KU LEUVEN

DEPARTMENT OF ECONOMICS

Revealed preference analysis with normal goods: application to cost of living indices

Laurens CHERCHYE, Thomas DEMUYNCK, Bram DE ROCK and Khushboo SURANA

FACULTY OF ECONOMICS AND BUSINESS

DISCUSSION PAPER SERIES DPS18.04

APRIL 2018

Revealed preference analysis with normal goods: Application to cost of living indices

Laurens Cherchye* Thomas Demuynck† Bram De Rock‡
Khushboo Surana§

April 13, 2018

Abstract

We present a revealed preference methodology for nonparametric demand analysis under the assumption of normal goods. Our methodology is flexible in that it allows for imposing normality on any subset of goods. We show the usefulness of our methodology for empirical welfare analysis through cost of living indices. An illustration to US consumption data drawn from the Panel Study of Income Dynamics (PSID) demonstrates that mild normality assumptions can substantially strengthen the empirical analysis. It obtains considerably tighter bounds on cost of living indices, and a significantly more informative classification of better-off and worse-off individuals after the 2008 financial crisis.
Keywords: nonparametric demand analysis, revealed preferences, normal goods, individual welfare analysis, cost of living indices
JEL code: C14, D01, D11, D60

1 Introduction

Changing price-income regimes can have a substantive impact on individual demand patterns. The empirical analysis of the associated welfare effects has attracted considerable attention in the applied welfare literature. In the current paper, we propose a structural method for such welfare analysis that is intrinsically nonparametric: it does not impose any parametric/functional structure on the individual utilities, but merely exploits the preference information that is directly revealed by the observed consumption behavior.

*Department of Economics, University of Leuven (KU Leuven). E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven.be. Laurens Cherchye gratefully acknowledges the European Research Council (ERC) for his Consolidator Grant 614221. Part of this research is also funded by the Research Fund KU Leuven and the Fund for Scientific Research-Flanders (FWO).

†ECARES, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: thomas.demuynck@ulb.ac.be. Thomas Demuynck acknowledges financial support by the Fonds de la Recherche Scientifique-FNRS under grant nr F.4516.18

‡ECARES, Université Libre de Bruxelles, and Department of Economics, University of Leuven (KU Leuven). Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges FRNS, FWO and EOS for their financial support.

§Department of Economics, University of Leuven (KU Leuven). E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: khushboo.surana@kuleuven.be.

Particularly, we demonstrate that mild normality assumptions on the demand for (a subset of) goods can obtain a significantly informative analysis of individual cost of living indices. We show this through an empirical illustration to household demand data taken from the Panel Study of Income Dynamics (PSID), in which we analyze the welfare effects of the 2008 financial crisis for a sample of singles in the US.

Welfare evaluation and counterfactual demand analysis. The structural analysis of welfare effects associated with changing prices and/or incomes requires predicting demand in counterfactual price-income regimes. This issue is standardly addressed by adopting a parametric approach, which assumes a specific functional form for the consumers' utility or expenditure functions.[1] The parameters of this functional form are then estimated from the observed consumption behavior, and these estimations can be used to interpolate or extrapolate demand in unobserved price-income situations. A main problem of this parametric approach is that it crucially relies on some a priori assumed functional form for the individual preferences, which is typically non-verifiable. This implies an intrinsic risk of specification error.

We can avoid this specification risk by adopting the nonparametric revealed preference approach that was initiated by Samuelson (1938) and Houthakker (1950), and further developed by Afriat (1967), Diewert (1973) and Varian (1982). Basically, this nonparametric approach develops testable implications for observed consumption patterns (prices and quantities) that must hold under rational demand behavior associated with *any* well-behaved utility function. These testable implications are then used as a basis for counterfactual demand predictions in the form of set identification (producing bounds on possible demand responses in new price-income regimes). By its very nature, this nonparametric approach avoids the possibility of erroneous conclusions following from a wrongly specified functional form.

Revealed preference analysis and normal goods. Although this nonparametric orientation of the revealed preference approach is conceptually appealing, its empirical usefulness is often put into question. Generally, an informative empirical analysis requires a rich data set with high price variation and low income variation. In many observational settings, however, the opposite holds true (i.e., low price variation combined with high income variation). In such cases, the nonparametric testable implications have little empirical bite and, correspondingly, the set identification results are not very informative (see, for example, Varian (1982) and Bronars (1987) for detailed discussions). As an implication, the revealed preference methodology is then of limited practical value.

In the current paper, we show that this lack of power can be remediated by assuming normality of the goods that are consumed. Normality is often a natural assumption to make. Basically, a good is normal if its income expansion path is increasing. A convenient feature of our method is that we can impose normality without needing to estimate the expansion path; our nonparametric testable implications apply to *any* expansion path that satisfies normal demand. Moreover, our method applies to settings with any number of goods, and can impose normality on any subset of these goods. The only assumption it makes is that normality holds for the observed prices, so avoiding the stronger hypothesis

[1] Popular functional forms in the literature are the Cobb-Douglas, the translog (Christensen, Jorgenson, and Lau, 1975), the almost ideal demand (Deaton and Muellbauer (1980)) and quadratic almost ideal demand specification (Banks, Blundell, and Lewbel (1997)).

that normality must apply to any (observed or unobserved) price.[2]

In a recent series of papers, Blundell, Browning, and Crawford (2003, 2007, 2008) and Blundell, Browning, Cherchye, Crawford, De Rock, and Vermeulen (2015) also used the assumption of normal demand for observed prices to deal with the power issue associated with empirical revealed preference analysis. However, we see at least two main differences between the method proposed by these authors and our novel method. First, they assume that normality holds for all goods simultaneously, whereas our method is equally applicable to normality for any subset of goods. Second, and more importantly, these authors exploit normality of demand by using (nonparametrically) estimated income expansion paths (assuming a repeated cross-sectional data set). As indicated above, our method avoids this prior estimation step (and associated statistical issues); it directly applies revealed preference restrictions (for normal demand) to the observed consumption choices. Interestingly, our empirical application shows that our method can yield an informative welfare analysis even with a short time series of (three) consumption observations per individual.

Empirical welfare analysis and cost of living indices. We show that our revealed preference method can be used for a meaningful welfare analysis on the basis of cost of living indices. We demonstrate this through an empirical application to data drawn from the PSID. We select a balanced panel from the 2007, 2009 and 2011 waves of the PSID to study the welfare effects of the 2008 financial crisis. A large number of studies has analyzed these welfare effects since the onset of the crisis. As the crisis led to a substantial rise in unemployment, the principal focus so far has been on the extensive margin of labor supply (see, for example, Verick (2009); Hurd and Rohwedder (2010); Goodman and Mance (2011); Deaton (2011)). By contrast, in our application we concentrate on individuals who remained employed after the crisis.

More specifically, our structural analysis assumes a model of rational labor supply for singles who spend their potential income on leisure, food, housing and other goods, hereby imposing normality on all consumption categories except from leisure. To assess the empirical bite of the testable implications associated with normality, we also compute the empirical results for the rational labor supply model without normal demand. Our results show that imposing normality entails a substantially more powerful empirical analysis. In particular, we obtain considerably tighter bounds on cost of living indices, and a significantly more informative classification of better-off and worse-off individuals after the 2008 crisis.

Outline. Section 2 develops the revealed preference characterization of utility maximization under normality assumptions. Section 3 introduces the cost of living index for our empirical welfare analysis, and defines the goodness-of-fit measure that we will use to evaluate the empirical fit of our normality assumptions. Section 4 presents our empirical application to PSID data. Section 5 concludes.

[2]See Cherchye, Demuynck, and De Rock (2018) for a discussion of the nonparametrically testable implications of normal demand under any (observed or unobserved) price, with a main focus on the specific demand setting with two consumed goods.

2 Rational demand with normal goods

Our main theoretical result defines the testable implications for the observed demand behavior to be consistent with rationality (i.e., utility maximization) and normality of (a subset of) the consumed goods. To this end, we first define the Generalized Axiom of Revealed Preference (GARP) in terms of Hicksian demand bundles that correspond to the observed prices and associated utility levels (for the given quantity bundles). Imposing normality boils down to restricting these Hicksian demand bundles at any observed price regime to be monotone in utility (Fisher, 1990). Basically, our testable revealed preference conditions verify whether there exists at least one possible specification of the utility levels and Hicksian demand bundles that satisfy this requirement. If so, we cannot reject the joint hypothesis of normality and rational behavior.

Generalized Axiom of Revealed Preference (GARP). Throughout, we focus on a finite set T of observed prices and corresponding quantities. For each consumption observation $t \in T$, let $q_t \in \mathbb{R}^n_+$ and $p_t \in \mathbb{R}^n_{++}$ denote the (column) vectors of quantities and prices, respectively. This defines the data set $S = \{(p_t, q_t)\}_{t \in T}$. We say that S is "rationalizable" if there exists a utility function $u(.)$ such that, for each observation $t \in T$, q_t maximizes this function $u(.)$ over all affordable bundles for the given prices p_t and outlay $x_t = p_t q_t$. Throughout, we will assume utility functions that are continuous and strictly monotone.

Definition 1 *A data set $S = \{(p_t, q_t)\}_{t \in T}$ is rationalizable if there exists a continuous and strictly monotone utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ such that, for all $t \in T$ and $x_t = p_t q_t$,*

$$q_t \in \arg\max u(q) \;\; s.t. \;\; p_t q \leq x_t.$$

Varian (1982) has shown that the Generalized Axiom of Revealed Preference (GARP) defines a necessary and sufficient condition for a data set S to be rationalizable. Thus, checking rationalizability boils down to verifying whether or not the set S satisfies GARP. To formally define this GARP requirement, we will need the following concepts.

Definition 2 *Consider a data set $S = \{(p_t, q_t)\}_{t \in T}$. We say that q_t, $t \in T$, is* directly revealed preferred *to the bundle q_v, $v \in T$, if $p_t q_t \geq p_t q_v$. We denote this as $q_t R^D q_v$. Next, we say that q_t is* strictly directly revealed preferred *to q_v if $p_t q_t > p_t q_v$. We denote this as $q_t P^D q_v$. Finally, we say that q_t is* revealed preferred *to q_v if there exists a (possibly empty) sequence $u, s, \cdots, r \in T$ such that*

$$q_t R^D q_u, q_u R^D q_s, \ldots, q_r R^D q_v.$$

We denote this as $q_t R q_v$.

Thus, the quantity bundle q_t is directly revealed preferred to the bundle q_v (i.e., $q_t R^D q_v$) if q_v was affordable when bundle q_t was chosen (i.e., $p_t q_t \geq p_t q_v$). If the inequality is strict (i.e., $p_t q_t > p_t q_v$), then q_t is strictly directly revealed preferred to q_v (i.e., $q_t P^D q_v$). Finally, from the direct revealed preference relations, we can define the more general concept of (direct or indirect) revealed preference relations by exploiting transitivity of preferences (i.e., $q_t R q_v$ follows from $q_t R^D q_u, q_u R^D q_s, \ldots, q_r R^D q_v$).

We can now define GARP.

Definition 3 *A data set $S = \{(p_t, q_t)\}_{t \in T}$ satisfies GARP if, for all $t, v \in T$, $q_t R q_v$ implies not $q_v P^D q_t$.*

In words, a data set S satisfies GARP if, for any two observed bundles q_t and q_v, $q_t R q_v$ implies that q_v is not strictly directly revealed preferred to q_t (i.e., not $q_v P^D q_t$). Intuitively, GARP excludes that bundle q_t is revealed preferred to q_v while, at the same time, q_t was affordable at a strictly lower cost when q_v was purchased.

In what follows, we will focus on a less standard reformulation of the GARP condition in Definition 3. This alternative formulation will be instrumental for our characterization of rationalizable consumer behavior under normal demand. It is contained in the following result.[3]

Proposition 1 *A data set $S = \{(p_t, q_t)\}_{t \in T}$ satisfies GARP if and only if there exist numbers $(u_t)_{t \in T}$ such that, for all $s, t \in T$,*

- *if $u_t \geq u_s$, then $p_s q_s \leq p_s q_t$,*
- *if $u_t > u_s$, then $p_s q_s < p_s q_t$.*

Thus, rationalizability of a data set S can be verified by checking the existence of "utility numbers" u_t that satisfy a series of "if–then" conditions. Intuitively, each number u_t represents the consumer's utility level associated with the bundle q_t. If the utility level at observation t is (strictly) above the utility level at observation s (i.e., $u_t \geq (>) u_s$), then the bundle q_t must be (strictly) more expensive than the bundle q_s at the prices p_s.

Normality-extended GARP (N-GARP) Let $M \subseteq \{1, ..., n\}$ be a subset of the goods that are consumed. We say that a data set S is *rationalizable by normal demand on the subset M* if there exists a well behaved utility function that (1) represents each observed bundle q_t as utility maximizing under (2) the additional requirement that, for each good $i \in M$, the income expansion path at the observed prices has a positive slope. Formally, we have the following definition.

Definition 4 *A data set $S = \{(p_t, q_t)\}_{t \in T}$ is rationalizable by normal demand on the subset M ($M \subseteq \{1, ..., n\}$) if there exists a continuous and strictly monotone utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ and functions $q_t : \mathbb{R}_+ \to \mathbb{R}^n_+$ such that, for all $t \in T$ and $x_t = p_t q_t$,*

- *$q_t(x) \in \arg\max u(q)$ s.t. $p_t q \leq x$,*
- *$q_t^i(x)$ is strictly monotone in x for all $i \in M$,*
- *$q_t = q_t(x_t)$.*

In this definition, the function $q_t(.)$ represents the income expansion path at the observed prices p_t, defining the quantities demanded by the consumer at the price-income pair (p_t, x) for any value of x. Definition 4 defines three conditions for the functions $u(.)$ and $q_t(.)$. The first condition states that, for all income levels x, $q_t(x)$ maximizes the function $u(.)$ over all affordable bundles at prices p_t and income x. The second condition imposes that $q_t^i(x)$ is increasing in x, meaning that good $i \in M$ is normal at prices p_t.

[3] This equivalent reformulation of GARP has been used in the literature on nonparametric production analysis. We refer to Varian (1984) (Theorem 2) for a formal proof of Proposition 1.

The last condition requires that $q_t(x_t)$ equals the observed demand q_t for the observed income/outlay x_t $(= p_t q_t)$ and prices p_t.

We next establish the revealed preference characterization of rationalizable behavior as defined in Definition 4. This provides nonparametric testable implications for the observed data set S to be by consistent with utility maximization under the additional assumption of normal demand. In particular, we can show that rationalizability under normal demand holds if and only if the data set S satisfies the normality-extended GARP (N-GARP).

Definition 5 *For $M \subseteq \{1, ..., n\}$, a data set $S = \{(p_t, q_t)\}_{t \in T}$ satisfies N-GARP if there exist numbers $(u_t)_{t \in T}$ and vectors $(q_{t,v})_{t,v \in T}$ $(q_{t,v} \in \mathbb{R}^n_+)$ such that, for all $r, s, t, v \in T$,*

- $q_{t,t} = q_t$,
- *if* $u_t \geq u_v$, *then* $p_r q_{r,v} \leq p_r q_{s,t}$,
- *if* $u_t > u_v$, *then* $p_r q_{r,v} < p_r q_{s,t}$,
- *if* $u_t \geq u_v$, *then* $q^i_{r,v} \leq q^i_{r,t}$ *for all* $i \in M$.

The following proposition contains our main theoretical result.[4]

Proposition 2 *A data set $S = \{(p_t, q_t)\}_{t \in T}$ is rationalizable by normal demand on the subset M $(M \subseteq \{1, ..., n\})$ if and only if it satisfies N-GARP.*

Similar to Proposition 1, we obtain that rationalizability imposes the existence of utility numbers u_t that satisfy a series of if-then conditions. In our N-GARP definition, each vector $q_{t,v}$ represents the Hicksian demand bundle at prices p_t for the utility level associated with the bundle q_v (captured by the number u_v). Rationalizability requires the numbers u_t and vectors $q_{t,v}$ to satisfy the four conditions in Definition 5. The first condition states, for each observation $t \in T$, that the Hicksian demand $q_{t,t}$ at the given prices (p_t) and utility level (u_t) must equal the observed Marshallian demand q_t. The second and third condition impose GARP (as formulated in Lemma 1) on the sets $(p_t, q_{t,v})_{t,v \in T}$, which consist of observed prices p_t and Hicksian demand vectors $q_{t,v}$. The final condition requires that the Hicksian quantities for each good $i \in M$ are monotonically increasing in utility, which corresponds to normal demand (Fisher, 1990).

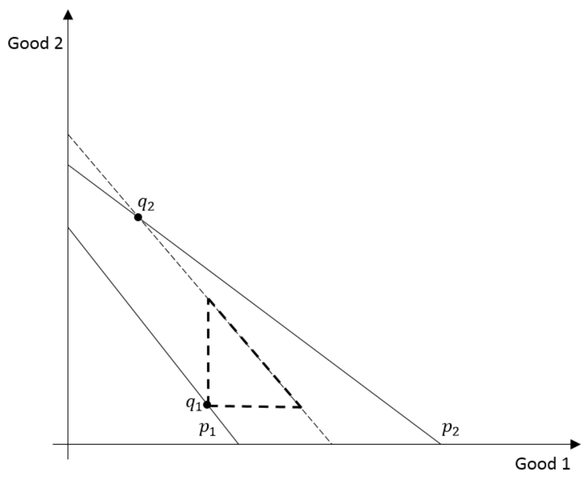
When comparing the conditions in Proposition 1 with those in Definition 5, it is clear that N-GARP generally implies stronger rationalizability requirements than GARP. N-GARP reduces to GARP (only) in the limiting case that does not impose normality for any good. We illustrate the difference between N-GARP and GARP in Example 1, which contains a data set that satisfies GARP but violates N-GARP. It indicates that imposing normality can yield a more powerful revealed preference analysis. This is an attractive feature, as normality assumptions are often little debatable and thus easy to make.

Finally, in Appendix B we show that the N-GARP condition in Definition 5 can be reformulated in terms of inequality constraints that are linear in unknowns and characterized by (binary) integer variables. These linear inequality constraints are easily operationalized, which is convenient from an application point of view.[5]

[4] Appendix A contains the proof of Proposition 2.

[5] For example, we used the software package IBM ILOG CPLEX Optimization Studio for our empirical application in Section 4. Our CPLEX codes are available upon request.

Figure 1: Example data set that violates N-GARP but not GARP



Example 1. We illustrate the difference between N-GARP and GARP by means of a simple numerical example using a data set S with two goods ($n = 2$) and two observations ($T = \{1, 2\}$):

$$p_1 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, p_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, q_1 = \begin{bmatrix} 8 \\ 1 \end{bmatrix}, q_2 = \begin{bmatrix} 4 \\ 10 \end{bmatrix}.$$

Figure 1 depicts the two quantity bundles and associated budget sets. From this figure, it is easy to verify that the set S satisfies GARP. In particular, the budget lines do not cross, which automatically implies consistency with GARP. More formally, referring to Proposition 1, we have $p_1q_1 = 36, p_1q_2 = 56, p_2q_1 = 29$ and $p_2q_2 = 62$. Then, using $u_1 = 0.1$ and $u_2 = 0.2$ obtains that all conditions in Proposition 1 are satisfied.

Next, we can show that the same data set S violates N-GARP for $M = \{1, 2\}$, i.e. both goods are assumed to be normal goods. In particular, we prove that there do not exist numbers u_1, u_2 and vectors $q_{1,1}, q_{1,2}, q_{2,1}, q_{2,2}$ that simultaneously meet the four conditions in Definition 5. To see this, we begin by noting that the first N-GARP condition imposes

$$q_{1,1} = q_1 = \begin{bmatrix} 8 \\ 1 \end{bmatrix}, \text{ and } q_{2,2} = q_2 = \begin{bmatrix} 4 \\ 10 \end{bmatrix}. \tag{1}$$

In addition, the second N-GARP condition (using that $u_2 \geq u_2$) imposes

$$p_1q_{1,2} \leq p_1q_{2,2} \text{ and } p_2q_{2,2} \leq p_2q_{1,2}. \tag{2}$$

Combining (1) and (2) obtains (using superscripts to indicate the quantities of goods 1 and 2)

$$4q^1_{1,2} + 4q^2_{1,2} \leq 56,$$
$$62 \leq 3q^1_{1,2} + 5q^2_{1,2}.$$

These two inequalities together imply

$$62 \leq 3q^1_{1,2} + 5q^2_{1,2} \leq 3q^1_{1,2} + 5(14 - q^1_{1,2}) \quad \Leftrightarrow \quad q^1_{1,2} \leq 4. \tag{3}$$

On the other hand, because $p_1 q_{1,1} = 36 < p_1 q_{2,2} = 56$, the third N-GARP condition in Proposition 5 requires

$$u_1 < u_2.$$

Then, the fourth N-GARP condition imposes (using that goods 1 and 2 are both normal)

$$q_{1,1} \leq q_{1,2}.$$

Combined with (1), this entails

$$q^1_{1,2} \geq 8,$$

which contradicts (3). Thus, we conclude that N-GARP is violated.

We can also graphically illustrate this N-GARP violation in Figure 1. To see this, we first note that the Hicksian demand $q_{1,2}$ should lie below the dashed line associated with the budget $p_1 q_2$. Also, if both goods are normal at the prices p_1, it must hold that $q_{1,2}$ contains more of both goods 1 and 2 than q_1 (i.e., $q_{1,2} \geq q_1$). Taken together, we conclude that $q_{1,2}$ is situated in the triangular region formed by the thick-dashed lines. Then, the conclusion that N-GARP is violated follows from the observation that no $q_{1,2}$ in this region is consistent with rationalizability of the consumption observation (p_2, q_2). Specifically, any such $q_{1,2}$ is strictly less expensive than the bundle q_2 at prices p_2. As an implication, for the outlay $p_2 q_2$ and prices p_2 associated with the quantity bundle q_2, the consumer could have chosen bundles strictly better than $q_{1,2}$. This implies that $q_{1,2}$ and q_2 cannot yield the same utility value for a strictly monotone utility function.

3 Cost of living and goodness-of-fit

In this section, we introduce some additional concepts and tools that will be useful for our following application. First, we show how our testable conditions for normal demand can be used to identify cost of living indices for comparing individual welfare in alternative price-income regimes. Next, in their original formulation, our revealed preference conditions for rational behavior under normal demand define "exact" tests: data either satisfy the requirements or not. In our empirical application, we will use an Afriat type Critical Cost Efficiency Index (CCEI) to assess how closely behavior complies with rational behavior. This index will serve as a *goodness-of-fit* measure that has a specific interpretation as capturing the *economic significance* of violations of our testable implications.

Cost of living indices. An important application of empirical demand analysis consists of comparing consumers' welfare in alternative price-income regimes. More specifically, for two consumption observations (p_t, q_t) and (p_r, q_r), we not only wish to know which combination is (revealed) "better" by the consumer, but also "how much better". As utility theory is ordinal in nature, there is no unique answer to this last question. A popular method makes use of the money metric utility concept that was introduced by Samuelson (1974). In what follows, we will use this money metric representation of individual utility to compute cost of living indices associated with different price-income situations. Technically, we adapt the nonparametric method that was developed by Varian (1982), based on the GARP concept in Definition 3.[6] We will show that our N-GARP

[6] Varian (1982) refers to the money metric utility function as income compensation function. He considers welfare comparisons between price-income situations that are possibly unobserved. In the

characterization in Definition 5 easily allows for computing lower and upper bounds on individuals' cost of living indices. This effectively set identifies these indices using the assumption of rationalizability under normal demand.

The money metric utility function gives the minimum expenditure required in observation t (with price-income pair (p_t, x_t)) to attain the same utility level as under some reference price-income regime (p_r, x_r). Formally, it is defined as

$$\mu(p_t; p_r, x_r) \equiv e(p_t, v(p_r, x_r)),$$

with $e(p, u)$ the expenditure function quantifying the minimum income required to attain utility u at prices p, and $v(p, x)$ the indirect utility function giving the maximum utility level at prices p and income x. In our set-up, the vector $q_{t,r}$ represents Hicksian demand at price p_t and utility level u_r, which itself equals $v(p_r, x_r)$. Thus, we can simply write

$$\mu(p_t; p_r, x_r) = p_t q_{t,r}.$$

Then, using our N-GARP characterization of rationalizable consumer behavior under normal demand, we can define upper (or lower) bounds on $\mu(p_t; p_r, x_r)$ by maximizing (or minimizing) $p_t q_{t,r}$ subject to the conditions in Definition 5. This implies optimization problems with a linear objective and linear inequality constraints that are characterized by integer variables (see also Appendix B). It defines an interval set of possible values for $\mu(p_t; p_r, x_r)$ under the given normality assumptions.

In a following step, we can compare the welfare of some evaluated individual in consumption observation t and reference observation r by using the cost of living index

$$c_{t,r} = \frac{x_t - \mu(p_t; p_r, x_r)}{x_t} = \frac{x_t - p_t q_{t,r}}{x_t}.$$

In this expression, the numerator $x_t - \mu(p_t; p_r, x_r)$ defines the compensating variation associated with the price change from p_r to p_t. It measures the difference between the individual's potential income in the decision situation t (i.e., x_t) and the income needed by the same individual under the prices p_t to be equally well off as in the reference situation r (i.e., $\mu(p_t; p_r, x_r)$). To obtain the cost of living index $c_{t,r}$, we divide this compensating variation by the available income in observation t. This compares the individual's welfare in t relative to r. If $c_{t,r}$ exceeds zero, the individual is better off in t than in r. Conversely, if $c_{t,r}$ is below zero, the individual is worse off in t than in r.

Similar to before, our nonparametric characterization of rationalizable demand behavior allows us to nonparametrically identify upper and lower bounds on $c_{t,r}$ (using set identification of $\mu(p_t; p_r, x_r)$). In our empirical application, we will conclude that an individual is better off in situation t than in situation r if the lower bound of $c_{t,r}$ is above zero. It means that *any* specification of the individual utilities that rationalizes the observed consumption behavior entails a value for $c_{t,r}$ that exceeds zero. Similarly, we can conclude that the individual is worse off in t than in r if the upper bound of $c_{t,r}$ is below zero. Finally, if the lower and upper bounds have opposite signs, we cannot reject the hypothesis that the individual is equally well off in both decision situations: we are unable to robustly (i.e., for any specification of the rationalizing utilities) conclude that the individual is better or worse off in t than in r.

current paper, our focus is on comparing observed price-income situations. Under specific assumptions regarding unobserved prices, it is fairly easy to extend our following reasoning to welfare comparisons that involve unobserved price-income regimes.

Goodness-of-fit. The revealed preference characterization in Definition 5 allows us to define sharp tests for rationalizable consumption behavior: either the data satisfy the testable N-GARP conditions or they do not. When the data do not satisfy these exact conditions, it is often interesting to empirically evaluate the degree of violation. For example, it may happen that the data are close to satisfying the exact rationalizability conditions, and we may want to include such *almost rationalizable* data in our further empirical analysis. To this end, we extend Afriat (1973)'s notion of Critical Cost Efficiency Index (CCEI) to our specific setting. Intuitively, the CCEI quantifies the goodness-of-fit of the rationalizability conditions in terms of minimal adjustments of the observed expenditure levels that are needed to exclude violations of the nonparametric rationalizability conditions. In other words, it quantifies the error that must be accounted for such that the (corrected) data satisfy the rationality restrictions.[7]

Formally, to apply the CCEI concept to our N-GARP characterization, we introduce a parameter $e \in [0,1]$. Correspondingly, we adjust the last three (if–then) conditions in Definition 5 for which $r = v$, while keeping the other conditions intact. That is, we only change the conditions for which $q_{r,v}$ is equal to the observed bundle q_v. This obtains the following adapted conditions (for all $r, s, t, v \in T$):

- if $u_t \geq u_v$, then $ep_v q_v \leq p_v q_{s,t}$,
- if $u_t > u_v$, then $ep_v q_v < p_v q_{s,t}$,
- if $u_t \geq u_v$, then $eq^i_{r,v} \leq q_{v,t}$ for all $i \in M$.

For a given data set S, the CCEI equals the highest value of e such that the consumption observations satisfy these adjusted rationalizability conditions.[8] Obviously, higher CCEI-values signal a better fit of the rationalizability conditions. Next, as argued by Apesteguia and Ballester (2015, Section V), the CCEI has two properties that are specifically attractive from a practical point of view. First, it satisfies continuity, which means that it never increases with the number of observations. Second, the CCEI satisfies rationality, which implies that it equals one if and only if the data are (exactly) rationalizable.

Let e^* represent the CCEI of a given data set S. Then, we can define the adjusted data set $S^* = (p_t, e^* q_t)_{t \in T}$ which, by construction, satisfies the N-GARP restrictions in Definition 5. For this adjusted set S^*, we can compute cost of living indices by using the nonparametric procedure outlined above. In the following section, our main empirical analysis will do so for the individuals with CCEI-values $e^* \geq 0.99$, which means that the observed behavior is sufficiently close to rationalizability.

4 Illustrative application

To evaluate the welfare effects of the 2008 financial crises, we make use of a balanced panel drawn from the 2007, 2009 and 2011 waves of the Panel Study of Income Dynamics

[7]The CCEI was originally introduced by Afriat (1973) and further developed by Varian (1990). Choi, Kariv, Müller, and Silverman (2014) used the CCEI in a large scale field experiment as a measure of consumers' decision making quality. Intuitively, they interpret low CCEI-values as revealing optimization errors arising from imperfect decision making quality. We may use a similar interpretation of the CCEI results in our empirical application in Section 4. See also Apesteguia and Ballester (2015) and Halevy, Persitz, and Zrill (2018) for related discussions.

[8]See Appendix B for more information concerning the computation of the CCEI.

(PSID). By considering only three PSID waves, we can show that our methodology enables an informative empirical analysis even for short time series of consumption observations.[9] Moreover, it seems more reasonable to assume stable individual preferences over a shorter time period. In Appendix D we demonstrate the robustness of our main qualitative conclusions for a longer panel containing four consumption observations per individual (adding the 2013 PSID wave to our original data set). This extra analysis also allows us to study the impact of the crisis over a longer time period.

Data and set-up. The PSID, which was initiated in 1968, is a widely used survey of a national representative sample of 18,000 individuals living in 5000 families in the United States. The data set contains information on income, wealth, health, marriage, childbearing, child development, education and other socio-demographic variables. Since 1999, the panel also provides additional expenditure information on a detailed set of consumption categories (see Blundell, Pistaferri, and Saporta-Eksten (2016) for more details).

Our empirical analysis specifically focuses on the welfare effects of the 2008 crisis for singles (with and without children). Thus, we exclude couples from our investigation, which also conveniently avoids preference aggregation issues associated with the welfare analysis of multi-person households.[10] We concentrate on individuals who are situated on the intensive margin of labor supply, that is, our subjects are actively working on the labor market in each period under study. We excluded the self-employed to avoid issues regarding the imputation of wages and the separation of consumption from work-related expenditures. After excluding observations with missing information (e.g. on wages, labor hours or consumption expenditures), we end up with a sample of 821 individuals.

Table 4 in Appendix C reports summary statistics for our sample. We assume that individuals spend their full potential income on four consumption categories: food, housing, leisure and other goods. We compute leisure quantities by assuming that each individual needs 8 hours per day for personal care and sleep. Leisure equals the available time that could have been spent on market work but was not (i.e., leisure per week = (24-8)*7 - market work). We calculate the individuals' weekly expenditures (i.e., nominal dollars per week) on the three remaining consumption categories (food, housing and other goods) as the reported annual expenditures divided by 52. The price of leisure equals the individual's hourly wage for market work. The prices of food, housing and other goods are region-specific consumer price indices that have been constructed by the Bureau of Labor Statistics.

For our empirical analysis, we take it that the normality assumption is arguably debatable for leisure. We only assume normality for the consumption categories food,

[9] In principle, it is possible to use our methodology with only two consumption observations per individual. However, it can be shown that, in such a case, the N-GARP-based lower bounds on the cost of living indices always equal the GARP-based lower bounds, by construction. Thus, by using three consumption observations per individual, we can illustrate the usefulness of normality assumptions for obtaining lower bounds that are more informative than the GARP-based bounds.

[10] Practical welfare analysis of multi-person households often adopts a unitary assumption, which models these households as single decision makers. However, this unitary assumption has been rejected by a large number of empirical studies (see, for example, Browning and Chiappori (1998) and Dauphin, El Lahga, Fortin, and Lacroix (2011)). This suggests the extension of our analysis towards collective household models, with multi-person households consisting of multiple decision makers, as an interesting avenue of follow-up research. Such an extension can build on Cherchye, De Rock, and Vermeulen (2007, 2011), who developed the revealed preference characterization of rational consumption for collective households.

housing and other goods. We effectively do believe it plausible that these expenditures are normal, all the more because they pertain to aggregate consumption categories. We will conduct a goodness-of-fit analysis (using the CCEI) as well as a welfare analysis (on the basis of cost of living indices) for the N-GARP condition associated with these normality assumptions. We will compare these N-GARP-based results with the GARP-based results that make no use of any normality assumption (recalling that N-GARP reduces to GARP if no good is assumed to be normal). In Appendix D, we also present a robustness check that considers the scenario in which all four consumption categories are assumed to be normal. Evidently, stronger normality assumptions generally imply a more powerful N-GARP-based analysis. Importantly, however, our principal qualitative conclusions remain unaffected.

In our following exercises, we will conduct separate N-GARP-based and GARP-based analyses for all 821 individuals that we observe. Using our notation of Section 2, this defines a data set S with 3 observations (i.e., $T = \{2007, 2009, 2011\}$) and 4 goods (i.e., $n = 4$) for every single in our sample. By analyzing each individual separately, we fully account for preference heterogeneity across individuals.

Goodness-of-fit. We begin by using Afriat's Critical Cost Efficiency Index (CCEI) to check data consistency with N-GARP and GARP for the sample of singles under study. Basically, the GARP-based CCEI results reveal how well the assumption of utility maximization fits the observed behavior, while the N-GARP-based CCEI results indicate the empirical fit of our normality assumptions (for food, housing and other goods) in addition to utility maximization. As explained in Section 3, the CCEI evaluates model fit in terms of necessary adjustments of observed expenditures to obtain data consistency with the (N-GARP and GARP) rationalizability conditions that are subject to evaluation. CCEI-values are situated between zero and one, with higher values signaling a better fit.

Table 1 summarizes our CCEI results. The first row shows the number of individuals who satisfy the exact N-GARP and GARP conditions (corresponding to CCEI $= 1$). The second row reports the number of individuals who are very close to rationalizability (characterized by CCEI ≥ 0.99). Generally, the CCEI-values for the N-GARP condition are below the CCEI-values for the GARP condition. This should not be surprising because, as explained above, the N-GARP condition is more stringent than the GARP condition. Importantly, we find that the average CCEI-value is very high for N-GARP as well as GARP: it equals 0.9912 for N-GARP and 0.9987 for GARP. However, we also observe that the behavior of some individuals turns out to be quite far from exact rationalizability. For example, the minimum CCEI-value equals 0.6774 for N-GARP and 0.7451 for GARP.

Overall, the results in Table 1 provide rather strong empirical support for N-GARP (as well as GARP) applied to our sample of individuals. In most cases, we need only (very) small expenditure adjustments to obtain consistency with the rationalizability conditions. In our following welfare analysis, we will focus on the subsample of 702 individuals with a N-GARP-based CCEI-value greater than or equal to 0.99. As explained above, such a high CCEI-value signals behavior that is very close to exactly rationalizable, which empirically motivates using the assumption of rationality (with normal demand) for our welfare analysis. Appendix D contains a robustness analysis that only includes the 584 individuals of which the N-GARP-based CCEI-value equals 1 (i.e., exactly rationalizable behavior). Comfortingly, this additional analysis yields the same main findings.

Table 1: Critical Cost Efficiency Index (CCEI)

	N-GARP	GARP
CCEI = 1	584 (71.13%)	782 (95.25%)
CCEI $\geq$ 0.99	702 (85.51%)	803 (97.81%)
mean	0.9912	0.9987
std. dev.	0.0297	0.0124
min	0.6774	0.7451
25%	0.9980	1.0000
50%	1.0000	1.0000
75%	1.0000	1.0000
max	1.0000	1.0000

Cost of living indices. We quantify the welfare effects of the 2008 crisis by calculating cost of living indices. For each individual in our sample, we estimate the difference in cost of living between 2007 and 2011. More formally, we define this as the difference between the actual income in 2011 and the income that would be required in the same year (at 2011 prices) to be equally well equally well off as in 2007:

$$c_{2011,2007} = \frac{x_{2011} - p_{2011}q_{2011,2007}}{x_{2011}},$$

We use the nonparametric set identification procedure outlined above. Particularly, we compute GARP-based and N-GARP-based lower and upper bound on $c_{2011,2007}$ by using the rationalizability restrictions associated with GARP (in Definition 3) and N-GARP (in Definition 5), respectively.

Table 2 gives a summary of our results for the sample of individuals under study. Columns 2-3 summarize our N-GARP-based bounds and columns 5-6 our GARP-based bounds. Correspondingly, Δ_n in column 4 and Δ_g in column 7 represent the differences between the respective upper and lower bounds. Finally, column 8 reports on the relative difference between Δ_n and Δ_g. This measures the extent to which the N-GARP-based bounds are tighter than the GARP-based bounds. In a sense, it quantifies the identifying power that specifically follows from our normality assumptions.

We observe that the N-GARP-based bounds are substantially tighter than the GARP-based bounds. The mean (median) difference between the N-GARP-based lower and upper bounds is 7.00 % (2.88%), which is much below the difference of 14.40% (9.36%) between the GARP-based bounds. Moreover, the relative difference between Δ_n and Δ_g amounts to no less than 50% for about half of our sample, again showing a significant increase of identifying power when imposing normality.

As a following exercise, Figures 2 and 3 depict the empirical cumulative distribution functions (CDFs) of our N-GARP-based and GARP-based lower and upper bounds for $c_{2011,2007}$. In line with our results in Table 2, the N-GARP-based CDFs are much closer to each other than the GARP-based CDFs. From all this, we may safely conclude that our (mild) normality assumptions do yield a considerably more informative welfare analysis. Further inspection of Table 2 and Figures 2-3 reveals that, for our specific data, this improvement in identifying power is mostly driven by lower upper bounds (and to a lesser degree by higher lower bounds).

Table 2: Bounds on $c_{2011,2007}$

	N-GARP-based			GARP-based			
	lower	upper	Δ_n	lower	upper	Δ_g	$\frac{\Delta_g - \Delta_n}{\Delta_g}$
mean	-0.0368	0.0332	0.0700	-0.0376	0.1064	0.1440	0.4686
std. dev.	0.2878	0.2473	0.1313	0.2878	0.2787	0.1678	0.3913
min	-3.0441	-2.4921	0.0000	-3.0441	-2.4888	0.0000	0.0000
25%	-0.1199	-0.0417	0.0083	-0.1199	0.0000	0.0417	0.0109
50%	-0.0051	0.0072	0.0288	-0.0061	0.0379	0.0936	0.4999
75%	0.0835	0.1310	0.0758	0.0835	0.2547	0.1933	0.8647
max	0.8305	0.8973	2.0989	0.8302	0.8993	2.2848	1.0000

Figure 2: CDF of N-GARP-based bounds

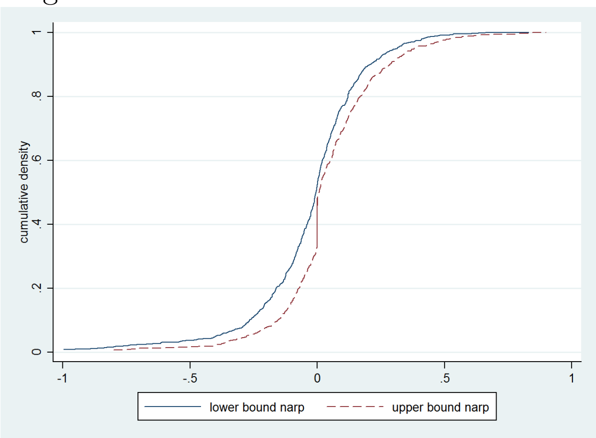
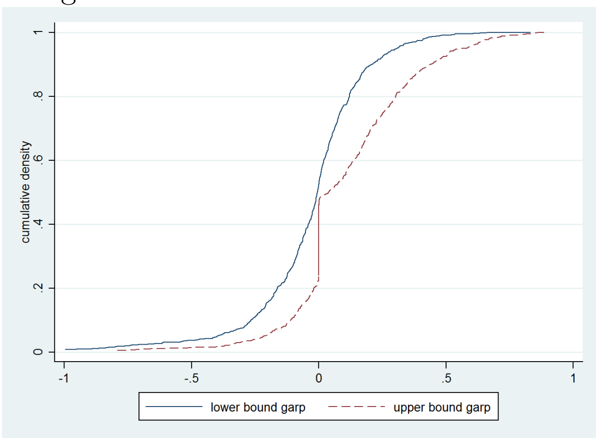


Figure 3: CDF of GARP-based bounds



Better-off and worse-off individuals. As explained in Section 3, we can state that an individual is better off in 2011 than in 2007 if the lower bound of $c_{2011,2007}$ (LB) exceeds zero, while the individual is worse off in 2011 if the upper bound of $c_{2011,2007}$ (UB) is below zero. These better-off and worse-off classifications are robust in that they hold for any specification of the individual utilities that rationalize the observed consumption behavior. Finally, if the lower and upper bounds have opposite signs (i.e., LB ≤ 0 and UB ≥ 0), we cannot robustly conclude that the individual is better or worse off in 2011.

Rows 2-4 of Table 3 give the fractions of individuals that are classified as better-off, worse-off and cannot-say according to our N-GARP-based (column 3) and GARP-based (column 4) bounds for $c_{2011,2007}$. In column 5, we report the associated Z-score for tests of equal proportions in the N-GARP-based and GARP-based categories. Using our N-GARP-based bounds, we classify 33.19% of our individuals as worse off and 47.86% as better off, with a residual 18.95% falling in the cannot-say category. By contrast, our GARP-based bounds classify only 22.36% as worse off and 47.58% as better off, now leaving 30.06% in the cannot-say category. From our Z-score results, we learn that the fraction of individuals in the worse-off category is significantly higher in the N-GARP-based classification than in the GARP-based classification. Correspondingly, the fraction of individuals in the cannot-say category is significantly lower in the N-GARP-based classification than in the GARP-based classification. This shows that using normality assumptions obtains a significantly more informative classification of individuals after the 2008 crisis. Particularly, the N-GARP restrictions for rational behavior enable a considerably better identification of the individuals who suffered from a welfare loss after the 2008 crisis.

Overall, Table 3 provides further support for our earlier conclusion that (mild) normality assumptions can substantially improve the informative value of nonparametric welfare analysis. Moreover, our cost of living estimates reveal quite some heterogeneity across individuals. In Appendix D, we investigate this further by relating these cost of living estimates to observable individual characteristics. A main finding is that individuals with higher potential incomes in 2007 have been hit more severely by the crisis.[11] Next, we also observe that having children correlates significantly with our estimated welfare effects. At this point, it is worth recalling that our empirical analysis considers singles who remained employed after the crisis. This contrasts with existing studies, which mainly focused on the extensive margin of labor supply.

Table 3: Worse-off and better-off individuals

		N-GARP	GARP	Z-score
UB ≤ 0	Worse off in 2011	33.19	22.36	4.53***
LB ≥ 0	Better off in 201	47.86	47.58	0.11
LB ≤ 0 and $0 \leq$ UB	Cannot-say	18.95	30.06	-4.84***

*** p<0.01, ** p<0.05, * p<0.1

[11] This is mainly driven by the fact that, in our sample, people with higher initial wages suffered from more severe wage drops after the crisis. See our discussion of Table 12 in Appendix D.

5 Conclusion

We presented a revealed preference characterization of rational consumer behavior under the assumption of normal demand. The characterization is easily operationalized in practice, and it is flexible in that it can impose normality on any subset of goods. We have also shown the use of our characterization to analyze the welfare effects (in terms of cost of living indices) of changing price-income regimes. As normality is often a plausible assumption to make, this provides a useful toolkit to remediate the lack of power that is frequently associated with empirical revealed preference analysis.

We used our novel methodology to evaluate the welfare impact of the 2008 financial crisis for individuals situated on the intensive margin of labor supply. Particularly, we studied the labor supply behavior of a sample of singles drawn from the PSID. Our main focus was on comparing the goodness-of-fit and identifying power of our nonparametric characterization of utility maximization, with and without normality assumptions. We found that the goodness-of-fit results were hardly affected when imposing normality, providing good empirical support for our normality hypotheses. Next, and more importantly, we showed that using mild normality assumptions yields a substantially more powerful empirical welfare analysis: it obtained considerably sharper set identification of individuals' cost of living indices, and a significantly more informative classification of better-off and worse-off individuals after the 2008 crisis.

References

Afriat, S. N., 1967. The construction of utility functions from expenditure data. International Economic Review 8 (1), 67–77.

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

Apesteguia, J., Ballester, M. A., 2015. A measure of rationality and welfare. Journal of Political Economy 123, 1278–1310.

Banks, J., Blundell, R., Lewbel, A., 1997. Quadratic engel curves and consumer demand. The Review of Economics and Statistics 79, 527–539.

Blundell, R., Browning, M., Cherchye, L., Crawford, I., De Rock, B., Vermeulen, F., 2015. Sharp for SARP: Nonparametric bounds on the behavioral and welfare effects of price changes. American Economic Journal: Microeconomics 7, 43–60.

Blundell, R., Browning, M., Crawford, I., 2007. Improving revealed preference bounds on demand responses. International Economic Review 48, 1227–1244.

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

Blundell, R., Pistaferri, L., Saporta-Eksten, I., 2016. Consumption inequality and family labor supply. The American Economic Review 106, 387–435.

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

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

Browning, M., Chiappori, P.-A., 1998. Efficient intra-household allocations: a general characterization and empirical tests. Econometrica, 1241–1278.

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

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

Cherchye, L., Demuynck, T., De Rock, B., 2018. Normality of demand in a two-goods setting. Journal of Economic Theory 173, 361–382.

Choi, S., Kariv, S., Müller, W., Silverman, D., 2014. Who is (more) rational? The American Economic Review 104, 1518–1550.

Christensen, L. R., Jorgenson, D. W., Lau, L. J., 1975. Transcendental logarithmic utility functions. American Economic Review 65, 367–383.

Dauphin, A., El Lahga, A.-R., Fortin, B., Lacroix, G., 2011. Are children decision-makers within the household? The Economic Journal 121, 871–903.

Deaton, A., 2011. The financial crisis and the well-being of americans: 2011 OEP Hicks lecture. Oxford Economic Papers 64, 1–26.

Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. The American Economic Review 70, 312–326.

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

Fisher, F. M., 1990. Normal goods and the expenditure function. Journal of Economic Theory 51, 431–433.

Goodman, C. J., Mance, S. M., 2011. Employment loss and the 2007–09 recession: an overview. Monthly Labor Review 134, 3–12.

Halevy, Y., Persitz, D., Zrill, L., 2018. Parametric recoverability of preferences. Journal of Political Economy forthcoming.

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

Hurd, M. D., Rohwedder, S., 2010. Effects of the financial crisis and great recession on american households. Tech. rep., National Bureau of Economic Research.

Nishimura, H., Ok, E. A., Quah, J. K.-H., 2017. A comprehensive approach to revealed preference theory. The American Economic Review 107, 1239–1263.

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

Samuelson, P. A., 1974. Complementarity: An essay on the 40th anniversary of the hicks-allen revolution in demand theory. Journal of Economic literature 12, 1255–1289.

Varian, H. R., 1982. The nonparametric approach to demand analysis. Econometrica, 945–973.

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

Varian, H. R., 1990. Goodness-of-fit in optimizing models. Journal of Econometrics 46, 125–140.

Verick, S., 2009. Who is hit hardest during a financial crisis? the vulnerability of young men and women to unemployment in an economic downturn. Tech. rep.

A Proof of Proposition 2

Proof. We begin by showing necessity of our N-GARP conditions in Definition 5, i.e. any observed demand originating from utility maximization under normality must satisfy the conditions in Proposition 2. In a following step, we show sufficiency of the N-GARP conditions by using the auxiliary results stated in Lemmas 1 and 2 below. For simplicity, our following proof considers the case where all the goods are normal. The proof for the case where a strict subset of goods is normal develops directly analogously.

Necessity. Let $S = (p_t, q_t)_{t \in T}$ be rationalizable under normal demand (on the set $M = \{1, ..., n\}$) by the utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ and expansion paths $q_t : \mathbb{R}_+ \to \mathbb{R}^n_+$ that are monotone in x for all goods i.

For all $t \in T$, define $u_t = u(q_t)$ and, for all $t, v \in T$, define $q_{t,v}$ as the bundle on the intersection of expansion path $q_t(x)$ and the indifference curve through q_v. Now, by definition, we have that the intersection of $q_t(x)$ with the indifference curve through q_t is q_t. This gives the first N-GARP condition in Definition 5, i.e. $q_{t,t} = q_t$ for all $t \in T$.

For the second N-GARP condition, let $u_t \geq u_v$ and assume (towards a contradiction) that $p_r q_{r,v} > p_r q_{s,t}$. This means that $q_{r,v} P^D q_{s,t}$ and implies that the utility level at $q_{r,v}(= u_v)$ is strictly above the utility level at $q_{s,t}(= u_t)$, a contradiction. Thus, we have $p_r q_{r,v} \leq p_r q_{s,t}$. We can derive the third N-GARP condition in a directly similar way.

Finally, for the fourth N-GARP condition, we observe that, if $u_t = u(q_t) \geq u_v = u(q_v)$, then the indifference curve through q_v lies below the indifference curve through q_t. This implies that, for any monotone income expansion path $q_r(x)$, we have $q^i_{r,t} \geq q^i_{r,v}$ for every normal good i.

Sufficiency. Suppose the data set $S = \{(p_t, q_t)\}_{t \in T}$ is consistent with the N-GARP conditions in Definition 5 (for the set $M = \{1, ..., n\}$). We want to construct a utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ and expansion paths $q_t : \mathbb{R}_+ \to \mathbb{R}^n_+$ (which are monotone in x for each good i) that generate the observed demand.

Let $(u_t)_{t \in T}$ be the utility numbers and $(q_{t,v})_{t,v \in T}$ the bundles that satisfy the N-GARP restrictions for the given set S. If $u_t = u_v$ for $t \neq v$, then for any $r \in T$ it is always possible to remove one of the bundles $q_{r,t}$ or $q_{r,v}$ from the N-GARP conditions such that the restrictions still hold. Hence, without loss of generality, we can remove the duplicates in the set $(u_t)_{t\in T}$ for which $u_t = u_v$. As such, we obtain a reduced set $(u_t)_{t\in R\subseteq T}$ and $(q_{t,v})_{t\in T, v\in R\subseteq T}$ that satisfy the N-GARP conditions. Now, let us reorder the elements in R such that $u_1 < u_2 < \cdots < u_{|R|}$.

As a first step, we construct the income expansion paths $q_t(x) = q(p_t, x)$, in the following way (for any x and any $t \in R$):

- If $x > p_t q_{t,|R|}$, then $q_t(x) = \gamma q_{t,|R|}$ with $\gamma = \frac{x}{p_t q_{t,|R|}}$.
  We say that $q_t(x)$ is of level $|R| + 1$. Observe that $p_t q_t(x) = x$.
- If $x \leq p_t q_{t,1}$, then $q_t(x) = \gamma q_{t,1}$ with $\gamma = \frac{x}{p_t q_{t,1}}$.
  We say that $q_t(x)$ is of level 1. Again, observe that $p_t q_t(x) = x$.
- If $p_t q_{t,1} < x \leq p_t q_{t,|R|}$, then the ordering of the observations and the fourth condition of N-GARP imply that there exists a unique $v \in R$ such that $p_t q_{t,v-1} < x \leq p_t q_{t,v}$. As such, there exists a unique $\alpha \in (0, 1]$, such that
  $$x = \alpha(p_t q_{t,v}) + (1 - \alpha)(p_t q_{t,v-1}).$$
  Given this $\alpha \in (0, 1]$, define
  $$q_t(x) = \alpha q_{t,v} + (1 - \alpha) q_{t,v-1}.$$
  In this case, we will say that $q_t(x)$ is of level v. Also, $p_t q_t(x) = x$.

Observe that the $q^i_t(x)$ are monotone in x for all goods i, continuous and generate the observed demand by construction.

As a following step, we obtain our result by showing that, for any $N \in \mathbb{N}$, any sequence of income levels $x_1, x_2, \cdots, x_N$ and any sequence of observations $t_1, \ldots, t_N \in T$, the set $(p_{t_i}, q_{t_i}(x_i))_{i\leq N}$ satisfies GARP. Suppose (towards a contradiction) that the result does not hold. Then, there is a $N \in \mathbb{N}$, a sequence $x_1, x_2, \cdots, x_N$ of income levels and a sequence $t_1, t_2, \cdots, t_N$ of observations that violate GARP. That is,

$$\begin{aligned}
p_{t_1} q_{t_1}(x_1) &\geq p_{t_1} q_{t_2}(x_2), \\
p_{t_2} q_{t_2}(x_2) &\geq p_{t_2} q_{t_3}(x_3), \\
&\vdots \\
p_{t_N} q_{t_N}(x_N) &\geq p_{t_N} q_{t_1}(x_1),
\end{aligned}$$

with at least one strict inequality. From Lemma 1, we know that the level of the bundles along the cycle cannot increase. Also, it cannot strictly decrease as this would mean that somewhere along the cycle it must strictly increase. This means that the level of all bundles should be the same, say r. We distinguish three cases for r:

- If $r = |R| + 1$, then there are $\gamma_1, \ldots, \gamma_N$ such that
  $$\begin{aligned}
  q_{t_1}(x_1) &= \gamma_1 q_{t_1,R}, \\
  q_{t_2}(x_2) &= \gamma_2 q_{t_2,R}, \\
  &\cdots \\
  q_{t_N}(x_N) &= \gamma_N q_{t_N,R}.
  \end{aligned}$$
  By Lemma 2, we have $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n \geq \gamma_N \geq \gamma_1$, with at least one strict inequality, a contradiction.

- If $r = 1$, then there are $\gamma_1, \ldots, \gamma_N$ such that

$$
\begin{aligned}
q_{t_1}(x_1) &= \gamma_1 q_{t_1,1}, \\
q_{t_2}(x_2) &= \gamma_2 q_{t_2,1}, \\
&\ldots \\
q_{t_N}(x_N) &= \gamma_N q_{t_N,1}.
\end{aligned}
$$

  Again, by Lemma 2, we have $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_N \geq \gamma_1$, with at least one strict inequality, a contradiction.

- If $1 < r < |R| + 1$, then there are $\alpha_1, \ldots, \alpha_N \in (0, 1]$ such that

$$
\begin{aligned}
q_{t_1}(x_1) &= \alpha_1 q_{t_1,r} + (1 - \alpha_1) q_{t_1,r-1}, \\
q_{t_2}(x_2) &= \alpha_2 q_{t_2,r} + (1 - \alpha_2) q_{t_2,r-1}, \\
&\ldots \\
q_{t_N}(x_N) &= \alpha_N q_{t_N,r} + (1 - \alpha_N) q_{t_N,r-1}.
\end{aligned}
$$

  By Lemma 2, we have $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_N \geq \alpha_1$, with at least one strict inequality, a contradiction.

Thus, we conclude that, for any $N \in \mathbb{N}$, any sequence $x_1, x_2, \cdots, x_N$ of income levels and any sequence $t_1, t_2, \cdots, t_N$ of observations, the set $(p_{t_i}, q_{t_i}(x_i))_{i \leq N}$ satisfies GARP. Then, Proposition 3 of Nishimura, Ok, and Quah (2017) implies that there exists a continuous and strictly increasing utility function that rationalizes our constructed expansion paths. ■

Lemma 1 *If $p_t q_t(x) \geq p_t q_v(y)$, then the level of $q_v(y)$ is not strictly higher than the level of $q_t(x)$.*

Proof of Lemma 1. Let $q_v(y)$ be of level r and $q_t(x)$ be of level s. Assume (towards a contradiction) that Lemma 1 does not hold, that is, $r > s$. Then,

- If $r(= |R| + 1) > s(= 1)$, then $p_t q_{t,1} \leq p_t q_{v,|R|}$, so

$$p_t q_t(x) \leq p_t q_{t,1} \leq p_t q_{v,|R|} < p_t q_v(y),$$

  a contradiction.

- If $r(= |R| + 1) > s > 1$, then $p_t q_{t,s} \leq p_t q_{v,|R|}$ and $p_t q_{t,s-1} < p_t q_{v,|R|}$. As such, if $q_t(x) = \alpha q_{t,s} + (1 - \alpha) q_{t,s-1}$ with $\alpha \in (0, 1]$, then

$$p_t q_t(x) = \alpha(p_t q_{t,s}) + (1 - \alpha)(p_t q_{t,s-1}) \leq p_t q_{v,|R|} < p_t q_v(y),$$

  a contradiction.

- If $|R| + 1 > r > s = 1$, then $p_t q_{t,1} \leq p_t q_{v,r-1}$ and $p_t q_{t,1} < p_t q_{v,r}$. As such, if $q_v(y) = \beta q_{v,r} + (1 - \beta) q_{v,r-1}$ with $\beta \in (0, 1]$, then

$$p_t q_t(x) \leq p_t q_{t,1} < \beta p_t q_{v,r} + (1 - \beta) p_t q_{v,r-1} = q_v(y).$$

- If $|R| + 1 > r > s > 1$, then $p_t q_{t,s} \leq p_t q_{v,r-1}$, $p_t q_{t,s} < p_t q_{v,r}$, $p_t q_{t,s-1} < p_t q_{v,r-1}$ and $p_t q_{t,s-1} < p_t q_{v,r}$. This implies that any convex combination of $p_t q_{t,s}$ and $p_t q_{t,s-1}$ must always be strictly smaller than any convex combination of $p_t q_{v,r-1}$ and $p_t q_{v,r}$. As such, if $q_t(x) = \alpha q_{t,s} + (1-\alpha) q_{t,s-1}$ and $q_v(y) = \beta q_{v,r} + (1-\beta) q_{v,r-1}$ with $\alpha, \beta \in (0, 1]$, then

$$\begin{aligned} p_t q_t(x) &= \alpha p_t q_{t,s} + (1-\alpha) p_t q_{t,s-1} \\ &\leq \beta p_t q_{v,r} + (1-\beta) p_t q_{v,r-1} = p_t q_v(y), \end{aligned}$$

  a contradiction.

■

Lemma 2 *Let $p_t q_t(x) \geq p_t q_v(y)$, with the level of $q_t(x)$ the same as the level of $q_v(y)$. Then:*

- *If both $q_t(x)$ and $q_v(y)$ are of level $|R| + 1$, and $q_t(x) = \gamma q_{t,|R|}, q_v(y) = \delta q_{v,|R|}$, we have $\gamma \geq \delta$. In addition, if $p_t q_t(x) > p_t q_v(y)$, then $\gamma > \delta$.*
- *If both $q_t(x)$ and $q_v(y)$ are of level 1, and $q_t(x) = \gamma q_{t,1}, q_v(y) = \delta q_{v,1}$, we have $\gamma \geq \delta$. In addition, if $p_t q_t(x) > p_t q_v(y)$, then $\gamma > \delta$.*
- *If both $q_t(x)$ and $q_v(y)$ are of level r with $1 < r < |R| + 1$, and $q_t(x) = \alpha q_{t,r} + (1-\alpha) q_{t,r-1}$, $q_v(y) = \beta q_{v,r} + (1-\beta) q_{v,r-1}$ with $\alpha, \beta \in (0, 1]$, then we have $\alpha \geq \beta$. In addition, if $p_t q_t(x) > p_t q_v(y)$, then $\alpha > \beta$.*

Proof of Lemma 2. We look at the three cases separately:

Suppose that both $q_t(x)$ and $q_v(y)$ are of level $|R| + 1$. From the second N-GARP condition in Definition 5, we know that $p_t q_{t,|R|} \leq p_t q_{v,|R|}$. This implies

$$\begin{aligned} \delta p_t q_{v,|R|} &= p_t q_v(y) \\ &\leq p_t q_t(x) = \gamma p_t q_{t,|R|} \\ &\leq \gamma p_t q_{v,|R|}. \end{aligned}$$

So, $\delta \leq \gamma$ with a strict inequality if $p_t q_t(x) > p_t q_v(y)$.

Suppose that both $q_t(x)$ and $q_v(y)$ are of level 1. From the second N-GARP condition in Definition 5, we know that $p_t q_{t,1} \leq p_t q_{v,1}$. This implies

$$\begin{aligned} \delta p_t q_{v,1} &= p_t q_v(y) \\ &\leq p_t q_t(x) = \gamma p_t q_{t,1} \\ &\leq \gamma p_t q_{v,1}. \end{aligned}$$

So, $\delta \leq \gamma$ with a strict inequality if $p_t q_t(x) > p_t q_v(y)$.

Suppose that both $q_t(x)$ and $q_v(y)$ are of level r with $|R| + 1 > r > 1$. From the second N-GARP condition in Definition 5, we know that $p_t q_{t,r} \leq p_t q_{v,r}$ and $p_t q_{t,r-1} \leq p_t q_{v,r-1}$. As such,

$$\begin{aligned} \alpha(p_t q_{t,r}) + (1-\alpha)(p_t q_{t,r-1}) &= p_t q_t(x) \\ &\geq p_t q_v(y) \\ &= \beta(p_t q_{v,r}) + (1-\beta)(p_t q_{v,r-1}) \\ &\geq \beta p_t q_{t,r} + (1-\beta) p_t q_{t,r-1}. \end{aligned}$$

This is equivalent to the condition $(\alpha - \beta)(p_t q_{t,r} - p_t q_{t,r-1}) \geq 0$. The third N-GARP condition in Definition 5 implies that $p_t q_{t,r} > p_t q_{t,r-1}$. As such, it must be that $\alpha \geq \beta$, with a strict inequality if $p_t q_t(x) > p_t q_v(y)$. ■

B Practical implementation

Mixed integer programming formulation of N-GARP. The N-GARP conditions in Definition 5 can be reformulated in terms of linear inequalities that are characterized by (binary) integer variables.

Proposition 3 *A data set $S = \{(p_t, q_t)\}_{t \in T}$ satisfies the N-GARP conditions in Definition 5 if and only if there exist binary numbers $r_{t,v} \in \{0, 1\}$ and numbers $u_t \in [0, 1]$ such that, for all $r, s, t, v, s \in T$,*

- $q_{t,t} = q_t$,
- $u_t - u_v < r_{t,v}$,
- $(r_{v,t} - 1) < u_v - u_t$,
- $p_r q_{r,v} - p_r q_{s,t} < r_{v,t} A$,
- $A(r_{t,v} - 1) \leq (p_r q_{s,t} - p_r q_{r,v})$,
- $B(r_{t,v} - 1) \leq q^i_{r,t} - q^i_{r,v}$ *for all* $i \in M$.

where A is a fixed number greater than any possible value $p_r q_{r,v} (r, v \in T)$ and B is a fixed number greater than any $q^i_{r,v} (i \in M, r, v \in T)$. By default A and B are finite numbers.

Proof of Proposition 3. Necessity. Assume that the N-GARP conditions in Definition 5 are satisfied. Let us use the same solution and define $r_{t,v} = 1$ if and only if $u_t \geq u_v$. The the first three conditions above are satisfied by default. By the definition of A, the fourth condition is only binding if $r_{v,t} = 0$, which means that $u_t > u_v$. In this case, Definition 5 implies that $p_r q_{r,v} < p_r q_{s,t}$ and the condition holds. Similarly, the fifth condition is binding only if $r_{t,v} = 1$, which implies that $u_t \geq u_v$ and thus that $p_r q_{s,t} \geq p_r q_{r,v}$. Finally, the last condition only binds if $r_{t,v} = 1$, which implies that $u_t \geq u_v$, In this case the last condition of Definition 5 gives $q^i_{r,v} \leq q^i_{r,t}$. We can thus conclude that the conditions of Proposition 3 are feasible whenever Definition 5 is satisfied.
Sufficiency. Assume that there exists a solution for the conditions in Proposition 3. Then we can show that the conditions in Definition 5 are also satisfied for the same solution. The first condition in Definition 5 is satisfied by default. For the second condition, if $u_t \geq u_v$ then $r_{t,v} = 1$ by the second condition above and as such the fifth condition implies that $p_r q_{s,t} \geq p_r q_{r,v}$. This shows that the second condition of Definition 5 holds. Next, let $u_t > u_v$. If, towards a contradiction, $p_r q_{r,v} \geq p_r q_{s,t}$, then, by the fourth condition above, $r_{v,t} = 1$. This implies, by the third condition, that $u_v \geq u_t$, a contradiction. This shows that the third condition of Definition 5 holds. For the final condition, let $u_t \geq u_v$. Then, by the second condition above, $r_{t,v} = 1$ and, by the last condition, $q^i_{r,t} \geq q^i_{r,v}$, as was to be shown. ■

Computing the CCEI. The CCEI is found by solving the following optimization problem:

$$
\begin{aligned}
& \max \ e \\
\text{s.t. } & 0 \leq e \leq 1 \\
& \forall t \in T : 0 \leq u_t \leq 1 \\
& \forall t \in T : q_{t,t} = q_t \\
& \forall t, v, r, s \in T \text{ such that } r \neq v : u_t \geq u_v \rightarrow p_r q_{r,v} \leq p_r q_{s,t} \\
& \forall t, v, r, s \in T \text{ such that } r \neq v : u_t > u_v \rightarrow p_r q_{r,v} < p_r q_{s,t} \\
& \forall i \in M, \forall t, v, r \in T, \text{ such that } r \neq v : u_t \geq u_v \rightarrow q^i_{r,v} \leq q^i_{r,t} \\
& \forall t, v, r, s \in T \text{ such that } r = v : u_t \geq u_v \rightarrow e p_r q_{r,v} \leq p_r q_{s,t} \\
& \forall t, v, r, s \in T \text{ such that } r = v : u_t > u_v \rightarrow e p_r q_{r,v} < p_r q_{s,t} \\
& \forall i \in M, \forall t, v, r \in T \text{ such that } r = v : u_t \geq u_v \rightarrow e q^i_{r,v} \leq q^i_{r,t}.
\end{aligned}
$$

The if–then conditions can be reformulated in terms of linear restrictions with binary variables, following our reasoning leading up to Proposition 3. As a result, the above optimization problem can be reformulated as a mixed integer linear programming problem.

C Data

Table 4 provides a summary of the data set that we use in our empirical application. As explained in the main text, we assume that the individuals spend their full potential incomes on four different consumption categories: leisure, food, housing and other goods. Table 4 reports information on prices, quantities, incomes and some demographics for our sample of 821 singles. The subscripts 07, 09 and 11 refer to the years 2007, 2009 and 2011, respectively.

We compute leisure quantities by assuming that each individual needs 8 hours per day for personal care and sleep. Leisure equals the available time that could have been spent on market work but was not (i.e., leisure per week = (24-8)*7 - market work). Food expenditures include food at home, delivered and eaten away from home. Housing expenditures include mortgage and loan payments, rent, property tax, insurance, utilities, cable tv, telephone, internet charges, home repairs and home furnishing. Others expenditures include health, transportation, education and childcare. We calculate the individuals' weekly expenditures (i.e., nominal dollars per week) on the three remaining consumption categories (food, housing and other goods) as the reported annual expenditures divided by 52.

The price of leisure equals the individual's hourly wage for market work. The prices of food, housing and other goods are region-specific consumer price indices that have been constructed by the Bureau of Labor Statistics.

D Additional empirical results

In this appendix, we first provide several robustness checks of our empirical results discussed in Section 4 of the main text. These checks largely confirm our principal conclusions. In a following step, we conduct a regression analysis that relates our estimated cost

Table 4: Summary statistics

	mean	std. dev.	min	max
$qfood_{11}$	0.43	0.27	0.00	1.99
$qhouse_{11}$	1.20	2.06	0.00	56.28
$qother_{11}$	0.72	0.66	0.00	6.94
$qleisure_{11}$	71.35	11.00	16.00	111.00
$qfood_{09}$	0.41	0.26	0.00	2.13
$qhouse_{09}$	1.08	0.69	0.00	7.06
$qother_{09}$	0.82	1.24	0.00	22.86
$qleisure_{09}$	72.98	10.12	22.00	111.00
$qfood_{07}$	0.44	0.30	0.00	2.25
$qhouse_{07}$	1.17	1.38	0.00	22.60
$qother_{07}$	0.82	0.75	0.00	6.03
$qleisure_{07}$	70.31	12.15	12.00	105.00
$pfood_{11}$	226.53	4.00	220.43	233.20
$phouse_{11}$	213.27	17.03	199.98	248.68
$pother_{11}$	238.61	2.58	235.89	241.36
$pleisure_{11}$	20.55	17.58	0.50	180.85
$pfood_{09}$	217.00	4.35	211.09	224.35
$phouse_{09}$	211.90	16.48	197.21	243.76
$pother_{09}$	209.32	3.98	205.15	214.13
$pleisure_{09}$	19.66	15.32	2.05	165.52
$pfood_{07}$	201.09	4.44	195.48	207.76
$phouse_{07}$	204.13	15.99	193.38	236.25
$pother_{07}$	205.29	2.57	202.62	208.21
$pleisure_{07}$	16.46	11.95	2.15	149.29
$consumption_{07}$	1649.61	1070.51	289.31	13231.01
$consumption_{09}$	1919.21	1294.08	245.85	13179.54
$consumption_{11}$	1973.16	1442.99	181.25	13235.89
$fullincome_{07}$	1842.97	1338.09	240.80	16720.48
$fullincome_{09}$	2202.09	1715.59	229.60	18538.24
$fullincome_{11}$	2301.66	1968.76	56.00	20255.20
$nonlabor_{07}$	-193.36	513.20	-3489.47	4213.92
$nonlabor_{09}$	-282.88	617.01	-5358.70	4887.96
$nonlabor_{11}$	-328.50	802.93	-7999.98	10699.09
age_{07}	37.95	13.38	18.00	81.00
male	0.34	0.47	0.00	1.00
$homeowner_{07}$	0.36	0.48	0.00	1.00
have $children_{07}$	0.31	0.46	0.00	1.00
number. of $children_{07}$	0.54	0.96	0.00	6.00
years of $education_{07}$	13.53	2.10	6.00	17.00

of living indices to observable individual characteristics. This provides an (exploratory) investigation of who has been affected by the 2008 crisis.

Cost of living indices. As a first robustness check, Table 5 summarizes our N-GARP-based and GARP-based estimated bounds on $c_{2011,2007}$ for the 584 individuals whose behavior is exactly rationalizable under normal demand (i.e., N-GARP-based CCEI equals 1). We observe that the results are closely similar to the ones contained in Table 2 in the main text.

Next, Table 6 reports summary statistics for the estimated bounds on $c_{2011,2007}$ when assuming normality of all four consumption categories (food, housing, other goods and leisure). Similar to Table 5, we focus on the (415) individuals that are exactly rationalizable under these normality assumptions (i.e., N-GARP-based CCEI equals 1). We calculate three types of bounds for this sample of individuals: N-GARP-I bounds, which exploit normality for all four consumption categories, N-GARP-II bounds, which (only) exploit normality for the three categories food, housing and other goods (as in our main exercise), and GARP bounds, which do not use any normality assumption. As can be expected, the N-GARP-I bounds are tighter than the N-GARP-II bounds, and these N-GARP-II bounds are tighter than the GARP bounds. The relative improvement of the N-GARP-I bounds over the N-GARP-II and GARP bounds is quite substantial, which shows the empirical bite of the assumption that leisure is normal.

Table 5: Bounds on $c_{2011,2007}$ for individuals with N-GARP-based CCEI = 1

	N-GARP-based			GARP-based			
	min	max	Δ_n	min	max	Δ_g	$\frac{\Delta_g-\Delta_n}{\Delta_g}$
mean	-0.0483	0.0292	0.0775	-0.0491	0.0987	0.1479	0.4232
std. dev.	0.3084	0.2630	0.1382	0.3084	0.2928	0.1761	0.3789
min	-3.0441	-2.4921	0.0003	-3.0441	-2.4888	0.0013	0.0000
25%	-0.1278	-0.0492	0.0123	-0.1300	0.0000	0.0419	0.0000
50%	-0.0122	0.0000	0.0349	-0.0123	0.0000	0.0947	0.4008
75%	0.0820	0.1363	0.0852	0.0820	0.2492	0.1943	0.8182
max	0.8305	0.8973	2.0989	0.8302	0.8993	2.2848	0.9978

Table 6: N-GARP-I, N-GARP-II and GARP bounds on $c_{2011,2007}$

	N-GARP-I			N-GARP-II			GARP				
	min	max	Δ_{n1}	min	max	Δ_{n2}	min	max	Δ_g	$\frac{\Delta_g-\Delta_{n1}}{\Delta_g}$	$\frac{\Delta_g-\Delta_{n2}}{\Delta_g}$
mean	-0.052	0.000	0.052	-0.053	0.030	0.082	-0.054	0.112	0.166	0.721	0.471
std. dev.	0.284	0.230	0.133	0.286	0.217	0.155	0.286	0.254	0.193	0.243	0.378
min	-3.044	-0.945	0.000	-3.044	-0.945	0.000	-3.044	-0.792	0.002	0.000	0.000
25%	-0.157	-0.099	0.006	-0.157	-0.051	0.011	-0.163	0.000	0.049	0.588	0.035
50%	-0.012	-0.002	0.018	-0.012	0.000	0.037	-0.012	0.000	0.110	0.793	0.516
75%	0.081	0.114	0.053	0.079	0.132	0.086	0.078	0.277	0.211	0.915	0.845
max	0.659	0.897	2.099	0.659	0.897	2.099	0.659	0.899	2.285	0.999	0.998

Better-off and worse-off individuals. As a following robustness check of our results in Section 4, we consider the classification of worse-off, better-off and cannot-say individuals for two alternative scenarios: the first scenario uses the N-GARP-based and GARP-based classifications for the 584 individuals of which the N-GARP-based CCEI equals 1 (also included in Table 5); the second scenario uses the GARP-based classification for the 782 individuals whose behavior is exactly rationalizable when not imposing normality on any good (i.e., GARP-based CCEI equals 1).

The results for the two scenarios are summarized in Table 7. Comfortingly, we find that the results in Table 7 are generally close to the ones in Table 3 that we discuss in the main text. Again, it suggests that our main qualitative conclusions are robust.

		N-GARP-CCEI=1 (584 individuals)		GARP-CCEI=1 (782 individuals)
		N-GARP	GARP	GARP
UB ≤ 0	Worse off in 2011	33.56	22.6	22.38
LB ≥ 0	Better off in 2011	45.21	45.03	48.59
LB ≤ 0 and $0 \leq$ UB	Cannot-say	21.23	32.36	29.03

Table 7: Worse-off and better-off individuals for individuals with N-GARP-based CCEI=1 and GARP-based CCEI=1

Four PSID waves: 2007, 2009, 2011 and 2013. Next, we check robustness of our main findings for a longer panel containing four consumption observations per individual (adding the 2013 PSID wave to our original data set). The following Tables 8, 9 and 10 have a directly analogous interpretation as the Tables 1, 2 and 3 that we discussed in the main text.

Generally, we can conclude that the results in Tables 8, 9 and 10 are fairly close to those in Tables 1, 2 and 3. For our application, adding a consumption observation (i.e., PSID wave) per individual only moderately affects our goodness-of-fit and cost of living results.

Table 8: Critical Cost Efficiency Index (CCEI); 4 waves

	N-GARP	GARP
CCEI=1	368 (53.18%)	630 (91.04%)
CCEI $\geq$ 0.99	493 (71.24%)	665 (96.10%)
mean	0.9779	0.9975
std. dev.	0.0520	0.0160
min	0.6235	0.7456
25%	0.9849	1.0000
50%	1.0000	1.0000
75%	1.0000	1.0000
max	1.0000	1.0000

Who is affected by the crisis ? Generally, our cost of living estimates reveal quite some heterogeneity across individuals. In what follows, we investigate this further by relating the N-GARP-based cost of living estimates to observable individual characteristics. This can provide additional insight into which types of individuals (on the intensive

Table 9: Bounds on $c_{2011,2007}$; 4 waves

	N-GARP			GARP			
	min	max	Δ_n	min	max	Δ_g	$\frac{\Delta_g-\Delta_n}{\Delta_g}$
mean	-0.0715	0.0157	0.0872	-0.0730	0.0813	0.1543	0.4392
std. dev.	0.5201	0.2794	0.4051	0.5200	0.3014	0.4416	0.4422
min	-9.4503	-2.5142	0.0000	-9.4503	-2.4888	0.0000	-4.4851
25%	-0.1257	-0.0651	0.0080	-0.1260	-0.0094	0.0341	0.0363
50%	-0.0108	0.0007	0.0292	-0.0118	0.0000	0.0822	0.4481
75%	0.0806	0.1218	0.0764	0.0784	0.2254	0.1954	0.8285
max	0.8306	0.8378	8.7351	0.8303	0.8993	9.3705	1.0000

Table 10: Worse-off and better-off individuals; 4 waves

classification by bounds of:		N-GARP	GARP
UB ≤ 0	Worse off in 2011	38.74	27.59
LB ≥ 0	Better off in 2011	46.45	45.44
LB ≤ 0 and $0 \leq$ UB	Cannot-say	14.81	26.98

margin of labor supply) were particularly hit by the crisis. We conduct three regression exercises: our first exercise uses interval regression and explicitly takes the (difference between) lower and upper bounds into account, our second exercise is a simple OLS regression that uses the average of the lower and upper bounds as the dependent variable, and our last exercise is a logit regression that explains the probability of being better-off (versus worse-off) after the 2008 crisis (using our N-GARP-based classification as worse-off or better-off to define the dependent variable). Further, to distinguish between short-run and longer-run effects of the crisis, we ran our regressions for two cost of living indices: $c_{2009,2007}$ (capturing the short run effect) and $c_{2011,2007}$ (capturing the longer run effect). We use the N-GARP-based bound estimates for the 702 individuals with a CCEI-value at least equal to 0.99 (with $c_{2011,2007}$-values summarized in Table 3).

Table 11 summarizes our findings. We see that individuals with higher labor incomes (i.e., wages) and nonlabor incomes are generally associated with lower cost of living indices, and are less likely to be better off in both the short run and the longer run when compared to their pre-crisis utility level. Next, while we find no significant short run effect related to region of residence (captured by the dummy variables North Central, South and West, using North East as the reference category) or industry (captured by the dummy variables construction and services), we do see that individuals residing in the West region are generally worse off in the longer run, while the opposite holds true for individuals working in the service sector.

Next, we observe that many individual characteristics that are statistically significant in the short run become insignificant in the longer run. For example, homeowners and single parents are better off than non-home owners and childless singles in the short run. However, these effects fade out in the longer run. Similarly, being a single male parent corresponds to a significantly negative crisis effect in the short run, but this effect disappears in the longer run.

Table 12 shows pairwise correlation coefficients between wages in 2007 and relative

Table 11: Welfare effects and individual characteristics

	interval		simple OLS		Logit	
	$c_{2009,2007}$	$c_{2011,07}$	$c_{2009,2007}$	$c_{2011,2007}$	$c_{2009,2007}$	$c_{2011,2007}$
fullincome$_{07}$	-0.000107***	-0.000115***	-0.000109***	-0.000118***	-0.00118***	-0.00136***
	(1.35e-05)	(1.53e-05)	(1.41e-05)	(1.62e-05)	(0.000199)	(0.000320)
nonlabor$_{07}$	-0.000362***	-0.000420***	-0.000371***	-0.000428***	-0.00416***	-0.00454***
	(4.54e-05)	(5.28e-05)	(4.72e-05)	(5.40e-05)	(0.000739)	(0.000823)
years of education$_{07}$	0.00298	0.00245	0.00266	0.00296	0.0642	-0.0162
	(0.00411)	(0.00450)	(0.00423)	(0.00475)	(0.0600)	(0.0671)
North Central	-0.0247	-0.0340	-0.0215	-0.0333	-0.426	-0.916**
	(0.0267)	(0.0222)	(0.0277)	(0.0232)	(0.420)	(0.450)
South	-0.00914	-0.00821	-0.00946	-0.00885	-0.202	-0.753*
	(0.0253)	(0.0206)	(0.0263)	(0.0218)	(0.389)	(0.430)
West	-0.0322	-0.0662**	-0.0301	-0.0615**	-0.686	-1.269***
	(0.0289)	(0.0266)	(0.0300)	(0.0275)	(0.441)	(0.468)
homeowner$_{07}$	0.0312**	0.0156	0.0324**	0.0196	0.360	0.262
	(0.0154)	(0.0159)	(0.0162)	(0.0174)	(0.255)	(0.264)
male	0.0153	0.000816	0.0155	0.00113	0.0334	-0.209
	(0.0168)	(0.0168)	(0.0177)	(0.0176)	(0.276)	(0.276)
have child$_{07}$	0.0647***	0.0371*	0.0665***	0.0332	0.872***	0.272
	(0.0178)	(0.0198)	(0.0184)	(0.0212)	(0.297)	(0.294)
male*have child$_{07}$	-0.145***	-0.000880	-0.145***	0.00703	-1.769**	-0.278
	(0.0455)	(0.0612)	(0.0459)	(0.0618)	(0.880)	(0.945)
age$_{07}$	0.000837	-0.000611	0.000899	-0.000643	0.00973	-0.0109
	(0.000692)	(0.000582)	(0.000710)	(0.000605)	(0.00919)	(0.00900)
construction$_{07}$	-0.00392	-0.0141	-0.000317	-0.00615	-0.350	-0.115
	(0.0228)	(0.0294)	(0.0236)	(0.0307)	(0.423)	(0.396)
services$_{07}$	0.0227	0.0295*	0.0251	0.0355**	0.0211	0.246
	(0.0154)	(0.0154)	(0.0161)	(0.0179)	(0.238)	(0.250)
constant	0.0416	0.114	0.0412	0.103	0.605	3.116***
	(0.0646)	(0.0694)	(0.0670)	(0.0737)	(1.009)	(1.011)
observations	628	628	628	628	476	453
R-squared			0.415	0.437		

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Pairwise correlation coefficients (significant: in bold)

		a	b	c	d	e	f	g
wage in 2007	a	**1**						
relative increase in wage	b	**-0.2296**	**1**					
		0						
relative increase in leisure hours	c	-0.0298	**0.0782**	**1**				
		0.4303	0.0383					
relative increase in leisure expenditures	d	**-0.1761**	**0.9002**	**0.3932**	**1**			
		0	0	0				
relative increase in food expenditures	e	-0.0556	0.0487	0.0159	0.0376	**1**		
		0.144	0.2008	0.6765	0.3232			
relative increase in house expenditures	f	-0.0429	0.0205	0.0357	0.0182	0.0195	**1**	
		0.2572	0.5883	0.3457	0.6301	0.609		
relative increase in other expenditures	g	**-0.0669**	0.0146	0.0236	0.0147	0.0553	0.0129	**1**
		0.0789	0.7016	0.5352	0.7003	0.149	0.7356	

changes in wages, leisure hours, expenditures on leisure, expenditures on food, housing expenditures and other expenditures (measuring the relative change in variable y as $\frac{y_{11}-y_{07}}{y_{07}}$). We see that people with higher initial wages (in 2007) generally experienced larger wage drops (and thus income drops) than people with lower initial wages. This explains the negative regression coefficient for the initial full income in Table 11: if a higher initial potential income corresponds to a greater loss in total income, it is also associated with a more pronounced utility loss.