LISER

LUXEMBOURG INSTITUTE OF SOCIO-ECONOMIC RESEARCH

n° 2017-01 January 2017

WORKING PAPERS

What types are there?

Sam **COSAERT** [1]

[1] LISER, Luxembourg

What types are there?

Sam Cosaert*

Abstract

Preferences differ in the population, and this heterogeneity may not be adequately described by observed characteristics and additive error terms. As a first contribution, this study shows that preference heterogeneity can be represented graphically by means of violations of the Weak Axiom of Revealed Preference (WARP), and that computing the minimum number of partitions necessary to break all WARP violations in the sample is equivalent to computing the chromatic number of this graph. Second, the study builds the bridge between revealed preference theory and cluster analysis to assign individuals to these partitions (i.e. preference types). The practical methods are applied to Dutch labour supply data, to recover reservation wages of individuals who belong to particular preference types.

Keywords: preference heterogeneity; chromatic number; revealed preference; labour supply; constrained clustering.

JEL codes: C14, C38, C44, D12, D13

*Luxembourg Institute of Socio-Economic Research (LISER) and Department of Economics, University of Leuven (KU Leuven). Porte des Sciences 11, L-4366 Esch-sur-Alzette, Luxembourg. E-mail: sam.cosaert@liser.lu

1 Introduction

In many datasets on consumption and labour supply, individuals are observed only once. The estimation of demand functions and labour supply functions then requires the pooling of individuals. Although these estimation techniques take variation in observed characteristics into account, the variation in demand and labour supply outcomes cannot be fully captured by observed characteristics. This is also reflected in the typically low R^2 : only a small fraction of the variation in outcomes is captured by variation in observed characteristics. In other words, additive error terms are too restrictive to properly address unobserved preference heterogeneity. Lewbel (2001), for instance, argued that implementing additive errors comes very close to a representative agent assumption.

Instead of pooling all individuals and adding error terms to the demand or labour supply functions, the current study follows a different approach. More specifically, this study partitions a dataset from the Longitudinal Internet Studies for the Social Sciences in the minimum number of sets so that the utility maximisation hypothesis holds simultaneously for all individuals within a set. The underlying principle is Occam's razor, which states that a simpler theory is preferable to a complex one, thereby promoting testability. For this reason, this study focuses on the minimum number of utility functions necessary to rationalise the relevant consumption and labour supply choices in the sample. The true number of utility functions may be higher, but this is irrelevant from the perspective of rationalising behaviour. By restricting the number of utility functions, more convincing recovery and prediction exercises may be executed, given that multiple individuals have the same utility function. To this end, a proposal by Kalai et al. (2002) is translated to deal with consumption and labour supply data from standard budget surveys.

Related literature Kalai et al. (2002) proposed to partition a combination of (general) choice sets in the minimum number of partitions so that the choices within each partition could be rationalised by the same 'rationale'. The concept of rationality referred, in their setting, to selecting the elements that corresponded to some unobserved binary relation. In this way, the authors presented a framework to deal with 'bounded rationality', in particular, violations of the 'Independence of Irrelevant Alternatives' (IIA) Axiom. Adding or removing elements from the choice sets may lead to inconsistent choices, which are explained by the fact that different choice sets are subject to different rationales.

Apesteguia and Ballester (2010) and Demuynck (2011) investigated the computational complexity of rationalising choices by means of a minimum number of rationales and a fixed number of rationales, respectively. Both studies concluded that the problem is computationally very difficult (e.g. NP–complete). For this reason, Apesteguia and Ballester (2010) linked the rationalisability question to a graph problem, and suggested that analyses of the minimum number of rationales could borrow from algorithms from the computer science and operations research literature.

The studies of Apesteguia and Ballester (2010), Demuynck (2011) and Kalai et al. (2002) are purely theoretical. Their framework is very general, in the sense that choice correspondences may be multi–valued and choice sets must not necessarily correspond to linear budget sets. Furthermore, the focus of Apesteguia and Ballester (2010), Demuynck (2011) and Kalai et al. (2002) is on bounded rationality, in particular, violations of the IIA assumption. The current study brings the idea of the 'minimum number of rationales' into practice, by computing the minimum number of utility functions necessary to describe all consumption and labour supply choices in a sample of heterogeneous agents.

Revealed preference To compute the minimum number of utility functions, axioms from revealed preference theory are used, in particular, the Weak Axiom of Revealed Preference (WARP). Revealed preference axioms–in the spirit of Samuelson (1938)–are attractive in the current setting. They impose consistency conditions on observed consumption and labour supply choices independent of a functional form for the preferences. Conveniently, Afriat (1967), Diewert (1973), and Varian (1983) formulated a so called restricted–domain version of revealed preference theory, which can be applied when only a finite number of price–quantity observations is available (as opposed to the entire demand or labour supply function, or choice correspondence). This is particularly useful to deal with a (finite) sample of individuals in which each individual made a specific consumption and labour supply choice (in terms of the consumption expenditures and the number of hours worked).

In this sense, the current study is similar in nature to the one by Crawford and Pendakur (2013), who computed *lower* and *upper* bounds on the (minimum) number of preference types among consumers. Based on a (random) sequence of observations, Crawford and Pendakur (2013)'s algorithms constructed partitions of the sample that are consistent with the utility maximisation hypothesis. Because the outcome of these algorithms depends on the (random) ordering of observations, the authors proposed to run the algorithms multiple times. Finally, Crawford and Pendakur (2013) used their sample to infer that the number of types (i.e. sets of individuals whose behaviour may be described by a single utility function) in the population is at most 12. It is worth noting that Crawford and Pendakur (2013) used the Generalised Axiom of Revealed Preference (GARP), which is necessary and sufficient for behaviour to satisfy the utility maximisation hypothesis. However, given that the GARP also implements transitivity, these conditions are known to be computationally very demanding.

In a setting with only two goods (i.e. if transitivity is not important, see Rose (1958)), it has been shown that WARP is equivalent to the Strong Axiom of Revealed Preference (SARP). More generally, Cherchye et al. (2015) formulated conditions on prices and incomes under which transitivity has no specific testable implications. Whenever these conditions hold, SARP is equivalent to WARP. Finally, the current study takes the hypothesis of utility maximisation as given. Rather than treating rationality as an assumption which is subject to empirical testing, rationality is the identifying assumption. The focus is on recovery of preference types rather than testing of the rationality hypothesis.

Contributions The current study complements the paper by Crawford and Pendakur (2013) in at least two ways.

First, a graph–theoretical representation of preference heterogeneity is provided for the current consumption and labour supply framework. This follows the suggestion by Apesteguia and Ballester (2010), who argued that the literature on graph theory may be a starting point to address the computational complexity of the problem. The graphical representation is based on the Weak Axiom of Revealed Preference (WARP). More specifically, the vertices in this graph correspond to all individuals and the edges correspond to pairwise inconsistencies of these individuals with WARP. It is shown that the *chromatic* number associated with this graphical representation is equivalent to the minimum partition necessary to remove WARP violations, and thereby also bounds the minimum number of utility functions in the sample (from below). Furthermore, the chromatic number equals the minimum number of utility functions when WARP is equivalent to SARP. There exist numerous algorithms to compute the chromatic number, both exactly and approximately (e.g. by a greedy algorithm). This opens the door for many new applications of revealed preference theory and operations research to deal with interpersonal preference variation. In this sense, the current study extends results by Talla Nobibon et al. (2011) who provided approximation algorithms to deal with intra–household preference heterogeneity (i.e. between two members of the household) rather than inter–household heterogeneity.

Second, Crawford and Pendakur (2013) focused on computing the number of preference types in the sample. The current study also deals with the assignment of individuals to specific preference types. To this end, the nonparametric (revealed preference) conditions (i.e. WARP) are complemented with an objective function that minimises the within–type variation in observed characteristics (i.e. the within–cluster sum of squares corresponding to demographic variables). This allows researchers to select a particular partitioning of the sample, depending on the desired characteristics and criteria, without violating the consistency requirements from revealed preference theory. As a result, this so called revealed–

preference–consistent clustering combines the empirical appeal of cluster analysis and the theoretical robustness of revealed preference.

Data The methods are applied to a sample from the Longitudinal Monitoring Survey for the Social Sciences (LISS), which contains data on consumption and labour supply choices of Dutch households. This application is interesting for several reasons. First, as in many labour supply studies, the household members have preferences over private consumption and leisure (see e.g. Blundell et al. (2007) and Cherchye and Vermeulen (2008)). The issue of transitivity is irrelevant in this two–goods setting. Second, labour supply data contain rich cross–sectional variation in wages, which can be better exploited by the novel methodology. Alternative methods such as stochastic revealed preference conditions (in the spirit of McFadden (2005) and Falmagne (1978)) project all outcomes per cross–section on a single budget hyperplane, as a result of which a considerable amount of information is lost. Finally, the labour supply application will clearly indicate that the preference types have a large impact on important estimates such as reservation wages. Indeed, the reservation wages vary considerably across the different preference types. This emphasises the relevance of the presented methods.

The rest of the paper unfolds as follows. Section 2 contains the main methodological contributions (Subsections 2.2–2.3). Section 3 presents an application to Dutch labour supply data from the LISS. Section 4 concludes.

2 Preference heterogeneity: graph–theoretical approach

Subsection 2.1 sets the stage by introducing revealed preference theory and the notion of the minimal level of preference heterogeneity.[1] The contributions of this study are in Subsections 2.2 and 2.3. Subsection 2.2 presents an equivalent graph–theoretical solution concept, that has received much attention in the operations research and computer science literature. Subsection 2.3 opens the door for the recovery of specific preference types by means of revealed–preference–consistent clustering.

[1] Strategies to recover the 'minimal level of heterogeneity' using revealed preference techniques have become increasingly popular, see e.g. Crawford and Pendakur (2013) and Adams et al. (2015). Adams et al. (2015) used revealed preference theory to recover the minimal amount of taste variation necessary to rationalise patterns of tobacco consumption. These authors also addressed 'taste variation' across the sample, but they imposed more structure on individual utility functions. In particular, taste heterogeneity is captured by linear perturbations to a base utility function.

2.1 Revealing preference heterogeneity

Set-up Assume a sample of individuals $i \in N$ and let C represent a collection of goods (which may include consumption and leisure). Per individual $i \in N$, the econometrician observes a vector of observed attributes a_i and a dataset $S_i = \{\mathbf{p}_i, \mathbf{q}_i\}$. Dataset S_i consists of a price vector $\mathbf{p}_i \in \mathbb{R}^{|C|}_{++}$, containing the strictly positive prices of $|C|$ commodities, and a consumption vector $\mathbf{q}_i \in \mathbb{R}^{|C|}_{+}$, containing i's consumption. The individual's total expenditure is given by $y_i = \mathbf{p}_i'\mathbf{q}_i$. The methods in this paper take rationality of all individuals $i \in N$ as given, i.e.

$$\mathbf{q}_i = \arg\max_{\mathbf{q}} U_i(\mathbf{q}) \text{ s.t. } \mathbf{p}_i'\mathbf{q} \leq y_i.$$

Given that utility functions $U_i(\cdot)$ are unobserved, rationality of the full sample $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ requires that there exist utility functions such that the above utility maximisation problem is solved for all individuals. However, the number of utility functions in the sample can be bounded, see Definition 1.

Definition 1 *$\tau-$Rationalisability.* Given a dataset $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ of $|N|$ rational individuals. Dataset S is *$\tau-$rationalisable* if there exists a set of utility functions $\tilde{U}_t(\cdot)$ with $t \leq \tau$ such that for each $n \in N$:

$$\exists t : \mathbf{q}_n = \arg\max_{\mathbf{q}} \tilde{U}_t(\mathbf{q}) \text{ s.t. } \mathbf{p}_n'\mathbf{q} \leq y_n.$$

Definition 1 introduces a formal measure τ of the number of utility functions in the sample. On the one hand, $1-$rationalisability requires that S is rationalised using only one utility function $\tilde{U}(\cdot) = U_n(\cdot)$ for all $n \in N$. This imposes strong restrictions on the shape of $\tilde{U}(\cdot)$. After all, $\tilde{U}(\cdot)$ must be such that the observations $\{\mathbf{q}_n\}_{n \in N}$ maximise $|N|$ different optimisation problems simultaneously–using the same utility function. In cases where $1-$rationalisability is violated, one should avoid pooling all individuals of the cross section. The resulting recovery and prediction of $\mathbf{q}$ in new price–income situations may be biased. At the other extreme, $|N|-$rationalisability puts very little structure on the utility functions. Each individual can belong to a distinct preference type $\tilde{U}_1(\cdot) \neq ... \neq \tilde{U}_{|N|}(\cdot)$. The resulting model is highly permissive, as there is only one observation per utility function. Due to the lack of empirical bite, such model is unapt for convincing recovery and prediction analyses.

For this reason, the focus of the current paper is on the minimum number τ^* of utility functions necessary for rationality in the sense of Definition 1. More generally, the focus on the minimal degree of heterogeneity fits with the idea of Occam's razor, in that the 'most simple and parsimonious' model is selected that still explains the data. A formal description of τ^* is in Definition 2.

Definition 2 τ^*. Given a dataset $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ of $|N|$ rational individuals. τ^* is the minimum number of preference types in the sample if S is τ^*-*rationalisable* but not $(\tau^* - 1)-$*rationalisable.*

Revealed preference In a first step, the abovementioned problem is translated in terms of the Weak Axiom of Revealed Preference (WARP). Samuelson (1938) and Houthakker (1950) have shown that consistency with WARP is a necessary condition for the existence of a well–behaved utility function that rationalises the underlying data, in the sense that the data are utility maximising. WARP is computationally much easier compared to the Strong Axiom of Revealed Preference (SARP) which also imposes transitivity on the individual preferences. Furthermore, WARP is equivalent to SARP (let $WARP = SARP$ indicate that transitivity has no separate testable implications) in two–goods settings, or in settings described in Cherchye et al. (2015).

The revealed preference axioms will allow us to bound the number of utility functions regardless of their parametric structure. Definition 3 formally presents the notions of partitioning, WARP–partitions and $\hat{\tau}$.

Definition 3 Given a dataset $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ of $|N|$ individuals.

- A combination of sets $V_1, ..., V_T$ provides *a partitioning* of N if
  - $\forall s, t \leq T : V_s \cap V_t = \varnothing$,
  - $\forall n \in N$ it holds that $n \in V_1 \cup ... \cup V_T$.
- A specific partition V_t of N is *a WARP–partition* if and only if $\forall\ i, j \in V_t : i$ and j satisfy WARP, that is,
$$\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j \;\Rightarrow\; \mathbf{p}_j'\mathbf{q}_j < \mathbf{p}_j'\mathbf{q}_i.$$
- $\hat{\tau}$ is the *minimum number* of WARP–partitions $V_1, ..., V_{\hat{\tau}}$ of N. Formally, N can be partitioned in $\hat{\tau}$ but not in $\hat{\tau} - 1$ WARP–partitions.

The first bullet point states that a partitioning is an *exact set cover* of N using subsets $V_1, ..., V_T$.

The second bullet point defines a WARP–partition as a subset V_t in which the Weak Axiom of Revealed Preference (WARP) holds for all $i, j \in V_t$. Assume that WARP is violated: $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ and $\mathbf{p}_j'\mathbf{q}_j \geq \mathbf{p}_j'\mathbf{q}_i$. Then $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ implies that the consumption bundle of j was affordable for i. Agent i preferred his bundle over the bundle of j. Similarly, $\mathbf{p}_j'\mathbf{q}_j \geq \mathbf{p}_j'\mathbf{q}_i$ implies that the consumption bundle of agent i was affordable for j. Thus, agent j preferred his bundle over the corresponding bundle of i. However, there is no utility function $\tilde{U}(\cdot)$ $(= U_i(\cdot) = U_j(\cdot))$ that simultaneously attributes (strictly) higher utility to i and j. Conditional on the assumption of utility maximisation, it must be the case that agents i and j have different preferences $U_i(\cdot) \neq U_j(\cdot)$. For this reason, a necessary condition for a set of individuals $i, j \in V_t$ to behave consistently with a homogeneous utility function is that V_t is a WARP–partition.[2]

The third bullet point defines $\hat{\tau}$ as the minimum number of WARP–partitions of N. Proposition 1 combines the fact that N cannot be partitioned in less than $\hat{\tau}$ WARP–partitions and that WARP is nonetheless necessary for the existence of a utility function so that the utility maximisation hypothesis holds simultaneously for all individuals within the set (Samuelson (1938) and Houthakker (1950)).

Proposition 1 Consider a dataset $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ of $|N|$ rational individuals and the corresponding values τ^* and $\hat{\tau}$. Then $\tau^* \geq \hat{\tau}$.

If $WARP = SARP$ this can be strengthened to $\tau^* = \hat{\tau}$. While τ^* is a relatively theoretical concept, $\hat{\tau}$ is defined in terms of observables. Yet, the computation of $\hat{\tau}$ is difficult in practice (as shown in Subsection 2.2). Let us now introduce the main contributions of this paper. Subsection 2.2 shows that $\hat{\tau}$ is equivalent to the so called chromatic number of the graphical representation of WARP violations. The computation of the chromatic number has been extensively studied in the operations research and computer science literature, i.e. several algorithms have been proposed. Next, Subsection 2.3 introduces a program to 'construct' WARP–partitions of the data that moreover minimise the intra–partition variation in observed characteristics. Otherwise stated, this program computes revealed–preference–consistent clusters.

2.2 The chromatic number as measure of preference heterogeneity

The following method recovers $\hat{\tau}$, i.e. the minimum number of WARP–partitions in the sample. This is based on a graph–theoretical representation of the revealed preference conditions.[3] In particular, it is shown that the *chromatic number* (explained below) applied to

[2]Alternatively, SARP requires that for all $i, k, ..., z, j \in V_t : \mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_k, ..., \mathbf{p}_z'\mathbf{q}_z \geq \mathbf{p}_z'\mathbf{q}_j \;\Rightarrow\; \mathbf{p}_j'\mathbf{q}_j < \mathbf{p}_j'\mathbf{q}_i$, thereby excluding cycles of observations. Houthakker (1950) has shown that consistency with SARP is both a necessary and a sufficient condition for the existence of a well–behaved utility function that is maximised by the underlying observations.

[3]This is not the first combination of revealed preference and graph theory. Recently, Dean and Martin (2015) used insights from the computer science and operations research literature to compute the 'mini-

the graph of WARP violations is exactly equal to $\hat{\tau}$.

Example Consider the following simple example. A sample consists of six independent individuals, represented as vertices in Figure 1. A pair of vertices (i, j) is connected if and only if the observations of the corresponding individuals i and j violate WARP. As a further implication, i and j belong to separate preference types.

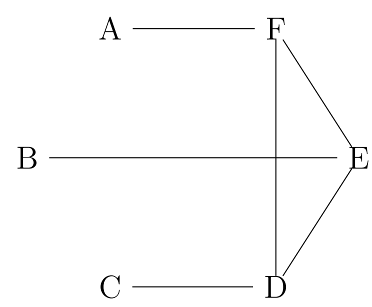

Figure 1: Graph with individuals as vertices and pairwise violations of revealed preference axioms as edges

As such, the graph in Figure 1 implies that individual A cannot belong to the same set as F, B cannot belong to the same set as E, C cannot belong to the same set as D, and finally, D, E, and F must be in three distinct sets. The question is how many WARP–partitions are minimally required to cover all individuals (vertices in Figure 1). By starting from individual A, it is possible to form a WARP–partition of three individuals (A, B, and C). Individuals D, E, and F are automatically excluded from this set, as it always contains at least one of their adjacent individuals. Thus, D, E, and F constitute three distinct preference types, resulting in a total of four types. However, this number overestimates the true minimum number of WARP–partitions. After all, we can match A and E, B and D, and C and F, resulting in a total of only three types.

This example shows that the minimum number of WARP–partitions is easily overestimated by algorithms that follow some random ordering of the data. In a sample of six people, it is fairly straightforward to solve the problem by visual inspection. In a sample of more than 100, the problem is very complex. A brute force algorithm would then require

mum' cost of breaking revealed preference cycles in consumer data. While these studies used the minimum set covering problem to compute the minimum cost associated with the observations to be removed from the data, the current study computes the chromatic number in order to bound the minimum number of distinct preference types in a sample. Similarly, Smeulders et al. (2014) and Smeulders et al. (2015) used graph–theoretical arguments to formulate efficient methods to test consistency with the collective model and compute goodness–of–fit measures, respectively.

the analysis of 2^{100} subsets of the data. The first contribution of this study is that it links the abovementioned partitioning problem to the computation of a number that has received much attention in the operations research and computer science literature.

The chromatic number Consider the following two–step procedure to identify (bounds on) τ^*. In the first step, graph G is constructed, in which each vertex i represents an individual and each connected pair of vertices (i, j) indicates a violation of WARP. Edges (i, j) indicate that i cannot belong to the same type as j. The second step computes the so called *chromatic number* of graph $G : \chi(G)$. The chromatic number gives the smallest number of colours (labels) necessary to obtain a (proper) *vertex colouring*. A (proper) vertex colouring is an assignment of colours to the different vertices in G in such a way that all adjacent vertices obtain different colours. The chromatic number is always bounded between 1 (i.e. in an edgeless graph) and $|N|$ (i.e. in a complete graph). The following shows, formally, that $\hat{\tau} = \chi(G)$.

Proof. Proof of $\hat{\tau} = \chi(G)$. Let G be constructed in line with the abovementioned approach.

1. Suppose that $\hat{\tau} > \chi(G)$. By definition of $\hat{\tau}$, this implies that there exists (at least) one pair of individuals (i, j) who got the same colour / label $t \leq \chi(G)$ and for whom $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ and $\mathbf{p}_j'\mathbf{q}_j \geq \mathbf{p}_j'\mathbf{q}_i$. By construction of G, the latter implies that i and j are adjacent. By definition, however, a vertex colouring cannot include adjacent individuals, so that the colouring in which i and j get the same label cannot exist. Thus, $\hat{\tau} \leq \chi(G)$.

2. Suppose on the other hand that $\hat{\tau} < \chi(G)$. By construction of G, a (proper) vertex colouring is necessary for the construction of WARP–partitions. By definition of the chromatic number, $\chi(G)$ is the smallest number of colours/labels to obtain a (proper) vertex colouring. Hence, $\hat{\tau} \geq \chi(G)$.

■

Thus, the chromatic number of graph G ($\chi(G)$) is equivalent to the minimum number of WARP–partitions ($\hat{\tau}$) in the data. Furthermore, $\hat{\tau}$ bounds τ^* by Proposition 1. The first contribution of this study is formally presented in Proposition 2.

Proposition 2 Given a dataset $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ of $|N|$ rational individuals and the corresponding value τ^*. Let $\chi(G)$ be the chromatic number associated with graph G, with G consisting of vertices $i, j \in N$ and edges (i, j) if $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ and $\mathbf{p}_j'\mathbf{q}_j \geq \mathbf{p}_j'\mathbf{q}_i$.

Then $\tau^* \geq \chi(G)$.

If $WARP = SARP$ this can be strengthened to $\tau^* = \chi(G)$. The intuition behind this proposition is straightforward. Suppose that the number of WARP violations is high. This will create a large number of edges in graph G. Adjacent nodes cannot be pooled, which raises the chromatic number $\hat{\tau} = \chi(G)$ and hence also τ^*. Three final remarks are in order here.

1. If $WARP \neq SARP$, $\chi(G)$ provides only a lower bound on τ^*. The reason is that edges e correspond to violations of WARP, that is, between any pair of individuals. By applying the vertex colouring problem, these violations are eliminated. However, this does not necessarily break violations of SARP. It is possible that individuals i, j, z are characterised by a violation of SARP (and not by a violation of WARP), which can occur when the number of goods in the analysis is more than two. In this case, the behaviour of i, j, z is such that $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ and $\mathbf{p}_j'\mathbf{q}_j \geq \mathbf{p}_j'\mathbf{q}_z$ and $\mathbf{p}_z'\mathbf{q}_z \geq \mathbf{p}_z'\mathbf{q}_i$. Notice however that consistency (i.e. with the GARP/SARP instead of WARP) of the individuals within preference types, and hence the prevalence of SARP violations, can easily be examined in a final step (i.e. after the vertex colouring). Furthermore, in the application in Section 3, $WARP = SARP$, so that $\chi(G)$ effectively corresponds to τ^*.

2. The computation of the chromatic number $\chi(G)$ for an arbitrary graph G is known to be NP–hard, see Karp (1972). Exact algorithms may be feasible for small to moderately sized graphs only. For this reason, the greedy algorithm has been proposed, which approximates the chromatic number (from above). Both algorithms are available in the MathGraph package, using the 'color' command–specifying either 'optimal' or 'greedy' (default). In the current application (Section 3, with $|N| = 106$) the greedy algorithm actually gives the true chromatic number. Furthermore, the equivalence of $\hat{\tau}$ and $\chi(G)$ implies that the literature on approximations of $\chi(G)$ is also useful in this revealed preference framework. It has for instance been shown that the chromatic number is bounded by the number of edges in G, $|E|$, in particular, $\chi(G) \cdot (\chi(G) - 1) \leq 2\,(|E|)$. As a consequence, $\hat{\tau} \cdot (\hat{\tau} - 1)$ is always bounded from above by the number of WARP violations multiplied by two. More generally, there is a vast literature on the computation of $\chi(G)$ for graphs G with specific properties.

3. A vertex colouring immediately provides an *exact set cover* of all individuals in the cross section. This does not shed light on the maximum number of individuals who can belong to one preference type, i.e. the largest possible WARP–partition. To identify the largest WARP–partition, the *maximum independent set* problem could be considered. This problem finds, in graph G, the largest possible set $\tilde{N}$ of vertices that are not adjacent.

2.3 Revealed–preference–consistent clustering

The previous subsection introduced a practical method to compute $\hat{\tau}$, by using the chromatic number applied to the graph of WARP violations. The current subsection addresses two further concerns. The first concern is that the WARP–partitions recovered in Subsection 2.2 are not necessarily unique. There may be a finite number of (proper) vertex colourings containing $\chi(G)$ colours, and therefore also a finite number of possible partitionings of N in WARP–partitions. More specifically, the number of partitionings of G in $\chi(G)$ sets is equivalent to the chromatic polynomial of G, $P(G, \chi(G))$, divided by the number of permutations of the colours, $\chi(G)!$. This number can be very large. In tree graphs–with exactly one path between each pair of vertices–for instance, the number of different partitionings is $\frac{\chi(G)\cdot(\chi(G)-1)^{|N|-1}}{\chi(G)!}$. The second concern is that observed characteristics, other than the chosen consumption and labour supply, are relevant when constructing preference types. The impact of preference heterogeneity on consumption and labour supply was filtered by breaking WARP violations and by allowing for multiple utility functions. The question remains whether this heterogeneity is (partly) captured by observed characteristics, or completely unobserved.

Let a_i reflect an $h-$dimensional vector of observed characteristics of individual i and let ϕ_t represent a vector of mean observed characteristics in partition V_t. The aim is to create clusters $V_1, ..., V_\tau$ in a way that minimises the within–cluster distance between a_n and ϕ_t ($d(a_n, \phi_t)$) under the restriction that the observations within each cluster can be described by a single, homogeneous utility function $U(\cdot)$. In this sense, the clusters $V_1, ..., V_\tau$ define preference types that are also as homogeneous as possible in terms of observed characteristics a_n. To bring this into practice, it is useful to adopt the following notation. Let s_{nt} denote binary variables that indicate whether individual n belongs to subset (or cluster) t. Each variable s_{nt} indicates whether individual n belongs to the $t-th$ (with $t \leq \tau$) cluster ($s_{nt} = 1$) or not ($s_{nt} = 0$).

As a first step, note that each WARP–partition satisfies an important necessary condition for the existence of a utility function that rationalises the within–cluster consumption and labour supply choices. Thus, the clusters V_t must at least be WARP–partitions. By definition, WARP–partitions cannot include pairwise violations of WARP. Therefore,

$$\begin{bmatrix} \mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j \\ \mathbf{p}_j'\mathbf{q}_j \geq \mathbf{p}_j'\mathbf{q}_i \end{bmatrix} \Rightarrow s_{it} + s_{jt} \leq 1. \tag{1}$$

In other words, if the pair of observations (i, j) is characterised by a WARP violation, the corresponding individuals i and j cannot belong to the same WARP–partition. Therefore, they must have distinct preference types. Formally, $i, j \in V_t$ is impossible.[4]

As a second step, each individual $n \in N$ must belong to exactly one partition. For this reason,

$$\sum_{t \leq \tau} s_{nt} = 1. \tag{2}$$

By definition of $\hat{\tau}$ it follows that conditions (1) and (2) cannot simultaneously hold unless $\tau \geq \hat{\tau}$, i.e. the number of subsets used to partition S must be (at least) as large as the minimum number of WARP–partitions (and hence the chromatic number $\chi(G)$ of the corresponding graph).

As a final step, a distance function $d(a_n, \phi_t)$ is added to the problem, in order to select the set of clusters / partitions $V_1, ..., V_\tau$ that minimise the variation in observed characteristics. A well–known parametrisation of this function is $d(a_n, \phi_t) = ||a_n - \phi_t||^2$, with $\sum_{t \leq \tau} \sum_{n \in V_t} (||a_n - \phi_t||^2)$ the within–cluster sum of squares (WCSS). The structure imposed on the distance function might seem at odds with the nonparametric formulation of utility functions and with the general, nonadditive forms of preference heterogeneity. It is important to note that the choice of the distance function is not crucial for the presented methodology to work. However, the chosen specification allows for a direct comparison with unconstrained $k-means$ clustering. Finally, the nonparametric methodology eliminated very general forms of preference heterogeneity. In order to select one (proper) vertex colouring as the desired set of preference types, a stronger selection criterion is desirable. Minimising the WCSS is equivalent to minimising the following objective function:

$$\sum_{n \in N} \sum_{t \leq \tau} \left(||a_n - \phi_t||^2 \right) s_{nt}.$$

The problem is now completely defined in terms of binary variables s_{nt}. Thus, clustering the 'nearest' individuals–by minimising the WCSS–in a way that is consistent with revealed preference theory, boils down to solving Problem 1.

[4] When $WARP \neq SARP$, conditions (1) could be modified as follows:

$$\begin{bmatrix} \mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_k \\ \mathbf{p}_k'\mathbf{q}_k \geq \mathbf{p}_k'\mathbf{q}_l \\ ... \\ \mathbf{p}_z'\mathbf{q}_z \geq \mathbf{p}_z'\mathbf{q}_i \end{bmatrix} \Rightarrow s_{it} + s_{kt} + s_{lt} + ... + s_{zt} \leq \alpha$$

with α the number of observations in the cycle minus one. Of course, implementing such set of conditions would make the program computationally much harder. As indicated earlier, $WARP = SARP$ in Section 3.

Problem 1 *RP–consistent clustering of $|N|$ individuals in τ clusters.* Individuals are characterised by dataset $S = \{\mathbf{p}_n, \mathbf{q}_n\}_{n \in N}$ and vectors of observed characteristics a_n ($\forall n \in N$), clusters are characterised by ϕ_t ($t \leq \tau$).

$$s^*_{nt} = \arg\min_{s_{nt}} \sum_{n \in N} \sum_{t \leq \tau} \left(||a_n - \phi_t||^2 \right) s_{nt}$$

$$s.t.$$

$$1: \quad \forall n \in N: \qquad \sum_{t \leq \tau} s_{nt} = 1;$$

$$2: \quad \forall i, j \in N, \forall t \leq \tau: \quad \begin{bmatrix} \mathbf{p}'_i \mathbf{q}_i \geq \mathbf{p}'_i \mathbf{q}_j \\ \mathbf{p}'_j \mathbf{q}_j \geq \mathbf{p}'_j \mathbf{q}_i \end{bmatrix} \Rightarrow s_{it} + s_{jt} \leq 1;$$

$$\text{with } s_{nt} \in \{0, 1\}.$$

The solution s^*_{nt} immediately yields WARP–partitions $V_t = \{n : s_{nt} = 1\}$ that are also as close as possible in terms of observed characteristics a_n. Given that a_n, $\mathbf{p}_n$ and $\mathbf{q}_n$ are observed for all $n \in N$, τ is given and ϕ_t is given in each iteration (see below), Problem 1 can be implemented in terms of a linear programming problem with binary variables.

1. How does this compare to standard $k - means$ clustering? The standard clustering approach minimises the sum of squared differences between elements in k clusters. Suppose first of all that dissimilar (in terms of a_i, a_j) observations i and j are more likely to violate WARP: $\mathbf{p}'_i \mathbf{q}_i \geq \mathbf{p}'_i \mathbf{q}_j$ and $\mathbf{p}'_j \mathbf{q}_j \geq \mathbf{p}'_j \mathbf{q}_i$. This implies that a considerable part of the relevant preference heterogeneity (reflected by WARP violations) is filtered out by partitioning the individuals on the basis of a_n. Formally, the solution to Problem 1 and the outcome of unconstrained clustering will be similar. Suppose on the other hand that there is no clear relationship between dissimilarity in a_n and the probability of violating WARP. This implies that the relevant preference heterogeneity may be present even within clusters of observationally equivalent (in terms of a_n) individuals. Formally, the solution to Problem 1 and the outcome of unconstrained clustering may be very different. There may be no homogeneous utility function describing the behaviour $\mathbf{q}_n$ of all individuals n in the unconstrained cluster. Recovery and prediction of behaviour in new price–income situations are then impossible.

2. In practice, researchers can determine ϕ_t by iterative refinement. In a first step, assume initial means ϕ_t that cover a wide enough part of the data, e.g. distinct percentiles of the variables' distributions. In a second step, solve Problem 1. This yields a solution

for s_{nt} and a set of WARP–partitions $V_1, ..., V_\tau$. Finally, define new values ϕ_t as the mean observed characteristics in V_t (i.e. $\phi_t = \frac{1}{|V_t|} \sum_{n \in V_t} a_n$) and re–run the algorithm until convergence. In what follows, age is treated as the unidimensional criterion. This not only provides insight in the relative magnitude of unobserved versus age–related preference heterogeneity, but also attaches a specific interpretation to the preference types. Of course, researchers can apply these techniques for any desired set of criteria.

3. Finally, Program 1 is solved relatively fast (i.e. in less than one minute) for the sample under consideration. For larger samples, it may be impossible to solve the linear programming problem with binary variables. In such case, researchers may apply a modified version of the standard $k-means$ clustering algorithm that takes constraints into account, see e.g. Wagstaff (2001). In each assignment step, observations are added to the closest cluster under the restriction that the relevant constraint (in casu condition (1)) is not violated. This is no longer guaranteed to result in a feasible allocation when $\tau = \chi(G)$. However, this can easily be checked.

3 Application to Dutch labour supply behaviour

3.1 Data

The labour supply data in this study come from the Longitudinal Internet Studies for the Social Sciences (LISS). This dataset contains detailed information on consumption and time–use choices by Dutch households. The sample consists of the collection phases in 2009, 2010, and 2012. Students, pensioners, self–employed and job seekers are dropped. After all, job seekers are unemployed due to labour demand restrictions rather than unemployed by choice. Attention is restricted to childless couples consisting of a household head and a wedded or unwedded partner. This gives 106 observations. Appendix A describes the data in more detail.

Table 1 summarises male and female wages, weekly income and consumption expenditure. The mean wage of men (14.46 EUR/hour) is clearly higher than the mean wage of women (12.28 EUR/hour). The difference in total labour income is even more outspoken (except for the highest percentiles). On the other hand, women's private expenditure (120.84 EUR/week) is higher than men's private expenditure (95.70 EUR/week). The non–labour income (per household member) is the difference between private expenditure and the labour income earned, per week. Therefore, each household member can be treated as a separate decision maker, in line with the collective labour supply model by Chiappori (1988, 1992). Cherchye et al. (2012) applied a collective labour supply model with home production to a similar sample from the LISS. In particular, the current application is similar to the ones by Blundell et al. (2007) and Cherchye and Vermeulen (2008). These authors also focused on the trade–off between private consumption and leisure.[5]

However, the current framework not only addresses preference heterogeneity *within* couples[6], but also general (nonadditive) forms of preference heterogeneity *across* couples.

	min	Q1	median	mean	std dev	Q3	max
male wage	2.06	11.64	14.04	14.46	3.67	17.54	23.10
female wage	5.85	9.68	11.95	12.28	3.40	13.99	22.93
male income	0	451.92	557.69	557.18	193.45	653.85	1,076.4
female income	0	230.77	303.49	328.94	168.36	423.08	1,105.8
male cons	4.67	60.67	88.08	95.70	59.55	117.83	472.50
female cons	26.83	74.67	98.58	120.84	90.74	129.50	665.00

Table 1: Summary statistics

Figure 2 plots the distribution of the number of hours worked by men and women in the sample. While men are most likely to work full time (± 40 hours per week), the probability density curve for women peaks around 24, 32 and 40. Women are more likely to engage in part–time work. A small fraction of women in the sample is not participating in the labour market.

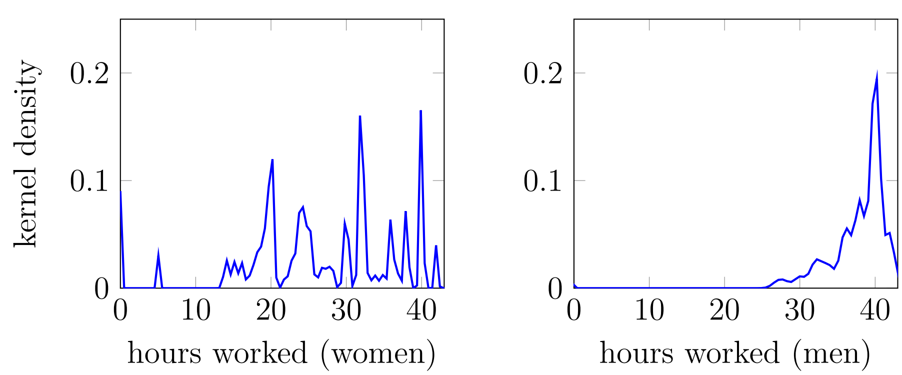

Figure 2: Probability density of the number of hours worked (per week) in the sample

Assuming that full working days consist of 8 hours, the sample contains individuals working 0 (less than 8 hours), 1 (less than 16 hours), 2 (less than 24 hours), 3 (less than

[5]Even in the presence of public consumption, this test is valid if individual utility is separable in private consumption and leisure.

[6]See Talla Nobibon et al. (2011) for algorithms to deal with intra–household preference heterogeneity when the nature of the commodities (private or public within the household) is unknown.

	$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$
male labour supply	0.04	0	0	0.08	0.58	0.31
female labour supply	0.06	0.06	0.31	0.30	0.23	0.05

Table 2: Labour supply decisions ranging from $x = 0$ to $x = 5$ days of paid market work: relative frequency

32 hours), 4 (less than 40 hours) and 5 (40 hours or more) days per week.[7] The results are in Table 2. About 6 per cent of female household members do not participate in the labour market whereas 4 per cent of male household members do not work. The difference in terms of full–time work is very marked. While 89 per cent of men work at least 4 days per week, 73 per cent of women work less than 4 days per week. Subsection 3.2 examines the differences in labour supply conditional on wages, non–labour income and individual preference heterogeneity in the sample. In this application, the commodities vector **q** consists of private consumption and leisure (= the number of available days minus x) and the price vector **p** consists of one (= the normalised price of the Hicksian aggregate) and the individual's wage.

3.2 Results

Given the abovementioned set–up with only two goods (consumption and leisure), $WARP = SARP$ and the minimum number of WARP–partitions $\hat{\tau}$ necessary to cover all individuals is equivalent to the minimum number of preference types τ^*. In what follows, the terms 'WARP–partitions' and 'preference types' are therefore interchangeable, although 'preference types' mainly refer to the revealed–preference–consistent clusters (which are WARP–partitions, by construction).

In a first step, the minimum number of male and female preference types is computed. Individuals are assigned to types in a way that minimises the age variation per type, but subject to the requirement that the minimum number of types is not exceeded, and that individuals who violate WARP cannot belong to the same type. In a second step, the age distribution within types is discussed in more detail. Finally, reservation wages are computed per preference type, to investigate the interhousehold heterogeneity in the willingness to work (full time).

[7]The numbers of hours worked in this sample are generally large because these numbers sum labour supply and commuting time.

Preference heterogeneity Per gender, the minimum number of preference types in the sample–in the sense of Definition 2–is computed.[8] In practice, the analysis is based on two graphs: one representing all men and the second representing all women as vertices. Pairwise violations of WARP are represented as edges in the graphs. The chromatic numbers of these graphs are computed by using the 'color' command in Matlab–MatGraph. The greedy algorithm is solved in a few seconds and provides the exact minimum number of (male and female) preference types.

Table 3 (row 1) presents the minimum number of preference types in the sample. There are (at least) four male preference types and three female preference types. Applying the algorithms of Crawford and Pendakur (2013) to the same data produces lower bounds of one and upper bounds of four.[9] Not surprisingly, the numbers in Table 3 lie within these bounds, but the novel approach is more precise (even the greedy algorithm). In the current setting, the novel numbers exactly identify the minimum number of utility functions in the sense of Definition 2. The increased precision shows that the minimum number of male and female preference types is distinct in the current sample. Finally, it turns out that no WARP–partition of men $\tilde{N}^m$ can cover more than 84 per cent of the sample. Similarly, no WARP–partition of women $\tilde{N}^f$ can cover more than 80 per cent of the sample (see remark 2 in Subsection 2.2).

men	women
4	3
[1,4]	[1,4]
0.23	0.29
0.23	0.31
0.38	0.40
0.17	

Table 3: Number of preference types (Crawford and Pendakur (2013) bounds between brackets) and corresponding fraction of the sample covered

In a second step, all individuals are assigned to WARP–partitions that also minimise the within–cluster sum of squared differences in age. Men are divided into four subsamples and women are divided into three subsamples, in a way that minimises the distance between the observed characteristics (i.e. age) of individuals who belong to the same preference type.[10]

[8]Recall that preference types are defined in the most general way. The utility functions of individuals of different preference types may have nothing in common, apart from the typical assumptions of monotonicity and concavity.

[9]For the implementation of the bounds by Crawford and Pendakur (2013) please see Boelaert (2015).

[10]The distance function is $d(age_n, \phi_t) = ||age_n - \phi_t||^2$. The initial means ϕ_t are increasing percentiles of the age distribution. Subsequently, updated means ϕ_t' equal the average age for individuals assigned to V_t.

Of course, researchers may incorporate any desired set of observed characteristics in the distance function. This has important consequences for the interpretation of the results. For the following comparison with unconstrained clustering (see below) and for the exposition of the preference types (in terms of the observed characteristic), it is useful to focus on one continuous characteristic in this application.

The bottom four rows in Table 3 report the fraction of the sample covered by the selected types. The male preference types cover 17–38 per cent of the men in the sample, while female preference types cover 29–40 per cent of the women in the sample. The reported sample shares show the probability of belonging to a particular type. Likewise, it is possible to compute the probabilities of observing specific household combinations of preference types, e.g. the likelihood of matches between male preference type 1 and female preference type 2. An example is discussed in Appendix B.

Comparison with unconstrained $k-means$ clustering The preference types reported in Table 3 minimise the within–cluster sum of squares (WCSS) conditional on consistency, of each type, with the Weak Axiom of Revealed Preference (WARP). By contrast, clusters from standard $k - means$ clustering minimise the WCSS under no restrictions (except the limit on the number of clusters). Therefore, one may expect the WCSS to be considerably higher in the revealed–preference–consistent clustering (operationalised in Problem 1) compared to unconstrained clustering. The extent to which the revealed–preference–consistent clusters and the standard unconstrained clusters are different, shows the impact of heterogeneity in unobserved factors (specifically: unobserved preference heterogeneity) on the individuals' labour supply and consumption choices.

Figure 3 presents the age distribution–in particular, the minimum, mean and maximum–per (male and female) preference type. A distinction is made between the RP–consistent clusters (in Table 3) and unconstrained $k - means$ clusters.[11] Let us first consider the unconstrained clusters (graphs on the bottom row). Given the unidimensional set–up, it comes as no surprise that the age distributions in different partitions are distinct (i.e. do not overlap). In other words, the oldest individual of a 'younger' preference type is never older than the youngest individual of an 'older' preference type. Let us then focus on the actual preference types (graphs on the top row). The graphs show that the resulting age distributions overlap. The highest age of a woman of the first type is also higher than the mean age of women of the latter types. This implies that variation in the observed consumption and labour supply choices cannot be explained on the basis of age (and gender)

[11] Note that the $k-means$ clusters are obtained by running the program in Problem 1 without restrictions 1–2.

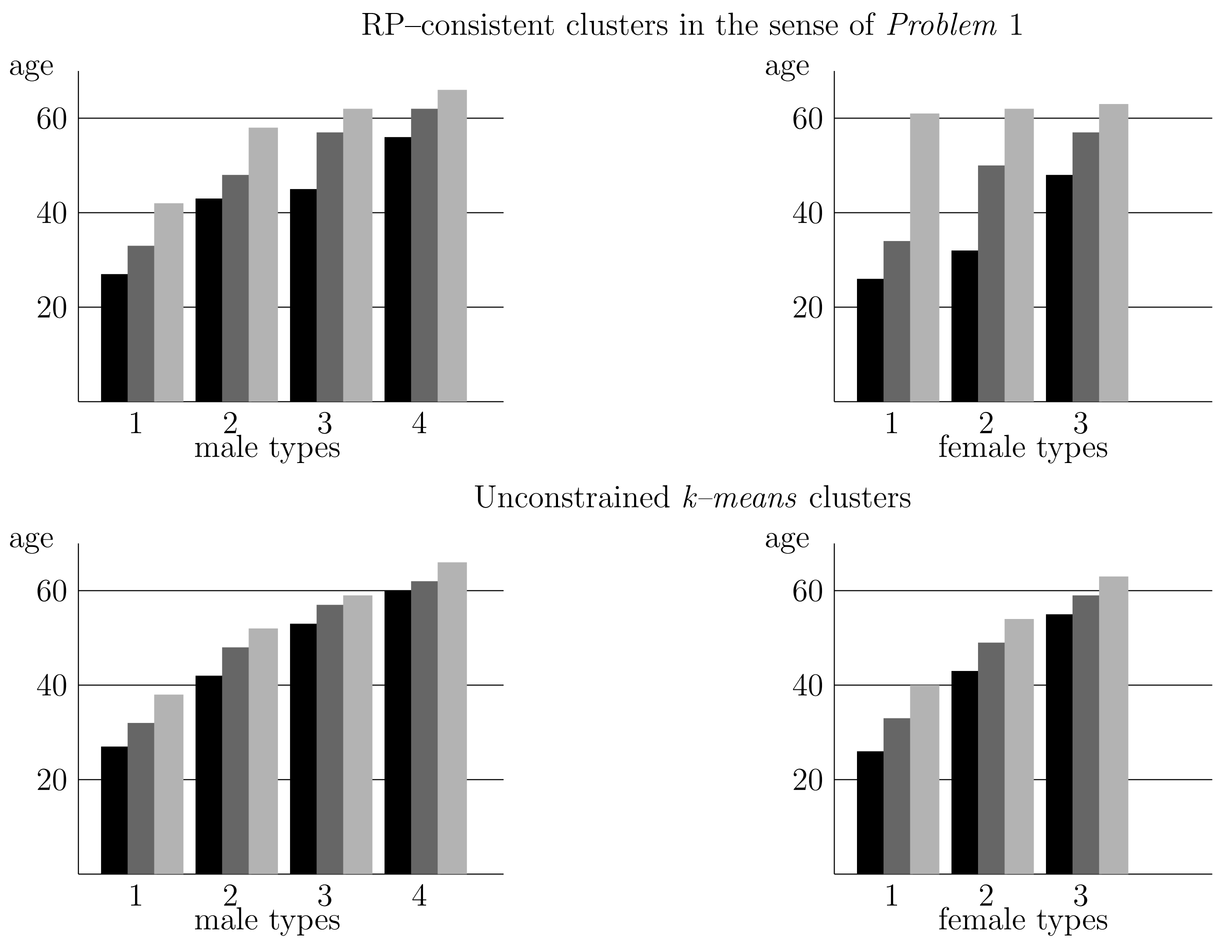

Figure 3: Graphical illustration of the min (black), mean (grey) and max (light grey) of the age distribution per preference type

alone, even if no parametric structure is imposed on the utility functions. As a result, the partitioning problem is highly relevant.

Reservation wages Based on the abovementioned preference types, one can compute the reservation wages associated with each preference type t, t'. The reservation wages associated with full time work are computed as lower bounds on the wages that rationalise full time work: $\underline{w}^m(t)$ and $\underline{w}^f(t')$. This computation is entirely nonparametric and follows Varian (1983). Details are in Appendix C. Attention is restricted to labour supply decisions at the intensive margin (i.e. the decision to work part time versus full time) for two reasons. First of all, a specific feature of the Dutch labour market is that women are less likely to work full time than men. This was also reflected by the summary statistics in Subsection 3.1. Second, the sample contains only a limited number of unemployed, due to the absence of wage information. The analysis of unemployed may therefore lack empirical support.

Reservation wages associated with full time work, per preference type, are presented in Table 4. The first two rows indicate the gender and preference type of individuals. The third row reports the share of the sample covered by each type. Finally, the bottom row shows the reservation wages $\underline{w}^m(t)$ and $\underline{w}^f(t')$ conditional on the median non–labour income in the sample.

	men (g = M)				women (g = F)		
t	1	2	3	4	1	2	3
% sample	23	23	38	17	29	31	40
$\underline{w}^g(t)$	16	10	15	13	11	18	13

Table 4: Reservation wages associated with full time work

The reservation wages below which men are not willing to work full time vary between 10 and 16 EUR/hour. Similarly, the wages below which women are not willing to work full time vary between 11 and 18 EUR/hour. The results show substantial interpersonal variation in reservation wages. This emphasises the need for methods that can deal with general forms of preference heterogeneity across the sample. The combination of reservation wages (bottom row in Table 4) and sample shares (third row in Table 4) allows researchers to bound the share of individuals who are willing to work full time for counterfactual wages and non–labour income (Figure 4).

The interpretation of Figure 4 is as follows. For wages below 11 EUR/hour, no woman in this sample can rationally supply her labour full time. By contrast, for wages between 11 and 13 EUR/hour, women of preference type 1 (29 per cent of the sample) *may be* willing to work full time. In order to rationalise full time work by all women, wages equal to or higher

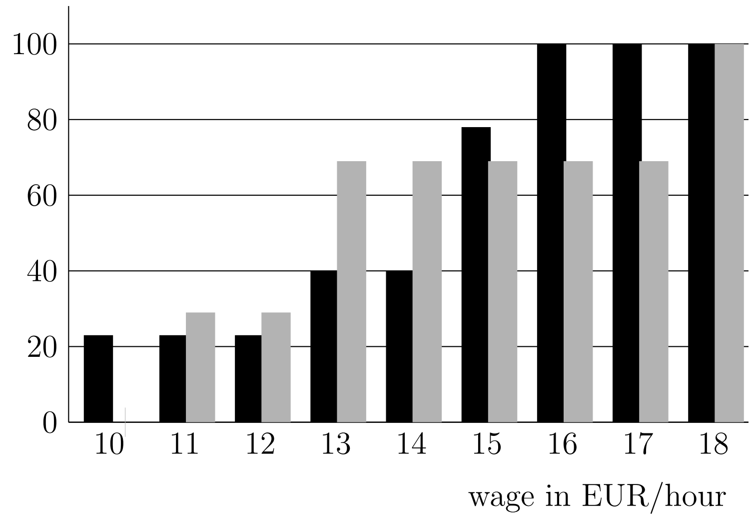

Figure 4: Upper bound on the predicted share of full time work for men (in black) and women (in grey)–conditional on wages, the median non–labour income and with general forms of unobserved preference heterogeneity

than 18 EUR/hour are required. Let us finally compare these results to the share of men who are willing to work full time. In order for all men to work five days per week, wages of 16 EUR/hour are necessary.

What can we learn from these results? The above application uses the methods from Subsections 2.2 and 2.3 for a relatively homogeneous sample. Attention is restricted to childless couples, excluding also students, pensioners, self–employed and job seekers. One may therefore argue that household members are relatively homogeneous in terms of their total available time. In this sample, it is found that 31 per cent of women are unwilling to work full time for wages below 18 EUR/hour–conditional on receiving the median non–labour income in the sample. This is in sharp contrast with the labour supply by men, who are all willing to work full time for wages of 16 EUR/hour, conditional on the same counterfactual situation. This indicates that the high share of part time work by women in the Dutch labour market is partly driven by gender–related preference heterogeneity. Interestingly, the results are relatively robust to debatable pooling assumptions across households, given that reservation wages are computed per preference type, and preference types may differ in terms of both observed and unobserved characteristics. If the number of *relevant* preference types in the population is indeed limited (see e.g. Crawford and Pendakur (2013)), it is possible to extend and refine Figure 4 to include all types. The corresponding graph would then shed light on labour supply elasticities and, potentially, optimal labour market policies. Finally, a more rigorous analysis of labour supply elasticities should also address heterogeneity in job market opportunities (beside heterogeneity in preferences), which is especially relevant when comparing childless couples to households with children, job seekers to employed, etc. In a first step, researchers may partition a sample on the basis of general household characteristics (such as the number of children). In a second step, the novel revealed–preference–consistent clustering could be applied to household members who face similar job market opportunities.

4 Conclusion

First, the idea of describing choice by multiple rationales (by Kalai et al. (2002)) is brought into practice, to compute the minimum number of utility functions necessary to rationalise consumption and labour supply choices in the cross–section. Complementary to Crawford and Pendakur (2013), who used approximation algorithms to compute the minimum number of partitions necessary to break violations of the Generalised Axiom of Revealed Preference (GARP), the current study uses insights from graph theory to efficiently compute the partitions necessary to break violations of the Weak Axiom of Revealed Preference (WARP). This follows a suggestion by Apesteguia and Ballester (2010), who argued that the problem of computing the minimum number of rationales is very complex, and that insights from graph theory may be helpful to address the problem. In a first step, a graph is constructed in which vertices represent individuals and edges represent pairwise violations of WARP. In a second step, it is shown that the minimum number of partitions necessary to break all WARP violations is equivalent to the chromatic number applied to this graph. The chromatic number always bounds the minimum number of utility functions in the sample from below, and it is equivalent to the minimum number of utility functions in the sample as long as $WARP = SARP$. Furthermore, a wide range of algorithms from the computer science and operations research literature–to compute the chromatic number–can be applied to solve this problem, both approximately (using a greedy algorithm) and optimally.

Second, the current paper deals with the recovery of sets of individuals with homogeneous preferences (in contrast to Crawford and Pendakur (2013) who focused mainly on computing the number of sets). To this end, variation in observed characteristics is used. In particular, the nonparametric (revealed preference) conditions are complemented with an objective function that minimises observed dissimilarities within each preference type. On the one hand, this objective function provides the additional structure that is necessary to select one particular partitioning of the sample. On the other hand, this novel revealed–preference–consistent clustering builds the bridge between cluster analysis and revealed preference theory (in casu: WARP). Indeed, dissimilarities in observed demographic variables within clusters are also minimised, but in a theoretically robust way. This contrasts with the traditional clustering approach, in which clusters need not be consistent with the utility maximisation hypothesis.

The methods are applied to consumption and labour supply choices by Dutch households (LISS). The cross–sectional variation in wages provides strong empirical bite. It turns out that there are four types of men and three types of women in the sample, which lies between the lower and upper bounds (one and four) generated by Crawford and Pendakur (2013)'s algorithm. Subsequently, this paper focused on the recovery of reservation wages–associated with full time work–per preference type. Towards this end, individuals were assigned to (three or four) preference types in a way that maintains consistency with the revealed preference axioms and minimises the difference in observables (in casu age). The results indicated considerable variation in the reservation wages across types.

Acknowledgements

I thank Laurens Cherchye, Thomas Demuynck and Ian Crawford, and also Bart Capéau, John Dagsvik, André Decoster, Matthieu Delpierre, Bram De Rock, Yuichi Kitamura, Arthur Lewbel, Lars Nesheim, Johannes Spinnewijn and Frederic Vermeulen for valuable suggestions. Furthermore, I am grateful to seminar participants in Kortrijk and Leuven (University of Leuven), Cergy–Pontoise (ADRES conference), Namur (University of Namur) and Esch–sur–Alzette (LISER). The LISS panel data were collected by CentERdata (Tilburg University, The Netherlands) through its MESS project funded by the Netherlands Organization for Scientific Research.

References

Adams, A., Blundell, R., Browning, M., Crawford, I., 2015. Prices versus preferences: taste change and revealed preference. Tech. rep., Institute for Fiscal Studies.

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

Apesteguia, J., Ballester, M., 2010. The computational complexity of rationalizing behavior. Journal of Mathematical Economics 46, 356–363.

Blundell, R., Chiappori, P., Magnac, T., Meghir, C., 2007. Collective labor supply: Heterogeneity and non-participation. Review of Economic Studies 74, 417–445.

Boelaert, J., 2015. Package revealedprefs. Tech. rep.

Cherchye, L., De Rock, B., Vermeulen, F., 2012. Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. American Economic Review 102, 3377–3405.

Cherchye, L., Demuynck, T., De Rock, B., 2015. Transitivity of preferences: when does it matter? Tech. rep., ECARES working paper 2015-44.

Cherchye, L., Vermeulen, F., 2008. Nonparametric analysis of household labor supply: Goodness-of-fit and power of the unitary and the collective model. Review of Economics and Statistics 90, 267–274.

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

Chiappori, P., 1992. Collective labor supply and welfare. Journal of Political Economy 100, 437–467.

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

Dean, M., Martin, D., 2015. Measuring rationality with the minimum cost of revealed preference violations. Review of Economics and Statistics forthcoming, 35.

Demuynck, T., 2011. The computational complexity of rationalizing boundedly rational choice behavior. Journal of Mathematical Economics 47, 425–433.

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

Falmagne, J. C., 1978. A representation theorem for finite random scale systems. Journal of Mathematical Psychology 18, 52–72.

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

Kalai, G., Rubinstein, A., Spiegler, R., 2002. Rationalizing choice functions by multiple rationales. Econometrica 70, 2481–2488.

Karp, R. M., 1972. Reducibility among combinatorial problems. Complexity of Computer Computations springer US, 85–103.

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

McFadden, D. L., 2005. Revealed stochastic preference: a synthesis. Economic Theory 26, 245–264.

Rose, H., 1958. Consistency of preference: The two-commodity case. Review of Economic Studies 25, 124–125.

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

Smeulders, B., Cherchye, L., Rock, B. D., Spieksma, F., Nobibon, F. T., 2015. Complexity results for the weak axiom of revealed preference for collective consumption models. Journal of Mathematical Economics 58, 82–91.

Smeulders, B., Spieksma, F., Cherchye, L., Rock, B. D., 2014. Goodness-of-fit measures for revealed preference tests: Complexity results and algorithms. ACM Transactions on Economics and Computation 2, 16.

Talla Nobibon, F., Cherchye, L., De Rock, B., Sabbe, J., Spieksma, F. C. R., 2011. Heuristics for deciding collectively rational consumption behavior. Computational Economics 38, 173–204.

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

Wagstaff, K., 2001. Constrained kmeans clustering with background knowledge. ICML 1, 577–584.

A Data

The dataset combines information from different studies available on the LISS website. Information on household characteristics is taken from the *Background Variables*, *Family and Household* and *Work and Schooling* sections. Specifically, the household members' ages are reported in *Background Variables. Work and Schooling* and *Economic Situation: Income* contain the necessary data to construct wages, by dividing the total labour income (from *Economic Situation: Income*) by the average number of hours worked (from *Work and Schooling*). This includes the income and hours worked of all jobs. If available, the wages of household members who do not work (i.e. when the number of hours worked is 0) are set equal to the corresponding wages in their last job. Endogeneity issues related to the construction of wages are avoided given that wages are based on the *average* number of hours worked whereas labour supply is based on the current number of hours worked (as reported in the time use section). Finally, the *Time Use and Consumption* section contains information on private consumption (by the household members) and their time use decisions. The respondents reported private consumption on an average monthly basis. This is converted to weekly data. Expenditure includes food and beverages, tobacco products, clothing, personal and medical care products and services, leisure time expenditure, schooling, gifts, and other expenditure. In addition, the *Time Use and Consumption* module collects data on time–use decisions, that is, leisure time and hours spent on market work in a particular week. For a detailed discussion of a similar sample from LISS, please see Cherchye et al. (2012).

B Matching

As an extension, the presented methodology can also be used to compute the probability that men and women of particular preference types match. Table 5 shows the likelihood that a man of type Mi forms a couple with a woman of type Fj in the sample. Intuitively, one would expect higher probability densities along the diagonal because of the age characteristic: people of similar age are more likely to match. This is confirmed in Table 5. For instance, 22 per cent of the households in the sample are characterised by the youngest male ($M1$) and the youngest female ($F1$) preferences. The reported probabilities may be valuable in applications of collective labour supply models in the spirit of Chiappori (1988, 1992). Household members have different preferences, but each pair of preference orderings is stable for a subset of the data. Similarly, the information in Table 5 seems useful for matching models and theories of assortative mating.

	F1	F2	F3
M1	0.22	0.01	0
M2	0.06	0.12	0.05
M3	0.02	0.14	0.22
M4	0	0.04	0.13

Table 5: Relative frequency of households in which men belong to the Mi preference type and women belong to the Fj preference type

C Nonparametric bounds on reservation wages

In this study, reservation wages indicate lower bounds on the wages that rationalise full time work. Given that reservation wages stem from individual preferences, these wages differ across the preference types. To compute nonparametric lower bounds on the wages that rationalise full time work, a procedure by Varian (1983) is followed. Let $\underline{\mathrm{w}}(t)$ represent the reservation wage associated with individuals of preference type t. Furthermore, $\bar{q}^1$ captures the median private expenditure and $\bar{q}^2$ is the available leisure in case of full time work.

For a 'representative' individual of type t who works full time (and has available leisure $\bar{q}^2$), compute $\underline{\mathrm{w}}(t)$ as the minimum wage p^2 that is still consistent with WARP (i.e. that still rationalises full time work).

$$\underline{\mathrm{w}}(t) = \min_{p^2} p^2$$

$$s.t.$$

$$\forall i \in \; V_t : q_i^1 + p_i^2 q_i^2 \geq \bar{q}^1 + p_i^2 \bar{q}^2 \Rightarrow \bar{q}^1 + p^2 \bar{q}^2 < q_i^1 + p^2 q_i^2.$$

The first part of the 'if–then' condition can be verified immediately, for it is fully defined by observations of individuals i and predetermined levels of $\bar{q}^1$ and $\bar{q}^2$–the latter representing available leisure when the individual works full time. All observations i for which this opening condition holds are 'revealed preferred' over $\bar{q}^1$ and $\bar{q}^2$. Rationality then requires that individuals minimise expenditure over this better–than set. As a result, p^2 must be such that the bundle $(\bar{q}^1, \bar{q}^2)$ is cheaper than the 'revealed preferred' bundles (q_i^1, q_i^2). Given $\bar{q}^2 \leq q_i^2$, p^2 is typically bounded from below.

LUXEMBOURG INSTITUTE OF
SOCIO-ECONOMIC RESEARCH

11, Porte des Sciences
Campus Belval
L-4366 Esch-sur-Alzette

T +352 58 58 55-1
F +352 58 58 55-700

www.liser.lu