Nonparametric Welfare Analysis for Discrete Choice: Levels and Differences of Individual and Social Welfare

Bart Capéau, Liebrecht De Sadeleer, Sebastiaan Maes, André Decoster

CESifo

Nonparametric Welfare Analysis for Discrete Choice: Levels and Differences of Individual and Social Welfare

Abstract

Empirical welfare analyses often impose stringent parametric assumptions on individuals' preferences and neglect unobserved preference heterogeneity. In this paper, we develop a framework to conduct individual and social welfare analysis for discrete choice that does not suffer from these drawbacks. We first adapt the broad class of individual welfare measures introduced by Fleurbaey (2009) to settings where individual choice is discrete. Allowing for unrestricted, unobserved preference heterogeneity, these measures become random variables. We then show that the distribution of these objects can be derived from choice probabilities, which can be estimated nonparametrically from cross-sectional data. In addition, we derive nonparametric results for the joint distribution of welfare and welfare differences, as well as for social welfare. The former is an important tool in determining whether those who benefit from a price change belong disproportionately to those who were initially well-off. An empirical application illustrates the methods.

Bart Capéau
Department of Economics
KU Leuven / Belgium
bart.capeau@kuleuven.be

Liebrecht De Sadeleer
Department of Economics
KU Leuven / Belgium
liebrecht.desadeleer@kuleuven.be

Sebastiaan Maes
Department of Economics
KU Leuven / Belgium
sebastiaan.maes@kuleuven.be

André Decoster
Department of Economics
KU Leuven / Belgium
andre.decoster@kuleuven.be

1 Introduction

Discrete choice random utility models (DC-RUMs) have a long tradition in both theoretical and applied microeconometric research. Since the pioneering work of McFadden (1974), these models have been applied to a wide range of problems in transportation, education, health care, industrial organisation, marketing, labour, and public finance. This success can be explained by DC-RUMs' ability to model individual demand among a discrete set of alternatives in a flexible way, allowing for the presence of unobserved preference heterogeneity. Some models within this class, such as the binary and multinomial logit models, also yield convenient closed-form choice probabilities, which considerably reduces the computational burden in both estimation and simulation (for a comprehensive overview, see Train, 2003). These parametric models are, therefore, widely used in empirical research.

The structural modelling of individual preferences in DC-RUMs renders this class of models especially suitable for the welfare analysis of price changes. In such an endeavour, the use of welfare measures (metrics) based on the expenditure function, the so-called money metric utilities (MMUs), is a well-established practice (for seminal contributions, see Diamond and McFadden, 1974; Dixit, 1975; King, 1983). Indeed, the reporting of deadweight losses, compensating variations (CV), and equivalent variations (EV), which are all related to the expenditure function, are ubiquitous in the applied welfare literature. MMUs facilitate interpersonal comparisons of welfare in the presence of individual preference heterogeneity over important life dimensions such as health, housing quality, emotional well-being, and material welfare.

In DC-RUMs, preferences are random from the point of view of the econometrician. Therefore, preference-based welfare measures also become stochastic objects, which complicates the analysis considerably (for an early discussion, see Small and Rosen, 1981). Over the last fifteen years, a methodological literature has emerged that derives closed-form expressions for the distribution of the CV and EV, which are both measures of *changes* in welfare, under ever less parametric assumptions on the nature of individuals' preferences.[1] For the class of additive DC-RUMs, Dagsvik and Karlström (2005) provide expressions for the distribution of the CV based on compensated (Hicksian) choice probabilities. The authors provide analytical results for models where unobserved heterogeneity is generalised extreme value distributed. Alternatively, de Palma and Kilani (2011) advance a direct approach for this class, in which they express this distribution in terms of uncompensated (Marshallian) choice probabilities. More recently, Bhattacharya (2015, 2018) showed that the marginal distributions of the CV and EV can be written as a functional of uncompensated choice probabilities,

---

[1] Before, no closed-form expressions existed, even for the expected values of the CV and EV. Therefore, researchers had to resort to approximations, except for the most simple of DC-RUMs in which individuals have constant marginal utility of income and unobserved heterogeneity is additive and generalised extreme value distributed (Small and Rosen, 1981; McFadden, 1999). These approximations are either biased (Morey et al., 1993), rather uninformative (Herriges and Kling, 1999), or computationally burdensome (McFadden, 1999).

even when unobserved heterogeneity is essentially unrestricted, and therefore possibly nonadditive.[2] His results readily imply that these objects are nonparametrically identified from cross-sectional data with sufficient relative price variation.

The analysis of *levels* of welfare in DC-RUMs, however, has received much less attention from econometricians. Characterising these levels is of first-order importance to applied welfare analysis for at least three reasons. First, knowledge on these levels enables researchers to rank individuals according to their well-being in any given situation, identifying those who are well-off from those who are poor. Second, in aggregating welfare levels across individuals, overall social welfare can be calculated and compared between two situations. Third, joint knowledge on levels and differences of welfare enables the investigation of the association between individuals' gains or losses from a price change and their position in terms of initial welfare. This allows for the assessment of, for example, whether the winners of a price change belong disproportionately to those who were initially well-off.

In recent years, the fairness literature has made substantial progress in developing a comprehensive theoretical framework that encompasses both the classical MMUs (Samuelson, 1974), adaptations of other measures like Pazner's (1979) ray utilities, and measures like the equivalent income and wage metrics (among others, see Pencavel, 1977; Fleurbaey, 2006; Fleurbaey, 2007; Fleurbaey and Gaulier, 2009; Fleurbaey, 2009; Fleurbaey and Blanchet, 2013; Decancq et al., 2015; and Fleurbaey and Maniquet, 2017). A large subset of these measures cardinalise preferences by associating their indifference sets with members of a family of nested opportunity sets; i.e. a lower ranked indifference set is associated with a smaller opportunity set. The sizes of those budget sets are argued to be an ethically more meaningful basis for interpersonal comparisons of well-being than income or reports on subjective satisfaction levels. Indeed, contrary to income and subjective satisfaction, such measures ensure that individuals with the same preferences and in a situation which makes them indifferent among each other are always considered to be equally well-off. We call these metrics nested opportunity set (NOS) measures and show how they relate to what is typically observed in cross-sectional and panel data.

Contributions   We adapt the class of NOS measures (Fleurbaey, 2009, Fleurbaey and Maniquet, 2017) to a setting where the choice set is discrete instead of continuous. This allows us to study welfare levels and differences in DC-RUMs, taking into account unrestricted, unobserved heterogeneity in individuals' preferences. We then prove that the marginal distribution of NOS measures can be recovered nonparametrically from cross-sectional data by evaluating the observed choice probabilities at counterfactual prices. This allows researchers to study levels of individual welfare in any given situation. Likewise, we show that the joint distribution of welfare levels and welfare differences can

---

[2]This result only holds for price changes. When quality changes occur, Bhattacharya (2018) demonstrates that functionals of the choice probabilities only set-identify the distribution of the CV (or EV).

be recovered nonparametrically from panel data by evaluating the observed transition probabilities at counterfactual prices.[3] Building on these two results, we are able to nonparametrically characterise levels and differences in aggregate welfare for any additively separable social welfare function.

In addition, we show that Samuelson's (1974) money metric utilities (MMUs) are within the class of the discrete NOS measures, and use them as a leading example to illustrate our approach. This also implies that we can specialise our results to cases where welfare differences are measured by the CV or EV. In doing so, we generalise the results of Dagsvik and Karlström (2005) and de Palma and Kilani (2011) to settings where unobserved heterogeneity is essentially unrestricted. As a by-product, we also present all our results conditional on the endogenous pre- or post-price change choices, which allows researchers to take the additional information conveyed by the observed choices into account. This conditioning might also be important from a political economy perspective.

These results allow researchers to conduct individual and social welfare analyses without resorting to stringent parametric assumptions on preferences or unobserved heterogeneity. Our identificationresults are constructive and can be implemented in empirical work using only nonparametric regression. We also demonstrate how Boole-Fréchet inequalities (Fréchet, 1935) and stochastic revealed preference restrictions can be exploited to construct bounds on the transition probabilities in the common event when only cross-sectional data is available. These bounds are functionals of the choice probabilities and are, as such, straightforward to implement. They can be readily used to set-identify the concepts that are expressed in terms of transition probabilities.

To illustrate the empirical usefulness of our results, we revisit the classical trade-off between leisure and consumption. For this purpose, we make use of micro-data from the 2018 wave of the German Socio-Economic Panel (SOEP), which contains detailed information on households' demographics, labour supply, wages, and out-of-work income. Single females' labour supply is modelled as a choice between three discrete alternatives: non-working, part-time employment, and full-time employment. Using a MMU, we present nonparametric, distributional estimates of individual and social welfare for the tax schedule that is currently in place in Germany. We also study the effects of replacing this nonlinear and progressive schedule with a basic income flat tax.

In the baseline, we find that for approximately 25% of the single females in our estimation sample, the distribution of the MMU is a step function. This means that we can determine their welfare level exactly. Aggregating these individual distributions in groups based on gross hourly wage quartiles, we find that the distribution of high-wage individuals who work full-time tends to first-order dominate that of those who are non-working or work part-time. The reverse is true for low-wage individuals. This is explained by a mismatch between these individuals' preferences and wages. Moreover,

---

[3]These transition probabilities are derived under the assumption that (unobserved) individual preferences are unaltered by the price change. Alternatively, Dagsvik (2002) and Delle Site and Salucci (2013) consider models where there is imperfect correlation in unobserved heterogeneity before and after the change.

irrespective of the baseline choice, the distribution of high-wage individuals dominates that of the groups with lower wages. For the latter, the distributions are more intermingled, which is suggestive evidence that besides wages, (unobserved) preferences do play an important role in assessing the welfare level of individuals.

For the entire population of single females, the distribution of welfare in the reform first-order dominates that in the baseline.[4] This does not imply, however, that every individual gains from the reform. Around 15% of single females find their welfare diminished; this is especially the case for those that are well-off in the baseline. Large gains are more prevalent among those that were initially poor. Among the bottom two-thirds of the population in terms of baseline welfare, almost 98% advance as a result of the reform. Approximately half of the single females gain up to 200 euro on a monthly basis.

Other related literature   Over the last decades, several semiparametric methods have been developed to relax functional form assumptions on either deterministic preferences or the distribution of unobserved heterogeneity in DC-RUMs (e.g., for early results see Manski, 1975; Matzkin, 1991; and Klein and Spady, 1993). Other contributions introduce entirely nonparametric methods that do not impose functional form restrictions on either of these components for this class of models, based on either shape restrictions (e.g., see Matzkin, 1993) or the availability of a large-support special regressor (e.g., see Lewbel, 2000 and Briesch et al., 2010). The approach we follow in this paper deviates from this literature as our objective is not to recover deterministic preferences and the distribution of unobserved heterogeneity, but instead to identify individual welfare measures which are functions of both these model primitives.

Another strand of literature focuses on the nonparametric identification of counterfactual choices and welfare under unobserved heterogeneity in models where demand is continuous instead of discrete. Most results exploit the smoothness of the underlying individual demand functions to arrive at Slutsky-like restrictions on average and quantile demands (e.g., see Dette et al., 2016; Hausman and Newey, 2016; Blundell et al., 2017; and Hoderlein and Vanhems, 2018). Other results exploit the axioms of revealed preference (RP) to attain identification under the presence of unobserved heterogeneity (e.g., see Blundell et al., 2014; and Cosaert and Demuynck, 2018). In contrast to our results, however, the availability of cross-sectional and short panel data is generally not sufficient to point-identify the distribution of welfare levels and differences in settings where demand is continuous and unobserved heterogeneity is unrestricted.

Finally, this paper contributes to the literature that applies NOS measures empirically. Using SOEP microdata, Decoster and Haan (2015) estimate a parameasure DC-RUM of labour supply and con-

---

[4]This is partially due to the fact that we constructed the reform to be revenue-neutral for the entire SOEP sample. Therefore, revenue neutrality does not necessarily hold for our subsample of single females.

struct rankings of households based on NOS measures. Carpantier and Sapata (2016) extend the approach of Decoster and Haan (2015) by integrating unobserved preference heterogeneity through a numerical procedure (comparable to the approach of Herriges and Kling (1999) for welfare differences). Our results show that the parametric assumptions imposed in these papers are not necessary to obtain identification.

The remainder of this paper is organised as follows. In the second section, we introduce the class of NOS welfare measures originally developed by Fleurbaey (2009) for settings where choice is continuous. In section three, our conceptual framework is laid out. We first specify the DC-RUM and impose some mild restrictions on individuals' preferences. We then adapt the class of NOS welfare measures to this discrete setting. Since preferences are random from the point of view of the econometrician, these welfare measures are also random variables. In the fourth section, we present our main results. We derive the distribution of the NOS measures, either conditional or unconditional on the initial and final choices. We also study welfare differences induced by a price change and derive the joint distribution of welfare in the initial choice and these welfare differences, again either conditional or unconditional on the initial and final choices. Moreover, we discuss how measures for social welfare can be constructed from these results. In section five, we discuss how the choice and transition probabilities by which the welfare concepts are identified, can be retrieved from cross-sectional data. In the sixth section, we illustrate our results by means of an application on female labour supply, using the German SOEP microdata. The final section contains concluding remarks.

2   NOS welfare measures in a continuous setting

In this section, we briefly explain and motivate the class of NOS welfare measures, which have been introduced by Fleurbaey (2009) for settings where choice is continuous.

These measures cardinalise preferences by associating each indifference set with a member of a family of nested opportunity sets, which is common for all individuals. A lower ranked indifference set is associated with a smaller set from that family, such that the size of the opportunity set acts as a measure of individual well-being, respecting that individual's preferences. Formally, let $\mathrm{B} \subseteq \mathbb{R}^n_+$ be the set of all bundles $\mathbf{b}$ an agent can possibly obtain, and let $\{B_\lambda \subsetneq \mathrm{B} \mid \lambda \in \Lambda \subseteq \mathbb{R}\}$ denote a family of nested budget sets indexed with a parameter λ such that $\lambda_i < \lambda_j \implies B_{\lambda_i} \subsetneq B_{\lambda_j}$. Given a well-behaved utility function $U(\mathbf{b}) : \mathrm{B} \to \mathbb{R}$, the NOS welfare measure evaluated in a bundle $\mathbf{b} \in \mathrm{B}$ is then defined as

$$W(\mathbf{b}) = \max_{\lambda} \left\{ \lambda \mid U(\mathbf{b}) \geq \max_{\mathbf{b}' \in B_\lambda} U(\mathbf{b}') \right\}, \tag{1}$$

that is, the largest value of λ — or, equivalently, the largest opportunity set — for which the individual still weakly prefers bundle $\mathbf{b}$ above all bundles $\mathbf{b}'$ in the virtual opportunity set B_λ.[5]

---

[5]Note that the discussion here is informal as additional assumptions are needed to guarantee the existence and the uniqueness of the NOS welfare measure. A more rigorous treatment is provided in Section 3.2.

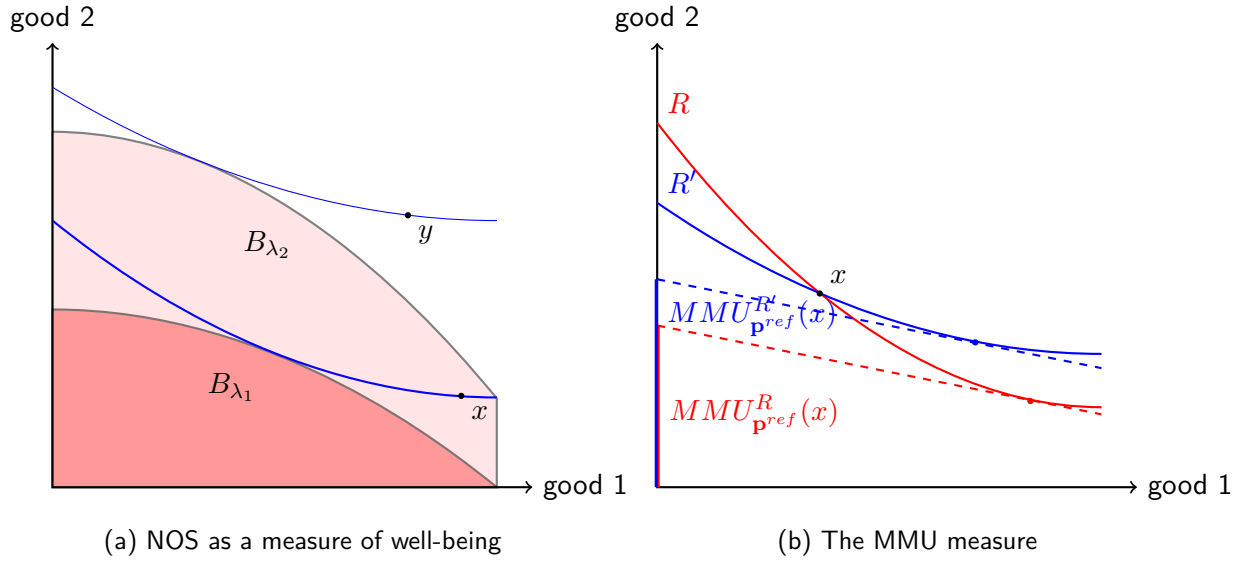(a) NOS as a measure of well-being   (b) The MMU measure

Figure 1: NOS welfare measures in a continuous choice setting

This definition is illustrated by means of a classical trade-off between two goods in Figure 1a. Suppose first an individual obtains a bundle x, and let the thick blue line denote her indifference curve passing through x. This indifference curve is associated with the opportunity set B_{λ_1}, which is shaded in dark red. In accordance with the definition in Equation (1), this set is designed such that the individual could obtain, at best, a bundle equally as good as x, when she would be faced with the opportunity set B_{λ_1}. Suppose now that the individual obtains a bundle y, which is better than x, according to her own preferences. Then the associated opportunity set B_{λ_2}, which is shaded in light red, is again such that the best bundle in this set is equally good as y, and includes the set B_{λ_1}. From this illustration, it is clear that the size of these opportunity sets serves as a measure of individual well-being that respects preferences, in the sense that the well-being level of an individual in situation y is higher than the well-being level in situation x, if and only if that individual prefers y to x. The size of an opportunity set is measured by its indexing parameter λ.

Example: the MMU measures One important set of welfare measures within the NOS class is the set of money metric utilities (MMUs) (Samuelson, 1974). In this case, the NOS are of the form

$$B_\lambda \equiv \left\{ \mathbf{b} \in B \,\middle|\, \mathbf{b}' \mathbf{p}^{ref} \leq \lambda \right\}, \tag{2}$$

where $\mathbf{p}^{ref}$ is a vector of reference prices that are fixed by the researcher.

In this specification, the indexing parameter λ can be interpreted as a monetary amount. Applying the definition in Equation (1), we find that the MMU measure evaluated in a bundle x, denoted by $MMU_{\mathbf{p}^{ref}}(x)$, measures well-being by the maximal monetary amount that can be granted to an individual faced with reference prices $\mathbf{p}^{ref}$, such that she would at most be equally well-off as in

7

point x. This coincides with the expenditure function representation of preferences, which is how MMUs were originally presented in the literature.

In Figure 1b, this is illustrated by means of a classical trade-off between two goods. The slope of the dashed lines equals *minus* the reference price p_1^{ref} for good 1, as good 2 serves as numeraire here. The welfare levels $MMU_{\mathbf{p}^{ref}}^{R}(x)$ of an individual with preferences R and $MMU_{\mathbf{p}^{ref}}^{R'}(x)$ of an individual with preferences R' are determined such that the best point in the opportunity set is a point on the indifference curve through their actual, common situation x (denoted by the red and blue point, respectively). With the chosen reference price, the individual with preference R' is better off than the one with preferences R.

Other examples of NOS measures are the ray utilities of Pazner (1979) and the equivalent income metrics introduced in Decancq et al. (2015). In the former, one fixes a reference bundle $\mathbf{b}^{ref}$ and determines the NOS as the sets of bundles that lie on the ray from the origin through this reference bundle, that is $B_\lambda = \left\{ \mathbf{b} \in B \mid \mathbf{b} = \alpha \mathbf{b}^{ref}, \alpha \leq \lambda \right\}$. In the latter, one measures well-being by the minimal amount of a numeraire good which is needed, given a reference value for the other commodities or dimensions of well-being in order to make an agent equally well-off as in her actual situation. We discuss the equivalent income metrics in more detail in Appendix A.1.

3 Conceptual framework

Our conceptual framework is similar to that of Bhattacharya (2015, 2018) and allows for unrestricted, unobserved heterogeneity in DC-RUMs. As this set-up does not impose any restrictions on *observed* individual characteristics, all results in this paper can be thought of as being conditional on these covariates.

3.1 Discrete choice model

Preferences Let Ω denote the universe of preference types and let the associated probability space $(\Omega, \mathcal{F}, \mathrm{Pr}_\omega[\{\omega \mid \cdot\}])$ represent the distribution of these preference types in the population.[6] Every preference type can be thought of as a different individual, who has idiosyncratic preferences over a finite (and common) set of alternatives $\mathcal{C}$, with $|\mathcal{C}| := n \in \mathbb{N}_0$, and a numeraire. These idiosyncratic preferences are assumed to be representable by a utility function $U_c^\omega(y - p_c) := U(y - p_c, c, \omega) : \mathbb{R} \times \mathcal{C} \times \Omega \to \mathbb{R}$, in which y denotes exogenous income and p_c the price of alternative $c \in \mathcal{C}$.[7] A given set of prices for all alternatives, $(p_c, c \in \mathcal{C})$, are recollected in a vector denoted by $\mathbf{p}$ and we will call $(\mathbf{p}, y)$ a budget set.

Note that our formulation of preferences is very flexible as it allows them to differ arbitrarily across

---

[6]In the remainder of the paper, we will abbreviate $\mathrm{Pr}_\omega[\{\omega \mid \cdot\}]$ as $\mathrm{Pr}_\omega[\{\cdot\}]$.

[7]For the sake of technical rigour, $U(y - p_c, c, \cdot)$ is required to be a $\mathcal{F}$-measurable function.

individuals. The only economically substantial restriction we will impose on this function is that utility is continuous and strictly increasing in the numeraire $y - p$.

Assumption 1 Individual preferences are represented by a utility function $U_c^\omega(y - p_c)$ that is continuous and strictly increasing in the numeraire for every preference type $\omega \in \Omega$ and every alternative $c \in \mathcal{C}$. Moreover, preferences satisfy the following regularity conditions: (R1) For each pair of alternatives $c, c' \in \mathcal{C}$, and for each fixed y and p_c, it holds that $U_c^\omega(y - p_c) > \lim_{p_{c'} \to \infty} U_{c'}^\omega(y - p_{c'})$ and that $U_c^\omega(y - p_c) < \lim_{p_{c'} \to -\infty} U_{c'}^\omega(y - p_{c'})$. (R2) For every budget set $(\mathbf{p}, y)$, the set of types which are indifferent between two or more alternatives in the choice set $\mathcal{C}$ has probability measure zero.

This assumption is ubiquitous in empirical work that employs (semi)parametric DC-RUMs. Monotonicity in the numeraire establishes the existence and uniqueness of our welfare measures and yields stochastic revealed preference conditions that we will exploit to obtain the identification results. Regularity condition (R1) ensures that when the price of a given alternative goes to infinity, it will never be preferred above another alternative with a finite price. Analogously, when the price of a given alternative goes to minus infinity, or residual income in that alternative goes to plus infinity, it will always be preferred above another alternative with a finite price. The negligibility of indifferences between alternatives (R2) ensures that no tie-breaking rule has to be established.

In addition, we also assume that the distribution of the preference types, denoted by $F(\omega)$, is independent of the budget set $(\mathbf{p}, y)$.

Assumption 2 The distribution of unobserved heterogeneity $F(\omega)$ is independent of prices $\mathbf{p}$ and exogenous income y: i.e. $F(\omega \mid \mathbf{p}, y) = F(\omega)$.

The exogeneity of budget sets is a strong, but standard, assumption in the literature on nonparametric identification of individual demand and welfare (e.g., see Hausman and Newey, 2016). Indeed, to the best of our knowledge, there are no theoretical results that allow for general forms of endogeneity in the presence of unrestricted, unobserved heterogeneity. Some forms of endogeneity, however, can be mitigated by using a control function approach (see Section 5).

Individual choice behaviour Finally, we assume that observed choice behaviour is actually generated by a DC-RUM (for a detailed technical overview on RUMs, see McFadden (1981, 2005)). This assumption entails that an individual ω chooses a given alternative i, if and only if this alternative yields the highest utility among the elements in her choice set $\mathcal{C}$.

Assumption 3 Let $J^\omega(\mathbf{p}, y) \equiv J(\mathbf{p}, y, \omega) : \mathbb{R}^{n+1} \times \Omega \to \mathcal{C}$ denote the individual demand function. It holds that $J^\omega(\mathbf{p}, y) = i \iff U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\}$.

Note that individual demand is single-valued with probability one as one can neglect indifferences

between alternatives by regularity condition (R2) in Assumption 1.

Choice and transition probabilities The individual choices induced by a DC-RUM are stochastic from the point of view of the econometrician, as the preferences types are non-observable. When this random variation is averaged out across types, one obtains a set $\{P_i(\mathbf{p}, y)\}_{i \in \mathcal{C}}$ of uncompensated (Marshallian) conditional choice probabilities,

$$
\begin{aligned}
P_i(\mathbf{p}, y) &= \Pr_\omega \left[\left\{ U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\} \right\} \right] \\
&= \int_\omega \mathbb{I} \left[U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\} \right] dF(\omega \mid \mathbf{p}, y) \qquad (3) \\
&= \int_\omega \mathbb{I} \left[J^\omega(\mathbf{p}, y) = i \right] dF(\omega),
\end{aligned}
$$

where $\mathbb{I}[\cdot]$ denotes the indicator function.[8] The last expression asymptotically coincides with the observed choice frequency for every alternative $i \in \mathcal{C}$, conditional on the budget set $(\mathbf{p}, y)$.[9] If cross-sectional data contains enough relative price and exogenous income variation, these objects are nonparametrically estimable.[10]

Another concept induced by DC-RUMs is the set $\{P_{i,j}(\mathbf{p}, \mathbf{p}', y)\}_{i,j \in \mathcal{C}}$ of uncompensated conditional transition probabilities. These probabilities are formally defined as

$$
\begin{aligned}
P_{i,j}(\mathbf{p}, \mathbf{p}', y) &= \Pr_\omega \left[\left\{ U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\} \right\} \cap \left\{ U_j^\omega(y - p_j') \geq \max_{c \neq j}\{U_c^\omega(y - p_c')\} \right\} \right] \\
&= \Pr_\omega \left[i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y) \right] \\
&= \int_\omega \mathbb{I} \left[J^\omega(\mathbf{p}, y) = i \right] \mathbb{I} \left[J^\omega(\mathbf{p}', y) = j \right] dF(\omega),
\end{aligned}
$$

which asymptotically coincide with the transition frequencies from alternative i to alternative j after an exogenous price change from $\mathbf{p}$ to $\mathbf{p}'$.[11] Naturally, if there is no price change, there are no transitions between different choices. In principle, these objects are nonparametrically estimable

---

[8]These choice probabilities are designated *conditional* as they depend on a vector of prices and income. In the interest of brevity, this qualification will be dropped in the sequel.

[9]This concept is also known as the *average structural function* (e.g., see Blundell and Powell, 2004). The asymptotic equivalence follows from the law of large numbers as the choice probabilities are essentially conditional expectation functions.

[10]It is clear from the second equality in Equation (3) that these probabilities are composed of both the utility function U_c^ω and the distribution of unobserved heterogeneity F. As such, they are not sufficiently informative to separately identify these two model primitives. Fortunately, knowledge on such primitives is not necessary for our purposes.

[11]Note, however, that transition probabilities do not impose any temporal structure. In other words, $P_{i,j}(\mathbf{p}, \mathbf{p}', y) = P_{j,i}(\mathbf{p}', \mathbf{p}, y)$. Furthermore, the assumption that the exogenous income y is common to both situations with prices $\mathbf{p}$ and $\mathbf{p}'$ imposes no constraint. Indeed, if exogenous incomes are different when faced with prices $\mathbf{p}$ and $\mathbf{p}'$ (denoted by y and y'), we can redefine prices and incomes in order to obtain a common exogenous income. Defining $\mathbf{p}'' = \mathbf{p}' - y' + y$, we obtain

$$
\Pr_\omega \left[i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y') \right] = \Pr_\omega \left[i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}' - y' + y, y) \right] = P_{i,j}(\mathbf{p}, \mathbf{p}'', y).
$$

from panel data with at least two periods. In addition, Section 5.1 shows how transition probabilities can be set-identified when only cross-sectional data is available.

Implicit in our definition of the transition probabilities is the assumption that individuals' preferences are unaffected by the price change. The perfect correlation between the preference types before and after the price change implies that transition probabilities are not simply equal to the product of their marginals: i.e. $P_{i,j}(\mathbf{p}, \mathbf{p}', y) \neq P_i(\mathbf{p}, y) P_j(\mathbf{p}', y)$.

3.2 NOS welfare measures in a discrete choice setting

In Section 2, the family of NOS welfare measures is introduced in a setting of continuously divisible goods. In this subsection, we will redefine them rigorously for settings where choice is determined by a DC-RUM that satisfies Assumptions 1–3.

Nested opportunity sets in DC-RUMs The for all preference types $\omega \in \Omega$ common family of nested opportunity sets is defined as follows. Let there be a closed set $\Lambda \subseteq \mathbb{R}$, and define for every $\lambda \in \Lambda$ a non-empty opportunity set $B_\lambda \subset \mathbb{R} \times \mathcal{C}$, such that $B_\lambda \subsetneq B_{\lambda'}$ if $\lambda < \lambda'$, and such that with $y_c^\lambda := \sup\{y \mid (y, c) \in B_\lambda\}^{12}$ and $\mathbf{y}^\lambda := (y_1^\lambda, \ldots, y_c^\lambda, \ldots, y_n^\lambda)$:

(a)
$$\lambda < \lambda' \implies \begin{cases} \forall c \in \mathcal{C} : y_c^\lambda \leq y_c^{\lambda'}, \\ \exists c \in \mathcal{C} : y_c^\lambda < y_c^{\lambda'}, \end{cases} \tag{4}$$

(b) the function
$$\Lambda \to \mathbb{R}^n : \lambda \mapsto \mathbf{y}^\lambda \text{ is continuous,} \tag{5}$$

(c) for all options c',
$$\inf_{\lambda \in \Lambda} y_{c'}^\lambda = -\infty, \tag{6}$$

(d) and for at least one option c,
$$\sup_{\lambda \in \Lambda} y_c^\lambda = +\infty. \tag{7}$$

Then $(B_\lambda)_{\lambda \in \Lambda}$ is called a family of nested opportunity sets. Note that the family is common to all individuals. Conditions (4) and (5) ensure that the family $(B_\lambda)_{\lambda \in \Lambda}$ is continuously increasing, while conditions (6) and (7) ensure that for every bundle x and preference type ω, there exists a member of the family of which all bundles are considered worse than x by ω, and one which contains a bundle considered to be better than x by ω. This latter property will prove necessary to define the welfare measure.

---

¹²We define the supremum of the empty set in this context to be equal to $-\infty$.

The vector $\mathbf{y}^\lambda$ can be seen as the upper bound of B_λ. Below, we will see that welfare measures associated to families of nested opportunity sets with the same upper bounds are identical. Hence, we will only consider opportunity sets of the form

$$B_\lambda := \left\{ (y', c) \mid c \in \mathcal{C}, y' \in \mathbb{R}, y' \leq y_c^\lambda \right\}. \tag{8}$$

It is often more convenient to characterise the opportunity sets in terms of virtual prices $\widetilde{p}_c(\lambda) := y - y_c^\lambda$. In particular, we have that

$$B_\lambda := \left\{ (y', c) \mid c \in \mathcal{C}, y' \in \mathbb{R}, y' \leq y - \widetilde{p}_c(\lambda) \right\}. \tag{9}$$

For future reference, we denote the vector of virtual prices as follows: $\widetilde{\mathbf{p}}(\lambda) = (\widetilde{p}_1(\lambda), \ldots, \widetilde{p}_n(\lambda))$. Note that as $\mathbf{y}^\lambda$ is increasing in λ, as described in Equation (4), $\widetilde{\mathbf{p}}(\lambda)$ is decreasing in λ in the same way. Moreover, $\lambda \to \widetilde{\mathbf{p}}(\lambda)$ is continuous by (5), $\sup_{\lambda \in \Lambda} \widetilde{p}_{c'}(\lambda) = +\infty$ for all c' by (6) and $\inf_{\lambda \in \Lambda} \widetilde{p}_c(\lambda) = -\infty$ for at least one c by (7). The fact that those virtual prices can become negative might seem strange at first. However, in a discrete choice context, one can always redefine prices and exogenous income by increasing both by an equal amount of the numeraire. As a result, negative prices can be converted into positive prices.

The intuition behind the definition of nested opportunity sets can be illustrated using the famous 'Matryoshka dolls'. If one has several boxes of different dimensions, one can measure their size in different ways. One possible way is defining the size of these boxes by the biggest Matryoshka doll of the set fitting in the box. Sets of Matryoshka dolls of different shape constitute different 'metrics'. Analogously, we use families of nested opportunity sets to construct welfare measures. Families of opportunity sets with different shapes will define different metrics.

Welfare measures in DC-RUMs In the continuous setting, the NOS welfare measure evaluated in a bundle was defined as the largest value of λ, or the largest opportunity set B_λ, such that this bundle was weakly preferred over all bundles in B_λ. The same idea can be applied to a setting where choice is discrete. More precisely, we define a NOS welfare measure as

$$W^\omega(y - p_k, k) = \sup_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_{(y', c) \in B_\lambda} U_c^\omega(y') \right\}, \tag{10}$$

that is, the largest value of λ such that option k is weakly preferred over all bundles in B_λ. Note that the dependence on the preference type ω implies that this welfare measure is a random variable[13]. According to Assumption 1, the utility function is strictly increasing in the numeraire, which allows us to restate this definition in terms of the upper bound of the opportunity sets. Formally, we have

---

[13]As the utility function U is $\mathcal{F}$-measurable, W is also $\mathcal{F}$-measurable.

that

$$W^\omega(y - p_k, k) = \sup_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_{(y', c) \in B_\lambda} U_c^\omega(y') \right\}$$

$$= \sup_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c \max_{y' \leq y_c^\lambda} U_c^\omega(y') \right\} \tag{11}$$

$$= \sup_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y_c^\lambda) \right\}.$$

This expression highlights that the value of the welfare measure only depends on the upper bound of the opportunity sets and that the assumption that the opportunity sets are of the form of (8) is not a genuine restriction. Furthermore, by conditions (6) and (7) and Assumption (R1), there exists (i) a $\lambda_{\min} \in \Lambda$ such that $U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y_c^{\lambda_{\min}})$, and (ii) a $\lambda_{\max} \in \Lambda$ such that $U_k^\omega(y - p_k) < \max_c U_c^\omega(y_c^{\lambda_{\max}})$. This implies that the set $\left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y_c^\lambda) \right\}$ is not empty by (i), and bounded by (ii). Moreover, by continuity of the utility function and of the function $\lambda \mapsto \mathbf{y}^\lambda$, $\lambda \mapsto \max_c U_c^\omega(y_c^\lambda)$ is also continuous, which implies, together with the closedness of Λ, that (iii) $\left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y_c^\lambda) \right\}$ is closed. As this set is not empty, bounded and closed, one can conclude that the suprema in Equations (10) and (11) are in fact attained and can be replaced by maxima.

Equivalently, when opportunity sets are characterised in terms of virtual prices, we can write that

$$W^\omega(y - p_k, k) = \max_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y - \widetilde{p}_c(\lambda)) \right\}. \tag{12}$$

For notational convenience, the characterisation in terms of virtual prices $\widetilde{p}_c(\lambda)$ instead of the numeraire y_c^λ will be used in the remainder of the paper.

The key insight of this paper is that the statement '$W^\omega(y - p_k, k) \geq w$' is equivalent with k (at its original price) being the optimal choice among all options with a virtual vector of prices that is welfare measure specific. This result is made precise in Lemma 1. For didactic purposes, the proof is included here.

Lemma 1

$$\{ \omega \mid w \leq W^\omega(y - p_k, k) \} = \{ \omega \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y - \widetilde{p}_c(w)) \}$$

Proof. Take an arbitrary $\omega \in \Omega$ such that $w \leq W^\omega(y - p_k, k)$. Then there exists a $\lambda \geq w$ such that $U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y - \widetilde{p}_c(\lambda))$. As $\lambda \geq w$, $\widetilde{p}_c(\lambda) \leq \widetilde{p}_c(w)$ for all $c \in \mathcal{C}$ and, hence,

$$\max_c U_c^\omega(y - \widetilde{p}_c(\lambda)) \geq \max_c U_c^\omega(y - \widetilde{p}_c(w)),$$

because U_c^ω is an increasing function by Assumption 1. It follows that $U_k^\omega(y - p_k) \geq \max_c U_c^\omega(y - \widetilde{p}_c(w))$. The other inclusion follows immediately from the definition of $W^\omega(y - p_k, k)$. $\qquad\square$

This equivalence is obtained without imposing any assumption on preferences besides Assumptions 1 and 2, and is, therefore, entirely nonparametric.[14] As will be shown below, its main practical implication is that the entire distribution of objects based on NOS measures can be obtained by only evaluating choice and transition probabilities at virtual prices. This entails that these distributions can be identified from cross-sectional and panel data in a nonparametric way.

Figure 2 provides a graphical illustration. The choice set $\mathcal{C}$ consists of three options: i, j, and k. For each option, the amount of the numeraire, $y - p_i$, $y - p_j$, and $y - p_k$ is shown on the vertical axis. The blue points indicate the indifference set of the point $(y - p_k, k)$. Three members of a family of nested opportunity sets $B_{\lambda \in \Lambda}$ are shown in red. For example, all option-numeraire combinations in dark red belong to B_{λ_1}. For illustrative convenience, we choose $y_c^\lambda < y_c^{\lambda'}$ whenever $\lambda < \lambda'$, and this for all $c \in \mathcal{C}$. Finally, the figure shows how to calculate the NOS welfare measure for option k. The welfare measure is defined by the nested opportunity sets $B_{\lambda \in \Lambda}$ shown in red in the figure. The upper bound of B_{λ_3}, consisting of the points $y_c^{\lambda_3}$, is denoted by the red dots. It is clear that λ_3 is the maximand of Equation (10) because the red dot of $y_j^{\lambda_3}$ coincides with the blue point at position j. This means that $U_j^\omega(y_j^{\lambda_3}) = U_k^\omega(y - p_k)$, and hence $W^\omega(y - p_k, k) = \lambda_3$ in this example.
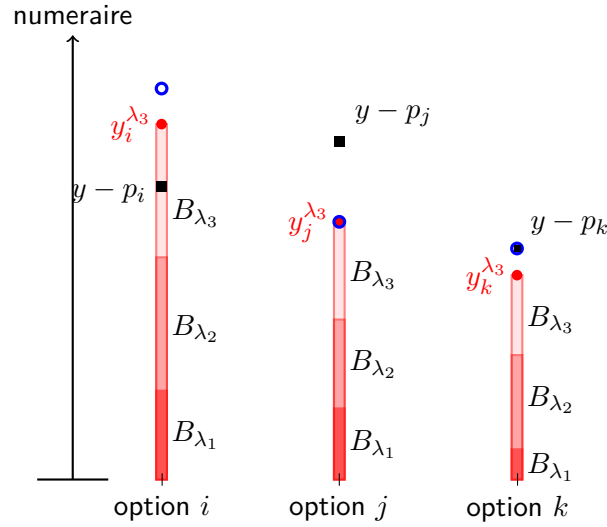

Figure 2: A graphical illustration of a NOS welfare measure in a discrete choice context

Example: the MMU measures Fix a set of n reference prices $\mathbf{p}^{ref}$, one for each option and let the upper bound of the opportunity sets be $y_c^\lambda = \lambda - p_c^{ref}$, or equivalently, $\widetilde{p}_c(\lambda) = y - \lambda + p_c^{ref}$. The crucial property is that the upper bounds increase by the same amount for every option: i.e. $y_c^{\lambda_1} - y_c^{\lambda_2} = \lambda_1 - \lambda_2$ for all $c \in \mathcal{C}$. The MMU evaluated in option k (with price p_k) at reference prices $\mathbf{p}^{ref}$ is then defined as

$$MMU_{\mathbf{p}^{ref}}^\omega(y - p_k, k) = \max_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega \left(y - (y - \lambda + p_c^{ref}) \right) \right\}. \tag{13}$$

---

[14]This result is similar in nature to the result Bhattacharya (2018) obtains for the marginal distribution of the CV.

This can also be defined implicitly as

$$U_k^\omega \left(y - p_k \right) = \max_c U_c^\omega \left(MMU_{\mathbf{p}^{ref}}^\omega \left(y - p_k, k \right) - p_c^{ref} \right).$$

Similar to the continuous case, this highlights the equivalence of the MMUs with the expenditure function representation of preferences, as each of them evaluates the expenditure function at a given set of reference prices.

From Equation (13), it can be seen that $MMU_{\mathbf{p}}^\omega \left(y - p_k, k \right) = y$ if $k = J^\omega \left(\mathbf{p}, y \right)$. When the reference prices coincide with the actual prices, the level of well-being according to the MMU of the optimal choice in the actual situation is equal to the actual amount of the numeraire (see also Corollary 2 below).

4 Distribution of the NOS welfare measures, welfare differences, and social welfare

As discussed before, the presence of unobserved preference heterogeneity entails that NOS welfare measures are random variables from the point of view of the econometrician. This randomness can be interpreted in two distinct ways. In the first interpretation, as the econometrician does not observe an individual's preference type, they can only derive the distribution of welfare for this particular individual and not its exact realisation. That is, the distribution reflects uncertainty for the econometrician. In the second interpretation, an observed individual in the sample represents the class of individuals in the population that share the same observable characteristics. In this case, the distribution reflects inequality in welfare among the members of this class. Our theoretical results are valid for both interpretations.

For notational convenience, we will present all our expressions in terms of the complementary cumulative distribution function (CCDF) instead of the more common cumulative distribution function (i.e. $1 - F(x)$ for a CDF F). The proofs in this section and the following ones are deferred to Appendix B.

4.1 Distribution of the NOS welfare measures

In this section the marginal distribution for the NOS measures is derived in terms of choice probabilities. We also provide distributional results joint with, and conditional on, the optimal observed choice.

Under Assumptions 1–3, which were introduced in Section 3, we can prove the following theorem.

Theorem 1 The joint distribution of the NOS welfare measure W, evaluated in an option k with

price p_k, and choosing j at prices $\mathbf{p}'$ and exogenous income y can be expressed in terms of transition probabilities as follows:

$$\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k), j = J^{\omega}(\mathbf{p}', y)\right] = P_{j,k}\Big(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\Big) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (14)$$

where $(p_k, \widetilde{\mathbf{p}}_{-k}(w)) = (\widetilde{p}_1(w), \ldots, \widetilde{p}_{k-1}(w), p_k, \widetilde{p}_{k+1}(w), \ldots, \widetilde{p}_n(w))$.

See proof on page 47.

The crucial insight here is that, by Lemma 1, the event $W^{\omega}(y - p_k, k) \geq w$ is translated into a statement on k being optimal under virtual prices. The joint distribution in Equation (14) can, therefore, be expressed in terms of transition probabilities, evaluated at both actual prices $\mathbf{p}'$ and p_k and virtual prices $\widetilde{\mathbf{p}}$.

Theorem 1 is formulated in the most general form; it considers a joint distribution, and not a marginal nor a conditional, and allows the price at which the welfare in alternative k is evaluated, p_k, to be different from the prices $\mathbf{p}'$. For example, if one wants to evaluate welfare levels after a price change from $\mathbf{p}$ to $\mathbf{p}'$ when only information on choices before the price changed is available, $\mathbf{p}'$ and $\mathbf{p}$ will typically not coincide. However, if one wants to evaluate welfare in a setting with only one actual price $\mathbf{p}'$, the latter can be set equal to the actual prices $\mathbf{p}$. Usually, one wants to evaluate welfare in an optimal bundle; then k can be set equal to j in Equation 14. In Corollary 1 below, we will derive some related distributions which are more directly relevant for applied work.

The exact formulation of $\widetilde{\mathbf{p}}(w)$ depends, as explained in Section 3.2, on the specific choice of the welfare measure. Nonetheless, we can give some intuition on the role of p_k and the overall course of the distribution of welfare. We know that the lower the price p_k, the higher is the residual numeraire $y - p_k$ in option k and hence, the more the indifference set containing $(y - p_k, k)$ is shifted upwards in the numeraire dimension. As the indifference set is an upper bound of the (virtual) nested opportunity sets, a higher indifference set implies higher welfare. Hence, the lower price p_k, the higher is the CCDF of welfare in option k and the more the distribution of welfare is shifted to the right.

Now, we examine the overall course of the CCDF in more detail by considering a typical plot of the CCDF for fixed prices p_k and $\mathbf{p}'$ in Figure 3a. When w is negative and large in absolute value, the $\widetilde{p}_c(w)$ are large (and positive). Hence $p_k \leq \widetilde{p}_k(w)$ and the CCDF approaches $P_j(\mathbf{p}')$ as expected. As w increases, $\widetilde{\mathbf{p}}_{-k}(w)$ decreases and the probability of choosing k at prices $(p_k, \widetilde{\mathbf{p}}_{-k}(w))$ decreases. Therefore, $\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k), j = J^{\omega}(\mathbf{p}, y)\right]$ decreases until w reaches its highest value at which point $p_k = \widetilde{p}_k(w)$, called w_k^*. There the CCDF drops to zero discontinuously, as the indicator function becomes zero. This means that w_k^* is an upper bound for welfare and that the probability distribution has a mass point.

As a corollary, we can immediately derive some associated distributions, such as the marginal and
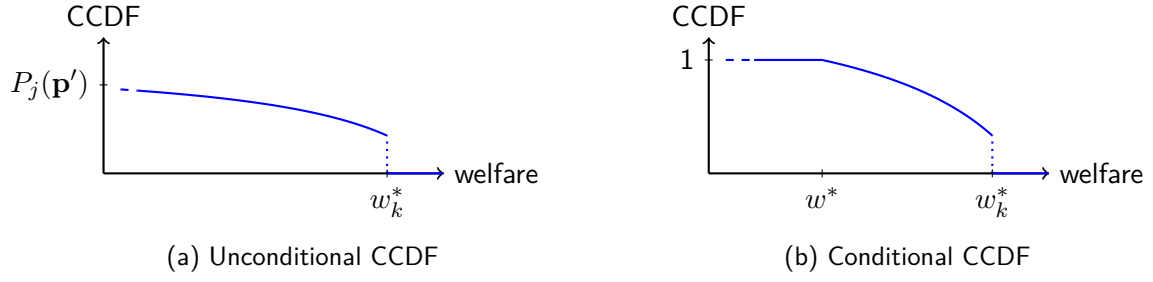
16

(a) Unconditional CCDF   (b) Conditional CCDF

Figure 3: The course of the (un)conditional CCDF of welfare

conditional CCDFs, which are more relevant in empirical applications.

Corollary 1

$$\Pr_{\omega}\left[w \leq W^{\omega}\left(y - p_k, k\right) \mid j = J^{\omega}(\mathbf{p}', y)\right] = \frac{P_{j,k}\left(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\right)}{P_j\left(\mathbf{p}', y\right)}\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (15)$$

$$\Pr_{\omega}\left[w \leq W^{\omega}\left(y - p_k, k\right) \mid k = J^{\omega}(\mathbf{p}, y)\right] = \frac{P_k\left(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\right)}{P_k\left(\mathbf{p}, y\right)}\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (16)$$

where $\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)) = \left(\min(p_1, \widetilde{p}_1(w)), \ldots, \min(p_n, \widetilde{p}_n(w))\right),$

$$\Pr_{\omega}\left[w \leq W^{\omega}\left(y - p_k, k\right)\right] = P_k\left((p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\right)\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (17)$$

and

$$\Pr_{\omega}\left[w \leq W^{\omega}\left(y - p_{J^{\omega}(\mathbf{p},y)}, J^{\omega}(\mathbf{p}, y)\right)\right] = \sum_k P_k\left(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\right)\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]. \qquad (18)$$

See proof on page 47.

We find again that the different derived distributions can be expressed in terms of choice and transition probabilities. Equations (15) and (16) can be used to assess the distribution of welfare when the researcher observes which bundle is optimal and wants to take this information into account. Equation (17) describes the marginal distribution of welfare evaluated in a specific bundle, not taking into account which bundle is optimal. Finally, Equation (18) specialises this result to a setting where welfare is evaluated in the optimal bundle.

A typical example of the distribution of welfare in bundle k conditional on bundle k being optimal is plotted in Figure 3b. As before, define for an option c, w_c^* to be the highest value of w such that $p_c = \widetilde{p}_c(w)$, and also define w^* to be $\min_c\{w_c^*\}$. Then we observe that for $w \leq w^*$, $p_c \leq \widetilde{p}_c(w)$ for all c, and hence, $\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)) = \mathbf{p}$. It follows that $\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k) \mid k = J^{\omega}(\mathbf{p}, y)\right] = 1$.

Hence, w^* is a lower bound for welfare in option k, conditionally on k being optimal. For $w > w^*$, $\Pr_\omega \left[w \leq W^\omega(y - p_k, k) \mid k = J^\omega(\mathbf{p}, y) \right]$ decreases continuously until w reaches w_k^* where $\Pr_\omega \left[w \leq W^\omega(y - p_k, k) \mid k = J^\omega(\mathbf{p}, y) \right]$ drops to 0, as seen in Figure 3a. Hence, w_k^* is an upper bound for welfare in option k, conditional on k being optimal, and the distribution has a mass point at w_k^*. If $w_k^* = w^*$, the distribution is thus a step function and, hence, the welfare level in bundle k, conditional on k being optimal at prices $\mathbf{p}$ and exogenous income y, is deterministic and equals w_k^*.

Example: the MMU measures For the MMUs, we obtain the following result.

Corollary 2 When using reference prices $\mathbf{p}^{ref}$, we have

$$\Pr_\omega \left[w \leq MMU^\omega_{\mathbf{p}^{ref}} (y - p_k, k), j = J^\omega(\mathbf{p}', y) \right] = P_{j,k}(\mathbf{p}', (p_k, y - w + \mathbf{p}^{ref}_{-k}), y) \mathbb{I} \left[p_k \leq y - w + p_k^{ref} \right].$$

When $p_k = p'_k$, and the reference prices equal the actual prices $\mathbf{p}'$ and k is the optimal choice, this simplifies to

$$\Pr_\omega \left[w \leq MMU^\omega_{\mathbf{p}'} (y - p'_k, k), k = J^\omega(\mathbf{p}', y) \right] = P_k(\mathbf{p}', y) \mathbb{I} [w \leq y]$$

and, hence,

$$\Pr_\omega \left[w \leq MMU^\omega_{\mathbf{p}'} (y - p'_k, k) \mid k = J^\omega(\mathbf{p}', y) \right] = \mathbb{I} [w \leq y],$$
$$\Pr_\omega \left[w \leq MMU^\omega_{\mathbf{p}'} \left(y - p'_{J^\omega(\mathbf{p}', y)}, J^\omega(\mathbf{p}', y) \right) \right] = \mathbb{I} [w \leq y].$$

See proof on page 49. Both the MMU in the optimal bundle and the MMU in bundle k, conditional on k being optimal, are, therefore, deterministic and equal the initial exogenous income y when reference equal actual prices.

4.2 Joint distribution of welfare levels and welfare differences

In this section, we derive the joint distribution of welfare levels and welfare differences. Joint knowledge on levels and differences of welfare enables investigation of the association between individuals' gains or losses from a price change and their position in terms of initial welfare. Assessment can be carried out on, for example, whether the winners of a price change belong disproportionately to those who were well-off in the first place. A price change is defined as an exogenous shift in prices from $\mathbf{p}$ to $\mathbf{p}'$. As discussed in Section 3.1, we will assume throughout that the unobserved preference type ω is unaltered by the price change.

Firstly, we study the general case in which welfare differences are defined on the basis of changes in NOS welfare measures (evaluated in optimal choices). We show that the joint distribution of levels and differences can be expressed in terms of transition probabilities. Secondly, we specialise our

results to the joint distribution of the MMU and the CV, which is a popular choice among applied welfare economists.[15] In doing so, we extend the results of Dagsvik and Karlström (2005) and de Palma and Kilani (2011) to a nonparametric setting.

4.2.1 Welfare differences in terms of NOS measures

As an intermediate step, we first derive the joint distribution of welfare before and after a price change in Proposition 1.

Proposition 1 The joint distribution of welfare in the optimal bundle i, before a price change, and welfare in the optimal bundle j, after the price change, is as follows:

$$
\begin{aligned}
&\Pr_\omega[w \le W_0^\omega(y - p_i, i), z \le W_1^\omega(y - p_j', j), i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] \\
&= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), \mathbf{min}(\mathbf{p}', \widetilde{\mathbf{p}}(z)), y\Big) \mathbb{I}\left[p_i \le \widetilde{p}_i(w)\right] \mathbb{I}\left[p_j' \le \widetilde{p}_j(z)\right].
\end{aligned}
\tag{19}
$$

See proof on page 49. Proposition 1 shows that this joint distribution can be written in terms of transition probabilities, evaluated at initial, final, and virtual prices. Using this proposition, the joint distribution of welfare levels and differences can be derived.

Theorem 2 The function h is defined by:

$$
\begin{aligned}
h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, s) &= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(\max(w, x))), \mathbf{min}(\mathbf{p}', \widetilde{\mathbf{p}}(s)), y\Big) \mathbb{I}\left[p_j' \le \widetilde{p}_j(s)\right] \\
&= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w), \widetilde{\mathbf{p}}(x)), \mathbf{min}(\mathbf{p}', \widetilde{\mathbf{p}}(s)), y\Big) \mathbb{I}\left[p_j' \le \widetilde{p}_j(s)\right].
\end{aligned}
$$

Then, the joint distribution of the stochastic welfare measure and the difference before and after the price change of this measure becomes,

$$
\begin{aligned}
&\Pr_\omega[w \le W_0^\omega(y - p_i, i), W_1^\omega(y - p_j', j) - W_0^\omega(y - p_i, i) \le z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] = \\
&- \int_{-\infty}^{+\infty} \partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, x + z) \mathbb{I}\left[p_i \le \min(\widetilde{p}_i(w), \widetilde{p}_i(x))\right] \, dx.
\end{aligned}
\tag{20}
$$

See proof on page 50. Unfortunately, it seems that this expression cannot be simplified. However, even though expression (20) seems technically complicated, only the transition probabilities are used as input. This object is nonparametrically identified from panel data.

4.2.2 Welfare differences in terms of the CV

We can specialise our results to the joint distribution of welfare and the CV, which is a popular choice among applied welfare economists. The CV refers to the (possibly negative) amount of the

---

[15]The derivation for the EV is similar and can be found in Appendix A.2.

numeraire an individual wants to give up after a price change to be equally well-off as before this change. For an individual of type ω, the compensating variation CV^ω is implicitly defined as

$$\max_c\{U_c^\omega(y - p_c)\} = \max_c\{U_c^\omega(y - p_c' - CV^\omega)\}, \tag{21}$$

where, as before, $\mathbf{p}$ are initial prices and $\mathbf{p}'$ final prices.[16] In fact this definition of the CV is equivalent to $MMU_{\mathbf{p}'}^\omega(y - p_{J^\omega(\mathbf{p}',y)}', J^\omega(\mathbf{p}',y)) - MMU_{\mathbf{p}'}^\omega(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p},y))$, i.e. the difference between the MMU with the final prices as reference price vector, in the optimal bundle after the price change, and the same MMU in the optimal bundle before the price change.[17] Note that the CV for a composition of two or more price changes cannot be calculated from the CV for each price change separately. In our more general general approach of measuring a change in welfare by the difference between two valuations of a welfare metric, this problem is inherently nonexistent.

Note that the results in Theorems 3 and 4, and in Corollaries 3 and 4 below, can in fact be seen as applications of Theorem 2. However, to give more insight, we also give direct proofs in Appendix B.

Distribution of the CV In order to derive the distribution of the CV when the choice is equal to option i under initial prices and option j under final prices, we can follow a similar strategy as Bhattacharya (2015) and de Palma and Kilani (2011). Analogously to Lemma 1, the condition $CV^\omega \leq z, i = J^\omega(\mathbf{p},y), j = J^\omega(\mathbf{p}',y)$ can be translated in i being the optimal bundle when faced with a counterfactual price vector.

Lemma 2 We have

$$\begin{aligned}
&\left\{\omega \mid CV^\omega \leq z, i = J^\omega(\mathbf{p},y), j = J^\omega(\mathbf{p}',y)\right\} \\
&= \left\{\omega \mid U_i^\omega(y - p_i) \geq \max_c\{U_c^\omega(y - p_c' - z)\}, i = J^\omega(\mathbf{p},y), j = J^\omega(\mathbf{p}',y)\right\}.
\end{aligned} \tag{22}$$

See proof on page 51.

With Lemma 2, we can state the following theorem.

---

[16]Note that in our definition, the CV has the opposite sign of that in Bhattacharya (2015), but our results are completely equivalent.

[17]Indeed, defining CV^ω by $MMU_{\mathbf{p}'}^\omega(y - p_{J^\omega(\mathbf{p}',y)}', J^\omega(\mathbf{p}',y)) - MMU_{\mathbf{p}'}^\omega(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p},y))$, we get $MMU_{\mathbf{p}'}^\omega(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p},y)) = y - CV^\omega$ by Corollary 2. Moreover, as i is the optimal bundle before the price change, we get

$$\begin{aligned}
\max_c\{U_c^\omega(y - p_c)\} &= U_{J^\omega(\mathbf{p},y)}^\omega(y - p_{J^\omega(\mathbf{p},y)}) \\
&= \max_c\{U_c^\omega(MMU_{\mathbf{p}'}^\omega(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p},y)) - p_c')\} \\
&= \max_c\{U_c^\omega(y - p_c' - CV^\omega)\}.
\end{aligned}$$

Theorem 3 The joint distribution of the CV and the optimal choices before and after the price change is as follows:

$$\Pr_{\omega}[CV^{\omega} \leq z, i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)] = P_{i,j}(\min(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)\mathbb{I}\left[p_i \leq p_i' + z\right]. \quad (23)$$

See proof on page 51.

We observe that $\Pr_{\omega}[CV^{\omega} \leq z, i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)]$ is bounded from below by $p_i - p_i'$. This is as expected; if the initial optimal bundle was i and the price of p_i drops to p_i', the numeraire must drop with at least this amount to be equally well-off as in the initial situation. This means that the minimal compensation, in terms of the joint distribution, is $p_i - p_i'$. Moreover, for $z \geq \max_k\{p_k - p_k'\}$, $\Pr_{\omega}[CV^{\omega} \leq z \mid i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)] = 1$. This means that the maximal compensation, in terms of the conditional distribution, cannot be higher than the maximal price difference, which is also as expected.

The next corollary follows immediately and may again be more useful to the applied researcher.

Corollary 3

$$\Pr_{\omega}[CV^{\omega} \leq z \mid i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)] = \frac{P_{i,j}(\min(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)}{P_{i,j}(\mathbf{p}, \mathbf{p}', y)}\mathbb{I}\left[p_i \leq p_i' + z\right], \quad (24)$$

$$\Pr_{\omega}[CV^{\omega} \leq z \mid i = J^{\omega}(\mathbf{p}, y)] = \frac{P_i(\min(\mathbf{p}, \mathbf{p}' + z), y)}{P_i(\mathbf{p}, y)}\mathbb{I}\left[p_i \leq p_i' + z\right], \quad (25)$$

$$\Pr_{\omega}[CV^{\omega} \leq z \mid j = J^{\omega}(\mathbf{p}', y)] = \sum_i \frac{P_{i,j}(\min(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)}{P_j(\mathbf{p}', y)}\mathbb{I}\left[p_i \leq p_i' + z\right], \quad (26)$$

and

$$\Pr_{\omega}[CV^{\omega} \leq z] = \sum_i P_i(\min(\mathbf{p}, \mathbf{p}' + z), y)\mathbb{I}\left[p_i \leq p_i' + z\right].^{18} \quad (27)$$

Equation (27) gives an expression for the marginal distribution of CV. Equations (24), (25) and (26), which present conditional distributions, can be used to calculate the distribution of CV when (i) the optimal bundles before and after price change are known, (ii) only before the price change is known, and (iii) only after the price change is known.

Joint distribution of the MMU and the CV We now apply Theorem 2 to the case where one chooses the MMU with final prices as the reference price vector, as a welfare measure. The difference in welfare before and after the price change is then equal to the CV.

---

[18] Note that Equation (27) is the main result of Bhattacharya (2015).

Theorem 4 The joint distribution of the MMU and the CV is as follows:

$$\Pr_\omega[w \le MMU^\omega_{\mathbf{p}'}(y - p_i, i), CV^\omega \le z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$$
$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), \mathbf{p}', y\Big)\mathbb{I}\left[p_i \le p'_i + \min(z, y - w)\right]. \tag{28}$$

See proof on page 52.

Again, Corollary 4 follows immediately.

Corollary 4

$$\Pr_\omega\Big[w \le MMU^\omega_{\mathbf{p}'}, CV^\omega \le z \mid i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\Big]$$
$$= \frac{P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), \mathbf{p}', y\Big)}{P_{i,j}(\mathbf{p}, \mathbf{p}', y)}\mathbb{I}\left[p_i \le p'_i + \min(z, y - w)\right], \tag{29}$$

$$\Pr_\omega[w \le MMU^\omega_{\mathbf{p}'}(y - p_i, i), CV^\omega \le z \mid i = J^\omega(\mathbf{p}, y)]$$
$$= \frac{P_i\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), y\Big)}{P_i(\mathbf{p}, y)}\mathbb{I}\left[p_i \le p'_i + \min(z, y - w)\right], \tag{30}$$

$$\Pr_\omega[w \le MMU^\omega_{\mathbf{p}'}\Big(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p}, y)\Big), CV^\omega \le z \mid j = J^\omega(\mathbf{p}', y)]$$
$$= \sum_i \frac{P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), \mathbf{p}', y\Big)}{P_j(\mathbf{p}', y)}\mathbb{I}\left[p_i \le p'_i + \min(z, y - w)\right], \tag{31}$$

and,

$$\Pr_\omega[w \le MMU^\omega_{\mathbf{p}'}\Big(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p}, y)\Big), CV^\omega \le z]$$
$$= \sum_i P_i\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), y\Big)\mathbb{I}\left[p_i \le p'_i + \min(z, y - w)\right]. \tag{32}$$

The joint cumulative distribution can again be written as (a sum of) choice or transition probabilities. Each choice and transition probability is calculated using up to three price vectors: the initial price vector $\mathbf{p}$, the final price vector $\mathbf{p}'$, and a translation of the $\mathbf{p}'$ vector for the combined MMU and CV part.

4.3 Social welfare

A classical additively, separable Bergson-Samuelson social welfare function (SWF) takes the form

$$SWF = \int h(u)\, dG_U(u), \tag{33}$$

where u is the value of a utility function representing the well-being of an individual in a particular state of the world, h is a concave function expressing the inequality aversion, and G_U is the CDF of the well-being distribution in the population in a given state of the world.[19] For example, in the utilitarian case, we have that $h(u) = u$.

Since the seminal work of Sen (1970), it is well known that in order to be able to meaningfully formulate a SWF, stringent assumptions on the measurability and degree of interpersonal comparability of such utility functions are required.[20] As a first option, researchers sought for conditions to make welfare prescriptions in terms of the income distributions instead (see, e.g. Roberts, 1980). These conditions turned out to be restrictive; preferences should be identical and homothetic, and the social welfare function in terms of incomes should be homothetic too.

A second option is to use the NOS welfare measures as a representation of preferences. Fortunately, they are known to satisfy a set of attractive principles of interpersonal comparability (see Fleurbaey and Maniquet, 2017; 2018). We can, therefore, use these measures directly as building blocks in the SWF in Equation (33). More specifically, the equivalent to the Bergson-Samuelson SWF in our framework reads as

$$SWF = \int \int h(w)\, dF_W(w \mid \mathbf{p}, y)\, dG(\mathbf{p}, y), \tag{34}$$

where G is the CDF of the joint distribution of prices and exogenous income in the population, which can be observed from the data, and $F_W(w \mid \mathbf{p}, y)$ is the conditional CDF of the NOS measure W, and equals $\Pr_\omega \left[W^\omega \left(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p}, y) \right) \leq w \right]$.[21]

Proposition 2 illustrates how the results on the distribution of welfare levels in Corollary 1 lead to the calculation of social welfare as defined in Equation (34), using only choice probabilities.

Proposition 2 The conditional CDF of individual welfare in the optimal bundle can be calculated using choice probabilities:

$$F_W(w \mid \mathbf{p}, y) = 1 - \sum_k P_k\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\Big) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w) \right]. \tag{35}$$

---

[19]A function f is additively separable when it can be written in the form $f(x_1, \ldots, x_n) = \sum_i f_i(x_i)$.

[20]In a context of demand under a linear budget constraint, an indirect utility function in terms of prices and income may be a natural candidate to serve as utility function. However, such utility functions are only determined up to a positive monotone transformation and economists did not find agreement on which representative is the most suitable to serve as a basis for making interpersonal utility comparisons (for an overview of the debate, see Hammond, 1991).

[21]As social welfare is a population level concept, we rely on the second interpretation of the randomness in the welfare measure (see the discussion at the beginning of Section 4).

See proof on page 54.

Hence, social welfare can be computed from these probabilities. The joint distribution of prices and exogenous income G can be estimated separately using standard nonparametric tools.

Moreover, this expression can be used to identify if a price change, for example, due to a policy reform, has a desirable effect on social welfare. Indeed, the difference in social welfare can be calculated as follows:

$$
\begin{aligned}
SWF' - SWF &= \int\int h(w)\,dF_W(w \mid \mathbf{p}', y)\,dG'(\mathbf{p}', y) - \int\int h(w)\,dF_W(w \mid \mathbf{p}, y)\,dG(\mathbf{p}, y) \\
&= \int\int h(w)\,dF_W(w \mid \mathbf{p} + \Delta\mathbf{p}, y)\,dG'(\mathbf{p} + \Delta\mathbf{p}, y) \\
&\quad - \int\int h(w)\,dF_W(w \mid \mathbf{p}, y)\,dG(\mathbf{p}, y) \\
&= \int\int h(w)\,d\Big(F_W(w \mid \mathbf{p} + \Delta\mathbf{p}, y) - F_W(w \mid \mathbf{p}, y)\Big)\,dG(\mathbf{p}, y).
\end{aligned}
\tag{36}
$$

where G (G') is the joint distribution of initial (final) prices and exogenous income, and $\Delta\mathbf{p} = \mathbf{p}' - \mathbf{p}$. With Equations (36) and (35), one can asses the desirability of a potential price change without parametric assumptions and only using choice probabilities and the initial distribution of prices and exogenous income.

Interestingly, in the spirit of Roberts (1980), we can derive conditions under which the expression for the SWF can be formulated in terms of incomes alone. In particular, when prices are equal for everyone and one uses the MMU with reference prices equal to those common prices, as individual welfare measure, one obtains a price independent SWF in terms of income.

Corollary 5 When prices are equal for everyone and when one uses the MMU with reference prices equal to those common prices as the welfare measure, the SWF can be written solely in terms of income.

See proof on page 54.

5   Discussion on implementation

5.1   Set-identifying transition probabilities from cross-sectional data

As mentioned before, the transition probabilities are nonparametrically identifiable and estimable from panel data that contains sufficient relative price and exogenous income variation. This immediately implies that all the results from previous subsections are also nonparametrically identified in such a data setting. One simply has to evaluate the estimated transition probabilities at virtual price vectors.

In many empirical applications, however, researchers only have access to (repeated) cross-sectional data. This type of data nonparametrically identifies the choice probabilities, but not the associated transition probabilities. However, by exploiting Boole-Fréchet (Fréchet, 1935) and stochastic revealed preference inequalities, one can derive bounds on the now unobserved transition probabilities based on the observed choice probabilities.

Proposition 3 Suppose Assumptions 1 – 3 hold. Then the transition probabilities $\{P_{i,j}(\mathbf{p}, \mathbf{p}', y)\}$ are set identified from the choice probabilities $\{P_i\}$ with bounds

$$P_{i,i}^L(\mathbf{p}, \mathbf{p}', y) = \max \left\{ P_i(\mathbf{p}, y) + P_i(\mathbf{p}', y) - 1, P_i\big((\max\{p_i, p_i'\}, \mathbf{min}\{\mathbf{p}_{-i}, \mathbf{p}'_{-i}\}), y \big) \right\},$$
$$P_{i,i}^U(\mathbf{p}, \mathbf{p}', y) = \min \{ P_i(\mathbf{p}, y),\ P_i(\mathbf{p}', y) \}.$$

For $i \neq j$, $P_{i,j}(\mathbf{p}, \mathbf{p}', y) = 0$ if $p_i \geq p_i'$ and $p_j \leq p_j'$ and

$$P_{i,j}^L(\mathbf{p}, \mathbf{p}', y) = \max\{P_i(\mathbf{p}, y) + P_j(\mathbf{p}', y) - 1,\ 0\},$$
$$P_{i,j}^U(\mathbf{p}, \mathbf{p}', y) = \min \{ P_i(\mathbf{p}, y), P_j(\mathbf{p}', y) \},$$

elsewhere.

See proof on page 55.

The Boole-Fréchet inequalities ensure that the transition probabilities are weakly smaller than their associated marginal choice probabilities $P_i(\mathbf{p}, y)$ and $P_j(\mathbf{p}', y)$. When $P_i(\mathbf{p}, y) + P_j(\mathbf{p}', y) - 1 > 0$ they also deliver nontrivial lower bounds. The stochastic revealed preference inequalities, which stem from the strong monotonicity of the utility function (see Assumption 1), provide additional identificational power in two particular instances. Firstly, by evaluating the choice probabilities at the least-favourable price vector $(\max\{p_i, p_j'\}, \mathbf{min}\{\mathbf{p}_{-i}, \mathbf{p}'_{-i}\})$, they yield an informative lower bound for the transition probabilities in the no-transition case where $i = j$. Secondly, when i becomes weakly less expensive and $j \neq i$ becomes weakly more expensive, the transition probability should equal zero, as it is irrational for individuals to make this transition within the context of our model.

5.2 Estimating choice probabilities

Given the exogeneity of budget sets presupposed in Assumption 3, the choice probabilities can be readily estimated using nonparametric regression, as they are essentially conditional expectation functions. Standard tools, such as kernel and series based regression, are available in most modern statistical software. One particular attractive feature of the Nadaraya-Watson kernel estimator is that the estimated choice probabilities add up to one for all price vectors when the same bandwidth is selected for every choice probability function.

When only samples of modest size are available, it might be useful to impose additional struc-

ture to mitigate the curse of dimensionality. In particular, in a setting with high-dimensional re-gressors, which arises when there are many goods or many choice and individual level characteristics, a (semi)parametric estimator can be used to increase efficiency at the expense of functional form mis-specification. In particular, our empirical illustration in Section 6 will make use of a semiparametric estimator that can be interpreted as a sieve approximation.

Shape restrictions   One point that needs further attention is that sampling noise might cause the non- or semiparametric estimates to be inconsistent with the monotonicity condition in Assumption 1 over some ranges of the data. When this condition is violated, the CDFs of our distributional results might be decreasing over some sections of their support. To avoid these inconsistencies, researchers can impose the following shape restrictions on the estimated choice probabilities:

Proposition 4   Suppose Assumptions 1-3 hold. Then the estimated choice probabilities $\{P_i\}$ should satisfy the following shape restrictions at all $(\mathbf{p}, y)$ for all $i, j \in \mathcal{C}$:

$$\frac{\partial P_i(\mathbf{p}, y)}{\partial p_i} \leq 0,$$
$$\frac{\partial P_i(\mathbf{p}, y)}{\partial p_j} \geq 0,$$

See proof on page 56. It is easy to see that, even under the presence of unrestricted, unobserved heterogeneity, utility maximisation yields restrictions on the partial derivatives of the choice probab-ilities. In particular, rationality implies that the choice probability for any given good is decreasing in its own price and increasing in the price of the other alternatives. In addition, the adding-up requirement for probabilities also imposes a restriction across choice probabilities.

Endogeneity of prices and income   In some circumstances, it might be unreasonable to assume that the budget set $(\mathbf{p}, y)$ is independent of the preference type ω. When a vector of instruments is available, however, some forms of endogeneity can be handled by using a standard control function approach (Blundell and Powell, 2004).

Ordered choice and outside good   Similar to Bhattacharya (2015), it is important to stress that identification generally fails in settings where choice is ordered and where the prices of alternatives are multiples of one another. In such a setting, there is no relative price variation in the data that identifies the effect of a price change in some alternative(s) while keeping the prices of the other alternatives fixed.

Moreover, in some applications, there is an outside good that exhibits no independent price variation, which also hinders the direct empirical implementation of our approach. However, this difficulty can be circumvented by exploiting variation in the exogenous income y. Suppose good $c_o \in \mathcal{C}$ is the

outside good for which one has to evaluate the effect of a price change $\Delta p_o = p'_o - p_o$. By a change of variables, it then always holds that $P_i(\mathbf{p}', y) = P_i(\mathbf{p}' - \Delta p_o, y - \Delta p_o)$. Note that the expression at the right-hand side does not require price variation for c_o, as $p'_o - \Delta p_o = p_o$ by construction.

5.3 Average welfare

A well-known implication of Fubini's theorem is that the mean of any random variable X, given that it exists, can be directly derived from its cumulative density function F_X, i.e.

$$\mathbb{E}(X) = \int_0^\infty (1 - F_X(u)) du - \int_{-\infty}^0 F_X(u) du.$$

This result allows us to calculate average welfare from any of the distributional results derived above. Note that when only bounds on the distribution of interest are available (see Section 5.1), the expected value can be bounded by

$$\int_0^\infty (1 - F_X^U(u)) du - \int_{-\infty}^0 F_X^U(u) du \leq \mathbb{E}(X) \leq \int_0^\infty (1 - F_X^L(u)) du - \int_{-\infty}^0 F_X^L(u) du,$$

where F_X^L and F_X^U denote the CDF of the lower and upper bound respectively. This result allows us to calculate average welfare from any of the distributional results derived above.

6   Empirical illustration

We highlight the empirical applicability of our results by revisiting the classical trade-off between leisure and consumption. The goal of this illustration is two-fold. Firstly, we demonstrate how the results in this paper enable researchers to assess the distribution of welfare within and across different groups in society. Secondly, we show how our results allow to evaluate the effects of an income tax reform on individual and social welfare. Thereby, we concentrate on two particular aspects: (i) a comparison of the welfare distribution before and after the reform, and (ii) the extent to which the winners and losers are (un)equally spread across the initial welfare distribution.

For this purpose, we make use of microdata from the 2018 wave of the German Socio-Economic Panel (SOEP), which contains detailed information on households' demographics, labour supply, wages, and out-of-work income. We model single females' labour supply as a choice between three discrete alternatives: non-working (NW), part-time employment (PT), and full-time employment (FT). As an income tax reform, we consider the introduction of a basic income flat tax in Germany, which would replace the current nonlinear tax schedule.

6.1 Setting and implementation

German tax system and a basic income flat tax reform The German personal income tax system is distinctly progressive. Taxes and social security contributions are paid on both earned, capital, and transfer income. After a basic tax-free allowance (8,820 euro in 2017), statutory marginal tax rates increase almost continuously from 14 to 45%. The system also has deductions for work related expenses, and allowances for lone parents and childcare expenses. There are no refundable tax credits; taxes, therefore, cannot become negative. Parents with dependent children are eligible for child benefits. For those who are not able to work, a subsistence income level is guaranteed by social assistance, which includes allowances for housing and heating costs. These benefits are means tested for income and wealth, and depend on the composition of the household.

As a policy reform, we consider the introduction of a basic income flat tax. In this exercise, the current nonlinear tax schedule is replaced by one where only a single rate is applied to every individual's taxable income. This does not yield a fully proportional tax schedule, however, as we leave the current deductions and reductions unchanged. The rate is set to 42%, which makes the reform revenue neutral from the point of the government.[22] In addition, we remove the means test for social assistance with regard to earned income, such that it acts as a basic income.

Data selection and estimation From the SOEP sample, we construct a subsample with single females that are available to the labour market. That is, we restrict the analysis to those below 60 years old. To reduce the effect of outliers, we also drop individuals with gross hourly wages outside of 4-90 euro and gross yearly asset income above 12,000 euro. Individuals with missing working hours are also discarded. Our final subsample consists of 1,922 single females; Appendix C.1 contains descriptive statistics for this subsample.

We map observed working hours into three discrete alternatives: non-working (i.e. hours strictly lower than 5 hours/week); part-time employment (i.e. hours higher than 5 hours/week and strictly lower than 32 hours/week); and full-time employment (i.e. hours higher than 32 hours/week). For each of these alternatives, we calculate disposable income by means of a tax-benefit calculator.[23] The disposable income for the unemployment alternative is at least at the level guaranteed by social assistance.

We model the choice probabilities for alternatives PT and FT semiparametrically, as we estimate for each a flexible binary logit model that contains cubic polynomials in the disposable income d of all

---

[22]Revenue neutrality is here defined without taking labour supply responses into account. It is also calculated on the whole sample, not just on the subsample of single females for which we conduct the welfare analysis. We find that behavioural reactions have little impact on tax revenues.

[23]Missing wages for the non-working are imputed using a Heckman-type selection model, with variables on the number of children acting as exclusion restrictions, as they are only included in the selection equation and not in the wage equation. For each individual, monthly disposable income is calculated for 0 (i.e. NW), 20 (i.e. PT), and 40 (i.e. FT) hours of work per week.

three alternatives and a linear index with demographic variables, such as age, years of education, number of children, and region. The choice probability for NW is defined as the complement. By means of an arbitrary normalisation, we fix an individual's exogenous income to $y = d_{FT}$, that is to the amount of disposable income she would obtain when working full-time. This is convenient, as it ensures that all prices are non-negative: i.e. $p_{NW} = d_{FT} - d_{NW}$, $p_{PT} = d_{FT} - d_{PT}$, and $p_{FT} = 0$. For each alternative, the shape restrictions in Proposition 4 are imposed by means of a penalty function. This penalty function also ensures that the choice probability for NW is nowhere negative in the support of the data. For more details on the estimation procedure, we refer to Appendix C.2.

Welfare measure and reference prices All our results are calculated on the basis of a MMU (see Equation (13) for a definition). We fix the reference price for each alternative at the sample median of the difference in disposable income between working full-time and that respective alternative: i.e. $p_{NW}^{ref} = \mathsf{med}(d_{FT} - d_{NW})$, $p_{PT}^{ref} = \mathsf{med}(d_{FT} - d_{PT})$, and $p_{FT}^{ref} = 0$. The virtual prices, therefore, become $\widetilde{p}_c(w) = d_{FT} - w + p_c^{ref}$, for c in $\{NW, PT, FT\}$.

6.2   Results

Individual and grouped welfare distributions We first study the distribution of individual welfare in the baseline, conditional on the chosen alternative (i.e. Equation (16) in Corollary 1). Figure 4 shows estimates of this distribution for all females in our subsample, partitioned in quartiles of gross hourly wages. Hourly wages reflect potential earning capacity and can be thought of as a proxy for ability.

As was noted at the outset of Section 4, each individual distribution either reflects the econometrician's uncertainty about the welfare level of an individual with such observable characteristics (i.e. choice, prices, exogenous income, and demographic variables) or it reflects the distribution of actually realised welfare levels in the population of single females with such observable characteristics. In either case, possible differences in welfare for an individual with given observable characteristics are due to unobserved preference heterogeneity. For simplicity of exposition, we will maintain the second interpretation in the remainder of this empirical illustration.

Visual inspection of Figure 4 reveals that these distribution functions have the expected shape (see Figure 3b).[24] On the one hand, there is a critical welfare level w^* below which the virtual prices of all three alternatives are higher than the actual price, and the welfare level will surely exceed that level. On the other hand, the welfare level w_{c^*}, for which the virtual price of the actually chosen alternative equals its actual price, is the highest welfare one can obtain. When $w_{c^*}^* = w^*$, the distribution degenerates to a step function, and we can determine the individual's welfare level

---

[24]Note that Figure 3b plots the CCDF, while the figures we present here are CDFs.
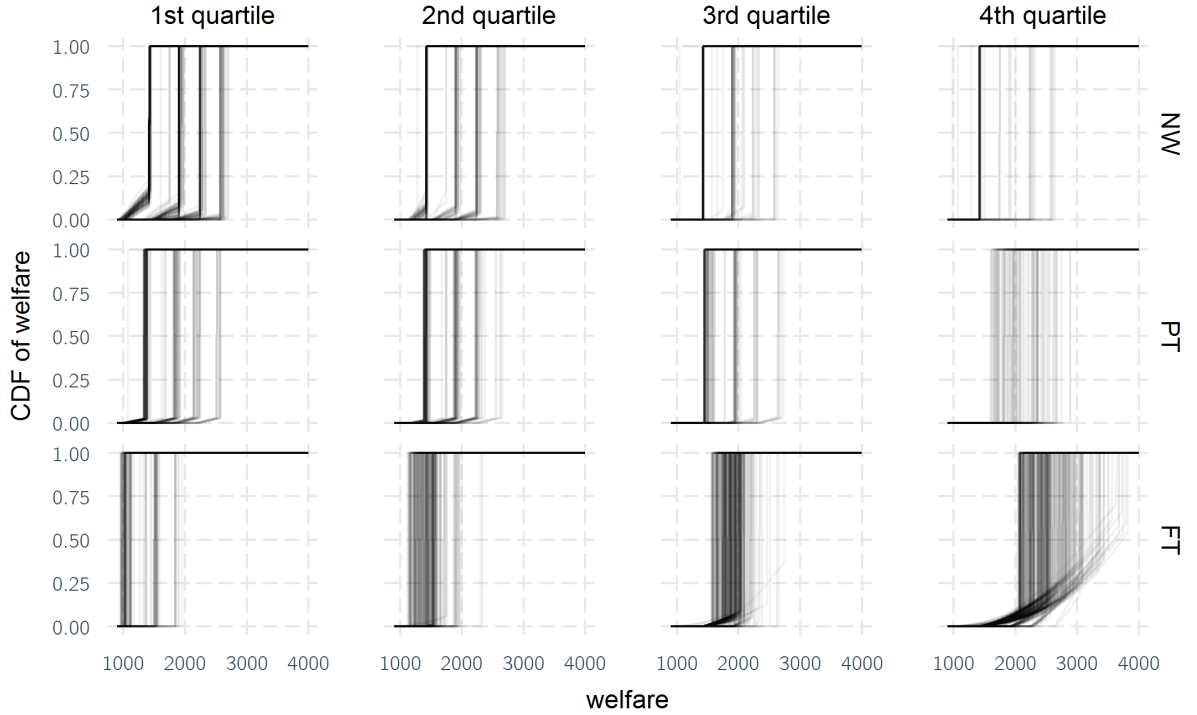
Figure 4: Individual welfare distributions for all females by alternative chosen and wage quartile

exactly, whatever her preferences are. This happens to be the case for approximately 25% of the females in our subsample. It predominantly occurs for low-wage individuals who choose FT and high-wage individuals who choose NW.

Figure 10 in Appendix C.5 also displays the individual *un*conditional welfare distributions, that is the welfare distribution of individuals with the same observable characteristics (i.e. prices, exogenous income, and demographic variables), whatever their optimal choice would have been (i.e. Equation (18) in Corollary 1). While these distributions turn out to exhibit several mass points, the deterministic cases seem to vanish. This is to be expected, as conditioning on observed choices introduces information that restricts the set of preference types. Therefore, the conditional distributions are 'less stochastic' than their associated marginal distributions.

In Figure 5, we present the welfare distribution for groups based on the alternative chosen and wage quartile. These grouped distributions are obtained by aggregating the individual distributions within each of the panels of Figure 4.[25] The figure, thus, represents the welfare distribution in society for

---

[25]In Figure 4, we plotted individual CDFs, i.e. for an individual ℓ, $\Pr_\omega[W^\omega \leq w \mid \mathbf{p}_\ell, y_\ell, \mathbf{x}_\ell, i = J^\omega(\mathbf{p}_\ell, y_\ell, \mathbf{x}_\ell), q]$. When aggregating these distributions to a group level, we lower the level of conditioning by integrating out over prices, exogenous income, and demographic characteristics. At the sample level, we obtained the aggregation as follows:

$$\Pr_\omega[W^\omega \leq w \mid i = J^\omega(\mathbf{p}, y, \mathbf{x}), q] = \sum_\ell \Pr_\omega[W^\omega \leq w \mid \mathbf{p}_\ell, y_\ell, \mathbf{x}_\ell, i = J^\omega(\mathbf{p}_\ell, y_\ell, \mathbf{x}_\ell), q = q_\ell] \frac{1}{\#\{\ell \mid i = J^\omega(\mathbf{p}_\ell, y_\ell, \mathbf{x}_\ell), q = q_\ell\}},$$

where $\mathbf{x}$ is a vector that contains demographic variables, and q indicates the wage quartile. Asymptotically, this is

each of these twelve groups. The welfare distribution of high-wage (i.e. fourth quartile) individuals who choose FT tends to first-order dominate the welfare distribution of those who choose alternative PT, and the latter dominates that of NW. Notwithstanding some exceptions at the bottom part of the distribution, the opposite is true for the low-wage individuals (i.e. first quartile). This can be explained as follows. Low-wage individuals have a relatively low gain in disposable income from choosing PT or FT compared with NW. As a consequence, low-wage individuals who choose FT must have more intense preference for income relative to leisure than other low-wage individuals. But because their disposable income is relatively low, this implies that their welfare is relatively low. One could say that, for those individuals, their preferences are less adapted to their wages. The reverse is the case for persons with high gross hourly wages.
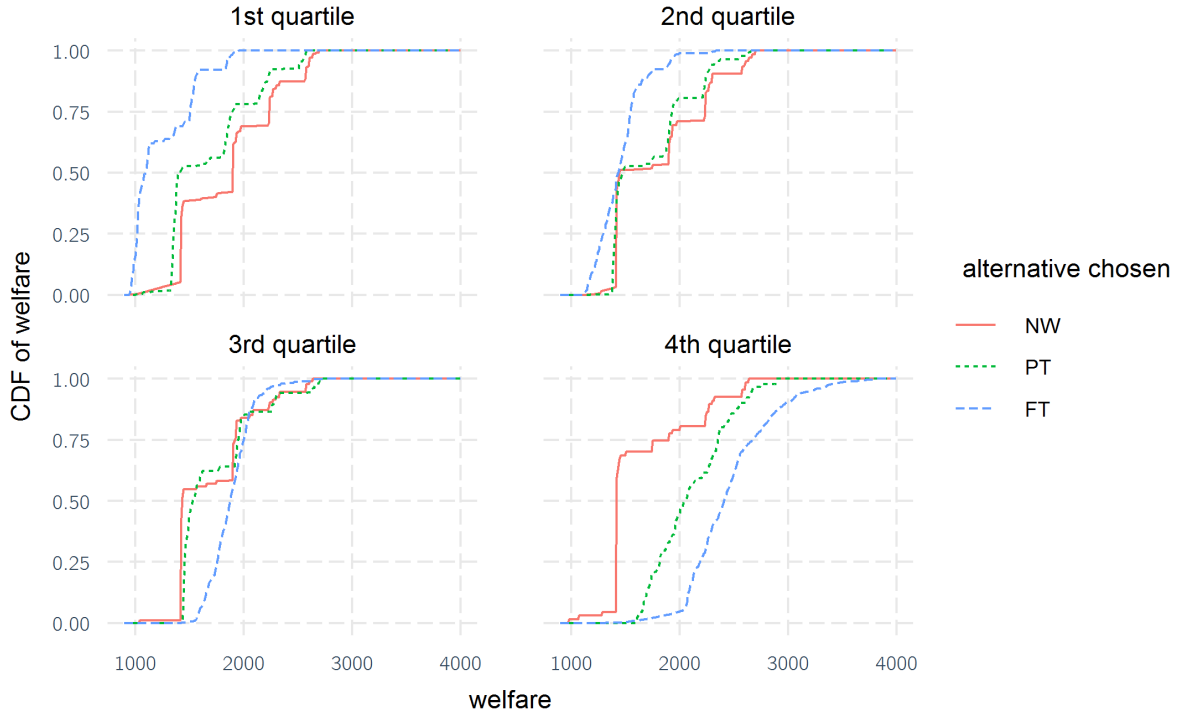

Figure 5: Grouped welfare distributions by alternative chosen and wage quartile

In Figure 6, we further aggregate these distributions by integrating out the chosen alternative. The group with the highest wages tends to first-order dominate the other groups.[26] Substantially higher wages thus lead to increased welfare, despite the large degree of unobserved preference variation we allow for. However, systematic preference differences between the populations of different wage quartiles, due to different composition of demographic variables (age, education, household size)

equivalent to the population concept

$$\int_{\mathbf{p},y,\mathbf{x}} \Pr_{\omega}[W^{\omega} \leq w \mid \mathbf{p}, y, \mathbf{x}, i = J^{\omega}(\mathbf{p}, y, \mathbf{x}), q]\, dG(\mathbf{p}, y, \mathbf{x} \mid i = J^{\omega}(\mathbf{p}, y, \mathbf{x}), q).$$

[26]Inequality also seems to be higher in the highest wage group compared to the others.

among the wage quartiles, might play a role too in explaining the welfare dominance of the fourth quartile. In contrast, welfare levels obtained by individuals belonging to the lower three wage quartiles turn out to be more intermingled. This suggests that, besides wages, both systematic and unobserved preference differences do play an important role in assessing the welfare of an individual.
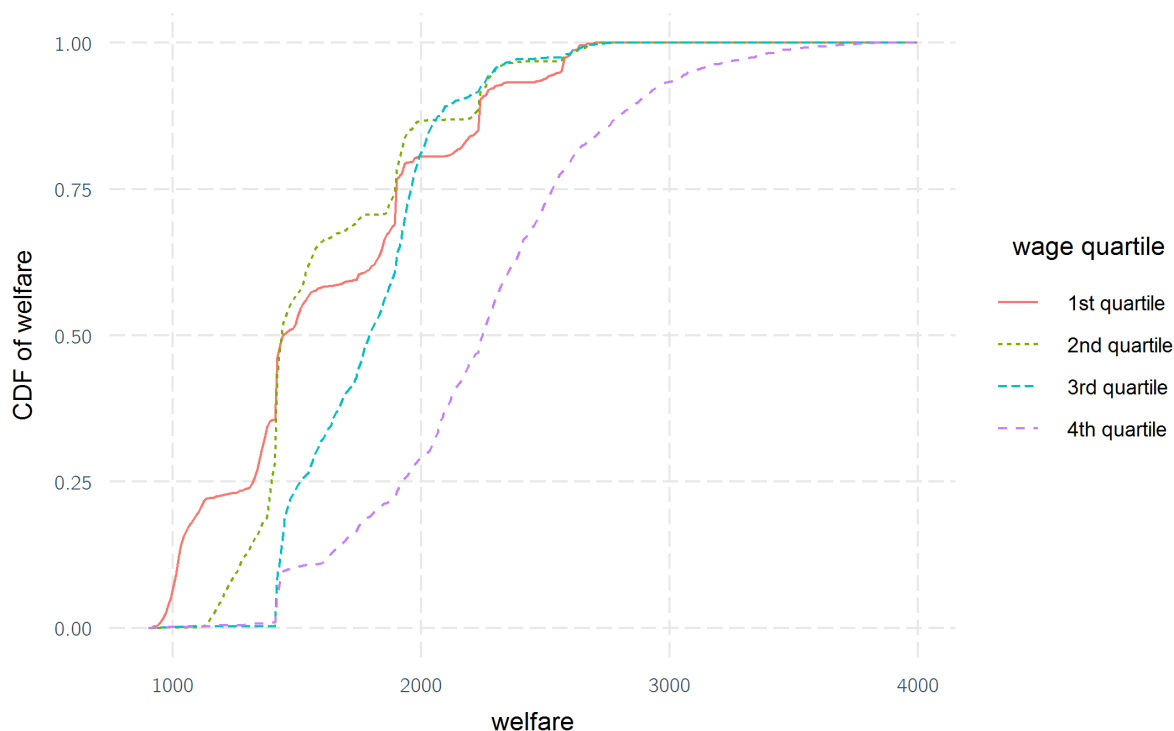

Figure 6: Grouped welfare distribution by wage quartile

Social welfare and reform   Our next point is to discuss the effects of the simulated reform, where the existing nonlinear tax system is replaced with a basic income flat tax.

Figure 7 compares the overall welfare distribution for the baseline and the reform. The welfare distribution is computed by further aggregating the distributions for the different wage groups of Figure 6 into one overall welfare distribution for the entire population of single females. The reform welfare distribution tends to first-order dominate the baseline welfare distribution.[27] It is well known that this implies that all the members of the class of SWFs we propose in Equation (34) will unanimously prefer the reform welfare distribution over that for the baseline. For completeness, we present some estimates for different inequality aversion parameters of the Atkinson SWF in Table 4 of Appendix C.5.

---

[27]This may be partly due to the fact that the reform is not budget neutral for this specific subsample. We leave a detailed analysis of this result for future research, as it is beyond the scope of this illustration.
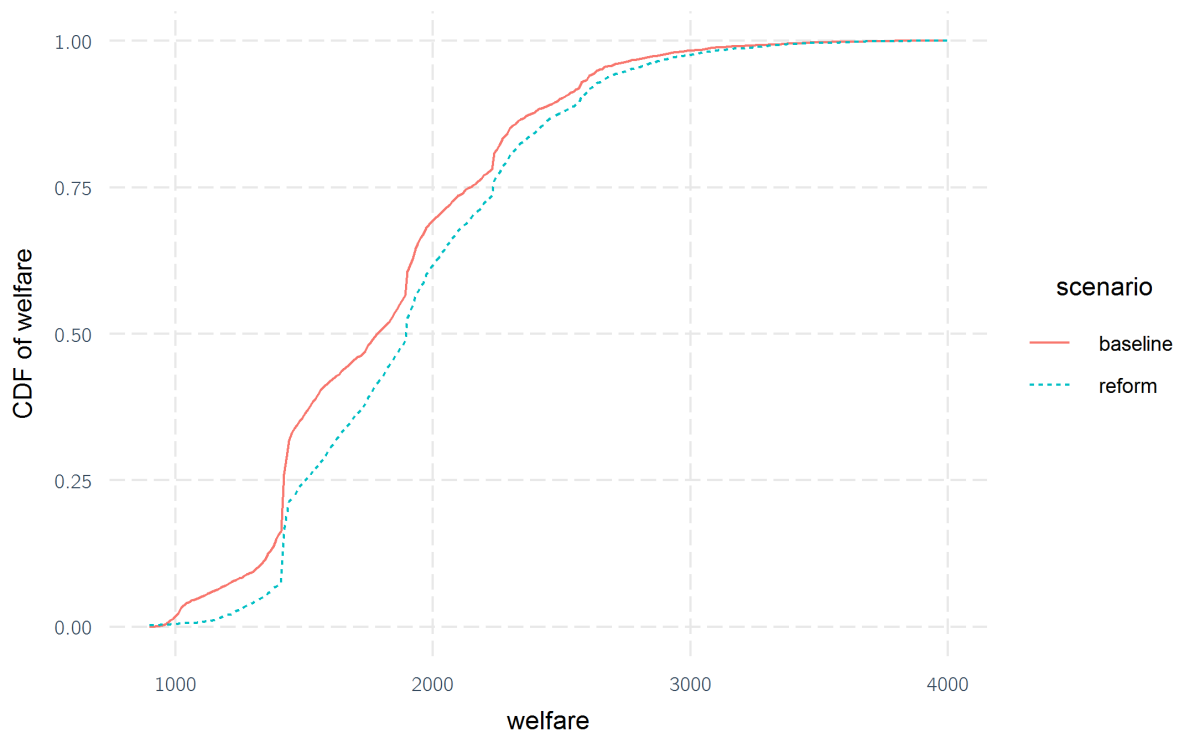
Figure 7: Welfare distribution in the baseline and reform

Winners and losers The presence of first-order dominance does not imply that everybody gains.[28] We further analyse, therefore, the extent to which the winners and losers are (un)equally spread across the baseline welfare distribution. For this purpose, we approximate the over all individuals and their baseline and reform choices aggregated joint distribution of welfare and welfare differences (i.e. Equation (20)) numerically.[29] Finally, we transform this joint distribution into the distribution of welfare differences conditional on the baseline welfare level. We refer to Appendix C.4 for more details on this procedure.

Figure 8 depicts a smoothed version of the 10th, 50th, and 90th iso-percentile contours (Figures 13 and 14 in Appendix C.5 show a more detailed picture). Each point (w, z) on the qth contour indicates the minimal welfare gain z (or a loss, if z is negative) that is reached by $q\%$ of the population with baseline welfare level w. First, observe that, despite the first-order dominance, there are a considerable amount of losers. For example, welfare levels at which the 90th iso-percentile curve lies below the zero point on the vertical axis, indicate that at least 10% of the persons with this baseline welfare level exhibit a loss. This occurs especially for welfare levels above 2,000 euro (this is above the mean, see Table 4 in Appendix C.5). Secondly, among the richer persons in the baseline, in terms of

---

[28] This is because first-order dominance is a combination of both Pareto-dominance and anonymity.

[29] Notice that we cannot use the simplified versions based on the CV (i.e. Equation (28) of Theorem 4) or the EV (i.e. Equation (48) of Theorem 6). Indeed, in our application the actual baseline and reform prices are individual specific. Using these as reference prices would imply that the resulting welfare measure does not any longer comply with our definition, which requires that the nested opportunity sets are common to all individuals.

well-being, there is a majority of losers. Thirdly, there is a small tendency for the number of losers to increase with welfare. However, the distribution of losses and gains varies a lot across baseline welfare levels, as might be seen from the more detailed pictures in Appendix C.5.
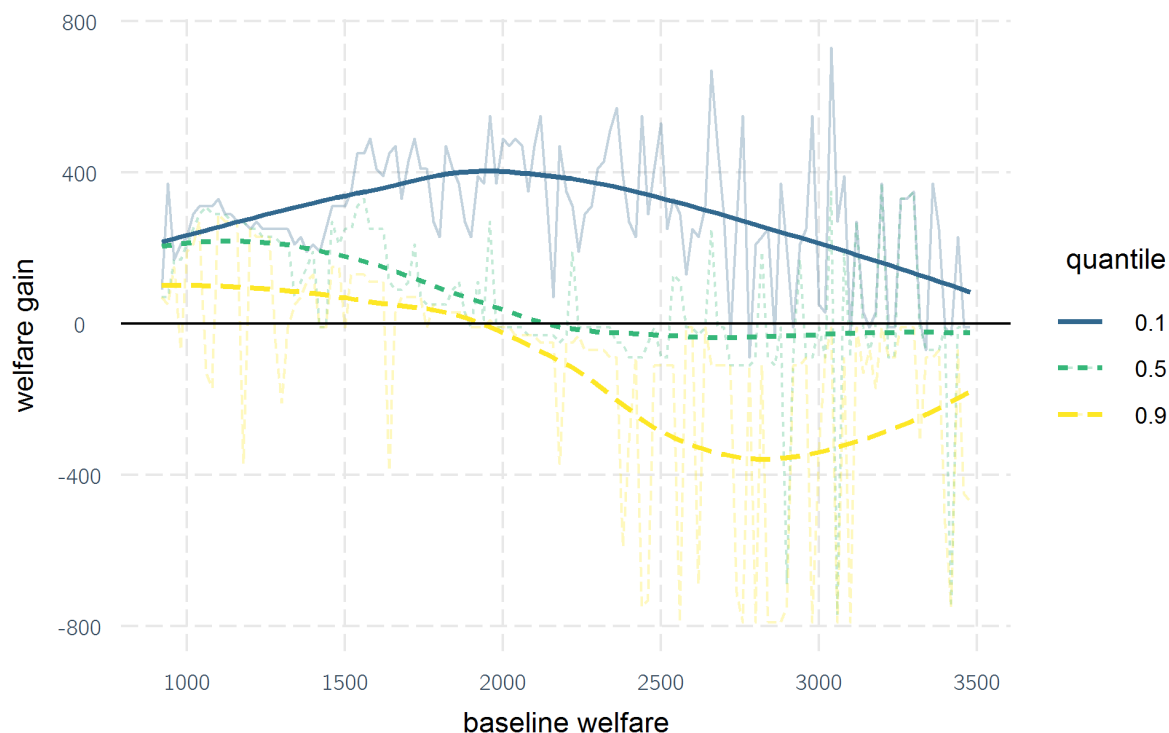

Figure 8: Distribution of welfare gains and losses conditional on baseline welfare

In Table 1, we distribute the population over three, roughly equal, groups of initial welfare levels, and three categories of winners and losers (losers, small gains, and big gains). About 90% of the losers turn out to belong to the initially best-off group. Overall, the losers form a minority of about 15%. Large gains are more prevalent for the initially worst-off third. Moderate gains occur at an equal frequency among the initially worst-off third and the middle group, and occur almost twice as much these groups compared to the initially best-off third. Over 40% of the initially best-off third are losers, while the latter account for only 3% in the middle group, and just 2% in the initially worst-off.

Table 1: Distribution of the winners and losers in terms of baseline welfare (in %)

Baseline welfare (euro)	Welfare gain (euro)			Row sums
	$(-1000, 0]$	$(0, 200]$	$(200, 1000]$	
$(800, 1500]$	0.6	19.1	14.8	34.5
$(1500, 2000]$	0.9	19.1	12.2	32.2
$(2000, 4000]$	13.1	11.2	8.9	31.7
Column sums	14.6	49.4	35.9	

7 Concluding remarks

In this paper, we provided a coherent framework to conduct individual and social welfare analysis for discrete choice. Allowing for unrestricted, unobserved preference heterogeneity, we argue that individual welfare measures become random variables from the point of view of the econometrician. For the broad class of NOS measures, we developed nonparametric methods to retrieve their distributions from observational data. In particular, we proved that all relevant marginal, conditional, and joint distributions can be expressed in terms of choice and transition probabilities, which are nonparametrically point-identified from cross-sectional and panel data, respectively. We also showed how transition probabilities can be set-identified when only cross-sectional data is available, which is important in empirical applications.

To illustrate the empirical usefulness of our results, we revisited the classical trade-off between leisure and consumption, using microdata from the 2018 wave of the German Socio-Economic Panel (SOEP). We modelled single females' labour supply as a choice between three discrete alternatives: non-working, part-time employment, and full-time employment. Using a MMU, we present nonparametric distributional estimates of individual and social welfare for the nonlinear and progressive tax schedule that is currently in place in Germany. In particular, we found that (unobserved) preferences do play an important role in assessing the welfare levels of single females: the welfare distributions of the lowest three wage quartiles are intermingled. Only the distribution of welfare of high-wage individuals tends to first-order dominate that of the other three groups. We also studied the welfare effects of replacing the current schedule with a basic income flat tax. We found that around 15% of the single females would see their welfare diminished; this is especially the case for those that are well-off in the baseline. Large gains were more prevalent among those that were initially poor. Among the bottom two-thirds of the population in terms of baseline welfare, almost 98% would advance as a result of the reform.

There are several avenues for future research. Firstly, one could extend our results to settings where, besides prices, other attributes of the alternatives are changed. In the same strand of thinking, the welfare cost of the introduction, and removal, of some alternatives could be studied. This will likely lead to set-identification, instead of point-identification, of the distributions of interest.[30] Secondly, another methodological innovation could allow for measurement and optimisation errors in the formal analysis. Depending on the specific application, a significant part of the variation in outcomes can be driven by these errors, which might bias welfare estimates. Lastly, future research is needed to assess the sensitivity of empirical welfare estimates, with respect to the choice of the welfare measure and the corresponding reference prices.

---

[30]Bhattacharya (2018) shows that this is indeed the case for the marginal distribution of the CV and EV.

References

Debopam Bhattacharya. Nonparametric Welfare Analysis for Discrete Choice. *Econometrica*, 83(2): 617–649, 2015. ISSN 0012-9682. doi: 10.3982/ECTA12574.

Debopam Bhattacharya. Empirical welfare analysis for discrete choice: Some general results. *Quantitative Economics*, 9(2):571–615, 2018. ISSN 1759-7323. doi: 10.3982/QE931.

Richard Blundell and James L. Powell. Endogeneity in Semiparametric Binary Response Models. *Review of Economic Studies*, 71:655–679, 2004.

Richard Blundell, Dennis Kristensen, and Rosa Matzkin. Bounding quantile demand functions using revealed preference inequalities. *Journal of Econometrics*, 179(2):112–127, April 2014. ISSN 03044076. doi: 10.1016/j.jeconom.2014.01.005.

Richard Blundell, Joel Horowitz, and Matthias Parey. Nonparametric Estimation of a Nonseparable Demand Function under the Slutsky Inequality Restriction. *The Review of Economics and Statistics*, 99(2):291–304, May 2017. ISSN 0034-6535, 1530-9142. doi: 10.1162/REST_a_00636.

Richard A. Briesch, Pradeep K. Chintagunta, and Rosa L. Matzkin. Nonparametric Discrete Choice Models With Unobserved Heterogeneity. *Journal of Business & Economic Statistics*, 28(2):291–307, April 2010. ISSN 0735-0015, 1537-2707. doi: 10.1198/jbes.2009.07219.

Jean François Carpantier and Christelle Sapata. Empirical welfare analysis: when preferences matter. *Social Choice and Welfare*, 46(3):521–542, 2016. ISSN 01761714. doi: 10.1007/s00355-015-0927-y.

Sam Cosaert and Thomas Demuynck. Nonparametric Welfare and Demand Analysis with Unobserved Individual Heterogeneity. *The Review of Economics and Statistics*, 100(2):349–361, May 2018. ISSN 0034-6535, 1530-9142. doi: 10.1162/REST_a_00672.

John K. Dagsvik. Discrete Choice in Continuous Time: Implications of an Intertemporal Version of the IIA Property. *Econometrica*, 70(2):817–831, 2002.

John K. Dagsvik and Anders Karlström. Compensating Variation and Hicksian Choice Probabilities in Random Utility Models That Are Nonlinear in Income. *The Review of Economic Studies*, 72 (1):57–76, 2005.

André de Palma and Karim Kilani. Transition choice probabilities and welfare analysis in additive random utility models. *Economic Theory*, 46(3):427–454, April 2011. ISSN 0938-2259, 1432-0479. doi: 10.1007/s00199-009-0513-6.

Koen Decancq, Marc Fleurbaey, and Erik Schokkaert. Happiness, equivalent incomes and respect for individual preferences. *Economica*, 82:1082–1106, 2015.

André M.J. Decoster and Peter Haan. Empirical welfare analysis with preference heterogeneity. *International Tax and Public Finance*, 22(2):224–251, 2015. ISSN 15736970. doi: 10.1007/s10797-014-9304-5.

Paolo Delle Site and Marco Valerio Salucci. Transition choice probabilities and welfare analysis in random utility models with imperfect before–after correlation. *Transportation Research Part B: Methodological*, 58:215–242, December 2013. ISSN 01912615. doi: 10.1016/j.trb.2013.09.003.

Holger Dette, Stefan Hoderlein, and Natalie Neumeyer. Testing multivariate economic restrictions using quantiles: The example of Slutsky negative semidefiniteness. *Journal of Econometrics*, 191 (1):129–144, March 2016. ISSN 03044076. doi: 10.1016/j.jeconom.2015.07.004.

Peter Diamond and Daniel McFadden. Some uses of the expenditure function in public finance. *Journal of Public Economics*, 3(1):2–21, 1974. ISSN 0047-2727. doi: 10.1016/0047-2727(74) 90020-6.

Avinash Dixit. Welfare effects of tax and price changes. *Journal of Public Economics*, 4(2):103–123, 1975.

Marc Fleurbaey. Social welfare, priority to the worst-off and the dimensions of individual well-being. In Francesco Farina and Ernesto Savaglio, editors, *Inequality and Economic Integration*, pages 236–277. Routledge, 2006.

Marc Fleurbaey. Two criteria for social decisions. *Journal of economic Theory*, 134(1):421–447, 2007.

Marc Fleurbaey. Beyond GDP: The quest for a measure of social welfare. *Journal of Economic Literature*, 47(4):1029–1075, 2009. ISSN 00220515.

Marc Fleurbaey and Didier Blanchet. *Beyond GDP: Measuring welfare and assessing sustainability*. Oxford University Press, 2013.

Marc Fleurbaey and Guillaume Gaulier. International comparisons of living standards by equivalent incomes. *Scandinavian Journal of Economics*, 111(3):597–624, 2009.

Marc Fleurbaey and François Maniquet. Fairness and well-being measurement. *Mathematical Social Sciences*, 90:119–126, 2017. ISSN 01654896. doi: 10.1016/j.mathsocsci.2016.11.003. URL http://dx.doi.org/10.1016/j.mathsocsci.2016.11.003.

Marc Fleurbaey and François Maniquet. Optimal income taxation theory and principles of fariness. *Journal of Economic Literature*, 56(3):1029–1079, 2018. doi: 10.1257/jel.20171238. URL https://doi.org/10.1257/jel.20171238.

Maurice Fréchet. Généralisation du théorème des probabilités totales. *Fundamenta Mathematicae*, 25(1):379–387, 1935.

Peter J. Hammond. Interpersonal comparisons of utility: Why and how they are and should be made. In J Elster and JE Roemer, editors, *Interpersonal Comparisons of Well-being*, pages 200–253. Cambridge University Press, Cambridge, 1991.

Jerry A. Hausman and Whitney K. Newey. Individual Heterogeneity and Average Welfare. *Econometrica*, 84(3):1225–1248, 2016. ISSN 0012-9682. doi: 10.3982/ECTA11899.

Joseph A. Herriges and Catherine L. Kling. Nonlinear Income Effects in Random Utility Models. *Review of Economics and Statistics*, 81(1):62–72, February 1999. ISSN 0034-6535, 1530-9142. doi: 10.1162/003465399767923827.

Stefan Hoderlein and Anne Vanhems. Estimating the distribution of welfare effects using quantiles. *Journal of Applied Econometrics*, 33(1):52–72, January 2018. ISSN 08837252. doi: 10.1002/jae.2587.

Mervyn King. Welfare analysis of tax reforms using household data. *Journal of Public Economics*, 21(2):183–214, 1983.

Roger W. Klein and Richard H. Spady. An Efficient Semiparametric Estimator for Binary Response Models. *Econometrica*, 61(2):387–421, March 1993. ISSN 00129682. doi: 10.2307/2951556.

Arthur Lewbel. Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97(1):145–177, July 2000. ISSN 03044076. doi: 10.1016/S0304-4076(00)00015-4.

Charles F. Manski. Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3:205–228, 1975.

Rosa L. Matzkin. Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models. *Econometrica*, 59(5):1315, September 1991. ISSN 00129682. doi: 10.2307/2938369.

Rosa L. Matzkin. Nonparametric identification and estimation of polychotomous choice models. *Journal of Econometrics*, 58(1-2):137–168, July 1993. ISSN 03044076. doi: 10.1016/0304-4076(93)90116-M.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In P Zarembka, editor, *Frontiers in Econometrics*. Academic Press, New York, NY, 1974.

Daniel McFadden. Econometric models of probabilistic choice. In Charles F Manski and Daniel McFadden, editors, *Structural analysis of discrete data with econometric applications*, pages 198–272. MIT press Cambridge, MA, 1981.

Daniel McFadden. Computing Willingness-to-Pay in Random Utility Models. In James Melvin, Raymond Riezman, and James Moore, editors, *Trade, Theory and Econometrics*, volume 15. Rout-

ledge, August 1999. ISBN 978-0-415-14831-3 978-0-203-07420-6. doi: $10.4324/9780203074206.$ $\text{ch}15.$

Daniel McFadden. Revealed Stochastic Preference: A Synthesis. *Economic Theory*, 26(2):245–264, 2005.

Edward R. Morey, Robert D. Rowe, and Michael Watson. A Repeated Nested-Logit Model of Atlantic Salmon Fishing. *American Journal of Agricultural Economics*, 75(3):578, August 1993. ISSN 00029092. doi: $10.2307/1243565.$

Elisha Pazner. Equity, nonfeasible alternatives and social choice: A reconsideration of the concept of social welfare. In JJ Laffont, editor, *Aggregation and revelation of preferences*. Elsevier, North-Holland, Amsterdam and New York, 1979.

John H Pencavel. Constant-utility index numbers of real wages. *The American Economic Review*, 67(2):91–100, 1977.

Ian Preston and Ian Walker. Welfare measurement in labour supply models with nonlinear budget constraints. *Journal of Population Economics*, 12(3):343–361, 1999.

Kevin Roberts. Price-independent welfare prescriptions. *Journal of Public Economics*, 13(3):277–297, 1980.

Paul A Samuelson. Complementarity: An essay on the 40th anniversary of the hicks-allen revolution in demand theory. *Journal of Economic literature*, 12(4):1255–1289, 1974.

Amartya K Sen. *Collective Choice and Social Welfare*. Holden–Day, San Francisco, 1970.

Kenneth A. Small and Harvey S. Rosen. Applied Welfare Economics with Discrete Choice Models. *Econometrica*, 49(1):105–130, January 1981. ISSN 00129682. doi: $10.2307/1911129.$

Kenneth E Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, 2003.

A Additional results

A.1 Additional NOS measures: the equivalent income metric and the wage metric

Fleurbaey (2006) discusses two additional NOS metrics: the equivalent income (EI) metric and the wage rate (WR) metric.[31] These measures are especially useful in situations where there is a good, such as leisure time or health for example, the consumption of which is bounded above.

A.1.1 Continuous setting

In the two-good continuous setting, we will call this capacity-constrained good as good 1 and the other good, which acts as the numeraire, as good 2.

Equivalent income metric The EI metric measures an individual's well-being by the amount of good 2 necessary to enjoy the full amount of good 1, denoted by T, and to be equally well-off as per her actual situation, denoted by x. Denote this amount by $EI^R(x)$, where R denotes the individual's preferences. The associated opportunity set is then the set of bundles $B_\lambda \equiv \{(g_1, g_2) | g_2 \leq \lambda, g_1 = T\}$.

This is illustrated in Figure 9a for two persons, an individual with preferences R, who more intensely values good 1, and an individual with preferences R', who more intensely values good 2. Suppose both receive the same bundle x. The individual with preferences R, for example, is considered to be worse-off because she gets an equal compensation for not having good 1 at full capacity while she values that good more.

Wage metric Consider the price a person with preferences R and bundle x should earn per unit of good 1, such that she would be able to reach a point that is equally as good as x if offering good 1 (e.g. labour time, in which case the price is a wage, hence the name of the criterion) were her sole source of income. The level of this price or wage rate, denoted by $WR^R(x)$, is the WR metric of well-being. The associated opportunity set is $B_\lambda \equiv \{(c, \ell) | c \leq \lambda (T - \ell)\}$.

Figure 9b illustrates that according to this measure, the individual with preferences R', who values good 2 (e.g. consumption) relatively more than the individual with preferences R, is now considered to be worse-off than the other one, in contrast to the previous case.

A.1.2 Discrete setting

We now describe how these measures can be adjusted to the discrete choice setting and present some distributional results.

---

[31]This last measure was originally developed by Pencavel (1977) and is also mentioned in Preston and Walker (1999).
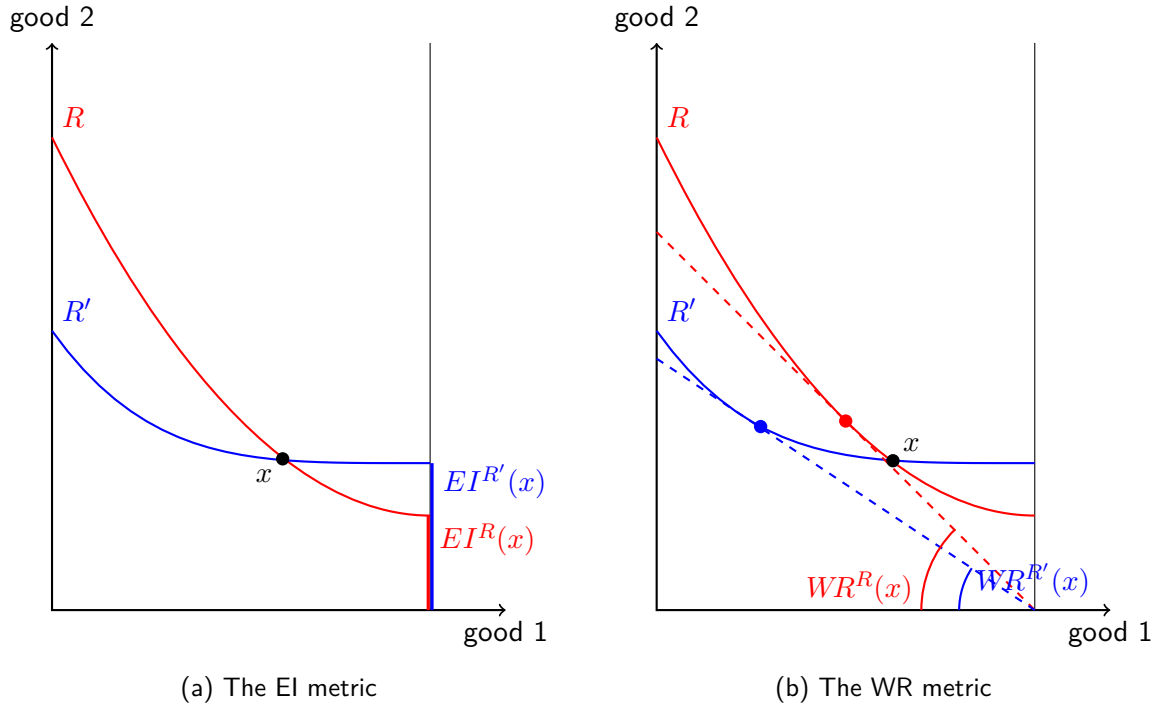
(a) The EI metric

(b) The WR metric

Figure 9: Additional NOS measures

Equivalent income metrics Fix a reference option, which we denote by 0. Well-being of an individual obtaining option k at price p_k is then defined as the amount of the numeraire good necessary to be equally well-off in the reference option 0 as in the point $(y - p_k, k)$. To obtain this, let $\Lambda = \mathbb{R}$, and let $y_c^\lambda = -\infty$ for all λ, when $c \neq 0$, and $y_0^\lambda = \lambda$. The family of opportunity sets can then be defined as $B_\lambda = \left\{ (y', 0) \mid y' \leq y_0^\lambda \right\}$. It is a collection of growing spikes for the 0-option. Equivalently, $\widetilde{p}_c(\lambda) = +\infty$ for all λ, when $c \neq 0$, and $\widetilde{p}_0(\lambda) = y - \lambda$. The EI metric $EI^\omega (y - p_k, k)$ is equal to:

$$
\begin{aligned}
EI^\omega (y - p_k, k) &= \max_\lambda \left\{ \lambda \mid U_k^\omega (y - p_k) \geq \max_c U_c^\omega (y - \widetilde{p}_c(\lambda)) \right\} \\
&= \max_\lambda \left\{ \lambda \mid U_k^\omega (y - p_k) \geq U_0^\omega (\lambda) \right\}.
\end{aligned}
\tag{37}
$$

Given a reference option 0, the equivalent income level of well-being of a type ω individual who obtains option k at price p_k, denoted by $EI^\omega (y - p_k, k)$, is thus implicitly defined by:

$$
U_0^\omega (EI^\omega (y - p_k, k)) = U_k^\omega (y - p_k).
\tag{38}
$$

Corollary 6 When using an EI metric, we have for $k = 0$

$$
\Pr_\omega \left[w \leq EI^\omega (y - p_0, 0), j = J^\omega(\mathbf{p}, y) \right] = P_j(\mathbf{p}, y) \mathbb{I} \left[w \leq y - p_0 \right],
\tag{39}
$$

41

and, hence,

$$\Pr_\omega\left[w \le EI^\omega\left(y - p_0, 0\right)\right] = \Pr_\omega\left[w \le EI^\omega\left(y - p_0, 0\right) \mid j = J^\omega(\mathbf{p}, y)\right] = \mathbb{I}\left[w \le y - p_0\right]. \tag{40}$$

For $k \ne 0$, we have

$$\Pr_\omega\left[w \le EI^\omega\left(y - p_k, k\right) \mid k = J^\omega(\mathbf{p}, y)\right] = \frac{P_k\Big(\big(\min(y - w, p_0), \mathbf{p}_{-0}\big), y\Big)}{P_k(\mathbf{p}, y)} \tag{41}$$

and

$$\Pr_\omega\left[w \le EI^\omega\left(y - p_k, k\right)\right] = \Pr_\omega\left[U_k^\omega(y - p_k) \ge U_0^\omega(w)\right]. \tag{42}$$

Moreover,

$$\begin{aligned}
\Pr_\omega\Big[w \le &EI^\omega\Big(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p}, y)\Big)\Big] \\
&= \sum_{k \ne 0} P_k\Big(\big(\min(y - w, p_0), \mathbf{p}_{-0}\big), y\Big) + P_0(\mathbf{p}, y)\mathbb{I}\left[w \le y - p_0\right] \\
&= \mathbb{I}\left[w \le y - p_0\right] + \mathbb{I}\left[w > y - p_0\right]\sum_{k \ne 0} P_k\Big(\big(y - w, \mathbf{p}_{-0}\big), y\Big)
\end{aligned} \tag{43}$$

The marginal and conditional EI in bundle 0 are deterministic and equal to $y - p_0$. The marginal CDF of EI in bundle k equals the probability of choosing k, when the options are k with price p_k and option 0 with price $y - w$.

Proof. When using an EI metric, we have for general $k \in \mathcal{C}$

$$\begin{aligned}
\Pr_\omega&\left[w \le EI^\omega\left(y - p_k, k\right), j = J^\omega(\mathbf{p}, y)\right] \\
&= \Pr_\omega\left[U_k^\omega(y - p_k) \ge \max_i U_i^\omega(y - \widetilde{p}_i\left(w\right)), U_j^\omega(y - p_j) \ge \max_{c \ne j} U_c^\omega(y - p_c)\right] \\
&= \Pr_\omega\left[U_k^\omega(y - p_k) \ge U_0^\omega(w), U_j^\omega(y - p_j) \ge \max_{c \ne j} U_c^\omega(y - p_c)\right]
\end{aligned} \tag{44}$$

as $\widetilde{p}_i\left(w\right) = +\infty$ for $i \ne 0$. Hence, for $k = 0$, this reduces to

$$\Pr_\omega\left[w \le EI^\omega\left(y - p_0, 0\right), j = J^\omega(\mathbf{p}, y)\right] = P_j(\mathbf{p}, y)\mathbb{I}\left[w \le y - p_0\right].$$

Equations (40) follow immediately.

For $k \neq 0$, (44) implies that

$$
\begin{aligned}
\Pr_\omega \left[w \leq EI^\omega \left(y - p_k, k \right) \mid k = J^\omega(\mathbf{p}, y) \right] &= \frac{\Pr_\omega \left[w \leq EI^\omega \left(y - p_k, k \right), k = J^\omega(\mathbf{p}, y) \right]}{P_k(\mathbf{p}, y)} \\
&= \frac{\Pr_\omega \left[U_k^\omega(y - p_k) \geq \max \left(U_0^\omega(w), \max_{c \neq k} U_c^\omega(y - p_c) \right) \right]}{P_k(\mathbf{p}, y)} \\
&= \frac{P_k \Big(\big(\min(y - w, p_0), \mathbf{p}_{-0} \big), y \Big)}{P_k(\mathbf{p}, y)},
\end{aligned}
$$

and

$$
\begin{aligned}
\Pr_\omega \left[w \leq EI^\omega \left(y - p_k, k \right) \right] &= \sum_j \Pr_\omega \left[w \leq EI^\omega \left(y - p_k, k \right), j = J^\omega(\mathbf{p}, y) \right] \\
&= \Pr_\omega \left[U_k^\omega(y - p_k) \geq U_0^\omega(w) \right].
\end{aligned}
$$

$\square$

Wage metric Let x_c reflect the number of units in option c of a good that is only available in discrete amounts (*e.g.* labour time). Let $\Lambda = \mathbb{R}_+$ and $y_c^\lambda = \lambda x_c$[32] where λ can be seen as the unit price of that good. Equivalently, we have $\widetilde{p}_c(\lambda) = y - \lambda x_c$. The wage metric $WR^\omega \left(y - p_k, k \right)$ equals

$$
WR^\omega \left(y - p_k, k \right) = \max_\lambda \left\{ \lambda \mid U_k^\omega(y - p_k) \geq \max_c U_c^\omega \left(\lambda x_c \right) \right\}. \tag{45}
$$

Let an option c be characterised by the availability of x_c units of a good. The WR level of well-being of a type ω individual who obtains option k at price p_k, denoted by $WR^\omega \left(y - p_k, k \right)$, is defined implicitly by:

$$
U_k^\omega \left(y - p_k \right) = \max_c U_c^\omega \left(WR^\omega \left(y - p_k, k \right) x_c \right). \tag{46}
$$

Corollary 7 When using the WR metric, we have

$$
\Pr_\omega \left[w \leq WR^\omega \left(y - p_k, k \right), j = J^\omega(\mathbf{p}, y) \right] = P_{j,k} \Big(\mathbf{p}, (p_k, y - w\boldsymbol{x}_{-\boldsymbol{k}}), y \Big) \mathbb{I} \left[w \leq (y - p_k)/x_k \right] \tag{47}
$$

---

[32]Strictly, this definition is not compatible with condition (6), as the infimum of y_c^λ is not $-\infty$, but 0, for all c. However, a straightforward adaption to $\mathbb{R}_+$ is possible.

and, hence,

$$\Pr_\omega\left[w \le WR^\omega\left(y - p_k, k\right) \mid j = J^\omega(\mathbf{p}, y)\right] = \frac{P_{j,k}\Big(\mathbf{p}, (p_k, y - w\boldsymbol{x}_{-\boldsymbol{k}}), y\Big)}{P_j(\mathbf{p}, y)}\mathbb{I}\left[w \le (y - p_k)/x_k\right]$$

$$\Pr_\omega\left[w \le WR^\omega\left(y - p_k, k\right) \mid k = J^\omega(\mathbf{p}, y)\right] = \frac{P_k\Big(\mathbf{min}(\mathbf{p}, y - w\boldsymbol{x}), y\Big)}{P_k(\mathbf{p}, y)}\mathbb{I}\left[w \le (y - p_k)/x_k\right]$$

$$\Pr_\omega\left[w \le WR^\omega\left(y - p_k, k\right)\right] = P_k\Big((p_k, y - w\boldsymbol{x}_{-\boldsymbol{k}}), y\Big)\mathbb{I}\left[w \le (y - p_k)/x_k\right]$$

$$\Pr_\omega\left[w \le WR^\omega\left(y - p_{J^\omega(\mathbf{p},y)}, J^\omega(\mathbf{p}, y)\right)\right] = \sum_k P_k\Big(\mathbf{min}(\mathbf{p}, y - w\boldsymbol{x}), y\Big)\mathbb{I}\left[w \le (y - p_k)/x_k\right].$$

Proof. As $\widetilde{p}_c(w) = y - wx_c$, we have

$$\begin{aligned}
\Pr_\omega\left[w \le WR^\omega\left(y - p_k, k\right), j = J^\omega(\mathbf{p}, y)\right] &= P_{j,k}\Big(\mathbf{p}, (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\Big)\mathbb{I}\left[p_k \le \widetilde{p}_k(w)\right] \\
&= P_{j,k}\Big(\mathbf{p}, (p_k, y - w\boldsymbol{x}_{-\boldsymbol{k}}), y\Big)\mathbb{I}\left[p_k \le y - wx_k\right] \\
&= P_{j,k}\Big(\mathbf{p}, (p_k, y - w\boldsymbol{x}_{-\boldsymbol{k}}), y\Big)\mathbb{I}\left[w \le (y - p_k)/x_k\right].
\end{aligned}$$

The conditional and marginal distributions follow immediately.

$\square$

A.2   Results for the equivalent variation

This section derives analogue results to Theorem 3, Corollary 3, and Theorem 4, but now for the EV instead of the CV.

For an individual of type ω, the equivalent variation EV^ω is defined as

$$\max_c\{U_c^\omega(y - p_c - EV^\omega)\} = \max_c\left\{U_c^\omega\left(y - p_c'\right)\right\},$$

i.e, the amount of money (possibly negative) an individual has to pay before the reform to be equally well-off as after the reform.

Theorem 5 For the distribution of the EV, we have the following results:

$$\Pr_\omega[EV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] = P_{i,j}(\mathbf{p}, \mathbf{min}(\mathbf{p} + z, \mathbf{p}'), y)\mathbb{I}\left[p'_j \leq p_j + z\right],$$

$$\Pr_\omega[EV^\omega \leq z \mid i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] = \frac{P_{i,j}(\mathbf{p}, \mathbf{min}(\mathbf{p} + z, \mathbf{p}'), y)}{P_{i,j}(\mathbf{p}, \mathbf{p}', y)}\mathbb{I}\left[p'_j \leq p_j + z\right],$$

$$\Pr_\omega[EV^\omega \leq z \mid i = J^\omega(\mathbf{p}, y)] = \sum_j \frac{P_{i,j}(\mathbf{p}, \mathbf{min}(\mathbf{p} + z, \mathbf{p}'), y)}{P_i(\mathbf{p}, y)}\mathbb{I}\left[p'_j \leq p_j + z\right],$$

$$\Pr_\omega[EV^\omega \leq z \mid j = J^\omega(\mathbf{p}', y)] = \frac{P_j(\mathbf{min}(\mathbf{p} + z, \mathbf{p}'), y)}{P_j(\mathbf{p}', y)}\mathbb{I}\left[p'_j \leq p_j + z\right],$$

$$\Pr_\omega[EV^\omega \leq z] = \sum_j P_j(\mathbf{min}(\mathbf{p} + z, \mathbf{p}'), y)\mathbb{I}\left[p'_j \leq p_j + z\right].$$

Proof. We have that

$$\left\{EV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}$$
$$= \left\{\max_c\{U_c^\omega(y - p_c - EV^\omega)\} \geq \max_c\{U_c^\omega(y - p_c - z)\}, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}$$
$$= \left\{\max_c\{U_c^\omega(y - p'_c)\} \geq \max_c\{U_c^\omega(y - p_c - z)\}, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\},$$

such that,

$$\Pr_\omega[EV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$$
$$= \Pr_\omega\left[U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - p_k), \quad U_j^\omega(y - p'_j) \geq \max_{l \neq j} U_l^\omega(y - p'_l),\right.$$
$$\left.\max_c U_c^\omega(y - p'_c) \geq \max_c\{U_c^\omega(y - p_c - z)\}\right]$$
$$= \Pr_\omega\left[U_j^\omega(y - p'_j) \geq \max_{k \neq j} U_k^\omega(y - \min(p_k + z, p'_k)), \quad U_i^\omega(y - p_i) \geq \max_{l \neq i} U_l^\omega(y - p_l)\right]\mathbb{I}\left[p'_j \leq p_j + z\right]$$
$$= P_{i,j}(\mathbf{p}, \mathbf{min}(\mathbf{p} + z, \mathbf{p}'), y)\mathbb{I}\left[p'_j \leq p_j + z\right].$$

The other equalities follow directly. □

Theorem 6 The joint distribution of the MMU, with initial prices as reference prices, and the EV is expressed as:

$$\Pr_\omega[w \leq MMU_\mathbf{p}^\omega(y - p_i, i), EV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$$
$$= P_{i,j}\Big(\mathbf{p}, \mathbf{min}(\mathbf{p}', \mathbf{p} + z), y\Big)\mathbb{I}\left[p'_j \leq p_j + z\right]\mathbb{I}\left[w \leq y\right]. \tag{48}$$

Proof. We have

$$\Pr_\omega[w \leq MMU_\mathbf{p}^\omega(y - p_i, i), EV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$$

$$= \Pr_\omega \left[U_i^\omega(y - p_i) \geq \max_{c'} U_{c'}^\omega(y - (p_{c'} + y - w)), \quad U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - p_k), \right.$$

$$\left. U_j^\omega(y - p_j') \geq \max_{l \neq j} U_l^\omega(y - p_l'), \quad U_j^\omega(y - p_j') \geq \max_c \{U_c^\omega(y - p_c - z)\} \right]$$

$$= \Pr_\omega \left[U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - \min(p_k, p_k + y - w)), \quad U_j^\omega(y - p_j') \geq \max_{l \neq j} U_l^\omega(y - \min(p_l', p_l + z)) \right]$$

$$\mathbb{I}\left[p_j' \leq p_j + z \right] \mathbb{I}\left[p_i \leq p_i + y - w \right]$$

$$= P_{i,j}\left(\mathbf{p}, (p_j', \min(\mathbf{p}_{-j}', \mathbf{p}_{-j} + z)), y \right) \mathbb{I}\left[p_j' \leq p_j + z \right] \mathbb{I}\left[w \leq y \right]$$

$$= P_{i,j}\left(\mathbf{p}, \min(\mathbf{p}', \mathbf{p} + z), y \right) \mathbb{I}\left[p_j' \leq p_j + z \right] \mathbb{I}\left[w \leq y \right].$$

$\square$

B   Proofs

Theorem 1   The joint distribution of the NOS welfare measure W, evaluated in an option k with price p_k, and choosing j at prices $\mathbf{p}'$ and exogenous income y can be expressed in terms of transition probabilities as follows:

$$\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k), j = J^{\omega}(\mathbf{p}', y)\right] = P_{j,k}\Big(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\Big) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (14)$$

where $(p_k, \widetilde{\mathbf{p}}_{-k}(w)) = (\widetilde{p}_1(w), \ldots, \widetilde{p}_{k-1}(w), p_k, \widetilde{p}_{k+1}(w), \ldots, \widetilde{p}_n(w))$.

Proof of Theorem 1.   Using Lemma 1, we have that

$$\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k), j = J^{\omega}(\mathbf{p}', y)\right]$$

$$= \Pr_{\omega}\left[U_k^{\omega}(y - p_k) \geq \max_c U_c^{\omega}(y - \widetilde{p}_c(w)), U_j^{\omega}(y - p_j') \geq \max_{c' \neq j} U_{c'}^{\omega}(y - p_{c'}')\right]$$

$$= \Pr_{\omega}\left[U_k^{\omega}(y - p_k) \geq \max_{c \neq k} U_c^{\omega}(y - \widetilde{p}_c(w)), U_j^{\omega}(y - p_j') \geq \max_{c' \neq j} U_{c'}^{\omega}(y - p_{c'}')\right] \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]$$

$$= P_{j,k}\Big(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\Big) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right].$$

$\square$

Corollary 1

$$\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k) \mid j = J^{\omega}(\mathbf{p}', y)\right] = \frac{P_{j,k}\Big(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\Big)}{P_j(\mathbf{p}', y)} \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (15)$$

$$\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k) \mid k = J^{\omega}(\mathbf{p}, y)\right] = \frac{P_k\Big(\min(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\Big)}{P_k(\mathbf{p}, y)} \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (16)$$

where $\min(\mathbf{p}, \widetilde{\mathbf{p}}(w)) = (\min(p_1, \widetilde{p}_1(w)), \ldots, \min(p_n, \widetilde{p}_n(w)))$,

$$\Pr_{\omega}\left[w \leq W^{\omega}(y - p_k, k)\right] = P_k\Big((p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\Big) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right], \qquad (17)$$

and

$$\Pr_{\omega}\left[w \leq W^{\omega}\Big(y - p_{J^{\omega}(\mathbf{p}, y)}, J^{\omega}(\mathbf{p}, y)\Big)\right] = \sum_k P_k\Big(\min(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\Big) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]. \qquad (18)$$

Proof of Corollary 1. (a) For the conditional distribution on option j being chosen, we have

$$\Pr_\omega\left[w \leq W^\omega(y - p_k, k) \mid j = J^\omega(\mathbf{p}', y)\right] = \frac{\Pr_\omega\left[w \leq W^\omega(y - p_k, k), j = J^\omega(\mathbf{p}', y)\right]}{P_j(\mathbf{p}', y)}$$

$$= \frac{P_{j,k}\left(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\right)}{P_j(\mathbf{p}', y)} \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right].$$

(b) For the conditional distribution on option k being chosen, we have

$$\Pr_\omega\left[w \leq W^\omega(y - p_k, k) \mid k = J^\omega(\mathbf{p}, y)\right] = \frac{\Pr_\omega\left[w \leq W^\omega(y - p_k, k), k = J^\omega(\mathbf{p}, y)\right]}{P_k(\mathbf{p}, y)}$$

$$= \frac{P_{k,k}\left(\mathbf{p}, (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\right) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]}{P_k(\mathbf{p}, y)}$$

$$= \left(\Pr_\omega\left[U_k^\omega(y - p_k) \geq \max_{c \neq k} U_c^\omega(y - p_c),\right.\right.$$

$$\left.\left. U_k^\omega(y - p_k) \geq \max_{c \neq k} U_c^\omega(y - \widetilde{p}_c(w)\right] \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]\right) / P_k(\mathbf{p}, y)$$

$$= \left(\Pr_\omega\left[U_k^\omega(y - p_k) \geq \max_{c \neq k} U_c^\omega(y - \min(p_c, \widetilde{p}_c(w)))\right]\right.$$

$$\left. \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]\right) / P_k(\mathbf{p}, y)$$

$$= \frac{P_k\left(\min(\mathbf{p}, (p_k, \widetilde{\mathbf{p}}_{-k}(w))), y\right) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]}{P_k(\mathbf{p}, y)}$$

$$= \frac{P_k\left(\min(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\right)}{P_k(\mathbf{p}, y)} \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right].$$

(c) For the marginal distribution in option k, we have

$$\Pr_\omega\left[w \leq W^\omega(y - p_k, k)\right] = \sum_j \Pr_\omega\left[w \leq W^\omega(y - p_k, k), j = J^\omega(\mathbf{p}', y)\right]$$

$$= \sum_j P_{j,k}\left(\mathbf{p}', (p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\right) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]$$

$$= P_k\left((p_k, \widetilde{\mathbf{p}}_{-k}(w)), y\right) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right].$$

(d) Finally, for the marginal distribution, in the optimal option, we have

$$\Pr_\omega\left[w \leq W^\omega\left(y - p_{J^\omega(\mathbf{p}, y)}, J^\omega(\mathbf{p}, y)\right)\right] = \sum_k \Pr_\omega\left[w \leq W^\omega(y - p_k, k), k = J^\omega(\mathbf{p}, y)\right]$$

$$= \sum_k P_k\left(\min(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\right) \mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right].$$

$\square$

Corollary 2 When using reference prices $\mathbf{p}^{ref}$, we have

$$\Pr_{\omega}\left[w \le MMU_{\mathbf{p}^{ref}}^{\omega}\left(y - p_k, k\right), j = J^{\omega}(\mathbf{p}', y)\right] = P_{j,k}(\mathbf{p}', (p_k, y - w + \mathbf{p}_{-k}^{ref}), y)\mathbb{I}\left[p_k \le y - w + p_k^{ref}\right].$$

When $p_k = p_k'$, and the reference prices equal the actual prices $\mathbf{p}'$ and k is the optimal choice, this simplifies to

$$\Pr_{\omega}\left[w \le MMU_{\mathbf{p}'}^{\omega}\left(y - p_k', k\right), k = J^{\omega}(\mathbf{p}', y)\right] = P_k(\mathbf{p}', y)\mathbb{I}\left[w \le y\right]$$

and, hence,

$$\Pr_{\omega}\left[w \le MMU_{\mathbf{p}'}^{\omega}\left(y - p_k', k\right) \mid k = J^{\omega}(\mathbf{p}', y)\right] = \mathbb{I}\left[w \le y\right],$$
$$\Pr_{\omega}\left[w \le MMU_{\mathbf{p}'}^{\omega}\left(y - p_{J^{\omega}(\mathbf{p}', y)}', J^{\omega}(\mathbf{p}', y)\right)\right] = \mathbb{I}\left[w \le y\right].$$

Proof of Corollary 2. The first equation follows from plugging $\widetilde{\mathbf{p}}(w) = y - w + \mathbf{p}^{ref}(w)$ into Equation (14). Moreover, using actual prices $\mathbf{p}'$ as reference prices and taking $p_k = p_k'$, $\mathbb{I}\left[p_k \le y - w + p_k^{ref}\right]$ implies that $w \le y$. Therefore,

$$\Pr_{\omega}\left[w \le MMU_{\mathbf{p}'}^{\omega}\left(y - p_k', k\right), k = J^{\omega}(\mathbf{p}', y)\right] = P_k\Big(\min(\mathbf{p}', y - w + \mathbf{p}'), y\Big)\mathbb{I}\left[w \le y\right)]$$
$$= P_k(\mathbf{p}', y)\mathbb{I}\left[w \le y\right)].$$

(49)

The last two equations then immediately follow from Bayes' theorem and summing over k. □

Proposition 1 The joint distribution of welfare in the optimal bundle i, before a price change, and welfare in the optimal bundle j, after the price change, is as follows:

$$\Pr_{\omega}[w \le W_0^{\omega}(y - p_i, i), z \le W_1^{\omega}(y - p_j', j), i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)]$$
$$= P_{i,j}\Big(\min(\mathbf{p}, \widetilde{\mathbf{p}}(w)), \min(\mathbf{p}', \widetilde{\mathbf{p}}(z)), y\Big)\mathbb{I}\left[p_i \le \widetilde{p}_i(w)\right]\mathbb{I}\left[p_j' \le \widetilde{p}_j(z)\right].$$

(19)

Proof of Proposition 1.

$$\Pr_\omega[w \leq W_0^\omega(y - p_i, i), z \leq W_1^\omega(y - p_j', j), i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$$

$$= \Pr_\omega\left[U_i^\omega(y - p_i) \geq \max_{c'} U_{c'}^\omega(y - \widetilde{p}_{c'}(w)), \quad U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - p_k), \right.$$

$$\left. U_j^\omega(y - p_j') \geq \max_{l \neq j} U_l^\omega(y - p_l'), \quad U_j^\omega(y - p_j') \geq \max_c U_c^\omega(y - \widetilde{p}_c(z)), \right]$$

$$= \Pr_\omega\left[U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - \min(p_k, \widetilde{p}_k(w))), \quad U_j^\omega(y - p_j') \geq \max_{l \neq j} U_l^\omega(y - \min(p_l', \widetilde{p}_l(z))) \right]$$

$$\mathbb{I}\left[p_i \leq \widetilde{p}_i(w)\right] \mathbb{I}\left[p_j' \leq \widetilde{p}_j(z)\right]$$

$$= P_{i,j}\Big((p_i, \mathbf{min}(\mathbf{p}_{-i}, \widetilde{\mathbf{p}}_{-i}(w))), (p_j', \mathbf{min}(\mathbf{p}_{-j}', \widetilde{\mathbf{p}}_{-j}(z)), y \Big) \mathbb{I}\left[p_i \leq \widetilde{p}_i(w)\right] \mathbb{I}\left[p_j' \leq \widetilde{p}_j(z)\right]$$

$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), \mathbf{min}(\mathbf{p}', \widetilde{\mathbf{p}}(z)), y \Big) \mathbb{I}\left[p_i \leq \widetilde{p}_i(w)\right] \mathbb{I}\left[p_j' \leq \widetilde{p}_j(z)\right].$$

$\square$

Theorem 2 The function h is defined by:

$$h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, s) = P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(\max(w, x))), \mathbf{min}(\mathbf{p}', \widetilde{\mathbf{p}}(s)), y \Big) \mathbb{I}\left[p_j' \leq \widetilde{p}_j(s)\right]$$

$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w), \widetilde{\mathbf{p}}(x)), \mathbf{min}(\mathbf{p}', \widetilde{\mathbf{p}}(s)), y \Big) \mathbb{I}\left[p_j' \leq \widetilde{p}_j(s)\right].$$

Then, the joint distribution of the stochastic welfare measure and the difference before and after the price change of this measure becomes,

$$\Pr_\omega[w \leq W_0^\omega(y - p_i, i), W_1^\omega(y - p_j', j) - W_0^\omega(y - p_i, i) \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] =$$

$$- \int_{-\infty}^{+\infty} \partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, x + z) \mathbb{I}\left[p_i \leq \min(\widetilde{p}_i(w), \widetilde{p}_i(x))\right] \, dx. \tag{20}$$

Proof of Theorem 2. Fix i and j and define $g(w, z) = \Pr_\omega[w \leq W_0^\omega(y - p_i, i), z \leq W_1^\omega(y - p_j', j), i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$. Then we have

$$\Pr_\omega[w \leq W_0^\omega(y - p_i, i), W_1^\omega(y - p_j', j) - W_0^\omega(y - p_i, i) \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]$$

$$= - \int_{-\infty}^{+\infty} \partial_2 g\left(\max(w, x), x + z\right) \, dx$$

$$= - \int_{-\infty}^{+\infty} \partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, x + z) \mathbb{I}\left[p_i \leq \widetilde{p}_i(\max(w, x))\right] \, dx$$

$$= - \int_{-\infty}^{+\infty} \partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, x + z) \mathbb{I}\left[p_i \leq \min(\widetilde{p}_i(w), \widetilde{p}_i(x))\right] \, dx.$$

$\square$

Lemma 2 We have

$$
\left\{\omega \mid CV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}
$$
$$
= \left\{\omega \mid U_i^\omega(y - p_i) \geq \max_c\{U_c^\omega(y - p_c' - z)\}, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}. \tag{22}
$$

Proof of Lemma 2.

$$
\left\{\omega \mid CV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}
$$
$$
= \left\{\omega \mid \max_c\{U_c^\omega(y - p_c' - CV^\omega)\} \geq \max_c\{U_c^\omega(y - p_c' - z)\}, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}
$$
$$
= \left\{\omega \mid \max_c\{U_c^\omega(y - p_c)\} \geq \max_c\{U_c^\omega(y - p_c' - z)\}, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\}
$$
$$
= \left\{\omega \mid U_i^\omega(y - p_i) \geq \max_c\{U_c^\omega(y - p_c' - z)\}, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right\},
$$

where the second equality follows from (21) and the last from $i = J^\omega(\mathbf{p}, y)$. $\square$

Theorem 3 The joint distribution of the CV and the optimal choices before and after the price change is as follows:

$$
\Pr_\omega[CV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] = P_{i,j}(\min(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)\mathbb{I}\left[p_i \leq p_i' + z\right]. \tag{23}
$$

Proof of Theorem 3. We have

$$
\Pr_\omega[CV^\omega \leq z, i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)]
$$
$$
= \Pr_\omega\left[U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - p_k), \quad U_j^\omega(y - p_j') \geq \max_{l \neq j} U_l^\omega(y - p_l'),\right.
$$
$$
\left. U_i^\omega(y - p_i) \geq \max_c U_c^\omega(y - p_c' - z)\right]
$$
$$
= \Pr_\omega\left[U_i^\omega(y - p_i) \geq \max_{k \neq i} U_k^\omega(y - \min(p_k, p_k' + z)), \quad U_j^\omega(y - p_j') \geq \max_{l \neq j} U_l^\omega(y - p_l')\right]
$$
$$
\mathbb{I}\left[p_i \leq p_i' + z\right]
$$
$$
= P_{i,j}(\min(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)\mathbb{I}\left[p_i \leq p_i' + z\right].
$$

$\square$

Corollary 3

$$
\Pr_\omega[CV^\omega \leq z \mid i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)] = \frac{P_{i,j}(\min(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)}{P_{i,j}(\mathbf{p}, \mathbf{p}', y)}\mathbb{I}\left[p_i \leq p_i' + z\right], \tag{24}
$$

$$
\Pr_\omega[CV^\omega \leq z \mid i = J^\omega(\mathbf{p}, y)] = \frac{P_i(\min(\mathbf{p}, \mathbf{p}' + z), y)}{P_i(\mathbf{p}, y)}\mathbb{I}\left[p_i \leq p_i' + z\right], \tag{25}
$$

$$\Pr_{\omega}[CV^{\omega} \leq z \mid j = J^{\omega}(\mathbf{p}', y)] = \sum_{i} \frac{P_{i,j}(\mathbf{min}(\mathbf{p}, \mathbf{p}' + z), \mathbf{p}', y)}{P_j(\mathbf{p}', y)} \mathbb{I}\left[p_i \leq p_i' + z\right], \tag{26}$$

and

$$\Pr_{\omega}[CV^{\omega} \leq z] = \sum_{i} P_i(\mathbf{min}(\mathbf{p}, \mathbf{p}' + z), y)\mathbb{I}\left[p_i \leq p_i' + z\right].^{33} \tag{27}$$

Theorem 4 The joint distribution of the MMU and the CV is as follows:

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}(y - p_i, i), CV^{\omega} \leq z, i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)]$$
$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), \mathbf{p}', y\Big)\mathbb{I}\left[p_i \leq p_i' + \min(z, y - w)\right]. \tag{28}$$

Proof of Theorem 4. **A direct proof of Theorem 4**

We have

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}(y - p_i, i), CV^{\omega} \leq z, i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)]$$
$$= \Pr_{\omega}\Big[U_i^{\omega}(y - p_i) \geq \max_{c'} U_{c'}^{\omega}(y - (y - w + p_{c'})), \quad U_i^{\omega}(y - p_i) \geq \max_{k \neq i} U_k^{\omega}(y - p_k),$$
$$U_j^{\omega}(y - p_j') \geq \max_{l \neq j} U_l^{\omega}(y - p_l'), \quad U_i^{\omega}(y - p_i) \geq \max_{c} U_c^{\omega}(y - p_c' - z)\Big]$$
$$= \Pr_{\omega}\Big[U_i^{\omega}(y - p_i) \geq \max_{k \neq i} U_k^{\omega}(y - \min(p_k, p_k' + y - w, p_k' + z)), \quad U_j^{\omega}(y - p_j') \geq \max_{l \neq j} U_l^{\omega}(y - p_l')\Big]$$
$$\mathbb{I}\left[p_i \leq p_i' + z\right]\mathbb{I}\left[p_i \leq p_i' + y - w\right]$$
$$= P_{i,j}\Big((p_i, \mathbf{min}(\mathbf{p}_{-i}, \mathbf{p}_{-i}' + \min(z, y - w))), \mathbf{p}', y\Big)\mathbb{I}\left[p_i \leq p_i' + \min(z, y - w)\right]$$
$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), \mathbf{p}', y\Big)\mathbb{I}\left[p_i \leq p_i' + \min(z, y - w)\right].$$

Theorem 4 as implied by Theorem 2

When choosing the MMU with the final prices as reference prices, Theorem 2 implies:

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}(y - p_i, i), CV^{\omega} \leq z, i = J^{\omega}(\mathbf{p}, y), j = J^{\omega}(\mathbf{p}', y)]$$
$$= -\int_{-\infty}^{+\infty} \partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, x + z)\mathbb{I}\left[p_i \leq \min(p_i' + y - w, p_i' + y - x)\right] dx. \tag{50}$$

where the function h is defined by

$$h_{i,j,\mathbf{p},\mathbf{p}'}(w, x, s) = P_{i,j}\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + y - \max(w, x)), \mathbf{min}(\mathbf{p}', \mathbf{p}' + y - s), y\Big)\mathbb{I}\left[p_j' \leq p_j' + y - s\right]. \tag{51}$$

---

[33]Note that Equation (27) is the main result of Bhattacharya (2015).

Rewriting, (51), we obtain

$$h_{i,j,\mathbf{p},\mathbf{p}'}(w,x,s) = P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+y-\max(w,x)),\mathbf{min}(\mathbf{p}',\mathbf{p}'+y-s),y\Big)\mathbb{I}\left[p_j' \leq p_j'+y-s\right]$$

$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+y-\max(w,x)),\mathbf{p}',y\Big)\mathbb{I}\left[s \leq y\right],$$

and hence

$$\partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w,x,x+z) = -P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+y-\max(w,x)),\mathbf{p}',y\Big)\delta(x+z-y),$$

where δ is a Dirac delta function. Plugging this in in (50), we obtain

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}(y-p_i,i), CV^{\omega} \leq z, i = J^{\omega}(\mathbf{p},y), j = J^{\omega}(\mathbf{p}',y)]$$

$$= -\int_{-\infty}^{+\infty} \partial_3 h_{i,j,\mathbf{p},\mathbf{p}'}(w,x,x+z)\mathbb{I}\left[p_i \leq \min(p_i'+y-w, p_i'+y-x)\right]\,dx$$

$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+y-\max(w,y-z)),\mathbf{p}',y\Big)\mathbb{I}\left[p_i \leq \min(p_i'+y-w, p_i'+y-(y-z))\right]$$

$$= P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+\min(y-w,z)),\mathbf{p}',y\Big)\mathbb{I}\left[p_i \leq p_i'+\min(y-w,z)\right]$$

as in Theorem 4. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Corollary 4

$$\Pr_{\omega}\left[w \leq MMU_{\mathbf{p}'}^{\omega}, CV^{\omega} \leq z \mid i = J^{\omega}(\mathbf{p},y), j = J^{\omega}(\mathbf{p}',y)\right]$$

$$= \frac{P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+\min(z,y-w)),\mathbf{p}',y\Big)}{P_{i,j}(\mathbf{p},\mathbf{p}',y)}\mathbb{I}\left[p_i \leq p_i'+\min(z,y-w)\right], \quad (29)$$

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}(y-p_i,i), CV^{\omega} \leq z \mid i = J^{\omega}(\mathbf{p},y)]$$

$$= \frac{P_i\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+\min(z,y-w)),y\Big)}{P_i(\mathbf{p},y)}\mathbb{I}\left[p_i \leq p_i'+\min(z,y-w)\right], \quad (30)$$

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}\Big(y-p_{J^{\omega}(\mathbf{p},y)}, J^{\omega}(\mathbf{p},y)\Big), CV^{\omega} \leq z \mid j = J^{\omega}(\mathbf{p}',y)]$$

$$= \sum_i \frac{P_{i,j}\Big(\mathbf{min}(\mathbf{p},\mathbf{p}'+\min(z,y-w)),\mathbf{p}',y\Big)}{P_j(\mathbf{p}',y)}\mathbb{I}\left[p_i \leq p_i'+\min(z,y-w)\right], \quad (31)$$

and,

$$\Pr_{\omega}[w \leq MMU_{\mathbf{p}'}^{\omega}\Big(y - p_{J^{\omega}(\mathbf{p},y)}, J^{\omega}(\mathbf{p}, y)\Big), CV^{\omega} \leq z]$$

$$= \sum_{i} P_i\Big(\mathbf{min}(\mathbf{p}, \mathbf{p}' + \min(z, y - w)), y\Big)\mathbb{I}\left[p_i \leq p_i' + \min(z, y - w)\right]. \quad (32)$$

Proposition 2 The conditional CDF of individual welfare in the optimal bundle can be calculated using choice probabilities:

$$F_W(w \mid \mathbf{p}, y) = 1 - \sum_{k} P_k\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\Big)\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right]. \quad (35)$$

Proof of Proposition 2.

$$\begin{aligned}
F_W(w \mid \mathbf{p}, y) &= \Pr_{\omega}\Big[W^{\omega}\Big(y - p_{J^{\omega}(\mathbf{p},y)}, J^{\omega}(\mathbf{p}, y)\Big) \leq w\Big] \\
&= 1 - \Pr_{\omega}\Big[w \leq W^{\omega}\Big(y - p_{J^{\omega}(\mathbf{p},y)}, J^{\omega}(\mathbf{p}, y)\Big)\Big] \\
&= 1 - \sum_{k} P_k\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\Big)\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right],
\end{aligned} \quad (52)$$

where the last equality follows from Equation (18) in Corollary 1. □

Corollary 5 When prices are equal for everyone and when one uses the MMU with reference prices equal to those common prices as the welfare measure, the SWF can be written solely in terms of income.

Proof of Corollary 5. From Proposition 2 and the definition of the virtual prices in case of an MMU with actual prices $\mathbf{p}$ as reference prices ($\widetilde{\mathbf{p}}(w) = y - w + \mathbf{p}$), it follows that

$$\begin{aligned}
F_W(w \mid \mathbf{p}, y) &= 1 - \sum_{k} P_k\Big(\mathbf{min}(\mathbf{p}, \widetilde{\mathbf{p}}(w)), y\Big)\mathbb{I}\left[p_k \leq \widetilde{p}_k(w)\right] \\
&= 1 - \sum_{k} P_k\Big(\mathbf{min}(\mathbf{p}, y - w + \mathbf{p}), y\Big)\mathbb{I}\left[p_k \leq y - w + p_k\right] \\
&= 1 - \sum_{k} P_k\Big(\mathbf{p}, y\Big)\mathbb{I}\left[w \leq y\right] \\
&= \mathbb{I}\left[y \leq w\right].
\end{aligned}$$

Hence,

$$SWF = \int \int h(w) \, dF_W(w \mid \mathbf{p}, y) \, dG(\mathbf{p}, y)$$
$$= \int \int h(w) d\mathbb{I}\left[y \le w\right] \, dG(\mathbf{p}, y)$$
$$= \int h(y) \, dG(\mathbf{p}, y).$$

Notice that $\mathbf{p}$ in the argument of G is redundant, as prices are assumed to be identical for all persons in this case. This completes the proof. $\square$

Proposition 3 Suppose Assumptions $1 - 3$ hold. Then the transition probabilities $\{P_{i,j}(\mathbf{p}, \mathbf{p}', y)\}$ are set identified from the choice probabilities $\{P_i\}$ with bounds

$$P_{i,i}^L(\mathbf{p}, \mathbf{p}', y) = \max\left\{P_i(\mathbf{p}, y) + P_i(\mathbf{p}', y) - 1, P_i\left(\left(\max\{p_i, p_i'\}, \mathbf{min}\{\mathbf{p}_{-i}, \mathbf{p}'_{-i}\}\right), y\right)\right\},$$
$$P_{i,i}^U(\mathbf{p}, \mathbf{p}', y) = \min\left\{P_i(\mathbf{p}, y), \ P_i(\mathbf{p}', y)\right\}.$$

For $i \ne j$, $P_{i,j}(\mathbf{p}, \mathbf{p}', y) = 0$ if $p_i \ge p_i'$ and $p_j \le p_j'$ and

$$P_{i,j}^L(\mathbf{p}, \mathbf{p}', y) = \max\{P_i(\mathbf{p}, y) + P_j(\mathbf{p}', y) - 1, \ 0\},$$
$$P_{i,j}^U(\mathbf{p}, \mathbf{p}', y) = \min\left\{P_i(\mathbf{p}, y), P_j(\mathbf{p}', y)\right\},$$

elsewhere.

Proof of Proposition 3. We will first derive upper and lower bounds that are implied by elementary probability theory. Denoting by A the set $\{\omega | i = J^\omega(\mathbf{p}, y)\}$ and B the set $\{\omega | j = J^\omega(\mathbf{p}', y)\}$, we have $P(A \cap B) = P_{i,j}(\mathbf{p}, \mathbf{p}'; y)$, $P(A) = P_i(\mathbf{p}; y)$ and $P(B) = P_j(\mathbf{p}'; y)$.

For the lower bound, note that

$$1 \ge P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{53}$$

and hence $P(A \cap B) \ge P(A) + P(B) - 1$ which translates into

$$P_{i,j}(\mathbf{p}, \mathbf{p}'; \ y) \ge P_i(\mathbf{p}; \ y) + P_j(\mathbf{p}'; \ y) - 1. \tag{54}$$

For the upper bound, note that $P(A \cap B) \le P(A)$ and $P(A \cap B) \le P(B)$ and hence

$$P_{i,j}(\mathbf{p}, \mathbf{p}'; \ y) \le \min(P_i(\mathbf{p}; \ y), P_j(\mathbf{p}'; \ y)) \tag{55}$$

We will now exploit the monotonicity condition imposed on the utility function U_c^ω to construct tighter bounds based on revealed preference restrictions. First consider the no-transition case.

Note therefore that if

$$U_i^\omega(y - \max\{p_i, p_i'\}) > U_k^\omega(y - \min\{p_k, p_k'\}),\tag{56}$$

then $U_i^\omega(y - p_i) > U_k^\omega(y - p_k)$ and $U_i^\omega(y - p_i') > U_k^\omega(y - p_k')$ and hence

$$P_i((\max\{p_i, p_i'\}, \mathbf{min}\{\mathbf{p}_{-i}, \mathbf{p}_{-i}'\}; y) = \Pr_\omega\left[\cap_{k \neq i}\{U_i^\omega(y - \max\{p_i, p_i'\}) > U_k^\omega(y - \min\{p_k, p_k'\})\}\right]\tag{57}$$

is a lower bound of $P_{i,i}(\mathbf{p}, \mathbf{p}'; y)$.

Finally, for the transition case, some transitions are ruled out by monotonicity. Indeed, if $p_i \geq p_i'$ and $p_j \leq p_j'$, good i becomes weakly less and good j weakly more expensive after the price change. By monotonicity, it holds that $U_i^\omega(y - p_i) \leq U_i^\omega(y - p_i')$ and $U_j^\omega(y - p_j) \geq U_j^\omega(y - p_j')$ and hence, if moreover $U_i^\omega(y - p_i) > U_k^\omega(y - p_k)$ for all $k \neq i$ and $U_j^\omega(y - p_j') > U_k^\omega(y - p_k')$ for all $k \neq j$, then

$$U_i^\omega(y - p_i') \geq U_i^\omega(y - p_i) > U_j^\omega(y - p_j) > U_i^\omega(y - p_i')\tag{58}$$

which is a contradiction. Hence, if $p_i \geq p_i'$ and $p_j \leq p_j'$, then $P_{i,j}(\mathbf{p}, \mathbf{p}', y) = 0$. $\qquad\square$

Proposition 4 Suppose Assumptions 1-3 hold. Then the estimated choice probabilities $\{P_i\}$ should satisfy the following shape restrictions at all $(\mathbf{p}, y)$ for all $i, j \in \mathcal{C}$:

$$\frac{\partial P_i(\mathbf{p}, y)}{\partial p_i} \leq 0,$$

$$\frac{\partial P_i(\mathbf{p}, y)}{\partial p_j} \geq 0,$$

Proof of Proposition 4. Remember that the choice probabilities are defined as

$$P_i(\mathbf{p}, y) = \Pr_\omega\left[\left\{U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\}\right\}\right].$$

Given Assumption 1, the first restriction immediately follows from observing that an increase in p_i would shrink the set

$$\{\omega \mid U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\}$$

and, hence, lowers the probability

$$\Pr_\omega\left[\left\{U_i^\omega(y - p_i) \geq \max_{c \neq i}\{U_c^\omega(y - p_c)\}\right\}\right].$$

Analogously, an increase in p_j with $j \neq i$ would increase $P_i(\mathbf{p}, y)$. $\qquad\square$

C Empirical illustration

C.1 Sample description

Table 2: Descriptive statistics estimation subsample

Variable	N	Min	Q₁	Median	Mean	Q₃	Max	SD	IQR	#NA
Weekly hours worked	1922	0.0	0.0	30.0	25.2	40.0	80.0	18.3	40.0	0
Hourly gross wage	1492	4.1	9.5	14.0	14.9	19.0	39.7	6.9	9.5	430
Yearly income from assets	1922	-5.0	0.0	3.0	260.8	131.5	11998.0	1085.5	131.5	0
Age	1922	18.0	34.0	45.0	43.3	53.0	60.0	11.1	19.0	0
Years education	1922	7.0	10.5	11.5	12.2	14.0	18.0	2.8	3.5	0
Number of children (all)	1922	0.0	0.0	0.0	0.5	1.0	3.0	0.8	1.0	0
Number of children (0-1)	1922	0.0	0.0	0.0	0.0	0.0	2.0	0.1	0.0	0
Number of children (2-4)	1922	0.0	0.0	0.0	0.0	0.0	2.0	0.2	0.0	0
Number of children (5-7)	1922	0.0	0.0	0.0	0.1	0.0	2.0	0.3	0.0	0
East Germany	1922	0.0	0.0	0.0	0.2	0.0	1.0	0.4	0.0	0

C.2 Estimation procedure

Choice probabilities To model the choice probabilities for alternatives PT and FT, we specify for each a flexible binary logit model that contains cubic polynomials for the disposable income of all three alternatives, and a linear index with demographic variables, such as individual ℓ's age, years of education, number of children, and region. Formally, we have for $c \in \{PT, FT\}$ that

$$\Pr_{\omega}[c = J^{\omega}(\mathbf{d}_\ell) \mid \mathbf{x}_\ell] := \Lambda_c(\mathbf{d}_\ell, \mathbf{x}_\ell; \boldsymbol{\theta}_c) = \left(1 + \exp\left(-\left(\alpha_c + \sum_{c'}\sum_{m=1}^{3}\beta_{c,c',m}d_{\ell,c'}^m + \mathbf{x}_\ell'\boldsymbol{\gamma}_c\right)\right)\right)^{-1},$$

where $\mathbf{d}_\ell := (d_{NW,\ell}, d_{PT,\ell}, d_{FT,\ell})$ is a vector with disposable incomes, $\mathbf{x}_\ell$ a vector with demographic characteristics, and $\boldsymbol{\theta}_c = (\alpha_c, \boldsymbol{\beta}_c, \boldsymbol{\gamma}_c)$ a vector with parameters. Alternative NW is defined as the complement of these two probabilities,

$$\Pr_{\omega}[NW = J^{\omega}(\mathbf{d}_\ell) \mid \mathbf{x}_\ell] := 1 - \Pr_{\omega}[PT = J^{\omega}(\mathbf{d}_\ell) \mid \mathbf{x}_\ell] - \Pr_{\omega}[FT = J^{\omega}(\mathbf{d}_\ell) \mid \mathbf{x}_\ell]$$

$$= 1 - \Lambda_{PT}(\mathbf{d}_\ell, \mathbf{x}_\ell; \boldsymbol{\theta}_{PT}) - \Lambda_{FT}(\mathbf{d}_\ell, \mathbf{x}_\ell; \boldsymbol{\theta}_{FT}),$$

which ensures that for every pair $(\mathbf{d}_\ell, \mathbf{x}_\ell)$, the choice probabilities add up to one. The model is then estimated by nonlinear least squares

$$\widehat{\boldsymbol{\theta}}^{NLS} = \arg\min_{(\boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT})} \sum_{\ell} \Big[[Y_{NW,\ell} - (1 - \Lambda_{PT}(\mathbf{d}_\ell, \mathbf{x}_\ell; \boldsymbol{\theta}_{PT}) - \Lambda_{FT}(\mathbf{d}_\ell, \mathbf{x}_\ell; \boldsymbol{\theta}_{FT}))]^2$$

$$+ [Y_{PT,\ell} - \Lambda_{PT}(\mathbf{d}_\ell, \mathbf{x}_\ell); \boldsymbol{\theta}_{PT})]^2 + [Y_{FT,\ell} - \Lambda_{FT}(\mathbf{d}_\ell, \mathbf{x}_\ell); \boldsymbol{\theta}_{FT})]^2 + \pi(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) \Big],$$

where $\{Y_{c,\ell}, c \in \{NW, PT, FT\}\}$ are dummy variables that encode individual ℓ's observed choice. The last term, i.e. $\pi(\mathbf{x}_\ell; \boldsymbol{\theta}_{HT}, \boldsymbol{\theta}_{FT})$, contains a positive-valued penalty function that imposes non-negativity of $\Pr_{\omega}[NW = J^{\omega}(\mathbf{d}_\ell) \mid \mathbf{x}_\ell]$ and the shape constraints discussed in Section 5.2.

Penalty function The penalty function consists of three components. Since some of these components depend in a complex way on both the data and the parameters, we opt to impose these on a three dimensional grid $\mathcal{D}_G$ of disposable incomes: i.e. $\mathcal{D}_G \subset \mathcal{D} := \text{supp}(d_{NW}) \times \text{supp}(d_{PT}) \times \text{supp}(d_{FT})$.

The first component of the penalty function ensures that the choice probability of alternative NW is non-zero for every pair $(\mathbf{d}_g, \mathbf{x}_\ell) \in \mathcal{D}_G \times \mathcal{X}$: i.e. $1 - \Lambda_{PT}(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_{PT}) - \Lambda_{FT}(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_{FT}) \geq 0$. The contribution to the penalty function then is defined as:

$$\pi_1(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) := -|\mathcal{D}_G|^{-1} \sum_{\mathbf{d}_g \in \mathcal{D}_G} \min(0, 1 - \Lambda_{PT}(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_{PT}) - \Lambda_{FT}(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_{FT})).$$

The second component of the penalty function ensures that choice probabilities for alternatives PT and FT are increasing in their own disposable income and decreasing in the disposable income of the other alternatives. To be precise, we have for every $c \in \{PT, FT\}$ and for every pair $(\mathbf{d}_g, \mathbf{x}_\ell) \in \mathcal{D}_G \times \mathcal{X}$ that

$$\frac{\partial \Lambda_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c)}{\partial d_{g,c}} = (\beta_{c,c,1} + 2\beta_{c,c,2}d_{g,c} + 3\beta_{c,c,3}d_{g,c}^2)\Gamma_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c) \geq 0$$

$$\frac{\partial \Lambda_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c)}{\partial d_{g,c'}} = (\beta_{c,c',1} + 2\beta_{c,c',2}d_{g,c'} + 3\beta_{c,c',3}d_{g,c'}^2)\Gamma_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c) \leq 0, \quad \forall c' \neq c,$$

in which $\Gamma_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c) := \Lambda_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c)\left(1 + \exp\left(\alpha_c + \sum_{c'}\sum_{m=1}^3 \beta_{c,c',m}d_{g,c'}^m + \mathbf{x}_\ell'\boldsymbol{\gamma}_c\right)\right)^{-1}$, or equivalently that

$$\beta_{c,c,1} + 2\beta_{c,c,2}d_{g,c} + 3\beta_{c,c,3}d_{g,c}^2) \geq 0$$

$$(\beta_{c,c',1} + 2\beta_{c,c',2}d_{g,c'} + 3\beta_{c,c',3}d_{g,c'}^2) \leq 0, \quad \forall c' \neq c.$$

The contribution to the penalty function is defined as:

$$\pi_{2,c}(\boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) := -|\mathcal{D}_G|^{-1} \sum_{\mathbf{d}_g \in \mathcal{D}_G} \min(0, \beta_{c,c,1} + 2\beta_{c,c,2}d_{g,c} + 3\beta_{c,c,3}d_{g,c}^2))$$

$$\pi_{2,c,c'}(\boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) := |\mathcal{D}_G|^{-1} \sum_{\mathbf{d}_g \in \mathcal{D}_G} \max(0, \beta_{c,c',1} + 2\beta_{c,c',2}d_{g,c'} + 3\beta_{c,c',3}d_{g,c'}^2)).$$

Finally, the third part of the penalty function ensures that the choice probability of alternative NW is decreasing in the disposable income of the other alternatives.[34] For every $c, c' \in \{PT, FT\}$ with

---

[34] Note that the second part of the penalty function also ensures that the choice probability of alternative NW is increasing in its own disposable income.

$c \neq c'$ and for every pair $(\mathbf{d}_g, \mathbf{x}_\ell) \in \mathcal{D}_G \times \mathcal{X}$, we have that

$$\frac{\partial \Lambda_c(\mathbf{d}_g, \mathbf{x}_\ell)}{\partial d_{g,c}} + \frac{\partial \Lambda_{c'}(\mathbf{d}_g, \mathbf{x}_\ell)}{\partial d_{g,c}} = (\beta_{c,c,1} + 2\beta_{c,c,2}d_{g,c} + 3\beta_{c,c,3}d_{g,c}^2)\Gamma_c(\mathbf{d}_c, \mathbf{x}_\ell; \boldsymbol{\theta}_c)$$
$$+ (\beta_{c',c,1} + 2\beta_{c',c,2}d_{g,c} + 3\beta_{c',c,3}d_{g,c}^2)\Gamma_{c'}(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_{c'}) \geq 0.$$

The contribution to the penalty function is defined as:

$$\pi_{3,c,c'}(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) := -|\mathcal{D}_G|^{-1} \sum_{\mathbf{d}_g \in \mathcal{D}_G} \min(0, (\beta_{c,c,1} + 2\beta_{c,c,2}d_{g,c} + 3\beta_{c,c,3}d_{g,c}^2)\Gamma_c(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_c)$$
$$+ (\beta_{c',c,1} + 2\beta_{c',c,2}d_{g,c} + 3\beta_{c',c,3}d_{g,c}^2)\Gamma_{c'}(\mathbf{d}_g, \mathbf{x}_\ell; \boldsymbol{\theta}_{c'})).$$

Arranging all components, the composite penalty function is then

$$\pi(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) = \pi_1(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) + \sum_{c \in \{PT,FT\}} \left[\pi_{2,c}(\boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) + \sum_{c' \in \{NW,PT,FT\}, c' \neq c} \pi_{2,c,c'}(\boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) \right]$$
$$+ \pi_{3,PT,FT}(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}) + \pi_{3,FT,PT}(\mathbf{x}_\ell; \boldsymbol{\theta}_{PT}, \boldsymbol{\theta}_{FT}).$$

C.3   Estimates

Table 3 contains the estimates for the choice probabilities of alternatives PT and FT. The 90% confidence intervals are obtained by a bootstrap procedure, in which the model was re-estimated on 200 samples randomly drawn with replacement.

Table 3: Estimates choice probabilities

Parameter	PT Estimate	PT 90% CI	FT Estimate	FT 90% CI
Constant	-1.82	[-2.18, -1.25]	-2.41	[-4.00, -2.30]
$(d_0/1000)$	0.00	[-0.06, 0.00]	-0.01	[-0.02, 0.00]
$(d_0/1000)^2$	0.00	[-0.01, 0.08]	-0.08	[-0.20, 0.06]
$(d_0/1000)^3$	0.00	[-0.08, 0.00]	-0.04	[-0.13, 0.04]
$(d_{20}/1000)$	0.51	[0.01, 0.65]	-0.14	[-0.20, 0.00]
$(d_{20}/1000)^2$	-0.29	[-0.27, 0.00]	0.06	[-0.02, 0.14]
$(d_{20}/1000)^3$	0.06	[0.00, 0.07]	-0.01	[-0.06, 0.00]
$(d_{40}/1000)$	0.00	[-0.06, 0.00]	1.09	[0.95, 2.50]
$(d_{40}/1000)^2$	-0.05	[-0.22, 0.00]	-0.27	[-0.60, -0.16]
$(d_{40}/1000)^3$	0.01	[0.00, 0.03]	0.03	[0.01, 0.05]
Age	0.01	[0.00, 0.02]	-0.01	[-0.01, 0.00]
Years education	-0.01	[-0.03, 0.03]	0.13	[0.10, 0.15]
Number of children (0-1)	-2.72	[-3.50, -2.29]	-2.62	[-3.42, -2.26]
Number of children (2-4)	-0.67	[-1.29, -0.42]	-1.08	[-1.18, -0.55]
Number of children (5-7)	0.26	[-0.15, 0.45]	-0.29	[-0.50, -0.01]
Number of children (all)	0.17	[0.06, 0.39]	-0.20	[-0.49, -0.13]
East Germany	-0.06	[-0.37, 0.04]	0.13	[-0.02, 0.35]

C.4 Calculation distribution of welfare differences conditional on initial welfare

The practical implementation of the results in Theorem 2 poses some difficulties. Firstly, the distribution depends on transition probabilities. As was noted in Section 5, with only cross-sectional data available, these transition probabilities can only be set-identified. We, therefore, calculated upper and lower bounds for the joint distribution of baseline and reform welfare levels (i.e. Equation (19) of Proposition 1). As lower and upper bounds in the aggregate are not far apart (at most a 4 percentage point difference), we continue the analysis using only the upper bound. Secondly, Equation (20) requires integration over the derivative of a transition probability, which is quite cumbersome.

Implementation We first calculate the joint distribution of baseline and reform welfare by integrating out the optimal baseline and reform choices in Equation (19) of Proposition 1. The resulting joint distribution function of initial and post reform welfare is denoted by $H_0(w, s)$, that is:

$$
\begin{aligned}
H_0(w, s) = Pr\left(w \leq W_0, s \leq W_1\right) = \\
\sum_{i,j \in \{NW,PT,FT\}} \Pr_\omega\left[w \leq W_0^\omega(y - p_i, i), s \leq W_1^\omega(y - p_j', j), i = J^\omega(\mathbf{p}, y), j = J^\omega(\mathbf{p}', y)\right].
\end{aligned}
\tag{59}
$$

As we are interested in this distribution at the population level rather than at the individual level, we aggregate the distribution H_0 by defining

$$
H(w, s) = \int H_0(w, s) \, dG(\mathbf{p}, \mathbf{p}', y),
\tag{60}
$$

where G is the distribution of prices and exogenous income in the population. The joint distribution of baseline welfare and the welfare gain, $Pr\left(w \leq W_0, z \geq W_1 - W_0\right)$, is then calculated by:

$$
\begin{aligned}
Pr\left(w \leq W_0, z \geq W_1 - W_0\right) &= \int Pr\left(w \leq W_0, s - z \leq W_0, s = W_1\right) ds \\
&= -\int \partial_2 H(\max(w, s - z), s) \, ds.
\end{aligned}
\tag{61}
$$

The integral and derivative in this equation can be approximated numerically. Note that this is an approximation of Equation (20), aggregated across the population.

The distribution of gains and losses conditional on the initial welfare level can then be calculated as follows:

$$
Pr\left(z \leq W_1 - W_0 \,|\, w = W_0\right) = Pr\left(z + w \leq W_1 \,|\, w = W_0\right)
$$

$$
= \frac{Pr(z + w \leq W_1, w = W_0)}{Pr(w = W_0)}
\tag{62}
$$

$$
= \frac{\partial_1 H(w, z + w)}{\partial_1 H(w, -\infty)}.
$$

Again, derivatives can be computed numerically.

C.5 Additional tables and figures

Individual welfare distribution   Figure 10 shows the estimates of the individual unconditional welfare distributions (i.e. Equation (18) in Corollary 1). Similarly as for the individual conditional distributions in the main text (i.e. Figure 4), these distributions reflect either the econometrician's uncertainty about an individual's welfare level, or the distribution of the actual levels obtained by individuals in the population with the same observable characteristics (i.e. prices, exogenous income, and demographic variables). The difference is, however, that now we take into account that some may have found another optimal alternative instead of that chosen by the sampled individual, contrary to what was the case for the conditional individual welfare distributions in Figure 4 of the main text.
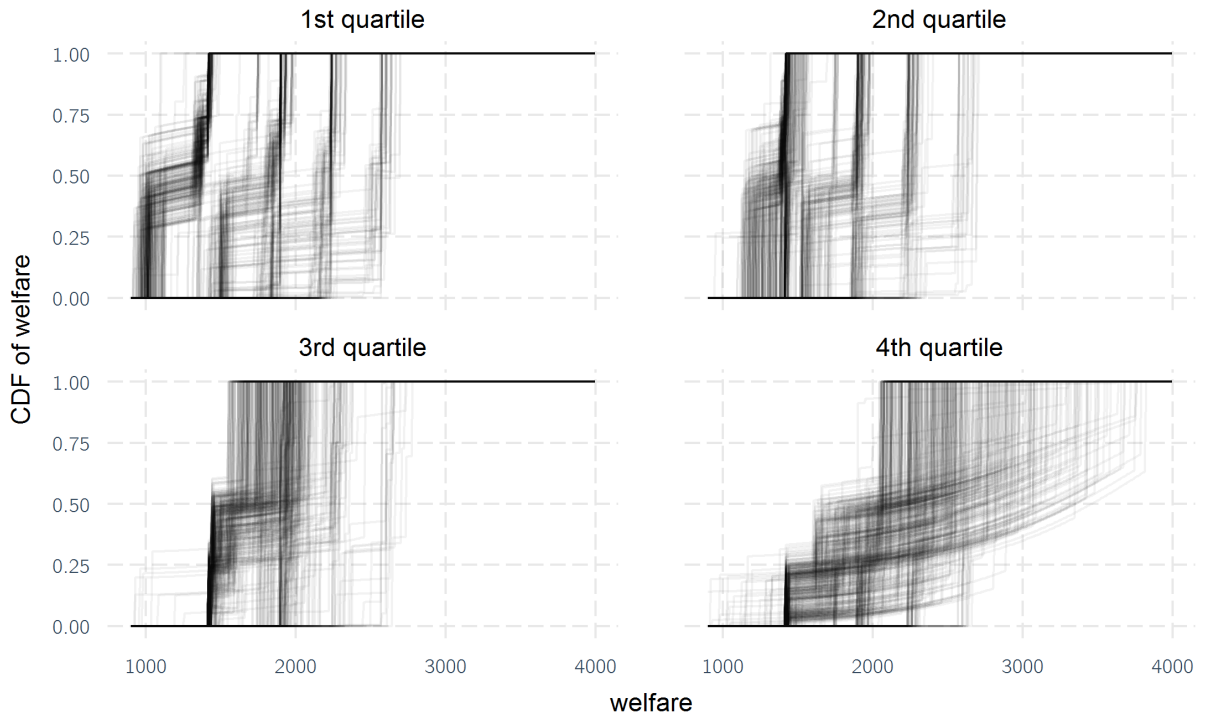

Figure 10: Unconditional individual welfare distribution by wage quartile

Social welfare   In Table 4, we present baseline and post-reform estimates of social welfare, i.e. Equation (34), for some members of the Atkinson-Kolm-Sen class of social welfare functions. This class is defined by specifying the concave transformation $h(w)$ in Equation (34) as $h(w) = \frac{w^{1-\varepsilon}}{1-\varepsilon}$ for $\varepsilon \geq 0$.[35] The calculations are performed using the result of Proposition 2.

As the class of Atkinson-Kolm-Sen functions obeys the first-order stochastic dominance criterion, i.e. all members of the class rank a distribution that first-order dominates another as the better one of

---

[35]The parameter ε is called the degree of inequality aversion. When ε equals zero, the social welfare function amounts to the average individual welfare level. As ε increases, more weight is given to individuals with lower welfare.

the two, it stands to reason that the reform distribution generates higher social welfare than the one for the baseline for all values of ε.

Table 4: Social welfare of the Atkinson form for baseline and reform (in euro)

ε	0	0.5	5
Baseline	1,815	1,783	1,530
Reform	1,925	1,897	1,673

Winners and losers Figures 11 and 12 plot the joint distribution of baseline welfare and welfare differences (i.e. Equation (61)). In Figure 11, the coordinates (w, z) of a point on the x-th iso-contour indicate the initial welfare level w and welfare gain z (or loss if z is negative) such that $x\%$ of the population obtains at least that initial welfare level w and does not gain more than z. In Figure 12, the initial welfare level w denotes the maximum level, rather than the minimum, that that number of people reach.
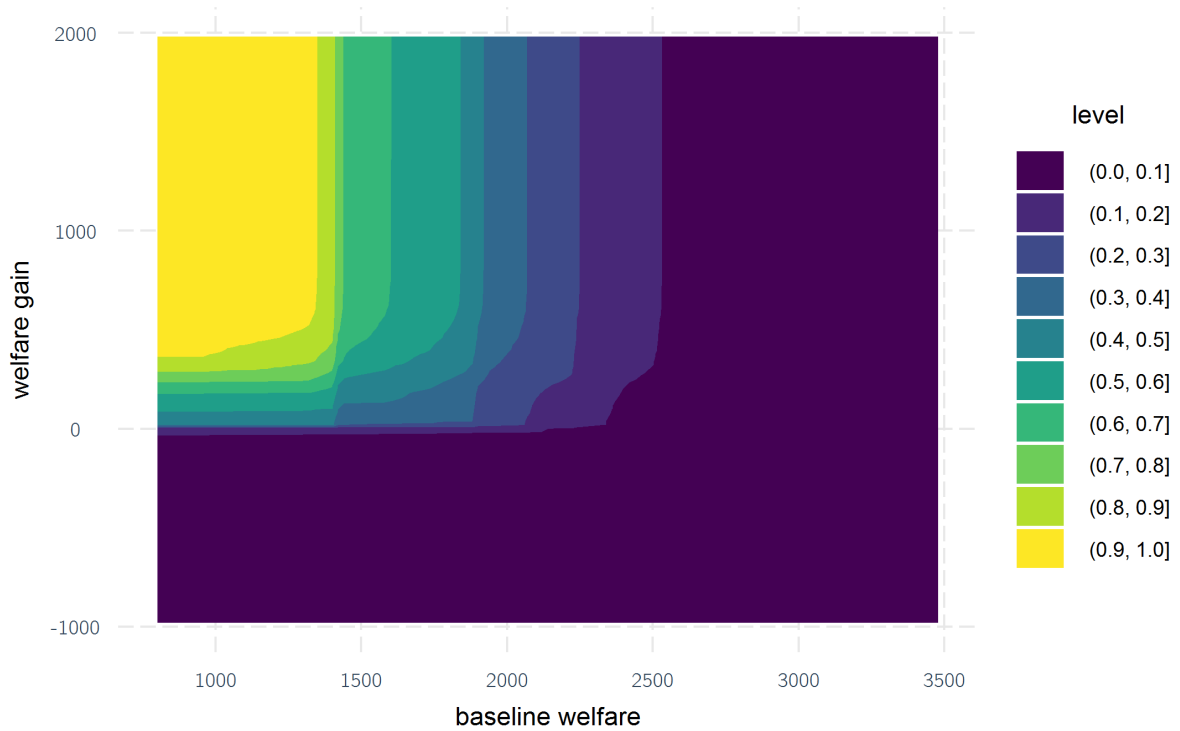

Figure 11: Joint distribution of baseline welfare and welfare gains/losses: $W_0 \geq w$

Figure 13 shows a more detailed, and less smoothed, version of the distribution of gains and losses conditional on baseline welfare (Figure 8 of the main text). The upper boundary of the yellow region tends to be lower than zero for higher welfare levels. For higher initial welfare levels, there are several regions where the median, that is the lower bound of the blue region, falls under the zero of the vertical axis. This confirms the findings highlighted in the main text.
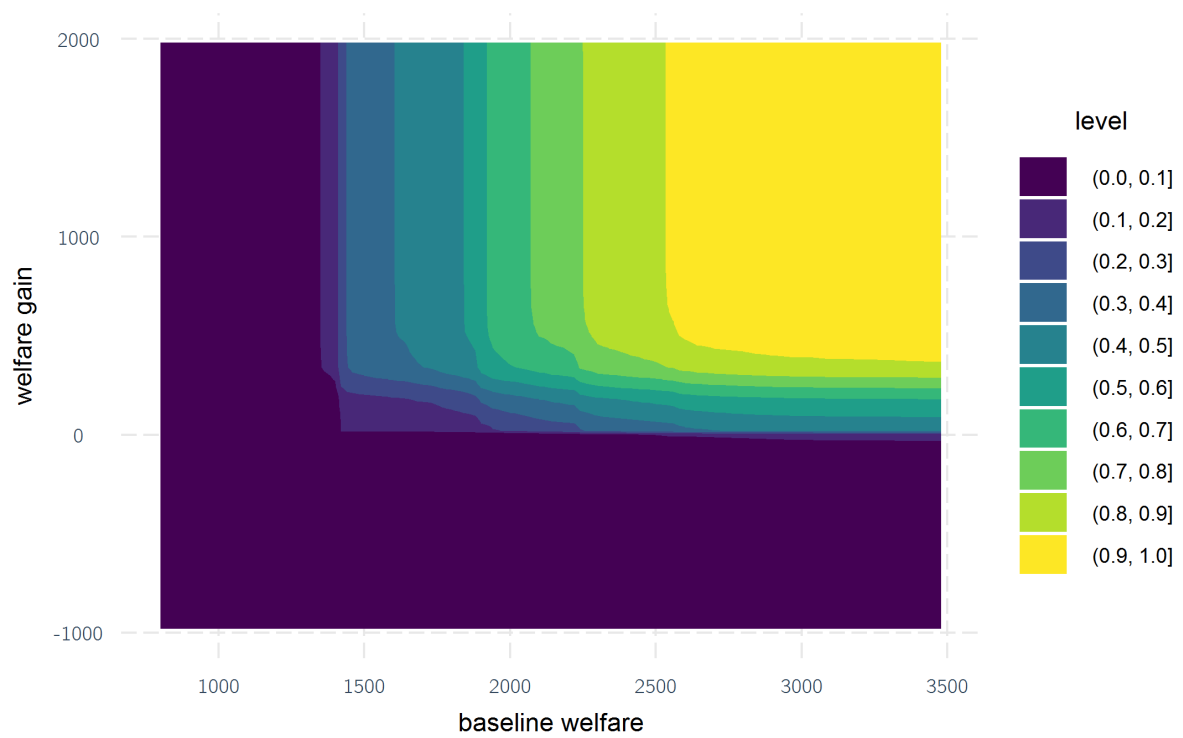
Figure 12: Joint distribution of baseline welfare and welfare gains/losses: $W_0 \le w$

Figure 14 shows the same conditional distribution of gains and losses, but now disaggregated per wage quartile and baseline choice. Gainers seem to be predominantly situated among individuals choosing alternatives PT and FT, especially in the lower half of the wage distribution, and among the initially poor persons with higher wages. Remarkably, high wage persons who initially choose NW tend to lose out.
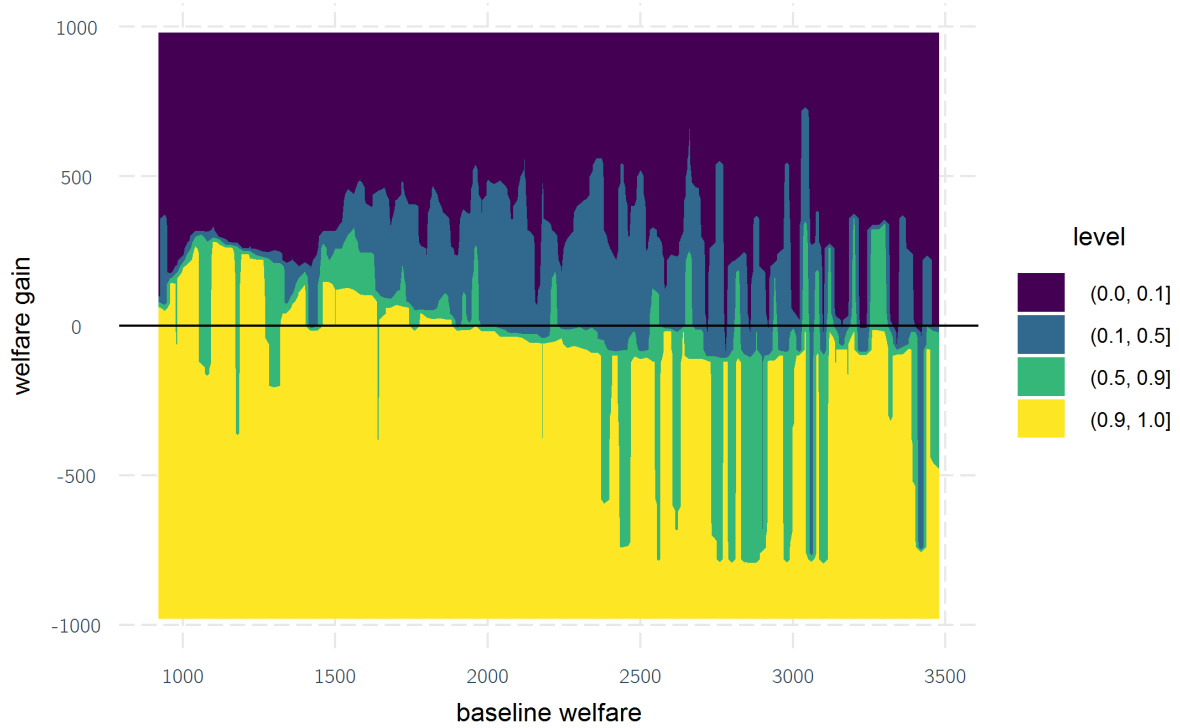
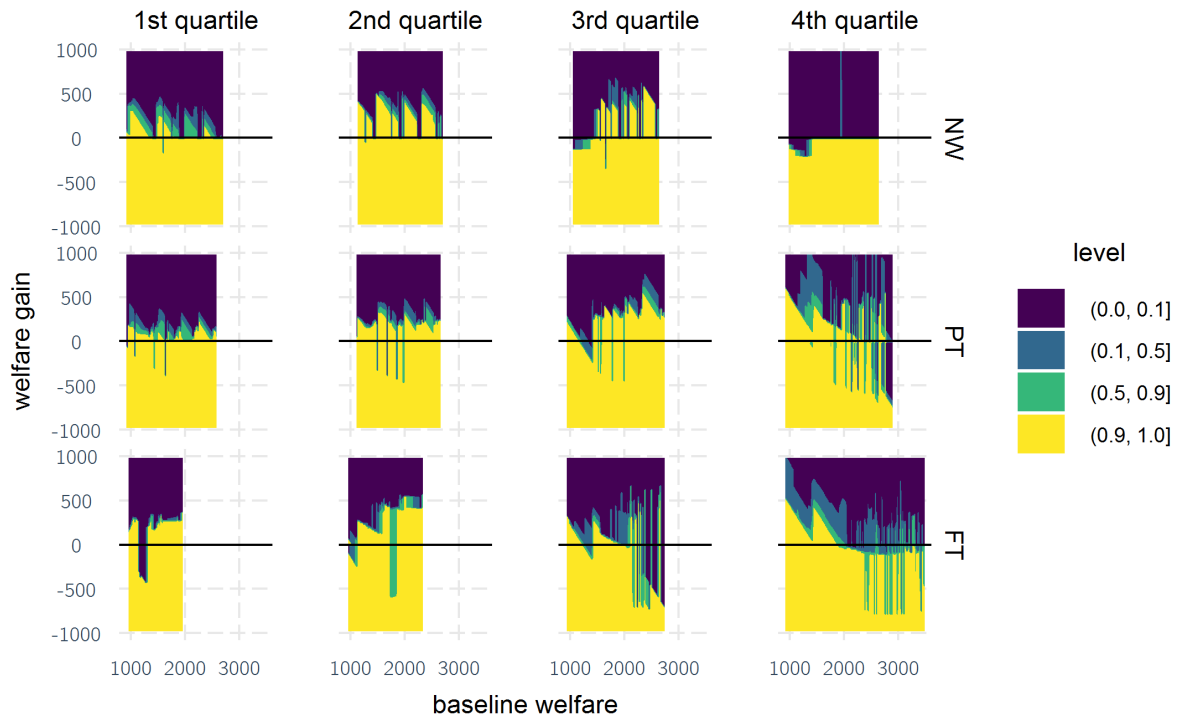Figure 13: Distribution of welfare gains and losses, conditional on baseline welfare

Figure 14: Distribution of welfare gains and losses, conditional on baseline welfare: by alternative chosen and wage quartile