \geq S_{t(k',\ell')}-S_{v(k,\ell)}\end{aligned}$$

**Case 5:** $(v(k,\ell), t(k',\ell'))$ and the $k$th literal in the $\ell$th clause and the $k'$th literal in the $\ell'$th clause do not have opposite signature.

$$\begin{aligned}\delta_{v(k,\ell)}\mathbf{q}_{v(k,\ell)}\left[\mathbf{y}_{t(k',\ell')}-\mathbf{y}_{v(k',\ell')}\right] &= \delta_{v(k,\ell)}\left[\begin{array}{l}\mathfrak{p}+(1-2) \\ +(1-1/\mathfrak{p}-2)\left|(\mathcal{G}_v-\{g(v,k,\ell)\})\cap\mathcal{O}(t,k',\ell')\right|\end{array}\right] \\ &\geq \frac{\mathfrak{p}-1-6r}{3+7r} = \frac{13+29r}{3+7r} > 3 \geq S_{t(k',\ell')}-S_{v(k,\ell)}\end{aligned}$$

## Appendix E: supplement to Appendix A

**Case 1:** $(t(k,\ell), t(k',\ell'))$

$$\begin{aligned}&\mathbf{p}_{t(k,\ell)}\left[\mathbf{x}_{t(k,\ell)}-\mathbf{x}_{t(k',\ell')}\right]+\frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)}-S_{t(k',\ell')}\right] \\ &= -\mathfrak{z}+(3-2)+\left[S_{t(k,\ell)}-S_{t(k',\ell')}\right] \\ &\leq -16-42r+1+3<0\end{aligned}$$

**Case 2:** $(t(k,\ell), v(k',\ell'))$ with $\ell \neq \ell'$.

$$\begin{aligned}&\mathbf{p}_{t(k,\ell)}\left[\mathbf{x}_{t(k,\ell)}-\mathbf{x}_{v(k',\ell')}\right]+\frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)}-S_{v(k',\ell')}\right] \\ &= -\mathfrak{z}+(3-1)+(1-1/\mathfrak{y}-2)+(2-1)+\left[S_{t(k,\ell)}-S_{v(k',\ell')}\right] \\ &\leq -16-42r+2+1+3<0\end{aligned}$$

**Case 3:** $(t(k,\ell), v(k,\ell))$

$$\begin{aligned}
&\mathbf{p}_{t(k,\ell)}\left[\mathbf{x}_{t(k,\ell)} - \mathbf{x}_{v(k,\ell)}\right] + \frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)} - S_{v(k,\ell)}\right] \\
&= -\mathfrak{z} + (3-1) + (1 - 1/\mathfrak{y} - 1) + \frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)} - S_{v(k,\ell)}\right] \\
&\leq -16 - 42r + 2 + 3 < 0
\end{aligned}$$

**Case 4:** $(t(k,\ell), v(k \oplus 2,\ell))$

$$\begin{aligned}
&\mathbf{p}_{t(k,\ell)}\left[\mathbf{x}_{t(k,\ell)} - \mathbf{x}_{v(k\oplus 2,\ell)}\right] + \frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)} - S_{v(k\oplus 2,\ell)}\right] \\
&= -\mathfrak{z} + (3-1) + (1 - 1/\mathfrak{y} - 2) + 1 + \frac{1}{\delta_{t(k,\ell)}}\left[S_{t(k,\ell)} - S_{v(k\oplus 2,\ell)}\right] \\
&\leq -16 - 42r + 2 + 1 + 3 < 0
\end{aligned}$$

**Case 5:** $(v(k,\ell), v(k',\ell'))$

$$\begin{aligned}
&\mathbf{p}_{v(k,\ell)}\left[\mathbf{x}_{v(k,\ell)} - \mathbf{x}_{v(k',\ell')}\right] + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{v(k',\ell')}\right] \\
&= -\mathfrak{y} + (2-1)\frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{v(k',\ell')}\right] \\
&\leq -11 - 29r + 1 + (3+7r)3 < 0
\end{aligned}$$

**Case 6:** $(v(k,\ell), t(k',\ell'))$ with $\ell \neq \ell'$.

$$\begin{aligned}
&\mathbf{p}_{v(k,\ell)}\left[\mathbf{x}_{v(k,\ell)} - \mathbf{x}_{t(k',\ell')}\right] + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k',\ell')}\right] \\
&= (1-3) + (3-2) - \mathfrak{y} + (2 - 1 + 1/\mathfrak{y}) + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k',\ell')}\right] \\
&\leq -11 - 29r + 1 + (3+7r)3 < 0
\end{aligned}$$

**Case 7:** $(v(k,\ell), t(k \oplus 1,\ell))$

$$\begin{aligned}
&\mathbf{p}_{v(k,\ell)}\left[\mathbf{x}_{v(k,\ell)} - \mathbf{x}_{t(k\oplus 1,\ell)}\right] + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k\oplus 1,\ell)}\right] \\
&= (1-3) + (3-2) - \mathfrak{y} + (2 - 1 + 1/\mathfrak{y}) + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k\oplus 1,\ell)}\right] \\
&\leq -11 - 29r + 1 + (3+7r)3 < 0
\end{aligned}$$

**Case 8:** $(v(k,\ell), t(k \oplus 2, \ell))$

$$
\begin{aligned}
&\mathbf{p}_{v(k,\ell)}\left[\mathbf{x}_{v(k,\ell)} - \mathbf{x}_{t(k\oplus 2,\ell)}\right] + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k\oplus 2,\ell)}\right] \\
&= (1-2) - \mathfrak{y} + (2 - 1 + 1/\mathfrak{y}) + \frac{1}{\delta_{v(k,\ell)}}\left[S_{v(k,\ell)} - S_{t(k\oplus 2,\ell)}\right] \\
&\leq -11 - 29r + 1 + (3+7r)3 < 0
\end{aligned}
$$