Nonparametric Welfare Analysis for Discrete Choice

Debopam Bhattacharya*
University of Oxford
December, 2013.

Abstract: We consider empirical measurement of exact equivalent/compensating variation resulting from price-change of a discrete good, using individual-level data. Our set-up comprises utility functions which are not required to be quasi-linear, parametrically specified or smooth and include unobserved heterogeneity of unknown dimension – thus allowing for extremely general preference-distributions. We show that for binary and multinomial choice, the marginal distribution of EV/CV are nonparametrically point-identified solely from average demand-functions, even when the distribution and dimension of unobserved heterogeneity are neither specified nor identified. We express welfare distributions as closed-form functionals of average demand and show that average EV for price-rise equals the change in average consumer surplus and is smaller than average CV for a normal good. Our point-identification results for multinomial choice complement Hausman-Newey’s (2013) partial identification results for welfare distributions resulting from price change of a continuous good.

Keyword: Binary choice, multinomial choice, taxes and subsidy, equivalent and compensating variation, unobserved heterogeneity, nonparametric identification.

JEL codes: D12, D11, C14, C25.

*I am grateful to Richard Blundell, Martin Browning, Tom Crossley, Eric Gautier, Arthur Lewbel, Whitney Newey and John Quah for helpful discussions and encouragement. All errors are mine.
1 Introduction

This paper concerns the empirical measurement of money-metric welfare in regard to goods which are consumed or, more generally, decisions which are made, in binary or multinomial form. The specific focus is on price variations brought about by taxes and subsidies which affect consumer utility and can give rise to deadweight loss. Examples include, inter alia, the effects of taxing unemployment-benefits on exiting unemployment, of reducing stamp-duty on the decision of whether to buy or rent a house and of price changes on choice of the mode of transportation. The setting is where the researcher observes realizations of the discrete decision at the individual-level from micro-datasets which also record the individual’s characteristics, including income, and prices faced by her in regard to the discrete decision.\(^1\) The goal is to estimate exact – rather than approximate – impact on individual welfare, measured in terms of income compensation, of a change in price brought about by taxes or subsidies and the associated deadweight loss. We incorporate unobserved, individual heterogeneity in utility functions and focus on recovering the distribution and average values of the impact of price change on individual welfare resulting from such heterogeneity.

**Overview of results:** Our key results for multinomial choice (binary choice being a special case) are: (i) the marginal distribution and, consequently, the average compensating variation (CV) and equivalent variation (EV) corresponding to a price change are nonparametrically point-identified from average demand functions, (ii) for a price rise, the average EV is identical to the change in average consumer surplus even if utility is not quasi-linear in income, (iii) the average CV exceeds the average EV if the good is normal and (iv) the above conclusions hold even if the dimension/distribution of heterogeneity are not identified from average demand. These identification results are fully nonparametric in the sense that they do not require one to specify the dimension or distribution of heterogeneity and the functional form of utilities – thus allowing for extremely general preference distributions in the population. The key theoretical implication of our results is that for binary and multinomial choice, the average demand function contains all the relevant information for point-identification of welfare distributions under extremely general and unrestricted forms of unobserved heterogeneity. A practical advantage of our results is that they express welfare distributions as closed-form functionals of average demand and thus can be easily calculated in applications.

**Related literature:** A large literature exists in econometrics on estimation of demand from

---

\(^1\)This is in contrast to demand analysis using market-level data, popular in empirical Industrial Organization.
individual level consumption data. Indeed, an important use of such demand estimates is the calculation of welfare effects of price change arising from taxes or subsidies. For demand of a good that is consumed in continuous quantities, such as gasoline, Hausman, 1981, in a seminal paper, formulated the nonparametric identification and parametric estimation of exact welfare effects of a price change. Vartia, 1983 provided an alternative computational approach to the same problem. Hausman and Newey, 1995, extended these analyses by formulating semiparametric estimation of the welfare effects and developing the corresponding theory of statistical inference. Their methods cannot be directly used in discrete choice settings where the effect of a price change on individual utilities depends in a fundamental way on the discreteness of choice possibilities as well as on general individual heterogeneity. Specifically, in discrete choice scenarios, corner solutions are generic and this makes it difficult to recover the compensation functions using the differential-equation based approach of the above papers.

In the discrete choice setting, Domencich and McFadden, 1975 (DM75, henceforth) pages 94-99, calculated the welfare effects of price changes under the strong assumption that utility is quasi-linear, i.e., additively separable in income, thereby equating Marshallian and Hicksian welfare measures. In subsequent work, Small and Rosen, 1981 (SR81, henceforth) investigated the measurement of welfare effects of price and quality change for discrete choice. In their empirical formulation, SR81 introduced additive scalar heterogeneity in utility functions but assumed that the discrete good is sufficiently unimportant to the consumer so that income effects from price or quality changes are negligible (c.f., SR81, page 124, assumptions a and b) – thereby equating Marshallian and Hicksian welfare measures. More recently, Herridges and Kling, 1999 (HK99) and Dagsvik and Karlstrom, 2005 (DK05), in their analysis of the same problem, allowed utility to be nonlinear in income and incorporated unobservables in utility but assumed that these unobservables have both a known dimension and follow a known parametric distribution for identifying and estimating the distribution of welfare effects of price changes. The HK99 and DK05 analysis also require the functional forms of utilities to be known up to finite dimensional parameters which are either non-stochastic, or stochastic with fully known distributions.

The purpose of the present paper is to establish nonparametric point-identification of the marginal distribution of welfare effects of price change in a multinomial choice setting, incorporating unobservable heterogeneity in the utility function, and assuming no knowledge of the dimension

---

and thus of the distribution of these unobservables. Indeed, in many applications, it is important to allow for multiple sources of unobserved heterogeneity and it is easy to see that restricting the dimension of heterogeneity can place arbitrary restrictions on the variation of individual preferences in the population. For instance, consider the canonical nonparametric binary demand equation

\[ q = 1 \{ \beta(p, y) + \varepsilon > 0 \}, \]

where \( \beta(p, y) \) is an unknown function of price \( p \) and income \( y \), and \( \varepsilon \) is a scalar additive heterogeneity with an unknown distribution.\(^3\) This model implies that if at a specific price and income \( (p, y) \) an individual \( i \) prefers to buy while individual \( j \) does not (implying \( \varepsilon_i > \varepsilon_j \)), then at no price-income combination \( (p', y') \), can we have a situation where individual \( i \) prefers not to buy while individual \( j \) prefers to buy – an extremely strong (and untestable) restriction. Thus allowing for heterogeneity of unrestricted dimension immediately extends the scope of the results to a much larger set of preference profiles.

The *continuous* consumption analog of the present paper is Hausman and Newey, 2013 (HN13). They consider unobserved individual heterogeneity of unspecified dimension in utility functions and focus on recovering average equivalent variation resulting from price change of a *continuous* consumption good, based on demand data. HN13 show that (a) the dimension of heterogeneity is not identified from average demand and (b) if one allows for heterogeneity of unspecified dimension, then one cannot point-identify average welfare but may obtain bounds on it.\(^4\) Kitamura and Stoye (2013) have also worked under heterogeneity of unspecified dimension and provided tests of consumer rationality in that framework (see also Hoderlein and Stoye, 2013). We are not aware of any other work which performs demand analysis at this level of generality regarding both individual heterogeneity and the form of utility functions.\(^5\) For general binary choice, Ichimura and Thompson, 1998 (IT98) and Gautier and Kitamura, 2012 (GK12) considered random coefficient models where the dimension of heterogeneity is specified to be equal to the number of regressors (plus one for the random intercept) and enter the individual outcome equation in a special way, viz., as scalar multipliers – one attached to each regressor. These authors provide conditions – including large support requirements for regressors – under which the distribution of these random coefficients

\[^3\]See Matzkin (1992) for a nonparametric and and Manski (1975) for a semiparametric analysis of this model.

\[^4\]The HN13 method, like HN95, also involves solving a partial differential equation, but with income elasticities replaced by deterministic bounds on them. Such partial differential equation based approaches are difficult to use for discrete choice where corner solutions to consumer optimization are generic. Our approach instead works directly from definitions of Hicksian welfare.

\[^5\]Hoderlein and Vanhems 2011, Blundell, Matzkin and Kristensen, 2013 and Lewbel and Pendakur, 2013 among others discuss demand and welfare estimation for continuous choice under restricted heterogeneity.
can be nonparametrically identified up to scale normalization. Such support requirements may be restrictive in specific demand applications. Besides, estimation of the heterogeneity distribution is complicated owing to ill-posed inverse problems. The present paper shows that for the purpose of welfare analysis, this exercise is not necessary and that even the dimension of heterogeneity does not need to be specified in advance.\footnote{Of course, the IT98 and GK13 are general mathematical results which are applicable to any random coefficient model of binary outcome, not necessarily related to utility-based welfare analysis.} This finding appears to be of significant practical importance because in demand applications, a key motivation for recovering the distribution of consumer heterogeneity is the calculation of welfare distributions resulting from price change.

The rest of the paper is organized as follows. In section 2, we analyze the leading case of binary choice; in section 3, we provide an example to show that CV and EV distributions are point-identified in a binary choice situation even when the dimension and distribution of unobserved heterogeneity are not; in section 4 we show that the results for the binary case generalize to multinomial choice. All proofs are collected in an appendix at the end of the paper.

2 Binary Choice

2.1 Set-up

Consider an individual with income $y$, who faces the choice between buying and not buying a binary good which costs $p$. Let $a$ represent the quantity of numeraire which the individual consumes in addition to the binary good. Suppose that the utility realized by the individual is given by

\[
\begin{align*}
U_1(a, \eta), & \text{ if choose } 1, \\
U_0(a, \eta), & \text{ if choose } 0,
\end{align*}
\]

(1)

Here $\eta$ is a possibly vector-valued, individual-specific taste-variable of unknown dimension, unobserved by the econometrician, which enter the utility functions in any arbitrary way. Income and prices faced by the individual are observed by the econometrician. In addition, he may observe a set of covariates. The latter will be suppressed in the exposition below for notational clarity, i.e., the entire analysis should be thought of as implicitly conditioned on these observed covariates.

Given total income $y$, the budget constraint is $a + pq = y$ where $q \in \{0, 1\}$ represents the binary choice. Replacing the budget constraint in (1), the consumer’s realized utility is given by

\[
\begin{align*}
U_1(y - p, \eta), & \text{ if choose } 1, \\
U_0(y, \eta), & \text{ if choose } 0.
\end{align*}
\]
A consumer of type $\eta$, income $y$ and facing price $p$ chooses option 1 (buy the good) iff $U_1(y - p, \eta) > U_0(y, \eta)$.

Define the average demand function at a hypothetical price and income $(p, y)$ as

$$
\bar{q}(p, y) = \Pr \{ U_1(y - p, \eta) > U_0(y, \eta) \}
= \int 1 \{ U_1(y - p, \eta) > U_0(y, \eta) \} dF(\eta), \tag{2}
$$

where $F(\cdot)$ denotes the marginal distribution of $\eta$. This is akin to the "average structural function" defined in Blundell and Powell, 2003. We will state our key identification results below in terms of $\bar{q}(\cdot, \cdot)$. When observed realizations of price and income are jointly independent of heterogeneity $\eta$ (conditional on other covariates), one can obtain $\bar{q}(p, y)$ by a nonparametric regression of the individual’s decision to buy on price and income – the so-called average Marshallian demand curve.\footnote{If $\eta$ is correlated with price or income in the data, then one can recover $\bar{q}(p, y)$ using a control function based approach (c.f., Blundell and Powell, 2003).}

Independence of heterogeneity and budget sets has been assumed in all pre-existing research on welfare estimation for discrete choice, including DM75, SR81 and DK05. It is also a maintained assumption in random coefficient models of binary choice which are more restrictive in regards to the form and dimension of heterogeneity than in our set-up (see footnote 9 below).

Now, we impose the following assumption on the utility functions – our only substantive assumption in this paper:

**Assumption 1** Suppose that for each $\eta$, $U_0(a, \eta)$ and $U_1(a, \eta)$ are continuous and strictly increasing in $a$. Suppose $U_1^{-1}(b, \eta)$ denotes the unique solution $x$ to the equation $U_1(x, \eta) = b$ and $U_0^{-1}(b, \eta)$ denotes the unique solution in $x$ to the equation $U_0(x, \eta) = b$.

The above specification is much more general than the additive, scalar heterogeneity structure used in DM75 and SR81 where utility is given by

$$
\begin{cases}
U_1(y - p) + \varepsilon_1, & \text{if choose } 1, \\
U_0(y) + \varepsilon_0, & \text{if choose } 0,
\end{cases}
$$

where the functional forms of $U_1(\cdot, \cdot)$ and $U_0(\cdot, \cdot)$ are known (up to estimable finite dimensional parameters), $\varepsilon_1$ and $\varepsilon_0$ are scalar random variables, independent of price and income and have a
known distribution. DK05 (c.f. section 5 of their paper) consider the mixed multinomial logit-type structure

$$\begin{cases} U_1(y - p, \beta) + \varepsilon_1, \text{ if choose 1}, \\ U_0(y, \beta) + \varepsilon_0, \text{ if choose 0}, \end{cases}$$

(3)

where utilities are of known functional form and smooth in parameters; the random coefficients $\beta$ are independent of scalar-valued additive errors $\varepsilon_1, \varepsilon_0$, and $(\beta, \varepsilon_1, \varepsilon_0)$ have a fully known joint probability distribution independent of $(p, y)$. DK05 derive expressions for the distribution of Hicksian welfare measures in terms of these known probability distributions and known utility functions.\(^8\)

Indeed, if (i) the dimension and distribution of unobserved heterogeneity are separately identified from demand data, (ii) functional forms of utilities are known and (iii) the key unobservables (denoted by $\varepsilon$) enter as additive scalar errors in utility functions and all other unobservables (denoted by $\beta$) are either non-existent or enter utilities in a known way and are independent of the scalar additive errors, then one can use the expressions in DK05 to calculate the distribution of Hicksian welfare measures. These conditions are arbitrary and highly restrictive. In particular, we know from HN13 corollary 2 that even in the continuous consumption case, the dimension of heterogeneity – let alone its distribution – is not identified from demand data. In section 3 of the present paper, we provide an example specialized to the binary choice case where the dimension and thus distribution of heterogeneity are not identified and thus the DK05 results cannot be applied, and yet welfare distributions are nonparametrically point-identified using our results because they are based solely on the average demand function. Indeed, our identification results require only that utilities are continuous and strictly increasing in the numeraire (assumption 1) and, unlike the papers cited above, do not require one to specify the functional form of the utilities (including how unobserved heterogeneity enter utilities), to impose differentiability, to specify the dimension and distribution of unobserved heterogeneity or to assume arbitrary independence conditions among

\(^8\)For example, DK05 page 63, after the proof of theorem 2, clarify that “...one can calculate the Hicksian choice probabilities readily, provided the cumulative distribution $F^B(\cdot)$ is known since only a one-dimensional integral is involved in the formula for $\mathbb{P}^B_B(j; w, u)$.” When the heterogeneity distribution $F^B(\cdot)$ is unknown, the theorems and corollaries of DK05 do not provide identification results; they express analytically the objects of interest – written on the LHS in equation (12), theorem 2, corollary 2 etc. – in terms of the model primitives, viz., the individual utility functions and the distribution of heterogeneity (and its partial derivatives) which appear on the RHS. They do not establish that the RHS expressions are nonparametrically identified from observed demand data without knowledge of the heterogeneity distribution (if they did, there should have been another equality in these results stating what that functional of observed data is in each case). This may be contrasted with equations (5) and (8) of the present paper, where the LHS contains the object of interest and the RHS contains functionals of observed demand.
different components of heterogeneity, as in model (3) above.\footnote{Our set-up is also more general than a pure random coefficient model, which postulates \( q = 1 \{ -p + \epsilon_1 y + \epsilon_0 > 0 \} \), where \( (\epsilon_0, \epsilon_1) \) is a 2-dimensional random vector independent of \( (p, y) \). For example, the random coefficient model implies that for fixed \( p \), if an individual prefers to buy at income \( y \) and also at income \( y' > y \), then she must also prefer to buy at any intermediate income \( \lambda y + (1 - \lambda) y' \) for \( \lambda \in (0, 1) \). This is because \[-p + \epsilon_1 (\lambda y + (1 - \lambda) y') + \epsilon_0 = \lambda \{ -p + \epsilon_1 y + \epsilon_0 \} + (1 - \lambda) \{ -p + \epsilon_1 y' + \epsilon_0 \},\] so that if each term on the RHS is positive, then so is the LHS. This is not imposed by our set-up because \( U_1 (y - p, \eta) > U_0 (y, \eta) \) and \( U_1 (y' - p, \eta) > U_0 (y', \eta) \) need not imply that \( U_1 (\lambda y + (1 - \lambda) y' - p, \eta) > U_0 (\lambda y + (1 - \lambda) y', \eta) \).}

We now turn to welfare analysis within our set-up.

2.2 Equivalent Variation

First, consider equivalent variation as a measure of welfare. We will evaluate welfare change for a ceteris paribus price increase from \( p_0 \) to \( p_1 \). That is, for fixed \( y \), we will calculate the amount of income to be subtracted from an \( \eta \) type individual with income \( y \) and facing prices \( p_0 \) in order that her (maximized) utility in this situation equals that when she were facing prices \( p_1 \) where \( p_1 > p_0 \). For lack of a better word, we will call this "compensation" but with the understanding that it is an amount of money being taken away.

Note that by definition, the EV, denoted by \( S^{EV} (y, p_0, p_1, \eta) \), will satisfy

\[
\max \left\{ U_0 (y - S^{EV} (y, p_0, p_1, \eta), \eta) , U_1 (y - S^{EV} (y, p_0, p_1, \eta) - p_0, \eta) \right\} = \max \left\{ U_0 (y, \eta), U_1 (y - p_1, \eta) \right\} . \tag{4}
\]

We will demonstrate that the expressions for \( S^{EV} (y, p_0, p_1, \eta) \) take the following forms for different ranges of \( \eta \), for given \( y, p_1 \) and \( p_0 \).

\begin{proposition}
Suppose Assumption 1 holds. Then (i) if \( U_1 (y - p_0, \eta) \leq U_0 (y, \eta) \), then \( S^{EV} (y, p_0, p_1, \eta) = 0 \); (ii) if \( U_1 (y - p_1, \eta) \leq U_0 (y, \eta) < U_1 (y - p_0, \eta) \), then \( S^{EV} (y, p_0, p_1, \eta) = y - p_0 - U_1^{-1} (U_0 (y, \eta), \eta) \); (iii) if \( U_1 (y - p_1, \eta) > U_0 (y, \eta) \), then \( S^{EV} (y, p_0, p_1, \eta) = p_1 - p_0 \).
\end{proposition}
These three cases correspond respectively to \(\eta\)s who do not buy at the lower price, those who switch from buying at lower price to not buying at the higher price and those who buy at both low and high price. While the zero EV is obvious for the first group, the other cases are not entirely obvious because one needs to understand how buying is affected when income is deducted from a situation of low price.

The above proposition is an intermediate result which leads us to the first identification result, where we provide a closed-form expression for the distribution of EV, induced by the distribution of the heterogeneity \(\eta\).

**Proposition 2** Suppose Assumption 1 holds. Consider a price rise from \(p_0\) to \(p_1\). Then for fixed income \(y\), the distribution of the equivalent variation is given by

\[
\Pr \{ S^{EV} (y, p_0, p_1, \eta) \leq a \} = \begin{cases} 
0, & \text{if } a < 0, \\
1 - \bar{q} (p_0, y), & \text{if } a = 0, \\
1 - \bar{q} (a + p_0, y), & \text{if } 0 < a < p_1 - p_0, \\
1, & \text{if } a \geq p_1 - p_0.
\end{cases}
\]  

(5)

where \(\bar{q} (\ldots)\) is defined above in (2).

(See appendix for proof)

**Interpretation:** An intuitive interpretation of the result is as follows. Consider an \((y, \eta)\) type individual whose reservation price is \(p_0 + t (y, \eta)\) where \(0 < t (y, \eta) < p_1 - p_0\). This means that she is indifferent between buying or not buying at price \(p_0 + t (y, \eta)\) when she has income \(y\) so that

\[
U_0 (y, \eta) = U_1 (y - (p_0 + t (y, \eta)), \eta).
\]

At any higher price, the RHS is smaller and she does not buy; at any lower price, the RHS is larger and she buys. But since \(y - (p_0 + t (y, \eta)) = (y - t (y, \eta)) - p_0\), we get that

\[
U_0 (y, \eta) = U_1 ((y - t (y, \eta)) - p_0, \eta),
\]

which means that if we take away an amount of the numeraire equal to \(t (y, \eta)\), then she would reach the same level of utility from buying at price \(p_0\) as she would when not buying. Recall that since \(t (y, \eta) < p_1 - p_0\), she was not buying at the higher price \(p_1\) and getting utility \(U_0 (y, \eta)\), which is
precisely the reference utility for calculation of her EV. The previous display therefore implies that the EV for a price increase from $p_0$ to $p_1$ is $t(y, \eta)$ for this consumer. That is, the EV equals the difference between the reservation price and the initial lower price $p_0$. Therefore, the probability that EV is less than $a$ equals the proportion of individuals with reservation price less than $p_0 + a$ which, by definition, equals the fraction of individuals who do not buy at prices higher than $p_0 + a$, and is thus given by $1 - \bar{q}(p_0 + a, y)^{10}$.

Having obtained the distribution of the EV, we now aggregate these functions w.r.t. the heterogeneity to obtain the average EV.

**Corollary 1** Suppose Assumption 1 holds. Then for a price increase from $p_0$ to $p_1$, the average equivalent variation is given by

$$
\mu^{EV}(y, p_0, p_1) = \int_{p_0}^{p_1} \bar{q}(p, y) \, dp,
$$

where $\bar{q}(p, y)$ is defined in (2). (see appendix for proof)

**Remark 1** Since $\bar{q}(p, y)$ is simply the average Marshallian demand observed in the data, the above conclusion shows that the change in Marshallian consumer surplus and the average Hicksian equivalent variations are equal. This result obtains although the utility functions were not specified to be quasi-linear. Indeed, for quasi-linear utility, both CV and EV are equal to the change in Marshallian consumer surplus which will not (necessarily) be the case here, as we shall see below.

### 2.3 Compensating Variation

Now consider measurement of compensating variation. Fix a level of income $y$ and consider a price rise from an initial value $p_0$ to $p_1$. The compensating variation $S^{CV}(y, p_0, p_1, \eta)$ measures the income $S$ to be given to an individual of type $\eta$ with income $y$ and facing price $p_1 > p_0$, so that their maximized utility in this situation is equal to their maximized utility when prices were $p_0$ and income was $y$, i.e.,

$$
\max \{ U_0(y + S, \eta), U_1(y + S - p_1, \eta) \} = \max \{ U_0(y, \eta), U_1(y - p_0, \eta) \}.\tag{7}
$$

We will demonstrate that the expressions for $S^{CV}(y, p_0, p_1, \eta)$ take the following form.

---

$^{10}$This reasoning is also consistent with the HN13 finding that one can only bound average EV in the continuous consumption case. Indeed, for continuous choice, for every $(y, \eta)$, there are uncountably infinite reservation prices – one at each possible value of consumption – and so this reservation price based analysis does not hold there.
**Proposition 3** Suppose Assumption 1 holds. Then, (i) if \( U_1(y - p_0, \eta) \leq U_0(y, \eta) \), then \( S^{CV}(y, p_{01}, \eta) = 0 \); (ii) if \( U_0(y, \eta) < U_1(y - p_0, \eta) \leq U_0(y + p_1 - p_0, \eta) \), then \( S^{CV}(y, p_{01}, \eta) = U^{-1}_0(U_1(y - p_0, \eta), \eta) - y \); (iii) if \( U_1(y - p_0, \eta) > U_0(y + p_1 - p_0, \eta) \), then \( S^{CV}(y, p_{01}, \eta) = p_1 - p_0 \) (see appendix for proof).

These three cases correspond respectively to \( \eta \)s who (i) do not buy at the lower price, (ii) those who switch from buying at lower price to not buying at the higher price but when compensated by the amount of price change would prefer not to buy and (iii) switchers who when compensated by the amount of price change would prefer to buy as well as those who buy at both low and high price. While the zero EV is obvious for the first group, the other cases are not entirely obvious because one needs to understand how buying is affected when income is raised from a situation of high price.

The above proposition, like proposition 1, is also an intermediate result which helps us reach our second identification result, where we establish that the distribution of CV, induced by the distribution of the heterogeneity \( \eta \) can also be expressed solely in terms of the average demand functional \( \bar{q}(\cdot, \cdot) \).

**Proposition 4** Suppose Assumption 1 holds. Consider a price rise from \( p_0 \) to \( p_1 \). Then the distribution of the compensating variation is given by

\[
\Pr \{ S^{CV}(y, p_0, p_1, \eta) \leq a \} = \begin{cases} 
0, & \text{if } a < 0, \\
1 - \bar{q}(p_0, y) & \text{if } a = 0, \\
1 - \bar{q}(a + p_0, y + a) & \text{if } 0 < a < p_1 - p_0, \\
1, & \text{if } a \geq p_1 - p_0,
\end{cases}
\]

where \( \bar{q}(p, y) \) is defined in (2) (see appendix for proof).

**Interpretation:** An intuitive interpretation of the result is as follows. Consider an \((y, \eta)\) type individual, who was buying the good at price \( p_0 \) but is not buying it at the higher price \( p_1 \). Suppose her utility from not buying at price \( r \geq p_0 \) when compensated by the amount of price rise \( r - p_0 \) exceeds her utility from buying, i.e.,

\[
U_0(y + r - p_0, \eta) > U_1((y + r - p_0) - r, \eta) = U_1(y - p_0, \eta).
\]

For CV, the RHS utility is the reference utility and therefore any compensation exceeding \((r - p_0)\) will lead such a person to not buy and enjoy a utility level exceeding the reference utility level.
Thus for such a person, the \( CV \) must not exceed \( r - p_0 \). Such individuals can be identified in the data as those who do not buy when they have income \( y + r - p_0 \) and price is \( r \) and thus

\[
\Pr \left( S_{CV}^{} (y, p_0, p_1, \eta) \leq r - p_0 \right) = 1 - \tilde{q} (r, y + r - p_0).
\]

Replacing \( r - p_0 = a \), one obtains \( \Pr \left( S_{CV}^{} (y, p_0, p_1, \eta) \leq a \right) = 1 - \tilde{q} (a + p_0, y + a) \) for \( a \geq 0 \). In other words, the \( CV \) must not exceed \( a \) for those who had switched initially from buying to not buying due to the price-rise from \( p_0 \) to \( p_0 + a \) but who would nonetheless, upon getting compensated income \( y + a \), strictly prefer to not buy and instead "run away" with the extra money when price is \( p_0 + a \). These individuals can be made as happy as they originally were by paying a compensation strictly less than \( a \) and letting them run away with this extra money without buying.

**Corollary 2** Suppose Assumption 1 holds. Then the average compensating variation \( CV \) for a price increase from \( p_0 \) to \( p_1 \) with income fixed at \( y \) is given by

\[
\mu_{CV}^{} (y, p_0, p_1) = \int_{p_0}^{p_1} \tilde{q} (p, y + p - p_0) \, dp. \tag{9}
\]

(see appendix for proof).

**Remark 2** If the binary good is normal, then \( \tilde{q} (p, y) \leq \tilde{q} (p, y + p - p_0) \) for fixed \( y \) and for all \( p \geq p_0 \). Hence, it follows from the expressions for \( CV \) and \( EV \) in (6) and (9) that for a price increase from \( p_0 \) to \( p_1 \), \( \mu_{EV}^{} (y, p_0, p_1) \leq \mu_{CV}^{} (y, p_0, p_1) \).

**Remark 3** For a per unit tax of \( \tau \), the average deadweight loss is given by

\[
\begin{align*}
DWL^{ax} (EV) &= \int_{p_0}^{p_0 (1+\tau)} \tilde{q} (p, y) \, dp - \tau p_0 \times \tilde{q} (p_0 (1 + \tau), y), \\
DWL^{ax} (CV) &= \int_{p_0}^{p_0 (1+\tau)} \tilde{q} (p, y + p - p_0) \, dp - \tau p_0 \times \tilde{q} (p_0 (1 + \tau), y).
\end{align*}
\]

**Remark 4** When a subsidy reduces prices from \( p_1 \) to \( p_0 \), the labelling of \( EV \) and \( CV \) reverses and we get that

\[
\mu_{CV}^{} (p_0, p_1, y) = \int_{p_0}^{p_1} \tilde{q} (p, y) \, dp, \quad \mu_{EV}^{} (p_0, p_1, y) = \int_{p_0}^{p_1} \tilde{q} (p, y + p - p_0) \, dp.
\]

**Remark 5** In order for the C.D.F. of \( CV \) to be weakly increasing, one needs to check whether for all \( a > 0 \) and for all \( y \), \( \tilde{q} (p_0 + a, y + a) \) is smaller than \( \tilde{q} (p_0, y) \). Similarly, for C.D.F. of \( EV \) to
be weakly increasing, one needs to check that for all \( a > 0 \), \( \bar{q}(p_0 + a, y) < \bar{q}(p_0, y) \), for all \( y \). Note that according to the random utility model considered here,

\[
\bar{q}(p_0, y) - \bar{q}(p_0 + a, y + a) = \Pr \{ U_0(y, \eta) - U_1(y - p_0, \eta) \leq 0 \} - \Pr \{ 0 > U_0(y + a, \eta) - U_1(y + a - (p_0 + a), \eta) \}
\]

\[= \Pr \{ U_0(y, \eta) - U_1(y - p_0, \eta) \leq 0 \} - \Pr \{ 0 > U_0(y + a, \eta) - U_1(y - p_0, \eta) \}, \]

which is non-negative since \( U_0(\cdot, \eta) \) is assumed to be strictly increasing and \( a > 0 \). Similarly,

\[
\bar{q}(p_0, y) - \bar{q}(p_0 + a, y)
\]

\[= \Pr \{ U_0(y, \eta) < U_1(y - p_0, \eta) \} - \Pr \{ U_0(y, \eta) < U_1(y - (p_0 + a), \eta) \}, \]

which is non-negative since \( U_1(\cdot, \eta) \) is assumed to be strictly increasing and \( a > 0 \). One may test these conditions after estimating an unrestricted \( \bar{q}(\cdot, \cdot) \) or impose these restrictions during estimation. The first condition, viz., \( \bar{q}(p_0 + a, y + a) \leq \bar{q}(p_0, y) \) can be interpreted as a revealed preference or Slutsky type inequality for binary choice.

**Remark 6** Maler, 1974, page 139 claims that individual demand for indivisible goods has income elasticity equal to zero and hence average CV and EV are both identical to the aggregate Marshallian consumer surplus. From (6) and (9) above, this assertion appears to be incorrect.

### 3 Relation with Distribution of Heterogeneity

It is important to note that the proofs of the four propositions above do not rely on whether the distribution of unobserved heterogeneity \( \eta \) is known or identified. In this section, we provide an example where the distribution and even the dimension of heterogeneity are not identified and yet the conditions of proposition 2 and 4 are satisfied, so that the distribution of welfare measures are point-identified.

**Example:** Suppose \( \eta \equiv (\eta_1, \eta_2) \) is jointly independent of \((p, y)\) and \( \eta_1 \perp \eta_2 \). Assume that the support of price distribution in the data is contained in \([0, p_H]\) and income is bounded below by \( y_L \) with \( y_L > p_H > 0 \). Let

\[
U_1(y - p, \eta) = y - p + \eta_1, \quad U_0(y, \eta) = (1 - \eta_2) y,
\]

where \( \eta_2 \) is distributed uniform \((0,1)\) and the support of \( \eta_1 \) – denoted by \( T \) – is contained in \((p_H - y_L, 0)\). Denote the c.d.f. of \( \eta_1 \) by \( G(\cdot) \). An individual of type \((y, \eta)\) and facing price \( p \) buys the good if and only if \( y - p + \eta_1 > (1 - \eta_2) y \).
Thus for any fixed $\eta = (\eta_1, \eta_2)$ in the support, the utility functions are continuous and strictly increasing in income. Thus proposition 2 and 4 apply and imply that the distribution of EV and CV arising from a price change are point-identified.

Now, consider average demand in this model. Since $p_H - y_L \leq \eta_1 \leq 0$ w.p.1, it follows that for any $p, y$ in the support of the data, we must have that $p - y < \eta_1 < p$, or

$$0 < \frac{p - \eta_1}{y} < 1, \text{ w.p. } 1. \tag{10}$$

Therefore, 

$$\bar{q}(p, y) = \Pr \{y - p + \eta_1 > (1 - \eta_2)y | p, y\}$$

$$= \Pr \{\eta_2 y + \eta_1 > p | p, y\}$$

$$= \Pr \{\eta_2 y + \eta_1 > p\}, \text{ by } \eta \equiv (\eta_1, \eta_2) \bot (p, y)$$

$$= \Pr \left\{\eta_2 > \frac{p - \eta_1}{y}\right\} \text{ since } y > 0$$

$$= \int_T \left(1 - \frac{p - \eta_1}{y}\right) dG(\eta_1), \text{ by (10), } \eta_1 \perp \eta_2 \text{ and } \eta_2 \sim U(0, 1)$$

$$= \left(1 - \frac{p}{y}\right) + \frac{1}{y} \int_T \eta_1 dG(\eta_1)$$

$$= \left(1 - \frac{p}{y}\right) + \frac{1}{y} E(\eta_1).$$

Thus, $\bar{q}(p, y)$ depends on the distribution of $\eta_1$ only through its expectation. Therefore, all distributions for $\eta_1$ with support contained in $(p_H - y_L, 0)$ and having the same expectation will give rise to the same average demand for each value of $p$ and $y$, implying that the distribution of $\eta_1$ cannot be identified from average demand. In particular, $\eta_1 \sim \text{Uniform } [p_H - y_L, 0]$ and $\eta_1 = \frac{pH - yL}{2}$ with probability 1 will both produce identical $\bar{q}(p, y)$ for all $(p, y)$ in the support of price and income. This implies that the dimension of heterogeneity is also not identified (dim($\eta_1, \eta_2$) is 1 when $\eta_1$ is degenerate and 2 when $\eta_1$ is uniform). Indeed, HN13 have shown that even for continuous choice, the dimension of heterogeneity cannot be identified from demand data (c.f. corollary 2 of HN13). Yet, as clarified above, propositions 2 and 4 of the present paper imply that the distribution of EV and CV are point-identified from observed $\bar{q}(\cdot, \cdot)$.

The above example demonstrates that identifiability of the heterogeneity distribution, or even correct specification of its dimension are not a requirement for identifiability of welfare distributions.
4 Multinomial Choice

We now extend the above results for binary outcomes to multinomial choice. Assume that a consumer with income $y$ and taste $\eta$ (again, possibly vector-valued and of unknown dimension) faces a mutually exclusive set of alternatives with alternative-specific prices – the classic example being choice of the mode of transportation (e.g., bus, train, walk etc.). The consumer can pick only one among the various alternatives. Let the set of alternatives be denoted by $\{0, 1, ..., J\}$ with $p_j$ denoting the price of the $j$th alternative for $j = 1, ..., J$ and the $0$th alternative denoting not choosing any of the $J$ alternatives. As before, assume that utility from not choosing any alternative is $U_0 (y, \eta)$ and choosing alternative $j$ produces utility $U_j (y - p_j, \eta)$.

**Assumption 2** $U_j (\cdot, \eta)$ is continuous and strictly increasing for each $\eta$, for $j = 0, 1, ..., J$.

Now, consider a price increase for alternative 1 from $p_{10}$ to $p_{11}$, with the prices of all other alternatives held fixed. Let $p_{-1} = (p_2, p_3, ..., p_J)$, and

$$U^* (y, p_{-1}, \eta) \equiv \max \{U_0 (y, \eta), U_2 (y - p_2, \eta), ..., U_J (y - p_J, \eta)\}.$$ 

Note that Assumption 2 implies that $U^* (\cdot, p_{-1}, \eta)$ is strictly increasing and continuous for each $\eta$. Let $U^{*-1} (a, p_{-1}, \eta)$ denote the unique solution for $x$ in the equation $U^* (x, p_{-1}, \eta) = a$.

Now, it can be seen that all our results from the binary case carry over with the utility $U_0 (y, \eta)$ replaced by $U^* (y, p_{-1}, \eta)$, since the latter does not involve the price of alternative 1. For example, for CV calculation, the relevant compensation function $S$ satisfies

$$\max \{U_0 (y + S, \eta), U_1 (y + S - p_{11}, \eta), U_2 (y + S - p_2, \eta), ..., U_J (y + S - p_J, \eta)\}$$

$$= \max \{U_0 (y, \eta), U_1 (y - p_{10}, \eta), U_2 (y - p_2, \eta), ..., U_J (y - p_J, \eta)\}.$$ 

Since $\max \{a, b, c\} = \max \{a, \max \{b, c\}\}$, the previous display can be rewritten as

$$\max \{U^* (y + S, p_{-1}, \eta), U_1 (y + S - p_{11}, \eta)\}$$

$$= \max \{U^* (y, p_{-1}, \eta), U_1 (y - p_{10}, \eta)\},$$
which is exactly analogous to (7) with $U_0(y, \eta)$ replaced by $U^*(y, p_{-1}, \eta)$. Following through the proof of proposition 3 with $U_0(y, \eta)$ replaced by $U^*(y, p_{-1}, \eta)$ yields

$$CV = \begin{cases} 
0 & \text{if } U_1(y - p_{10}, \eta) < U^*(y, p_{-1}, \eta) \\
U^{*\pi-1}(U_1(y - p_{10}, \eta), p_{-1}, \eta) - y, & \text{if } \begin{pmatrix} U^*(y, p_{-1}, \eta) \leq U_1(y - p_{10}, \eta) \\
< U_0(y + p_{11} - p_{10}, \eta) \end{pmatrix} \\
p_{11} - p_{10} & \text{if } U^*(y + p_{11} - p_{10}, p_{-1}, \eta) < U_1(y - p_{10}, \eta)
\end{cases}. $$

Now,

$$\Pr \left( U^{*\pi-1}(U_1(y - p_{10}, \eta), p_{-1}, \eta) - y < r \right) = \Pr( U_1(y - p_{10}, \eta) < U^*(y + r, p_{-1}, \eta)) = \Pr(U_1(y + r - (p_{10} + r), \eta) < U^*(y + r, p_{-1}, \eta)) = 1 - \tilde{q}_1(p_{10} + r, p_{-1}, y + r),$$

where $\tilde{q}_1(t, p_{-1}, y)$ denotes (analogous to (2)) the average demand function for alternative 1 when its price is $t$, prices of the other alternatives are held fixed at $p_{-1}$ and income is $y$. Thus we have – analogous to the binary case – that

$$\Pr(CV \leq r) = \begin{cases} 
0, & \text{if } r < 0 \\
1 - \tilde{q}_1(p_{10}, p_{-1}, y) & \text{if } r = 0 \\
1 - \tilde{q}_1(p_{10} + r, p_{-1}, y + r) & \text{if } 0 < r < p_{11} - p_{10} \\
1, & \text{if } r \geq p_{11} - p_{10}
\end{cases}, $$

$$E(CV) = \int_{p_{10}}^{p_{11}} \tilde{q}_1(r, p_{-1}, y + r - p_{10}) dr. \quad (11)$$

An exactly analogous argument yields that

$$\Pr(EV \leq a) = \begin{cases} 
0, & \text{if } a < 0 \\
1 - \tilde{q}_1(p_{10}, p_{-1}, y) & \text{if } a = 0 \\
1 - \tilde{q}_1(p_{10} + a, p_{-1}, y) & \text{if } 0 < a < p_{11} - p_{10} \\
1, & \text{if } a \geq p_{11} - p_{10}
\end{cases}, $$

$$E(EV) = \int_{p_{10}}^{p_{11}} \tilde{q}_1(r, p_{-1}, y) dr. \quad (12)$$

From (11) and (12), it is clear that welfare distributions for the multinomial case are also point-identified from the average demand function alone.
5 Summary and conclusion

To summarize, the key insight of this paper is that for binary and multinomial choice, the average demand function alone contains all the relevant information for nonparametric point-identification of welfare distributions under extremely general and unrestricted forms of unobserved heterogeneity and utility functions. This result continues to hold even if the dimension – and therefore the distribution – of underlying heterogeneity are neither specified nor identified. These results complement Hausman and Newey’s (2013) recent finding that for price change of a continuous good, averages of money-metric welfare measures are only set-identified under unrestricted heterogeneity. On the practical side, our results express the distribution of welfare as closed-form functionals of average demand and are thus easy to compute in real applications.

References


Appendix

Proof of proposition 1:

We denote the individual level EV $S_{EV}^{E}(y, p_0, p_1, \eta)$ by simply $S(y, \eta)$ to avoid cumbersome notation in the proof. Also for ease of reference, we rewrite condition (4) again:

$$
\max \{U_0(y - S(y, \eta), \eta), U_1(y - S(y, \eta) - p_0, \eta)\} = \max \{U_0(y, \eta), U_1(y - p_1, \eta)\}.
$$

Proof. By monotonicity of $U_1(\cdot, \eta)$ and $U_0(\cdot, \eta)$, we must have that $S(y, \eta) \geq 0$ in order for (4) to hold.

Now, in case (i), $0 \leq U_0(y, \eta) - U_1(y - p_0, \eta)$. So if $S(y, \eta) > 0$, then

$$
\max \{U_0(y - S(y, \eta), \eta), U_1(y - S(y, \eta) - p_0, \eta)\} < \max \{U_0(y, \eta), U_1(y - p_0, \eta)\} = U_0(y, \eta) \leq \max \{U_0(y, \eta), U_1(y - p_1, \eta)\},
$$

contradicting (4). This implies that $S(y, \eta) = 0$.

Now, consider case (ii). Given the restriction $0 \leq U_0(y, \eta) - U_1(y - p_1, \eta)$, the RHS of (4) is $U_0(y, \eta)$. Therefore from (4), $S(y, \eta)$ must satisfy

$$
U_0(y, \eta) = \max \{U_1(y - S(y, \eta) - p_0, \eta), U_0(y - S(y, \eta), \eta)\}. \quad (13)
$$

Now, equation (13) is equivalent to

$$
U_0(y, \eta) = U_1(y - S(y, \eta) - p_0, \eta). \quad (14)
$$

To see why, suppose the first term on the RHS of (13) is smaller than the second, i.e.,

$$
U_1(y - S(y, \eta) - p_0, \eta) < U_0(y - S(y, \eta), \eta). \quad (15)
$$

Then equation (13) implies

$$
U_0(y, \eta) = U_0(y - S(y, \eta), \eta)
$$

implying $S(y, \eta) = 0$ by strict monotonicity of $U_0(\cdot, \eta)$. But then, the first term on the RHS of equation (13), viz., $U_1(y - S(y, \eta) - p_0, \eta)$ equals $U_1(y - p_0, \eta)$ while the second term equals
$U_0(y, \eta)$; but because we know that $U_1(y - p_0, \eta) \geq U_0(y, \eta)$ (we are in case (ii) of the proposition), the inequality (15) is violated. Therefore, we must have that

$$U_1(y - S(y, \eta) - p_0, \eta) \geq U_0(y - S(y, \eta), \eta),$$

so that the maximum on the RHS of equation (13) must equal $U_1(y - S(y, \eta) - p_0, \eta)$, whence the conclusion (14) follows.

From (14), using monotonicity of $U_1(\cdot, \eta)$, we have that

$$S(y, \eta) = y - p_0 - U_1^{-1}(U_0(y, \eta), \eta).$$

Finally, consider case (iii): Given the restriction, $0 > U_0(y, \eta) - U_1(y - p_1, \eta)$, the RHS of (4) is $U_1(y - p_1, \eta)$. Now, suppose the LHS of (4) is $U_0(y - S(y, \eta), \eta)$. But since $S(y, \eta) \geq 0$, we must have that

$$U_0(y - S(y, \eta), \eta) \leq U_0(y, \eta) < U_1(y - p_1, \eta) = U_0(y - S(y, \eta), \eta), \text{ by (4)},$$

a contradiction. Therefore, the LHS of (4) must be $U_1(y - S(y, \eta) - p_0, \eta)$ and therefore by (4),

$$U_1(y - S(y, \eta) - p_0, \eta) = U_1(y - p_1, \eta),$$

whence, by strict monotonicity of $U_1(\cdot, \eta)$, we get that $S(y, \eta) = p_1 - p_0$. 

Proof of proposition 2:

Proof. The compensation must be non-negative and no larger than $p_1 - p_0$; otherwise (4) will be violated. EV is zero for those not purchasing at $p_0$ and hence EV has a point mass equal to the probability of no purchase at $p_0$, which is given by $1 - \bar{q}(p_0, y)$. So the only nontrivial step is for $0 < a < p_1 - p_0$. This case corresponds to case (ii) of proposition 1. Accordingly, for $0 < a < p_1 - p_0$,
the probability of compensation not exceeding \(a\) is given by

\[
\Pr(S(y, \eta) = 0, 0 \leq U_0(y, \eta) - U_1(y - p_0, \eta)) \\
+ \Pr \left( y - p_0 - U_1^{-1}(U_0(y, \eta), \eta) \leq a, \\
U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y, \eta) - U_1(y - p_1, \eta) \right)
\]

\[
= 1 - \bar{q}(p_0, y) \\
+ \Pr \left( 0 \leq U_0(y, \eta) - U_1(y - p_0 - a, \eta), \\
U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y, \eta) - U_1(y - p_1, \eta) \right)
\]

\[
= 1 - \bar{q}(p_0, y) \\
+ \Pr(U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y, \eta) - U_1(y - p_0 - a, \eta)), \text{ since } a < p_1 - p_0
\]

\[
= 1 - \bar{q}(p_0, y) \\
+ (p_1 - p_0) \Pr(U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y, \eta) - U_1(y - p_0 - a, \eta))
\]

This finishes the proof. ■

Proof of Corollary 1

From (5), the expected value of \(EV\) is obtained by integrating the individual \(EV\) w.r.t. the density of \(EV\) over the continuous part of its distribution and then adding on the discrete mass point times the probability mass which gives

\[
\mu^{EV}(y, p_0, p_1) = \int_0^{p_1 - p_0} \frac{\partial}{\partial a} \{1 - \bar{q}(a + p_0, y)\} \, da + (p_1 - p_0) \Pr(EV = p_1 - p_0)
\]

\[
= - \int_0^{p_1 - p_0} \frac{\partial}{\partial a} \{\bar{q}(a + p_0, y)\} \, da + (p_1 - p_0) \times \bar{q}(p_1, y)
\]

\[
\text{integrating by parts}
\]

\[
= (p_1 - p_0) \times \bar{q}(p_1, y) + \int_0^{p_1 - p_0} \bar{q}(a + p_0, y) \, da + (p_1 - p_0) \times \bar{q}(p_1, y)
\]

\[
= \int_0^{p_1 - p_0} \bar{q}(a + p_0, y) \, da \\
= \int_0^{p_1} \bar{q}(a + p_0, y) \, da
\]

Thus we get (6).

Proof of proposition 3
We denote the individual level CV \( S^{CV} (y, p_0, p_1, \eta) \) by simply \( S (y, \eta) \) to avoid cumbersome notation in the proof. Also for ease of reference, we rewrite condition (7) again:

\[
\max \{ U_0 (y + S (y, \eta), \eta), U_1 (y + S (y, \eta) - p_1, \eta) \} = \max \{ U_0 (y, \eta), U_1 (y - p_0, \eta) \}.
\]

**Proof.** First observe that by (7), we must have that \( S (y, \eta) \geq 0 \). Otherwise, the LHS of (7) must be strictly smaller than the RHS, by the monotonicity of \( U_0 (\cdot) \) and \( U_1 (\cdot) \). Now consider the following cases.

**Case (i):** \( 0 \leq U_0 (y, \eta) - U_1 (y - p_0, \eta) \)

Since \( 0 \leq U_0 (y, \eta) - U_1 (y - p_0, \eta) \), then the RHS of (7) is \( U_0 (y, \eta) \). If \( S (y, \eta) > 0 \), then the first term of LHS of (7) must be strictly larger than \( U_0 (y, \eta) \), by strict monotonicity of \( U_0 (\cdot, \eta) \). This would imply that the LHS of (7) must be strictly larger than \( U_0 (y, \eta) \) — a contradiction. Therefore, in this case, we must have \( S (y, \eta) = 0 \). Intuitively, this means that those \( \eta \) who were not buying at the initial price \( p_0 \) do not need to be compensated.

Now, suppose case (i) does not hold, so that RHS maximum is in fact \( U_1 (y - p_0, \eta) \), i.e., \( U_0 (y, \eta) - U_1 (y - p_0, \eta) < 0 \). This corresponds to those \( \eta \)'s who buy the good at price \( p_0 \). Now, there are two possibilities regarding which term is the maximum in the LHS of (7) — case (ii) corresponds to when the maximum is the first term and case (iii) to when the maximum is the second term.

**Case (ii):** Accordingly, first assume that the LHS maximum is \( U_0 (y + S (y, \eta), \eta) \). Then \( S (y, \eta) \) must satisfy

\[
U_0 (y + S (y, \eta), \eta) = U_1 (y - p_0, \eta) \implies S (y, \eta) = U_0^{-1} (U_1 (y - p_0, \eta), \eta) - y. \tag{16}
\]

In order for this to simultaneously satisfy that the LHS maximum of (7) is \( U_0 (y + S (y, \eta), \eta) \), we need

\[
\begin{align*}
\text{substituting } S(y, \eta) \text{ from (16)} \quad U_0 (y + S (y, \eta), \eta) & \geq U_1 (y + S (y, \eta) - p_1, \eta) \\
U_1 (y - p_0, \eta) & \geq U_1 (y + S (y, \eta) - p_1, \eta) \\
\implies y - p_0 & \geq y + S (y, \eta) - p_1 \\
\implies p_1 - p_0 & \geq S (y, \eta) = U_0^{-1} (U_1 (y - p_0, \eta), \eta) - y \\
\implies U_0 (y + p_1 - p_0, \eta) & \geq U_1 (y - p_0, \eta) \\
\implies 0 & \leq U_0 (y + p_1 - p_0, \eta) - U_1 (y - p_0, \eta).
\end{align*}
\]
Thus we arrive at the conclusion of case (ii), viz.,

\[ U_0(y) - U_1(y-p_0) < 0 \leq U_0(y+p_1-p_0, \eta) - U_1(y-p_0, \eta) \] and

\[ S(y, \eta) = U_0^{-1}(U_1(y-p_0, \eta), \eta) - y. \]

**Case (iii):** Finally, consider the remaining case where the maximum of the LHS of (7) is \( U_1(y + S(y, \eta) - p_1, \eta) \), whence we have

\[ U_1(y + S(y, \eta) - p_1, \eta) = U_1(y - p_0, \eta) \]

\[ \Rightarrow y + S(y, \eta) - p_1 = y - p_0 \]

\[ \Rightarrow S(y, \eta) = p_1 - p_0. \]

Replacing \( S(y, \eta) \) in the LHS of (7), in order to be consistent with our assumption that the LHS maximum is \( U_1(y + S(y, \eta) - p_1, \eta) \), we must have that

\[ U_1(y + S(y, \eta) - p_1, \eta) \geq U_0(y + S(y, \eta), \eta) \]

\[ \Leftrightarrow U_1(y - p_0, \eta) \geq U_0(y + p_1 - p_0, \eta) \]

which is precisely case (iii) of the proposition, viz., \( U_1(y - p_0, \eta) \geq U_0(y + p_1 - p_0, \eta) \).

---

**Proof of proposition 4**

**Proof.** First recall from (2) that for any \( a > 0 \), we have that

\[ q(p_0, y) = \Pr(0 > U_0(y, \eta) - U_1(y - p_0, \eta)), \]

\[ \bar{q}(a + p_0, y + a) = \Pr(0 > U_0(y + a, \eta) - U_1(y - p_0, \eta)), \]

where the probabilities are computed w.r.t. the distribution of \( \eta \).

Now, from proposition 3 and its proof, it is clear that for any \( \eta \), the compensation is non-negative and it equals zero for those not buying at price \( p_0 \). Hence the CV has no mass below 0 and a point mass of \( 1 - \bar{q}(p_0, y) \) at 0. Also, it is clear that the compensation cannot exceed \( p_1 - p_0 \) for any \( \eta \). Otherwise, (7) will be violated. Therefore the C.D.F. of CV must reach 1 at \( p_1 - p_0 \). So the only nontrivial case is \( 0 < a < p_1 - p_0 \). This corresponds to case (ii) of proposition 3. Accordingly, for \( 0 < a < p_1 - p_0 \), the probability of the compensation being no larger than \( a \) is
Substituting in (17), we get that for
\[
S(y, \eta) = 0, \quad U_0(y, \eta) - U_1(y - p_0, \eta) \geq 0
\]
\[
\Pr \left\{ \begin{array}{l}
U_0^{-1}(U_1(y - p_0, \eta), \eta) - y < a,
U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y + p_1 - p_0, \eta) - U_1(y - p_0, \eta)
\end{array} \right\}
\]
\[
= 1 - \tilde{q}(p_0, y)
\]
\[
+ \Pr \left\{ \begin{array}{l}
U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y + p_1 - p_0, \eta) - U_1(y - p_0, \eta)
\end{array} \right\}.
\] (17)

Now,
\[
\Pr \left( \begin{array}{l}
U_0^{-1}(U_1(y - p_0, \eta), \eta) - y < a,
U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y + p_1 - p_0, \eta) - U_1(y - p_0, \eta)
\end{array} \right)
\]
\[
= \Pr \left( \begin{array}{l}
0 < U_0(y + a, \eta) - U_1(y - p_0, \eta),
U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y + p_1 - p_0, \eta) - U_1(y - p_0, \eta)
\end{array} \right)
\]
\[
= \Pr (U_0(y, \eta) - U_1(y - p_0, \eta) < 0 \leq U_0(y + a, \eta) - U_1(y - p_0, \eta)), \text{ since } a < p_1 - p_0
\]
\[
= \Pr (0 \leq U_0(y + a, \eta) - U_1(y - p_0, \eta)) - \Pr \left( \begin{array}{l}
U_0(y, \eta) - U_1(y - p_0, \eta) \geq 0,
U_0(y + a, \eta) - U_1(y - p_0, \eta) \geq 0
\end{array} \right)
\]
\[
= \Pr (0 < U_0(y + a, \eta) - U_1(y - p_0, \eta)) - \Pr (0 \leq U_0(y, \eta) - U_1(y - p_0, \eta)) \text{ since } a \geq 0
\]
\[
= \Pr (0 < U_0(y + a, \eta) - U_1(y + a - (a + p_0), \eta)) - \Pr (0 \leq U_0(y, \eta) - U_1(y - p_0, \eta))
\]
\[
= \Pr (0 > U_0(y, \eta) - U_1(y - p_0, \eta)) - \Pr (0 > U_0(y + a, \eta) - U_1(y + a - (a + p_0), \eta))
\]
\[
= \tilde{q}(p_0, y) - \tilde{q}(a + p_0, y + a).
\]

Substituting in (17), we get that for $0 < a < p_1 - p_0$,
\[
\Pr \{ S(y, \eta) \leq a \}
\]
\[
= 1 - \tilde{q}(p_0, y) + \tilde{q}(p_0, y) - \tilde{q}(a + p_0, y + a)
\]
\[
= 1 - \tilde{q}(a + p_0, y + a),
\]
as desired. ■

**Proof of Corollary 2**
Proof. Essentially follows from the c.d.f. of CV by integrating over the continuous part and adding the discrete mass

\[
\mu_{CV}(y; p_0, p_1) = \int_{0}^{p_1-p_0} a \times \frac{\partial}{\partial a} \{1 - q(a + p_0, y + a)\} \, da + (p_1 - p_0) \Pr(S(y, \eta) = p_1 - p_0)
\]

\[
= - \int_{0}^{p_1-p_0} a \frac{\partial}{\partial a} q(a + p_0, y + a) \, da + (p_1 - p_0) \Pr(S(y, \eta) = p_1 - p_0)
\]

\[
= -a q(a + p_0, y + a) \bigg|_{0}^{p_1-p_0} + \int_{0}^{p_1-p_0} q(a + p_0, y + a) \, da
\]

\[\] + (p_1 - p_0) \times q(p_1, y + p_1 - p_0), \text{ integrating by parts}

\[
= \int_{0}^{p_1-p_0} q(a + p_0, y + a) \, da
\]

\[
= \int_{p_0}^{p_1} q(z, y + z - p_0) \, dz, \text{ substituting } z = a + p_0.
\]