# Incorporating Social Welfare in Program-Evaluation and Treatment Choice*

Debopam Bhattacharya[†] Tatiana Komarova

University Of Cambridge[‡] University of Manchester

October 25, 2023

**Abstract**

We introduce a notion of money-metric social welfare for discrete choice, under unrestricted heterogeneity and income-effects. It is the maximized indirect utility under normalization of the outside option. It also equals the amount of income necessary to achieve a given level of utility, while certain choices are prohibited. We show that the distribution of this quantity is non-parametrically identified as a closed-form functional of average structural demand for the outside option, making it useful for cost-benefit analysis and optimal targeting. An illustration with private tuition subsidies in India shows that the income-path of usage-maximizing subsidies differs significantly from welfare-maximizing ones.

Keywords: Discrete Choice, Unobserved Heterogeneity, Nonparametric Identification, Social Welfare, Indirect Utility, Cost-Benefit Analysis, Policy Interventions

JEL Codes C14 C25 D12 D31 D61 D63

*We are grateful to Meredith Crowley, Peter Hammond, Arthur Lewbel and Albert Park for discussions related to the topic of this paper.

†Bhattacharya acknowledges financial support from the European Research Council via a Consolidator Grant EDWEL, Project number 681565.

‡Address for correspondence: debobhatta@gmail.com

# 1 Introduction

Data-driven policy analysis has become central to modern economics, and has produced a large body of research on both program-evaluation cf. Heckman-Vytlacil 2007, Imbens-Wooldridge 2009, and optimal targeting of interventions cf. Manski 2004. Both these literatures have exclusively focused on functionals of *outcome* distributions as the key object of interest, and bypassed the classic public economics question of measuring program effects on unobservable *utilities* of individuals. For example, a college financial aid program would typically be evaluated in the econometric tradition via its treatment effect on aggregate enrolment or future earnings, etc., and the treatment choice problem would address the question of how to target a limited amount of subsidy funds to maximize this aggregate cf. Bhattacharya and Dupas 2012, Kitagawa and Tetenov 2018 etc. This approach ignores the question of how much and how differently individuals themselves value the subsidy – determined by their willingness-to-pay for college – and what is the subsidy's effect on aggregate utility, weighted by the social planner's distributional preferences. In particular, those subsidy eligibles who would attend college irrespective of the subsidy would contribute nothing to the average treatment effect, but their savings from subsidized tuition would raise their utility without changing their attendance behavior. Secondly, in many practical settings, multiple related outcomes of policy interest are likely to be affected by a single intervention. For example, a price subsidy for mosquito-nets (cf. Dupas 2014) can be evaluated in terms of aggregate take-up of nets, incidence of malaria, school absence of children, lost earnings and so forth. The utility based approach provides a natural way to aggregate these separate effects via how they determine the households' overall willingness-to-pay for the mosquito-net. Third, price-interventions for redistributions are often politically motivated. Efficiency-costs of such policies to society are therefore important metrics of political assessment.

Indeed, in public economics, cost-benefit analysis of an intervention was traditionally conducted by comparing the expenditure on it with the change it brings about in aggregate indirect utility, cf. Bergson 1938, Samuelson 1947 and Mirrlees 1971. However, when bringing these concepts to data, this literature ignored unobserved heterogeneity and imposed arbitrary functional-form restrictions on the utility functions of 'representative' consumers who were assumed to vary solely in terms of observables cf. Deaton 1984 and Ahmad and Stern 1984. Later work such as Feldstein 1999 and Saez 2001 have shown that in labour-supply models with consumption-leisure trade-off by heterogeneous agents, the optimal income-tax rate depends on individual heterogeneity via certain aggregates only, such as the average elasticity of taxable income w.r.t. the marginal tax rate, where the taxable income distribution is endogenously determined with the tax rate. Conceptually speaking, these results are *not* nonparametric identification results expressing the object of interest (consumer welfare, optimal tax-schedule etc.) in terms of quantities directly estimable from the data without functional-form assumptions about unobservables, cf. Saez 2001 page 219 and Section 5. Manski 2014 derives bounds on optimal income-tax schedule when consumers have heterogeneous Cobb-Douglas preferences.

The present paper proposes a notion of money-metric utility as an alternative goal for the evaluation and targeting of policy-interventions.[1] In the practically important setting of multinomial choice, the distribution of these utilities induced by unobserved heterogeneity can be expressed as closed-form functionals of solely the choice-probability for the outside option. This result, described in Theorem 1 below, assumes no knowledge of functional-form of utilities, nature of income-effects and dimension/distribution of unobserved preference heterogeneity. Conceptually, theorem 1 makes it unnecessary to estimate underlying preferences and their heterogeneity to perform cost-benefit analysis of interventions and infer the optimal way to target policies. Further, the knowledge of indirect utility distribution permits measurement of the efficiency-loss required to ensure a desired average outcome; for instance, in our mosquito-net example above, one can calculate the monetary value of the subsidy-induced market distortion (excess burden) necessary to reach an adoption target of say 80%, thus providing a theoretically justified numerical measure of the equity-efficiency

[1] As such, the paper does *not* contribute to the methodology of program-evaluation or treatment assignment.

trade-off involved.

An alternative way to evaluate welfare is via the expenditure function based Hicksian measures, viz. the equivalent and compensating variation (EV/CV). These are hypothetical income adjustments necessary to maintain individual utilities. While useful for measuring the distribution of change in individual utility, *adding* Hicksian measures across consumers to obtain a measure of social welfare change involves judgments that are known to be conceptually problematic. These include (i) an implicit assumption of a constant social marginal utility of income, i.e. that an additional dollar is valued by society identically no matter whether a rich or a poor person gets it, cf. Blackorby and Donaldson 1988, Dreze 1998, Banks et al 1996, (ii) ranking alternative interventions by their associated Hicksian compensations amounts to ranking based on *changes*, as opposed to *levels,* of individual satisfaction cf. Slesnick 1998, Chipman and Moore 1990, and (iii) comparing allocations via the aggregate compensation principle, as implied by adding EV/CV across consumers, leads to Scitovszky (1941) reversals, where two distinct allocations can both dominate and be dominated by each other in terms of social welfare. These problems with aggregate compensation criteria have led to widespread use of Bergson-Samuelson aggregate indirect utility for applied welfare analysis in public finance.[2] The 'logsum' formula for consumer surplus (cf. Train 2003, Chap 3.5), routinely used for welfare- analysis in empirical IO, is precisely the average indirect utility in the parametric multinomial logit (BLP) model. Interestingly, as shown below, there is a theoretical link in the discrete choice case between changes in aggregate indirect utility and Hicksian compensations for removal of alternatives. Further, based on the aggregate utility, one can define a microfounded measure of 'welfare inequality', analogous to the Atkinson index of income-inequality. Unlike CV/EV, however, the aggregate utility requires a normalization for empirical content, as is implicitly assumed in public finance. In the restrictive but popular special case of quasilinear preferences, under which demand is income-invariant,

[2]Aggregate indirect utility embodies interpersonal comparisons of preferences which, as noted by Hammond 1990 "... have to be made if there is to be any satisfactory escape from Arrow's impossibility theorem, with its implication that individualistic social choice has to be dictatorial ... or else that is has to restrict itself to solely recommending Pareto improvements."

the change in indirect utility approach coincides with the Hicksian/Marshallian ones.

Ahmad and Stern 1984, Mayshar 1990, Hendren and Sprung-Keyser 2020 and Hendren and Finkelstein 2020 investigated social cost-benefit analysis for *marginal* interventions, using the concept of 'marginal value of public funds' (MVPF), defined as the ratio of beneficiaries' marginal willingness-to-pay for a policy change at the status-quo to its marginal cost for the government. This approach does not cover non-marginal interventions and does not clarify how to account for unobserved preference heterogeneity across individuals targeted by such non-marginal interventions. Obviously, the larger the intervention, the poorer the resulting approximation by marginal cost-benefit analysis (see our application below). It is also not obvious how one would use the status-quo MVPF for optimal targeting of interventions. Garcia and Heckman 2022 propose the net social benefit for ranking different programs based on their opportunity costs. Our proposed measure of money metric social welfare offers a global (as opposed to local or marginal) alternative whose distribution can potentially be identified from demand data without requiring identification/estimation of latent unobservables characterizing individual preferences and heterogeneity. Practical calculation of our measure does, however, require estimation/extrapolation of average demand over the entire support of price and income. This can be difficult in practice due to data limitations, and a semiparametric or parametric specification of demand would be a feasible middle ground; alternatively, one can obtain bounds on welfare measures, dictated by the observed support of price and income in the data at hand. Endogeneity of price and/income, generic to almost all empirical demand analysis exercise, is also an important potential issue here. Significant literatures exist on this problem in both econometrics (cf. Newey 1987, Blundell and Powell 2004, 2009) and empirical IO (cf. Berry and Haile 2021). While endogeneity of price has received more attention (cost-factors being a popular choice of instruments), income endogeneity, say due to joint determination of individual income and preferences, remains challenging in micro-level analysis of demand. As in studies of labour supply, non-labour income or land-holding can potentially serve as possible instruments for income here.

Given that the main contribution of this paper is to propose a micro-founded concept of social welfare under preference heterogeneity, we do not explore the generic empirical issue of endogeneity in demand estimation any further here. Lastly, calculating the share of the outside option, the key ingredient of our welfare calculation, also requires the sample to include those individuals who buy none of the inside alternatives. Thus the typical dataset where our method can be used is a household survey, as opposed to, say, sellers' records.

The next section outlines the theory, presents our main result and provides related discussions including optimal treatment allocation, Section 3 outlines identification and estimation of our objects of interest, Section 4 presents an empirical illustration and Section 5 concludes.

# 2 Theory

This section introduces our key theoretical result, and outlines its main implications for social welfare analysis.

## 2.1 Setup and Main Result

There is a population of heterogeneous individuals, each facing a choice between $J+1$ exclusive, indivisible options. Examples include whether to attend college, whether to adopt a health-product, choice of phone-plan etc. Let $N$ represent the quantity of numeraire which an individual consumes in addition to the discrete good. If the individual has income $Y=y$, and faces a price $P_j=p_j$ for the $j$th option, then the budget constraint is $N+\sum_{j=1}^{J} Q_j p_j = y$, $\sum_{j=0}^{J} Q_j = 1$ where $Q_j \in \{0,1\}$, $j=0,...,J$ represents the discrete choice, with 0 denoting the outside option (set $p_0=0$). Individuals derive satisfaction from both the discrete good as well as the numeraire. Upon buying the $j$th option, an individual derives utility from it and from numeraire $y-p_j$, denoted by $U_j\left(y-p_j,\eta\right)$, where $\eta$ denotes unobserved (by us) preference heterogeneity of unspecified dimension and distribution. Upon not buying any of the $J$ alternatives, she enjoys utility from her outside option and the full numeraire $y$, given by $U_0(y,\eta)$. Observable characteristics of consumers and/or the alternatives are suppressed here for notational simplicity. We assume strict non-satiation in the numeraire, i.e. that $U_j(\cdot,\eta)$ and $U_0(\cdot,\eta)$ are strictly increasing in their first argument for each realization of $\eta$. On each budget set defined by the price vector $\mathbf{p} \equiv (p_1,...,p_J)$ and consumer income $y$, there is a 'structural' probability of choosing option $j$, denoted by $q_j(\mathbf{p},y)$; that is, if each member of the entire population were offered income $y$ and price $\mathbf{p}$, then a fraction $q_j(\mathbf{p},y)$ would buy the $j$th alternative, with $q_0(\mathbf{p},y)$ denoting not buying any of the $J$ alternatives, i.e.

$$q_j(\mathbf{p},y) = \int 1\left\{U_j(y-p_j,\eta) > \max_{\substack{k\in\{0,1,...,J\}\\ k\neq j}} U_k(y-p_k,\eta)\right\} dF(\eta), \tag{1}$$

where $F(\cdot)$ denotes the *marginal* distribution of $\eta$. This object is 'structural' in the sense that the integration is with respect to the marginal distribution of $\eta$, and not with respect to the conditional distribution of $\eta$, given $\mathbf{p}$ and $y$. It is thus analogous to the $Y_0/Y_1$ potential outcomes notion in the program evaluation literature, where the treatment of interest is binary. Note further that the above setup allows for completely general unobserved heterogeneity and income effects. Finally, the indirect utility function is given by

$$W(\mathbf{p},y,\eta) = \max\{U_0(y,\eta), U_1(y-p_1,\eta), ..., U_J(y-p_J,\eta)\}.$$

Note that this function is weakly decreasing in each price and strictly increasing in income. Therefore, a concave functional of $W(\mathbf{p},y,\eta)$ will correspond to assigning larger weights to those with lower utility and income.

**Normalization**: Since a monotone transformation of a utility function represents the same ordinal preferences and leads to the same choice, we need to normalize one of the alternative-specific utility functions in order to give empirical content to the indirect utility function. Since $U_0(\cdot,\eta)$ is supposed to be strictly increasing (non-satiated) and continuous in the numeraire, it is thus invertible. Then $U_j(y-p_j,\eta) \geq U_0(y,\eta)$ if and only if

$U_0^{-1}(U_j(y-p_j,\eta),\eta) \geq y$; also $U_0^{-1}(U_j(y-p_j,\eta),\eta) \geq U_0^{-1}(U_k(y-p_k,\eta),\eta)$ if and only if $U_j(y-p_j,\eta) \geq U_k(y-p_k,\eta)$. Therefore, $\mathcal{U}_0(y,\eta) \equiv y$ and $\mathcal{U}_j(y-p_j,\eta) \equiv U_0^{-1}(U_j(y-p_j,\eta),\eta)$ is an equivalent normalization of utilities representing exactly the same individual preferences as $\{U_j(y-p_j,\eta)\}$, $j=1,...,J$ and $U_0(y,\eta)$. This is analogous to the empirical IO convention that utility from the outside good be normalized to zero. Arbitrary functional-form specification for utilities (e.g. CES/CARA/CRRA) in traditional structural modelling assumes much more, in addition to an implicit normalization. Further, welfare-changes often result from removing/adding inside alternatives, normalizing utility of the *outside* option which remains unaffected by such changes, therefore seems natural. It also leads to an interpretation of indirect utility as a compensated income (see Sec 2.3 below). So, from now, we will work under this normalization.

**Theorem 1** *In the above setup, assume that for each $\eta$, the function $U_0(\cdot,\eta)$ is continuous and strictly increasing, and the utility differences $U_j(y-p_j,\eta)-U_0(y,\eta)$ are continuously distributed for $j=1,\ldots,J$, so the relevant ties occur with probability 0. Then the marginal distribution of indirect utility induced by the distribution of $\eta$ at fixed values of $\mathbf{p}$ and $y$ is nonparametrically identified from average demand.*

**Proof.** Using the normalization $\mathcal{U}_j(y-p_j,\eta) \equiv U_0^{-1}(U_j(y-p_j,\eta),\eta)$ and $\mathcal{U}_0(y,\eta) \equiv y$,

$$W(\mathbf{p},y,\eta) = \max\left\{y, U_0^{-1}(U_1(y-p_1,\eta),\eta), ..., U_0^{-1}(U_J(y-p_J,\eta),\eta)\right\}. \tag{2}$$

We wish to compute the distribution of $W(\mathbf{p},y,\eta)$ induced by the marginal distribution of $\eta$, for fixed $\mathbf{p}, y$. Since $W(\mathbf{p},y,\eta) \geq y$ a.s. by (2), we take $c > y$, and note that

$$\begin{aligned}
& \Pr\left[\max\left\{U_0^{-1}(U_1(y-p_1,\eta),\eta), ..., U_0^{-1}(U_J(y-p_J,\eta),\eta)\right\} \leq c\right] \\
= \;& \Pr\left[\max\left\{U_1(y-p_1,\eta), ..., U_J(y-p_J,\eta)\right\} \leq U_0(c,\eta)\right], \text{ since } U_0(\cdot,\eta) \text{ cont. and strictly} \nearrow \\
= \;& \Pr\left[\max\left\{U_1(c-(c-y+p_1),\eta), ..., U_J(c-(c-y+p_J),\eta)\right\} \leq U_0(c,\eta)\right] \\
= \;& q_0(c-y+p_1, ..c-y+p_J, c).
\end{aligned}$$

Therefore, the C.D.F. of $W(\mathbf{p}, y, \eta)$ generated by randomness in $\eta$ is given by

$$\Pr\left[W(\mathbf{p}, y, \eta) \leq c\right] = \begin{cases} 0 \text{ if } c < y, \\ q_0\left(c - y + p_1, ..c - y + p_J, c\right) \text{ if } c \geq y. \end{cases} \tag{3}$$

It is easily verified (see appendix) that $U_0(\cdot, \eta)$ being strictly increasing implies that the C.D.F. is non-decreasing in $c$. ■

To interpret (3) intuitively, note that for $c \geq y$, (3) is equivalent to

$$\Pr\left[W(\mathbf{p}, y, \eta) > c\right] = 1 - q_0\left(c - y + p_1, ..c - y + p_J, c\right).$$

Indeed, if $c \geq y$, the only $\eta$'s who attain a welfare value larger than $c$ must buy one of $\{1, ..., J\}$, since choosing 0 yields $y \leq c$, which explains the functional form $1 - q_0(\cdot)$. The arguments $(c - y + p_1, ..c - y + p_J, c)$ arise from the facts that reaching utility larger than $c$ requires that choosing the maximand $j \in \{1, ..., J\}$ and ending up with numeraire $y - p_j$ must yield higher utility than choosing 0 at income $c$ which yields utility $c$.

Lastly, note that by definition, $W(\mathbf{p}, y, \eta)$ is measured in units of money, which will be useful both in cost-benefit analysis and in comparison with Hicksian compensation.

## 2.2 Social Welfare Calculations

Theorem 1 gives us a way to calculate 'aggregate social welfare' (Bergson 1938, Atkinson 1970). Given price $\mathbf{p}$ for individuals with income $y$ and for an individual with unobserved heterogeneity $\eta$, the welfare level is given by $\frac{W(\mathbf{p}, y, \eta)^{1-\varepsilon}}{1-\varepsilon}$, where $0 \leq \varepsilon < 1$ denotes the planner's inequality aversion parameter. Therefore, the distribution of social welfare at fixed income $y$ across consumers follows from (3). In particular, the average (over unobserved heterogeneity) of social welfare (ASW henceforth) at income $y$ is

$$\mathcal{W}^{\varepsilon}(\mathbf{p}, y) \equiv E_{\eta}\left\{\frac{W(\mathbf{p}, y, \eta)^{1-\varepsilon}}{1-\varepsilon}\right\}, \tag{4}$$

where $E_\eta$ denotes expectation taken with respect to the marginal distribution of $\eta$. Note that (4) takes the form of an exact counterpart for measuring income inequality using compensated – instead of ordinary – income. For example, one can compute the analogs of the Gini coefficient or the Atkinson index for welfare inequality based on the distribution of $W(\mathbf{p}, y, \eta)$ simply by replacing ordinary income by the compensated income as defined in the LHS of (6) below. In the Appendix, we provide detailed steps for calculation of (4).

**Optimal Targeting Problem**: The optimal subsidy targeting problem maximizes aggregate welfare subject to a budget constraint on subsidy spending. Suppose in our multinomial setup, the planner considers subsidizing alternative 1, whose unsubsidized market price is $\bar{p}$. Let $M$ denote the aggregate subsidy budget, expressed in per capita terms, $F_Y(\cdot)$ the marginal distribution of income in the population, $\sigma(y)$ the amount of subsidy that a household with income $y$ will be entitled to, $\mathcal{T}$ denote the set of politically/practically feasible targeting rules $\sigma(\cdot)$, and $C(y, \sigma(y))$ the cost per capita of offering subsidy $\sigma(y)$ to individuals whose income is $y$; for example, in the multinomial case with alternative 1 being subsidized, $C(y, \sigma(y))$ equals $\int \sigma(y) \times q_1(\bar{p} - \sigma(y), \mathbf{p}_{-1}, y)\, dF_Y(y)$, where $\mathbf{p}_{-1}$ is the price vector for all the alternatives excluding option 1. Then the optimal subsidy solves

$$\arg\max_{\sigma(\cdot)\in\mathcal{T}} \int \mathcal{W}(\bar{p} - \sigma(y), \mathbf{p}_{-1}, y)\, dF_Y(y) \text{ s.t. } \int C(y, \sigma(y))\, dF_Y(y) = M. \tag{5}$$

Taxes can be incorporated into the analysis by allowing $\mathcal{T}$ to contain functions that take on negative values. In particular, a revenue-neutral welfare maximizing rule that taxes the rich, i.e. $\sigma(\cdot) < 0$ and subsidizes the poor i.e. $\sigma(\cdot) > 0$, will solve (5) with $M = 0$.

**Comparison with Income transfer**: Price subsidies, as opposed to a pure income transfer, entails a deadweight loss due to the distortionary effect of the price-intervention on behavior. In the context of problem (5), the aggregate value of this excess-burden can be computed as the difference between $M$ and the value function of problem (5). Indeed, the worldwide discussion of a universal basic income (cf. Banerjee et al 2019) can be informed by calculating the deadweight loss of various price-subsidies that the UBI would seek to replace.

**Hicksian Interpretation:** Our measure (2) can be interpreted as the Hicksian CV corresponding to elimination of all inside alternatives. To see this, consider an initial situation where none of the inside alternatives $1, ..., J$ is available ($p_1 = p_2 = ... = p_J = \infty$, denoted by the price vector $\infty_J$) and an the eventual situation when they become available at price vector $\mathbf{p}$. Using the utility functions $\mathcal{U}_j(y - p_j, \eta) = U_0^{-1}(U_j(y - p_j, \eta), \eta)$ and $\mathcal{U}_0(y, \eta) = y$, the former indirect utility is $y$ since other options are unavailable, and the latter indirect utility is

$$\max\left\{y, U_0^{-1}(U_1(y - p_1, \eta), \eta), ..., U_0^{-1}(U_J(y - p_J, \eta), \eta)\right\}.$$

Then the compensating variation $CV(y, \mathbf{p}, \infty_J, \eta)$ for this change solves

$$\begin{aligned} y + CV(y, \mathbf{p}, \infty_J, \eta) &= \max\left\{y, U_0^{-1}(U_1(y - p_1, \eta), \eta), ..., U_0^{-1}(U_J(y - p_J, \eta), \eta)\right\} \\ &= W(\mathbf{p}, y, \eta), \text{ by (2).} \end{aligned} \tag{6}$$

Thus the indirect utility at price $\mathbf{p}$ and income $y$ equals the compensated income at $y$ that equates individual utility when none of the alternatives $1, ..., J$ was available to the utility when they become available at price $\mathbf{p}$. It follows from (6) that the difference in individual indirect utility between two prices $\mathbf{p}^0$ and $\mathbf{p}^1$ equals

$$W\left(\mathbf{p}^1, y, \eta\right) - W\left(\mathbf{p}^0, y, \eta\right) = CV\left(y, \mathbf{p}^1, \infty_J, \eta\right) - CV\left(y, \mathbf{p}^0, \infty_J, \eta\right). \tag{7}$$

Note however that

$$W\left(\mathbf{p}^1, y, \eta\right) - W\left(\mathbf{p}^0, y, \eta\right) \neq -CV\left(y, \mathbf{p}^0, \mathbf{p}^1, \eta\right), \tag{8}$$

(proved in the Appendix); so asking if $\mathbf{p}^1$ is worse than $\mathbf{p}^0$ on the basis of ASW is not the same as asking if the average CV for a move from $\mathbf{p}^0$ to $\mathbf{p}^1$ is positive. Thus, comparing two situations on the basis of aggregate Hicksian compensation is different from comparing them based on the Bergson-Samuelson ASW criterion.

**Comparing ASW with CV**: Unlike aggregate CV, the ASW criterion does not assume that social marginal utility of income is constant across income, and does not suffer from conceptual ambiguities like Scitovszky reversal of social preferences (cf. Mas-Colell et al 1995 page 830-31), as illustrated in Figure 1.

Figure 1 shows two allocations $Q_1$ and $Q_2$ with the utility possibility frontiers $A_1Q_1DB_1$ and $A_2FQ_2B_2$ through them intersecting. Each frontier represents the utility combinations attainable via redistribution between individuals A and B, starting from any point on it. Then the allocation $D$ Pareto dominates $Q_2$ and can be attained from $Q_1$ via redistribution. Therefore $Q_1$ is *superior* to $Q_2$ via the aggregate compensation principle. At the same time, the allocation $F$ which can be attained via redistribution from $Q_2$ is Pareto superior to $Q_1$, implying that $Q_1$ is *inferior* to $Q_2$ via the compensation principle; so aggregate CV is again negative, thus leading to an ambiguity. These conceptual shortcomings of the Hicksian approach have led to widespread use of the Bergson-Samuelson ASW criterion in public finance applied research. In empirical IO, the widely used log-sum measure of consumer-welfare is precisely the ASW in the multinomial logit model, cf. Train 2003, Sec 3.5.

**Quasilinear Utilities**: If $U_j(y - p_j, \eta) = h_j(\eta) + y - p_j$ and $U_0(y, \eta) = h_0(\eta) + y$ for every $\eta$, i.e. utility is quasilinear in the numeraire for each individual, then it can be shown (see appendix for derivation) that

$$W\left(\mathbf{p}^1, y, \eta\right) - W\left(\mathbf{p}^0, y, \eta\right) = -CV\left(y, \mathbf{p}^0, \mathbf{p}^1, \eta\right), \tag{9}$$

and $W(\mathbf{p}, y, \eta)$ equals

$$\{y + \max\{0, h_1(\eta) - p_1 - h_0(\eta), ..., h_J(\eta) - p_J - h_0(\eta)\}\}, \tag{10}$$

so that the average marginal utility of income $\frac{\partial}{\partial y}\int W(\mathbf{p}, y, \eta)\, dF(\eta)$ equals 1, which does not depend on $y$. In this case, a social planner maximizing average social utility would be indifferent between giving a dollar to a rich versus a poor individual.

## 2.3 Special Case: Binary Choice

It is useful to illustrate our results through the important special case of binary choice, which is also the setting of our application. In this setting, with $J = 1$ and a subsidy on option 1 reducing its price from $\bar{p}$ to $\bar{p} - \sigma$,, the subsidy-induced **change in average social welfare** at income $y$ for a generic $\varepsilon \geq 0$ is given by

$$\Delta(\bar{p}, \sigma, y; \varepsilon) \equiv \int_0^\infty (z+y)^{-\varepsilon} \times [q_1(\bar{p} - \sigma + z, y + z) - q_1(\bar{p} + z, y + z)]\, dz; \tag{11}$$

under $\varepsilon = 0$, we have that

$$\Delta(\bar{p}, \sigma, y; 0) \equiv \int_0^\infty [q_1(\bar{p} - \sigma + z, y + z) - q_1(\bar{p} + z, y + z)]\, dz; \tag{12}$$

the **average CV** at income $y$ (cf. Bhattacharya 2015, eqn. 10) equals

$$S(\bar{p}, \sigma, y) \equiv \int_{\bar{p}-\sigma}^{\bar{p}} q_1(p, y + p - \bar{p} + \sigma)\, dp \overset{\text{subs } z = p - \bar{p} + \sigma}{=} \int_0^\sigma q_1(\bar{p} - \sigma + z, y + z)\, dz. \tag{13}$$

Furthermore,

$$\int_0^\infty q_1(\bar{p} + z, y + z)\, dz \overset{\text{sub } t = z + \sigma}{=} \int_\sigma^\infty q_1(\bar{p} + t - \sigma, y + t - \sigma)\, dt \tag{14}$$

and therefore, from (12), (13) and (14) we have that

$$\Delta(\bar{p}, \sigma, y; 0) - S(\bar{p}, \sigma, y) = \int_\sigma^\infty [q_1(\bar{p} - \sigma + z, y + z) - q_1(\bar{p} - \sigma + z, y - \sigma + z)]\, dz. \tag{15}$$

Now, the integrand in (15) is strictly positive (negative) for all $z$ if option 1 is normal (resp. inferior). Therefore, the only way $\Delta(\bar{p}, \sigma, y; 0) = S(\bar{p}, \sigma, y)$ is that $q_1(p, y)$ does not depend on $y$, which implies utilities are quasilinear, and therefore by (10), the social marginal utility of income equals 1.

The **average treatment effect**, the quantity most commonly used in program evaluation and the treatment choice literature, equals

$$T(\bar{p}, \sigma, y) \equiv q_1(\bar{p} - \sigma, y) - q_1(\bar{p}, y), \tag{16}$$

which is simply the integrand of (12) evaluated at the lower limit of the integral. Since this is measured as *quantity* of demand, a direct comparison with average or marginal subsidy cost is not possible. In contrast, the quantities $\Delta(\cdot,\cdot,\cdot)$ or $S(\cdot,\cdot,\cdot)$ provide theoretically justified monetary values of the choice, based on the choice-makers' own preference.

**Deadweight Loss (DWL)**: The average cost of the subsidy equals $\sigma \times q_1(\bar{p} - \sigma, y)$ in every case. Therefore, the DWL of the subsidy under $\varepsilon = 0$ is given by

$$\begin{aligned} \sigma \times q_1(\bar{p} - \sigma, y) - \Delta(\bar{p}, \sigma, y; 0) &= \int_{\bar{p}-\sigma}^{\bar{p}} [q_1(\bar{p} - \sigma, y) - q_1(t, y + t - \bar{p} + \sigma)]\, dt \\ &+ \int_{\bar{p}}^{\infty} [q_1(t, y + t - \bar{p}) - q_1(t, y + t - \bar{p} + \sigma)]\, dt. \end{aligned} \tag{17}$$

Note that the first term in (17) is positive because

$$\begin{aligned} &q_1(t, y + t - \bar{p} + \sigma) - q_1(\bar{p} - \sigma, y) = \Pr[U_1(y - \bar{p} + \sigma, \eta) \geq U_0(y + t - \bar{p} + \sigma, \eta)] \\ &- \Pr[U_1(y - \bar{p} + \sigma, \eta) \geq U_0(y, \eta)] \leq 0, \text{ for } t \geq \bar{p} - \sigma \end{aligned} \tag{18}$$

since $U_0(\cdot, \eta)$ is strictly increasing. The second term will be negative if the good is normal, and the DWL may be *negative* if the income effect is strongly positive. This is in contrast to the deadweight loss based on the CV which must necessarily be non-negative.

## 2.4 Treatment targeting

For binary choice, the constrained, optimal subsidy allocation problem takes the form

$$\max_{\sigma(\cdot)\in\mathcal{T}} \int B\left(\bar{p}, \sigma\left(y\right), y\right) dF_Y\left(y\right) \text{ s.t. } \int \sigma\left(y\right) \times q_1\left(\bar{p} - \sigma\left(y\right), y\right) dF_Y\left(y\right) \leq M, \tag{19}$$

where $M$ denotes the planner's budget constraint, $\mathcal{T}$ denotes the set of politically/practically feasible targeting rules, $F_Y\left(\cdot\right)$ is the marginal distribution of income in the population, and $B\left(\cdot,\cdot,\cdot\right)$ is one of $\Delta\left(\cdot,\cdot,\cdot\right)$, $S\left(\cdot,\cdot,\cdot\right)$ or $T\left(\cdot,\cdot,\cdot\right)$, defined in (12)-(16).

**Parameter Uncertainty**: Note that (19) seeks to maximize welfare of the individuals we observe. If instead, we treat our data as a random sample from a population, and wish to maximize welfare for that population, then we would need to take parameter uncertainty into account. This can be done by defining a loss function

$$\begin{aligned} L\left(\sigma\left(\cdot\right), \theta, c\right) &= -\int B\left(\bar{p}, \sigma\left(y\right), y, \theta\right) dF\left(y, \theta_1\right) \\ &\quad + c\left[M - \int \sigma\left(y\right) \times q_1\left(\bar{p} - \sigma\left(y\right), y, \theta_2\right) dF\left(y, \theta_1\right)\right]^2, \end{aligned} \tag{20}$$

where $c$ denotes the penalty incurred by the planner from violating the budget constraint, and $\theta_1$, $\theta_2$ denote the parameters determining the marginal distribution of income and the demand function e.g. logit coefficients, respectively.[3] Then define the optimal choice of $\sigma\left(\cdot\right)$ under a Bayesian criterion by solving

$$\min_{\sigma(\cdot)} \int L\left(\sigma\left(\cdot\right), \theta, c\right) dP_{post}\left(\theta|data\right) \tag{21}$$

where $P_{post}\left(\theta|data\right)$ refers to the posterior distribution of $\theta$ given the data. For computational simplicity, one can use the bootstrap distribution of $\theta$ to approximate the posterior corresponding to a flat prior (cf. Hastie et al 2009).

[3] Underspending can lead to reduction of budget in a following cycle. Also, one can treat over and under spending asymmetrically (cf. Tetenov 2012), which we avoid here for simplicity.

# 3 Identification and Estimation

Theorem 1 expresses the distribution of indirect utility in terms of the structural choice probability defined in (1). Learning the entire distribution of $W(\mathbf{p}, y, \eta)$ at fixed $\mathbf{p}, y$ would require one to estimate $q_0(c - y + p_1, ..c - y + p_J, c)$ for all values of $c \geq y$, or equivalently $q_0(z + p_1, ..z + p_J, z + y)$ for simultaneous, identical increment $z \in (0, \infty)$ to each price and income. In any finite dataset, of course there will be limited variation of prices and income. So one can use a flexible parametric model such as random coefficients to estimate $q_0(c - y + p_1, ..c - y + p_J, c)$ as is popular in empirical IO; shape restrictions on the choice probability functions (cf. Bhattacharya, 2021) can be imposed by restricting the support of the random coefficients. Any such parametric approximation would obviously impose additional restrictions on preference that are not required for Theorem 1 to hold.[4] Alternatively, one can remain nonparametric and work with bounds. In particular, $U_j(\cdot, \eta)$ being strictly increasing for each $j$ and

$$q_0(c - y + p_1, ..c - y + p_J, c) = \Pr\left[U_0(c, \eta) \geq \max\left\{U_1(y - p_1, \eta) ...U_J(y - p_J, \eta)\right\}\right]$$

yields nonparametric bounds on $q_0(c - y + p_1, ..c - y + p_J, c)$. Specifically, let $S = \{\mathbf{r}, z\}$ with $\mathbf{r} = (r_1, ..., r_J)$ be the set of price-income combinations observed in sample. Then lower and upper bounds on $q_0(c - y + p_1, ..c - y + p_J, c)$ are given by

$$\begin{aligned} LB(c; \mathbf{p}, y) &= \max\left\{q_0(\mathbf{r}, z) : (\mathbf{r}, z) \in S,\ z - r_j \geq y - p_j,\ z \leq c\right\} \\ UB(c; \mathbf{p}, y) &= \min\left\{q_0(\mathbf{r}, z) : (\mathbf{r}, z) \in S,\ z - r_j \leq y - p_j,\ z \geq c\right\}. \end{aligned}$$

These bounds are sharp because the only empirical implication of choice probabilities being generated by utility maximization are a set of shape restrictions on them, as shown in

[4] In particular, in a binary setting, a probit functional form with constant coefficients implicitly assumes additive scalar unobserved preference heterogeneity which implies rank invariance across consumers (cf. Bhattacharya 2021, page 463).

Bhattacharya 2021. It is precisely these shape restrictions that are being used to generate the bounds above.

Finally, if price and/or income are endogenous to individual preference, i.e. independence between utilities and budget set does not hold, then consistent estimation of $q_0$ would require the use of control function-type methods cf. Newey 1987, Blundell-Powell 2004.

# 4 Empirical Illustration: Private Tuition in India

Remedial tuition for children outside schools is ubiquitous in South Asia. This is provided on a for-profit basis either by external individuals or by school-teachers themselves in their private capacity, with no oversight or involvement of the schools. This creates perverse incentives for teachers to reduce their efforts inside the school classroom, cf. Jayachandran 2014. Thus children of richer households, who can afford the additional tuition-fees, benefit from the educational support outside school, whereas those from poorer households suffer the adverse consequences of lower-quality classroom-teaching. One possible way to address this problem is to tax private tuition for richer households and use the tax proceedings to subsidize poorer children. We investigate, empirically, the impact of this hypothetical policy intervention on social welfare, using the methods developed above.

We use micro-data from India's National Sample Survey 71st round, conducted in January-June 2014. The key variables and summary statistics are reported in Table 1. Our sample size is 51092. There are two important data issues here. Firstly, if a household does not purchase private, out-of-school tuition, we do not observe their potential spending had they bought it. This is a well-known empirical issue in discrete choice applications; we address it by using the average price of those opting for private tuition in the village/block of the reference household to impute that price. An intuitive justification is that households are likely to base their decision on information they gather from acquaintances, and tuition-rates are unlikely to vary much within a neighborhood. A second issue, given that the data are non-

Table 1: Application data summaries and welfare calculations.

Summary statistics.

| Variable | Mean | St. dev. | Median | Min | Max |
|---|---|---|---|---|---|
| choice | 0.4067 | 0.4912 | 0 | 0 | 1 |
| price | 3291.6 | 3689 | 2000 | 0 | 20000 |
| monthly income (per household member) | 9579.5 | 6379.2 | 8000 | 100 | 86700 |
| household size | 5.8799 | 2.4896 | 5 | 1 | 32 |
| age of child | 12.2658 | 3.5747 | 12 | 5 | 18 |
| male | 0.5395 | 0.4984 | 1 | 0 | 1 |

Welfare estimates.

| Welfare quantity | At median income | Averaged over income |
|---|---|---|
| average compensating variation | 1499.5 | 1501.8 |
| | (30.056) | (31.531) |
| average treatment effect | 0.0561 | 0.0561 |
| | (0.0134) | (0.0135) |
| change in average social welfare, $\varepsilon = 0$ | 1523.4 | 1520.9 |
| | (281.822) | (279.252) |
| change in average social welfare, $\varepsilon = 0.5$ | 10.51 | 10.43 |
| | (1.681) | (1.672) |
| change in average social welfare, $\varepsilon = 1$ | 0.0775 | 0.0787 |
| | (0.0132) | (0.0136) |

Welfare quantities in (12)-(16) calculated at $\bar{p} = 4200$ (75th percentile of the price distribution), $\bar{p} - \sigma = 900$ (25th percentile). Covariate values are at their median values. Bootstrap standard errors in parentheses are based on 400 replications.

experimental, is that prices are likely to be correlated with unobserved tuition-quality. We address this using 'Hausman-instruments' which are average price in other villages/blocks in the same strata (sampling areas larger than blocks but smaller than districts). Hausman instruments would potentially work by extracting the variation in local tuition-prices that is correlated with variation in private tuition rates in neighboring areas, caused by cost-of-living factors there. This variation is unlikely to be correlated with quality of local tuition. The first-stage F-statistic for the instruments has a p-value of $10^{-5}$.[5]

We model demand for private tuition as

$$q_1(p, y) = \Phi\left(\beta_p p + \sum_{m=1}^{M} \beta_{y,m} \mathcal{R}_{m,M+q}(y; q) + x'\beta\right),$$

where $p$ denotes price, $\mathcal{R}_{m,M+q}(y;q)$ are base B-splines of degree $q$ in monthly income $y$ and the covariate vector $x$ representing household size, the child's age and sex. We use Newey's (1987) two-step estimation approach that assumes joint normality of $u$ in the model for the latent variable

$$q_1^* = \beta_p p + \sum_{m=1}^{M} \beta_{y,m} \mathcal{R}_{m,M+q}(y; q) + x'\beta + u$$

and the error in the reduced-form equation for price. To impose shape constraints, in the second step, we require

$$\beta_p \leq 0, \qquad \beta_p + \sum_{m=2}^{M} \frac{\beta_{y,m} - \beta_{y,m-1}}{z_{m+q} - z_m} \mathcal{R}_{m-1,M+q-1}(y; q-1) \leq 0$$

with the second inequality imposed on a finite grid of income values. In the second inequality, $z_m$ denote knots on the support of income used in the construction of B-splines. The first (second) inequality guarantees that $q_1(p, y)$ is decreasing in price (in the direction $(1, 1)$, i.e. $\frac{\partial}{\partial p} q_1(p, y) + \frac{\partial}{\partial y} q_1(p, y) \leq 0$ (cf. Bhattacharya 2021)).

[5] As stated in the Introduction, individual income can potentially be endogenous to preferences. In this case, one would also need an instrument for income. Non-labour income or land-holding, which likely affects decisions via the aggregate income, can be potential instruments. Given that our application is primarily for illustration, we do not attempt to implement these steps here.

The left of Figure 2 plots demand as a function of price for fixed income, and as a function of income for fixed price for a household with a representative set of characteristics. The inverted U-shape of the income graph agrees with widespread anecdotal evidence that academic success is primarily a middle-class aspiration in India, cf. Varma 2007.

The middle of Figure 2 shows the ACV and change in ASW, net of average cost, at median income and price over a range of subsides. We approximated the change in ASW at a given income value $y$ by calculating integrals in the definition of CASW from 0 to $y_{max} - y$, where $y_{max}$ denotes the largest observed income. This approximation is quite accurate as the values of integrands around $y_{max} - y$ are decreasing, taking values less than 0.0003. A negative subsidy is a tax, and the corresponding net-benefit equals the tax-revenue less utility-loss. We consider taxes up to 20% of median price and subsidy up to $Med(p) - \min(p)$. In the same figure, we plot the net-benefit approximation by the MVPF, i.e.

$$\sigma \left( \frac{\partial}{\partial \sigma} \Delta \left( \bar{p}, \sigma, y; 0 \right) - \frac{\partial}{\partial \sigma} \left\{ \sigma \times q_1 \left( \bar{p} - \sigma, y \right) \right\} \right) \Big|_{\sigma=0} = -\sigma \left( \int_0^\infty \frac{\partial q_1 \left( \bar{p} + z, y + z \right)}{\partial p} dz + q_1(\bar{p}, y) \right).$$

This curve is a straight line through the origin, showing the declining accuracy of first-order approximation as $\sigma$ rises.

The rightmost panel of Figure 2 shows the difference between ACV and change in ASW, when income-effects are/aren't allowed. Note that

$$\Delta \left( \bar{p}, \sigma, y; 0 \right) - S \left( \bar{p}, \sigma, y \right) = \int_\sigma^\infty \left[ \bar{q}_1 \left( \bar{p} - \sigma + z, y + z \right) - \bar{q}_1 \left( \bar{p} - \sigma + z, y - \sigma + z \right) \right] dz.$$

This shows that the average social welfare for $\varepsilon = 0$ would exceed the average compensating variation if the good is normal, and this difference is larger when the larger is the income effect, i.e. for fixed price, average demand changes more rapidly with income.

In the given application the income-effect is strong; consequently, the ACV and change in ASW curves differ substantially. Secondly, while deadweight loss for ACV is necessarily positive, that for the ASW is actually *negative* (benefit exceeds cost) over a range of income, which empirically illustrates our discussion around eqn. (18) above.

Table 1 reports changes in ASW for $\epsilon = 0, 0.5, 1$, ACV and ATE calculated at $\bar{p}$ equal to the 75th percentile of the price distribution, $\bar{p} - \sigma$ equal to the 25th, covariates set equal to their median values, and $y$ set equal to median income $Med(y)$. We also include bootstrap standard errors.

A natural treatment-assignment problem in this case is to optimally subsidize the poor by taxing the rich in a budget-neutral way. The formal problem, analogous to (19) is

$$\max_{\sigma(\cdot)\in\mathcal{T}} \int B\left(\bar{p}, \sigma\left(y\right), y\right) dF_Y\left(y\right) \text{ s.t. } \int \sigma\left(y\right) \times q_1\left(\bar{p} - \sigma\left(y\right), y\right) dF_Y\left(y\right) = 0. \tag{22}$$

Our computational approach is to specify the optimal allocation function $\sigma(y)$ as a univariate B-spline in income (one would first choose the system of knots and the degree of the spline) and then solve the optimization problem in (22) with respect to the coefficients of the B-spline. Due to the partition of unity constraint on the base of B-splines, the restriction on all the coefficients of B-spline not exceeding a given $\bar{p}$ will ensure this restriction on the whole $\sigma(y)$. In an analogous way, suitable linear inequality restrictions on the spline coefficients will limit the degree of taxation shall $\sigma(y)$ happen to be negative. In a nutshell, $\mathcal{T}$ is the space of spline functions which can take both positive and negative values with coefficients satisfying certain linear inequality constraints.

The optimal allocation where $B\left(\cdot, \cdot, \cdot\right)$ corresponds to average treatment-effect, change in ASW with $\varepsilon = 0$ and average CV are shown in Figure 3.[6]

Three features stand out. First, the allocation maximizing the ACV differs from the one that maximizes the CASW at $\epsilon = 0$; this results from the large income-effect. Second, all three curves are downward sloping, because price-sensitivity of demand declines monotonically with income and the overall price-effect is much stronger than the income effect (see appendix Section 6.2.2 for details). Third, the ATE-maximizing allocation leads to the minimum disparity in subsidy/tax rates across the rich and poor, whereas maximizing the

[6] The results are fairly robust to variation in knot numbers in the definition of splines in $\mathcal{T}$.

CASW leads to the highest disparity where the poor receive the highest subsidy and the rich face the highest tax. The optimal ACV curve lies in between.

The right panel in Figure 3 shows CASW when we use the ATE maximizing allocations compared to the CASW at the CASW-optimal allocations. Evidently, using the ATE-optimal allocation leads to much smaller redistribution of welfare from the rich to the poor, relative to the CASW-optimal allocation. Most of this operates at the intensive margin since the two graphs cross zero close to each other. That is, almost the same set of individuals sees an increase and decrease in their welfare in the two optimal allocations, but the extent of welfare-gain is much higher for the poor when using the CASW-optimal allocations. The figure looks somewhat similar to the optimal subsidy graph because the price-effect on demand – and, hence, aggregate welfare – is much stronger than the income-effect. At high incomes, demand becomes less price-sensitive, which explains why the similarity declines there.

For the alternative version that incorporates sampling uncertainty, the loss-function analogous to (20) is

$$\begin{aligned} L\left(\sigma\left(\cdot\right),\theta,c\right) &= -\int B\left(\bar{p},\sigma\left(y\right),y,\theta_1\right)dF\left(y,\theta_2\right) \\ &\quad +c\left[\int \sigma\left(y\right)\times q_1\left(\bar{p}-\sigma\left(y\right),y,\theta_1\right)dF\left(y,\theta_2\right)\right]^2, \qquad (23) \end{aligned}$$

The optimal subsidy would solve (21) with this loss function (results not reported for brevity).[7]

# 5 Conclusion

We propose a notion of aggregate social welfare as an alternative to using outcomes in traditional econometric program-evaluation and statistical treatment-assignment problems. Our main result pertains to the practically important setting of multinomial choice. The key

[7] For details on the implementation we refer the reader to the Appendix.

insight is that the marginal distribution of suitably normalized individual indirect utility can be expressed as a closed-form functional of choice probabilities without functional-form assumptions on unobserved preference heterogeneity and income-effects. This leads to expressions for average weighted social welfare with weights reflecting planners' distributional preferences and the optimal targeting of interventions that maximize aggregate utility under fixed budget. We discuss practical issues of identification and estimation, connections with and advantages relative to aggregate Hicksian welfare-measures, and potential extension to ordered and continuous choice. We illustrate our results using the example of private tuition demand in India where optimal, income-contingent targeting of subsidies/taxes leads to very different paths depending on whether average uptake or average social welfare is being maximized. The main source of this difference is how price-elasticity of demand varies with income.

## References


1. Ahmad, E. and Stern, N., 1984. The theory of reform and Indian indirect taxes. Journal of Public economics, 25(3), pp.259-298.

2. Amemiya, T. 1978. The estimation of a simultaneous equation generalized probit model. Econometrica 46, pp.1193-1205.

3. Atkinson, A.B., 1970. On the measurement of inequality. Journal of economic theory, 2(3), 244-263.

4. Azam, M., 2016. Private tutoring: evidence from India. Review of Development Economics, 20(4), pp.739-761.

5. Banerjee, A., Niehaus, P. and Suri, T., 2019. Universal basic income in the developing world. Annual Review of Economics, 11, pp.959-983.

6. Banks, J., Blundell, R. and Lewbel, A., 1996. Tax reform and welfare measurement: do we need demand system estimation?. The Economic Journal, 106(438), 1227-1241.

7. Bergson, A., 1938. Reformulation of Certain Aspects of Welfare Economics. Quarterly Journal of Economics, 52.

8. Berry, S.T. and Haile, P.A., 2021. Foundations of demand estimation. In Handbook of industrial organization (Vol. 4, No. 1, pp. 1-62). Elsevier.

9. Bhattacharya, D., 2015. Nonparametric welfare analysis for discrete choice. Econometrica, 83(2), pp.617-649.

10. Bhattacharya, D., 2018. Empirical welfare analysis for discrete choice: Some general results. Quantitative Economics, 9(2), pp.571-615.

11. Bhattacharya, D., 2021. The empirical content of binary choice models. Econometrica, 89(1), pp.457-474.

12. Bhattacharya, D. and P. Dupas, 2012. Inferring welfare maximizing treatment assignment under budget constraints, Journal of Econometrics, March 2012,167(1), 168–196.

13. Blackorby, C. and Donaldson, D., 1988. Money metric utility: A harmless normalization?. Journal of Economic Theory, 46(1), 120-129.

14. Blundell, R. and Powell, J.L., 2009. Endogeneity in nonparametric and semiparametric regression models, Chap 8 in Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Cambridge University Press.

15. Blundell, R.W. and Powell, J.L., 2004. Endogeneity in semiparametric binary response models. The Review of Economic Studies, 71(3), pp.655-679.

16. Chipman, J. and Moore, J., 1990. Acceptable indicators of welfare change, consumer's surplus analysis, and the Gorman polar form. Preferences, Uncertainty, and Optimality: Essays in Honor of Leonid Hurwicz, Westview Press, Boulder.

17. Cohen, J. and Dupas, P., 2010. Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. Quarterly journal of Economics, 125(1).

18. De Boor, C. (1978). A practical guide to splines. New York, Springer-Verlag.

19. Deaton, A., 1984. Econometric issues for tax design in developing countries, reprinted in "The Theory of Taxation for Developing Countries". Washington, DC: World Bank, 1987.

20. Dreze, J., 1998. Distribution matters in cost-benefit analysis: Comment on K.A. Brekke. Journal of Public Economics, 70(3), pp.485-488.

21. Dupas, P., 2014. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. Econometrica, 82(1), pp.197-228.

22. Eilers, P.H. C. and Marx, B.D. (1996). Flexible Smoothing with *B*-splines and Penalties, *Statistical Science*, 11, pp. 89-102.

23. Feldstein MS. 1999. Tax avoidance and the deadweight loss of the income tax. Rev. Econ. Stat, 81:674–80

24. Finkelstein, A. and Hendren, N., 2020. Welfare Analysis Meets Causal Inference, Journal of Economic Perspectives, vol. 34, no. 4, pp. 146-67.

25. Fleurbaey, M. and Hammond, P.J., 2004. Interpersonally comparable utility. In Handbook of utility theory (pp. 1179-1285). Springer, Boston, MA.

26. García, J.L. and Heckman, J.J., 2022. On criteria for evaluating social programs (No. w30005). National Bureau of Economic Research.

27. Goldberg, P.K. and Pavcnik, N., 2007. Distributional effects of globalization in developing countries. Journal of economic Literature, 45(1), pp.39-82.

28. Hammond, P.J., 1990. Interpersonal comparisons of utility: Why and how they are and should be made, European University Institute.

https://cadmus.eui.eu/bitstream/handle/1814/342/1990_EUI%20WP_ECO_003.pdf?sequence=1

29. Hastie, T., Tibshirani, R., Friedman, J.H. and Friedman, J.H., 2009. The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

30. Hausman, J. 1981. Exact Consumer's Surplus and Deadweight Loss, *The American Economic Review*, Vol. 71, No. 4, 662-676.

31. Hausman JA. 1996. Valuation of new goods under perfect and imperfect competition. In The Economics of New Goods, eds. T. F. Bresnahan, RJ Gordon, chap. 5. Chicago: University of Chicago Press, 209–248.

32. Hausman, J.A. and Newey, W.K., 2016. Individual heterogeneity and average welfare. Econometrica, 84(3), pp.1225-1248.

33. Heckman, J.J. and Vytlacil, E.J., 2007. Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. Handbook of econometrics Vol. 6, 4779-4874.

34. Hendren, N. and Sprung-Keyser, B., 2020. A unified welfare analysis of government policies. The Quarterly Journal of Economics, 135(3), pp.1209-1318.

35. Herriges, J.A. and Kling, C.L., 1999. Nonlinear income effects in random utility models. Review of Economics and Statistics, 81(1), pp.62-72.

36. Imbens, G.W. and Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. Journal of economic literature, 47(1), pp.5-86.

37. Jayachandran, S., 2014. Incentives to teach badly: After-school tutoring in developing countries. Journal of Development Economics, 108, pp.190-205.

38. Kitagawa, T. and A. Tetenov 2018. Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice, Econometrica, 86(2), 591–616.

39. Kochar, A., 2005. Can targeted food programs improve nutrition? An empirical analysis of India's public distribution system. Economic development and cultural change, 54(1), pp.203-235.

40. Luenberger, D.G., 1969. Optimization by vector space methods. John Wiley & Sons.

41. Mas-Colell, A., Whinston, M.D. and Green, J.R., 1995. Microeconomic theory (Vol. 1). New York: Oxford university press.

42. McFadden, D., 1981. Econometric models of probabilistic choice. Structural analysis of discrete data with econometric applications, 198272.

43. Manski, Charles F. 2004. Statistical Treatment Rules for Heterogeneous Populations, Econometrica, 72(4), 1221–1246.

44. Manski, C. F. 2014. Choosing size of government under ambiguity: Infrastructure spending and income taxation. The Economic Journal, 124(576), pp.359-376.

45. Mayshar, J., 1990. On measures of excess burden and their application. Journal of Public Economics, 43(3), pp.263-289.

46. McFadden, D. 1981, Econometric Models of Probabilistic Choice, in Manski and McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, 198-272, MIT Press, Cambridge, Mass.

47. Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation. The review of economic studies, 38(2), pp.175-208.

48. Newey, W. K. 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. Journal of Econometrics, 36, pp. 231-250.

49. Pollak, Robert A. "Welfare comparisons and situation comparisons." journal of Econometrics 50, no. 1-2 (1991): 31-48.

50. Ramsey, F.P., 1927. A Contribution to the Theory of Taxation. The Economic Journal, 37(145), 47-61.

51. Saez E. 2001. Using elasticities to derive optimal income tax rates. Rev. Econ. Stud. 68:205–29

52. Samuelson, P.A. 1947. Foundations of Economic Analysis, Ch. VIII, "Welfare Economics".

53. Scitovszky, T., 1941. A note on welfare propositions in economics. The Review of Economic Studies, 9(1), pp.77-88.

54. Sen, A.K. 1970. Collective Choice and Social Welfare. San Francisco: Holden-Day.

55. Slesnick, D.T., 1998. Empirical approaches to the measurement of welfare. Journal of Economic Literature, 36(4), pp.2108-2165.

56. Small, K.A. and Rosen, H.S., 1981. Applied welfare economics with discrete choice models. Econometrica: Journal of the Econometric Society, pp.105-130.

57. Smith, Richard J., and Richard W. Blundell. "An exogeneity test for a simultaneous equation Tobit model with an application to labor supply." Econometrica: journal of the Econometric Society (1986): 679-685.

58. Stern, N., 1987. The theory of optimal commodity and income taxation in The theory of taxation for developing countries, The World Bank.

59. Tetenov, A., 2012. Statistical treatment choice based on asymmetric minimax regret criteria. Journal of Econometrics, 166(1), pp.157-165.

60. Train, K.E., 2009. Discrete choice methods with simulation. Cambridge university press.

61. Varma, P.K., 2007. The great Indian middle class. Penguin Books India.

62. Williams, H.C.W.L. 1977. On the Formulation of Travel Demand Models and Economic Measures of User Benefit, Environment & Planning A 9(3), 285-344.

# 6 Technical Appendix

**Proof that CDF in Theorem 1 is non-decreasing:** Note that for $c' > c \geq y$,

$$\begin{aligned}
& \Pr\left[W\left(\mathbf{p}, y, \eta\right) \leq c'\right] = q_0\left(c' - y + p_1, ..c' - y + p_J, c'\right) \\
= \; & \Pr\left[\max\left\{U_1\left(c' - \left(c' - y + p_1\right), \eta\right), ..., U_J\left(c' - \left(c' - y + p_J\right), \eta\right), \eta\right\} \leq U_0\left(c', \eta\right)\right] \\
= \; & \Pr\left[\max\left\{U_1\left(y - p_1, \eta\right), ..., U_J\left(y - p_J, \eta\right), \eta\right\} \leq U_0\left(c', \eta\right)\right] \\
\geq \; & \Pr\left[\max\left\{U_1\left(y - p_1, \eta\right), ..., U_J\left(y - p_J, \eta\right), \eta\right\} \leq U_0\left(c, \eta\right)\right], \text{ since } c' > c \text{ and } U_0\left(\cdot, \eta\right) \nearrow \\
= \; & \Pr\left[\max\left\{U_1\left(c - \left(c - y + p_1\right), \eta\right), ..., U_J\left(c - \left(c - y + p_J\right), \eta\right), \eta\right\} \leq U_0\left(c, \eta\right)\right] \\
= \; & q_0\left(c - y + p_1, ..c - y + p_J, c\right) = \Pr\left[W\left(\mathbf{p}, y, \eta\right) \leq c\right].
\end{aligned}$$

For $y > c' > c$, we have$\Pr\left[W\left(\mathbf{p}, y, \eta\right) \leq c\right] = 0 = \Pr\left[W\left(\mathbf{p}, y, \eta\right) \leq c'\right]$, and finally, for $c' \geq y > c$, we have

$$\Pr\left[W\left(\mathbf{p}, y, \eta\right) \leq c\right] = 0 \leq q_0\left(c' - y + p_1, ..c' - y + p_J, c'\right) = \Pr\left[W\left(\mathbf{p}, y, \eta\right) \leq c'\right].$$

**Calculation of (4)**: The calculation of $\mathcal{W}^{\varepsilon}\left(\mathbf{p}, y\right)$, defined in **(4)** is facilitated by the observations that $W\left(\mathbf{p}, y, \eta\right)$ has a point mass of $q_0\left(p_1, p_2, ..., p_J, y\right)$ at $y$ and a continuous distribution on $(y, \infty)$. As for the continuous part, note that for any continuously distributed random variable $X$ with smooth CDF $F_X\left(\cdot\right)$ and density $f_X\left(\cdot\right)$, we have that

$$\int_b^{\infty} x^{\alpha} f_X\left(x\right) dx = b^{\alpha}\left(1 - F_X\left(b\right)\right) + \alpha \int_b^{\infty} x^{\alpha - 1}\left(1 - F_X\left(x\right)\right) dx \tag{24}$$

using integration by parts. Therefore, from (3), (4) and (24), $\mathcal{W}^{\varepsilon}\left(\mathbf{p}, y\right)$ equals

$$
\begin{aligned}
& \frac{y^{1-\varepsilon}}{1-\varepsilon} \times q_0\left(p_1, p_2, \ldots, p_J, y\right)+\int_y^{\infty} \frac{c^{1-\varepsilon}}{1-\varepsilon} d F_{W(\mathbf{p}, y, \eta)}(c)\, d c \\
& \stackrel{\text{by (24)}}{=} \frac{y^{1-\varepsilon}}{1-\varepsilon} \times q_0\left(p_1, p_2, \ldots, p_J, y\right)+\frac{y^{1-\varepsilon}}{1-\varepsilon}\left[1-F_{W(\mathbf{p}, y, \eta)}(y)\right] \\
& +\int_y^{\infty} c^{-\varepsilon} \times\left[1-F_{W(\mathbf{p}, y, \eta)}(c)\right] d c \\
= & \frac{y^{1-\varepsilon}}{1-\varepsilon} \times q_0\left(p_1, p_2, \ldots, p_J, y\right)+\frac{y^{1-\varepsilon}}{1-\varepsilon}\left[1-q_0\left(p_1, p_2, \ldots, p_J, y\right)\right] \\
& +\int_y^{\infty} c^{-\varepsilon} \times\left[1-q_0\left(c-y+p_1, c-y+p_2, \ldots, c-y+p_J, c\right)\right] d c^{8} \\
= & \frac{y^{1-\varepsilon}}{1-\varepsilon}+\int_0^{\infty}(z+y)^{-\varepsilon} \times\left[1-q_0\left(z+p_1, z+p_2, \ldots, z+p_J, z+y\right)\right] d z . \qquad (25)
\end{aligned}
$$

For $\varepsilon = 0$, i.e. utilitarian planner preferences, (25) reduces to the line integral

$$
y+\int_0^{\infty}\left\{1-q_0\left(z+p_1, z+p_2, \ldots, z+p_J, z+y\right)\right\} d z \qquad (26)
$$

**Proof of Assertion (8)**: Note that the move from $\mathbf{p}^0$ to $\infty_J$ can be decomposed into that from $\mathbf{p}^0$ to $\mathbf{p}^1$ and then from $\mathbf{p}^1$ to $\infty_J$, i.e.

$$
\begin{aligned}
CV\left(y, \mathbf{p}^0, \infty_J, \eta\right) &= CV\left(y, \mathbf{p}^0, \mathbf{p}^1, \eta\right)+CV\left(\underbrace{y+CV\left(y, \mathbf{p}^0, \mathbf{p}^1, \eta\right)}_{\neq y}, \mathbf{p}^1, \infty_J, \eta\right) \\
&\neq CV\left(y, \mathbf{p}^0, \mathbf{p}^1, \eta\right)+CV\left(y, \mathbf{p}^1, \infty_J, \eta\right),
\end{aligned}
$$

because $y+CV\left(y, \mathbf{p}^0, \mathbf{p}^1, \eta\right) \neq y$, and therefore (8) holds.

**Quasilinear Utilities**: If $U_j\left(y-p_j, \eta\right)=h_j(\eta)+y-p_j$, and $U_0(y, \eta)=h_0(\eta)+y$, i.e. utility is quasilinear in the numeraire, then our normalization becomes $\mathcal{U}_0(y, \eta)=y$ and

[8]For standard parametric CDFs like probit or logit, the integral is bounded for $0 \leq \varepsilon \leq 1$.

$\mathcal{U}_j(y-p_j,\eta)=y+h_j(\eta)-h_0(\eta)-p_j$. Then

$$W(\infty_J,y,\eta) = \mathcal{U}_0(y,\eta)=y, \tag{27}$$

$$W(\mathbf{p},y,\eta) = y+\max\{0,h_1(\eta)-p_1-h_0(\eta),...,h_J(\eta)-p_J-h_0(\eta)\} \tag{28}$$

Therefore,

$$\begin{aligned}
& y+\max\{0,h_1(\eta)-p_{10}-h_0(\eta),...,h_J(\eta)-p_{J0}-h_0(\eta)\} \\
= & W\left(\mathbf{p}^0,y,\eta\right)=W\left(\mathbf{p}^1,y+CV\left(y,\mathbf{p}^0,\mathbf{p}^1,\eta\right),\eta\right) \\
= & y+CV\left(y,\mathbf{p}^0,\mathbf{p}^1,\eta\right)+\max\{0,h_1(\eta)-p_{11}-h_0(\eta),...,h_J(\eta)-p_{J1}-h_0(\eta)\}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
CV\left(y,\mathbf{p}^0,\mathbf{p}^1,\eta\right) &= y+\max\{0,h_1(\eta)-p_{10}-h_0(\eta),...,h_J(\eta)-p_{J0}-h_0(\eta)\} \\
&\quad -(y+\max\{0,h_1(\eta)-p_{11}-h_0(\eta),...,h_J(\eta)-p_{J1}-h_0(\eta)\}) \\
&= W\left(\mathbf{p}^0,y,\eta\right)-W\left(\mathbf{p}^1,y,\eta\right).
\end{aligned}$$

Thus, we have that under quasi-linear preferences, $W(\mathbf{p}^1,y,\eta)-W(\mathbf{p}^0,y,\eta)=-CV(y,\mathbf{p}^0,\mathbf{p}^1,\eta)$. Indeed, if utilities are quasilinear, then it follows that the social marginal utility of income $\frac{\partial\int W(\mathbf{p},y,\eta)dF(\eta)}{\partial y}$ equals

$$\frac{\partial}{\partial y}\int\{y+\max\{0,h_1(\eta)-p_1-h_0(\eta),...,h_J(\eta)-p_J-h_0(\eta)\}\}dF(\eta)=1, \tag{29}$$

which does not depend on $y$.

## 6.1 Some details of demand estimation

Denote the support of $y$ as $[y_{min},y_{max}]$. Suppose we use B-splines of degree $q$ with $M+1$ equally spaced knots on $[y_{min},y_{max}]$ (including the end points). Each of the end points in the systems of knots enters with the multiplicity $q$. Then we have $M+q$ base B-splines which we denote as $\mathcal{R}_{m,M+q}(\cdot;q), \quad m=1,\dots,M+q$. The definition of base B-splines through a recursive formula can be found, e.g., in de Boor 1978.

To incorporate covariates $x$ in the demand estimation, we collect all the covariates $x$ into an index $x'\beta$ and specify the demand function as $P\left(q=1\,|\,p,y,x\right)=\Phi\left(+x'\beta\right)$, where $\Phi$ is the C.D.F. of the standard normal distribution. This specification automatically imposes normalization constraints on probability (that is, the probability varying between 0 and 1). Moreover, a monotonicity of this function in a direction of $(p,y)$ is equivalent to the respective monotonicity of $\beta_p p+\sum_{m=1}^{M}\beta_{y,m}\mathcal{R}_{m,M+q}(y;q)+x'\beta$.

Using the well known formulas for the derivatives of B-splines (de Boor 1978), we obtain the following formulas for the derivative of $\sum_{m=1}^{M}\beta_{y,m}\mathcal{R}_{m,M+q}(y;q)$:

$$\frac{d\sum_{m=1}^{M}\beta_{y,m}\mathcal{R}_{m,M+q}(y;q)}{dy}=\sum_{m=2}^{M}\frac{\beta_{y.m}-\beta_{y,m-1}}{z_{m+q}-z_m}\mathcal{R}_{m-1,M+q-1}(y;q-1),$$

where $\mathcal{R}_{\tilde{m},M+q-1}(\cdot;q-1)$, $\tilde{m}=1,\dots,M+q-1$, are base polynomials of degree $q-1$ constructed on the system of knots $\{z_k\}$, $k=2,\dots,M+2q$ (the multiplicity of each boundary knot is now reduced by 1).

## 6.2 Treatment Targeting Problem

### 6.2.1 Uniqueness and second order condition for treatment targeting problem

**Uniqueness:** In this part of the appendix, we clarify some intuition for why the optimal subsidy problem has a unique solution and the second order conditions for constrained maximum. It is easiest to see this in a simple setting where $Y$ takes 2 values $y_1$ and $y_2$ w.p. $\pi_1$ and $\pi_2$ respectively. Then the optimization problem becomes

$$\max_{\sigma_1,\sigma_2}\pi_1 B\left(\sigma_1,y_1\right)+\pi_2 B\left(\sigma_2,y_2\right)$$

$$\text{s.t.} \qquad \pi_1 C(\sigma_1, y_1) + \pi_2 C(\sigma_2, y_2) = M$$

To investigate uniqueness, first consider indifference maps for the benefit and cost functions. For benefits: let $B'(\cdot)$ and $B''(\cdot)$ denote 1st and 2nd price derivative w.r.t. $\sigma$, then

$$\begin{aligned}
\pi_1 B(\sigma_1, y_1) + \pi_2 B(\sigma_2, y_2) &= \mathit{const.} \\
&\Rightarrow \pi_1 B'(\sigma_1, y_1)\, d\sigma_1 + \pi_2 B'(\sigma_2, y_2)\, d\sigma_2 = 0 \\
&\Rightarrow \left.\frac{d\sigma_2}{d\sigma_1}\right|_B = -\frac{\pi_1 B'(\sigma_1, y_1)}{\pi_2 B'(\sigma_2, y_2)} < 0 \\
&\Rightarrow \left.\frac{d^2\sigma_2}{d\sigma_1^2}\right|_B = -\frac{\pi_1 B''(\sigma_1, y_1)}{\pi_2 B'(\sigma_2, y_2)}
\end{aligned}$$

Similarly, for costs, $C(\sigma_1, y_1) = \sigma_1 \times q(\bar{p} - \sigma_1, y_1)$, let $C'(\sigma_1, y_1)$ and $C''(\sigma_1, y_1)$ denote 1st and 2nd derivative w.r.t. $\sigma_1$. For indifference maps

$$\begin{aligned}
\pi_1 C(\sigma_1, y_1) + \pi_2 C(\sigma_2, y_2) &= \mathit{const.} \\
&\Rightarrow \frac{d\sigma_2}{d\sigma_1} = -\frac{\pi_1 C'(\sigma_1, y_1)}{\pi_2 C'(\sigma_2, y_2)} < 0 \\
&\Rightarrow \left.\frac{d^2\sigma_2}{d\sigma_1^2}\right|_C = -\frac{\pi_1 C''(\sigma_1, y_1)}{\pi_2 C'(\sigma_2, y_2)}
\end{aligned}$$

For the application, from the net benefit curves (i.e. benefit minus cost curves plotted against $\sigma$) we can obtain that

$$B'(\sigma_2, y_2) - C'(\sigma_2, y_2) < 0; \;\; B''(\sigma_1, y_1) - C''(\sigma_1, y_1) < 0. \tag{30}$$

From the benefit curves themselves we can get that $B'(\sigma_2, y_2) > 0$, $B''(\sigma_1, y_1) > 0$.

Putting all of this together, we have that

$$0 < B'(\sigma_2, y_2) < C'(\sigma_2, y_2), \quad 0 < B''(\sigma_1, y_1) < C''(\sigma_1, y_1).$$

Therefore, from

$$\left.\frac{d^2\sigma_2}{d\sigma_1^2}\right|_B = -\frac{\pi_1 B''(\sigma_1, y_1)}{\pi_2 B'(\sigma_2, y_2)} \text{ and } \left.\frac{d^2\sigma_2}{d\sigma_1^2}\right|_C = -\frac{\pi_1 C''(\sigma_1, y_1)}{\pi_2 C'(\sigma_2, y_2)},$$

we conclude that both indifference curves are decreasing and concave when viewed from the origin, with the benefit indifference curve less concave than the constant cost curve. Therefore there is a single internal maxima at the point where the benefit indifference curve is tangent to the cost indifference curve.

**Second-order conditions**: The ideas behind the second-order conditions for this case will extend to a general case. We write the Lagrangian as

$$\pi_1 B(\sigma_1, y_1) + \pi_2 B(\sigma_2, y_2) + \lambda\left(M - \pi_1 C(\sigma_1, y_1) + \pi_2 C(\sigma_2, y_2)\right).$$

The FOC is then

$$\begin{aligned}\pi_1 B'(\sigma_1, y_1) &= \lambda \pi_1 C'(\sigma_1, y_1), \ \ \pi_2 B'(\sigma_2, y_2) = \lambda \pi_2 C'(\sigma_2, y_2) \\ &\Rightarrow \frac{B'(\sigma_1, y_1)}{C'(\sigma_1, y_1)} = \lambda = \frac{B'(\sigma_2, y_2)}{C'(\sigma_2, y_2)}.\end{aligned}$$

The Bordered Hessian of the Lagrangian is

$$H = \begin{bmatrix} B''(\sigma_1, y_1) - \lambda C''(\sigma_1, y_1) & 0 & -\pi_1 C'(\sigma_1, y_1) \\ 0 & B''(\sigma_2, y_2) - \lambda C''(\sigma_2, y_2) & -\pi_2 C'(\sigma_2, y_2) \\ -\pi_1 C'(\sigma_1, y_1) & -\pi_2 C'(\sigma_2, y_2) & 0 \end{bmatrix}$$

$$= \begin{bmatrix} B''(\sigma_1, y_1) - \frac{B'(\sigma_1, y_1) C''(\sigma_1, y_1)}{C'(\sigma_1, y_1)} & 0 & -\pi_1 C'(\sigma_1, y_1) \\ 0 & B''(\sigma_2, y_2) - \frac{B'(\sigma_2, y_2) C''(\sigma_2, y_2)}{C'(\sigma_2, y_2)} & -\pi_2 C'(\sigma_2, y_2) \\ -\pi_1 C'(\sigma_1, y_1) & -\pi_2 C'(\sigma_2, y_2) & 0 \end{bmatrix}$$

For our solution to FOC to be a local maximizer, we need $\det H(\cdot)$ to be positive, i.e.

$$
\begin{aligned}
&-\pi_2^2\left(C'(\sigma_2,y_2)\right)^2\times\left(B''(\sigma_1,y_1)-\frac{B'(\sigma_1,y_1)\,C''(\sigma_1,y_1)}{C'(\sigma_1,y_1)}\right)\\
&\qquad\qquad-\pi_1^2\left(C'(\sigma_1,y_1)\right)^2\times\left(B''(\sigma_2,y_2)-\frac{B'(\sigma_2,y_2)\,C''(\sigma_2,y_2)}{C'(\sigma_2,y_2)}\right)>0.
\end{aligned}
$$

Sufficient conditions for this are

$$
\frac{B''(\sigma_1,y_1)}{B'(\sigma_1,y_1)}<\frac{C''(\sigma_1,y_1)}{C'(\sigma_1,y_1)},\quad \frac{B''(\sigma_2,y_2)}{B'(\sigma_2,y_2)}<\frac{C''(\sigma_2,y_2)}{C'(\sigma_2,y_2)}.
$$

### 6.2.2 Shape of Optimal Subsidy as Function of Income

In this section we derive conditions that determine shapes of optimal ATE subsidies. In particular, we have a clear-cut condition of when these optimal subsidies as functions of income will be decreasing.

Suppose income $y$ takes 2 values $y_1 < y_2$ w.p. $\pi_1$ and $\pi_2$ respectively. The expected ATE maximization problem becomes

$$
\max_{\sigma_1,\sigma_2}\pi_1 q(\bar{p}-\sigma_1,y_1)+\pi_2 q(\bar{p}-\sigma_2,y_2)
$$

$$
\text{s.t.}\quad \pi_1\sigma_1 q(\bar{p}-\sigma_1,y_1)+\pi_2\sigma_2 q(\bar{p}-\sigma_2,y_2)=M.
$$

The Lagrangian is given by

$$
\pi_1 q(\bar{p}-\sigma_1,y_1)+\pi_2 q(\bar{p}-\sigma_2,y_2)+\lambda\left(M-\pi_1\sigma_1 q(\bar{p}-\sigma_1,y_1)-\pi_2\sigma_2 q(\bar{p}-\sigma_2,y_2)\right)
$$

yielding FOC

$$
-\pi_1\frac{\partial}{\partial p}q(\bar{p}-\sigma_1,y_1) = \lambda\pi_1\left(q(\bar{p}-\sigma_1,y_1)-\sigma_1\frac{\partial}{\partial p}q(\bar{p}-\sigma_1,y_1)\right) \tag{31}
$$

$$
-\pi_2\frac{\partial}{\partial p}q(\bar{p}-\sigma_2,y_2) = \lambda\pi_2\left(q(\bar{p}-\sigma_2,y_2)-\sigma_2\frac{\partial}{\partial p}q(\bar{p}-\sigma_2,y_2)\right) \tag{32}
$$

Taking ratios and using $\eta_1(\bar{p}, \sigma_1, y_1) \equiv -\frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_1, y_1)}{q(\bar{p}-\sigma_1, y_1)} > 0$, we get

$$\begin{aligned}
\frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_1, y_1)}{\frac{\partial}{\partial p} q(\bar{p}-\sigma_2, y_2)} &= \frac{q(\bar{p}-\sigma_1, y_1) - \sigma_1 \frac{\partial}{\partial p} q(\bar{p}-\sigma_1, y_1)}{q(\bar{p}-\sigma_2, y_2) - \sigma_2 \frac{\partial}{\partial p} q(\bar{p}-\sigma_2, y_2)} \\
&\implies \frac{\frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_1, y_1)}{q(\bar{p}-\sigma_1, y_1)}}{\frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_2, y_2)}{q(\bar{p}-\sigma_2, y_2)}} = \frac{1 - \sigma_1 \frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_1, y_1)}{q(\bar{p}-\sigma_1, y_1)}}{1 - \sigma_2 \frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_2, y_2)}{q(\bar{p}-\sigma_2, y_2)}} \\
&\implies \frac{\eta_1}{\eta_2} = \frac{1 + \sigma_1 \eta_1}{1 + \sigma_2 \eta_2} && (33) \\
&\implies \sigma_2 - \sigma_1 = \frac{\eta_2 - \eta_1}{\eta_2 \eta_1} = \frac{1}{\eta_1} - \frac{1}{\eta_2} && (34)
\end{aligned}$$

The above inequality is consistent with $\sigma_2 < \sigma_1$ (i.e. subsidy is progressive) if and only if $\eta_2 < \eta_1$, i.e.

$$\left| \frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_1, y_1)}{q(\bar{p}-\sigma_1, y_1)} \right| > \left| \frac{\frac{\partial}{\partial p} q(\bar{p}-\sigma_2, y_2)}{q(\bar{p}-\sigma_2, y_2)} \right| \tag{35}$$

for $y_1 < y_2$ and $\sigma_2 < \sigma_1$, i.e. the price sensitivity of demand measured by $\left| \frac{\partial}{\partial p} \ln q(p, y) \right| \Big|_{p=\tilde{p}, y=\tilde{y}}$ is decreasing in $\tilde{y}$ at fixed $\tilde{p}$ and is increasing in $\tilde{p}$ for fixed $\tilde{y}$. The derivations here clearly extend from two to multiple possible values of income.

For the application, Figure 4 depicts the absolute values of the derivatives $\left| \frac{\partial}{\partial p} \ln q(p, y) \right| \Big|$ as functions of incomes (for incomes from the 3rd till the 97th percentile) at different price percentiles (even though $\bar{p}$ is the median price, since the subsidy can be negative, we consider percentiles above the median). These absolute values are decreasing in $y$ and, moreover, there is an ordering of the curves with respect to price percentiles. Therefore, the optimal ATE subsidy cannot be increasing in $y$. Indeed, suppose it increased when we moved from $y_1$ to $y_2$, where $y_2 > y_1$. This would mean that we would move to a lower curve and, due to the decreasing nature of each curve, the inequality in (35) would hold, which leads to a contradiction. Therefore, the optimal ATE subsidy is decreasing in income, as illustrated in Figure 3. Because of the consistent shape pattern of curves across the incomes, similar findings apply also CASW and ACV criteria.

### 6.2.3 Empirical Approach to Optimal Subsidies with Sampling Uncertainty

If we treat data as a sample from the population for whom we wish to solve the optimal subsidy problem (22), then one would need to take parameter uncertainty into account. In order to do that, we can consider a parametric version of our demand estimation problem.

As described earlier, parameter uncertainty can be taken into account by defining the loss function (23) with $\theta = (\theta_1, \theta_2)$ denoting the parameters determining the demand function ($\theta_1$) and the marginal distribution of income ($\theta_2$). For the demand function, we can e.g. use the logit specification $\bar{q}_1(p, y, \theta_1) = \frac{1}{1+\exp(\theta_{10}+\theta_{11}p+\theta_{12}py+\theta_{13}y)}$. We can then define the optimal choice of $\sigma(\cdot)$ by solving $\min_{\sigma(\cdot)} \int L(\sigma(\cdot), \theta, c)\, dP_{post}(\theta|data)$, where $P_{post}(\theta|data)$ refers to the posterior distribution of $\theta$ given the data. Corresponding to 'flat' priors, $P_{post}(\theta_1|data)$ is taken to be the sampling distribution of the estimated demand parameters. With income following the log-normal distribution with mean $\mu$ and scale $\sigma$, i.e. $\ln(y)$ following $\mathcal{N}(\mu, \sigma^2)$, the posterior $P_{post}(\theta|data)$ can be computed by generating draws of $\theta_1$ from $N\left(\hat{\theta}_1, Covar\left(\hat{\theta}_1\right)\right)$. As for $\theta_2$, one can generate $\exp(y^*)$ where $y^*$ represents draws from the posterior of a normal $\mathcal{N}(\mu, \sigma^2)$ which is the same as $\mathcal{N}\left(\overline{\ln y}, s^2_{\ln y}\right)$, where $\overline{\ln y}$ stands for the sample mean of $\ln y$ and $s^2_{\ln y}$ stands for the sample variance of $\ln y$. The steps are summarized below:

1. Estimate the logit model for demand and store estimated coefficients $\hat{\theta}_1$ and estimated covariance matrix $Covar\left(\hat{\theta}_1\right)$
2. Estimate sample mean $\overline{\ln y}$ and variance $s^2_{\ln y}$ of $\ln(y)$
3. Draw $N$ values $\theta^*_{1j}$ from $\mathcal{N}\left(\hat{\theta}_1, Covar\left(\hat{\theta}_1\right)\right)$ and $ly^*_j$'s from $\mathcal{N}\left(\overline{\ln y}, s^2_{\ln y}\right)$ and calculate $y^*_j = \exp(ly^*)$
4. For each candidate $\sigma(\cdot)$, compute (8) by averaging over these many $y^*$'s and $\theta^*_1$, i.e.

$$-\frac{1}{N}\sum_{j=1}^{N} B\left(\bar{p}, \sigma\left(y^*_j\right), y^*_j, \theta^*_{1j}\right) + c\left[\frac{1}{N}\sum_{j=1}^{N} \sigma\left(y^*_j\right) \times \bar{q}_1\left(\bar{p} - \sigma\left(y^*_j\right), y^*_j, \theta^*_{1j}\right)\right]^2$$

5. Choose $\sigma(\cdot)$ (i.e. coefficients on the spline basis) to minimize this objective function.

### 6.2.4 General Approach to Demand Estimation under Shape Constraints

In our private tuition application we used the form of the demand function that implied that a monotonic transformation of demand is additively separable in price and income. In practice, researchers may decide to use more general demand estimation approaches inspired by generic nonparametric estimation approaches. Additionally, these approaches have to ensure that in the binary case, $\bar{q}_1(p, y)$ is decreasing in $p$ and increasing in $(y-p)$ for fixed $y$, i.e. $\bar{q}_1(\cdot,\cdot)$ is decreasing in directions $(1,0)$ and $(1,1)$ (as shown in Bhattacharya 2021).

A general estimation idea would be to use the tensor-product B-spline in $p$ and $y$. Denoting the support of $p$ as $[p_{min}, p_{max}]$ and the support $y$ as $[y_{min}, y_{max}]$, suppose that in the dimension $p$ we use B-splines of degree $q_1$ with $M_1+1$ equally spaced knots on $[p_{min}, p_{max}]$ (including the end points), and in the dimension $y$ we use B-splines of degree $q_2$ with $M_2+1$ equally spaced knots on $[y_{min}, y_{max}]$ (including the end points). Each of the end points in the systems of knots enters with the multiplicity $q_1$ for the first system and with the multiplicity $q_2$ for the second system of knots. Then in the dimension $p$ we have $M_1+q_1$ base B-splines denoted as $\mathcal{R}_{1;m,M_1+q_1}(\cdot;q_1), \quad m=1,\ldots,M_1+q_1$; in the dimension $y$ we have $M_2+q_2$ base B-splines denoted as $\mathcal{R}_{2;m,M_2+q_2}(\cdot;q_2), \quad m=1,\ldots,M_2+q_2$.

The tensor-product basis consists of $(M_1+q_1)(M_2+q_2)$ polynomials in the form

$$\mathcal{R}_{1;m_1,M_1+q_1}(p;q_1)\mathcal{R}_{2;m_2,M_2+q_2}(y;q_2), \quad m_1=1,\ldots,M_1+q_1,\ m_2=1,\ldots,M_2+q_2$$

(here calculated for specific values of $p$ and $y$) with associated coefficients denoted by $\{h_{m_1m_2}\}$, $m_j=1,\ldots,M_j+q_j$, $j=1,2$. A general tensor-product B-spline used to approximate a function of $(p, y)$ is a linear combination of these base tensor-product polynomials:

$$\mathcal{T}_{M_1,M_2}(p,y)=\sum_{m_1=1}^{M_1+q_1}\sum_{m_2=1}^{M_2+q_2} h_{m_1m_2}\mathcal{R}_{1;m_1,M_1+q_1}(p;q_1)\mathcal{R}_{2;m_2,M_2+q_2}(y;q_2).$$

To incorporate covariates $x$ in the demand estimation, can we collect all the covariates $x$ into an index $x'\beta$ and specify the demand function as $P\left(q=1\,|\,p,y,x\right)=\Phi\left(\mathcal{T}_{M_1,M_2}(p,y)+x'\beta\right)$, where $\Phi$ is the C.D.F of the standard normal distribution. This specification automatically imposes normalization constraints on probability (that is, the probability varying between 0 and 1). Moreover, a monotonicity of this function in a direction of $(p,y)$ is equivalent to the respective monotonicity of the tensor-product B-spline $\mathcal{T}_{M_1,M_2}(\cdot,\cdot)$.

To capture shape constraints through linear restrictions on coefficients $h_{m_1m_2}$ we write an expression for the directional derivative of $\mathcal{T}_{M_1,M_2}$. Denote the system of knots in the dimension $p$ as $\{z_k^{(1)}\}$, $k=1,\ldots,M_1+2q_1+1$, where $z_1^{(1)}=\ldots=z_{q_1}^{(1)}=p_{min}$ and $z_{M_1+q_1+1}^{(1)}=\ldots=z_{M_1+2q_1+1}^{(1)}=p_{max}$; and in the dimension $y$ as $\{z_k^{(2)}\}$, $k=1,\ldots,M_2+2q_2+1$, where $z_1^{(2)}=\ldots=z_{q_2}^{(2)}=y_{min}$ and $z_{M_2+q_2+1}^{(2)}=\ldots=z_{M_2+2q_2+1}^{(2)}=y_{max}$.

Using the well known formulas for the derivatives of B-splines (de Boor 1978), we obtain the following formulas for the partial derivatives of $\mathcal{T}_{M_1,M_2}$:

$$\frac{\partial\mathcal{T}_{M_1,M_2}(p,y)}{\partial p}=q_1\sum_{m_1=2}^{M_1+q_1}\sum_{m_2=1}^{M_2+q_2}\frac{h_{m_1m_2}-h_{m_1-1,m_2}}{z_{m_1+q_1}^{(1)}-z_{m_1}^{(1)}}\mathcal{R}_{1;m_1-1,M_1+q_1-1}(p;q_1-1)\mathcal{R}_{2;m_2,M_2+q_2}(y;q_2),$$
$$\frac{\partial\mathcal{T}_{M_1,M_2}(p,y)}{\partial y}=q_2\sum_{m_1=1}^{M_1+q_1}\sum_{m_2=2}^{M_2+q_2}\frac{h_{m_1m_2}-h_{m_1,m_2-1}}{z_{m_2+q_1}^{(2)}-z_{m_2}^{(1)}}\mathcal{R}_{1;m_1,M_1+q_1}(p;q_1)\mathcal{R}_{2;m_2-1,M_2+q_2-1}(y;q_2-1),$$

where $\mathcal{R}_{1;\tilde{m}_1,M_1+q_1-1}(\cdot;q_1-1)$, $\tilde{m}_1=1,\ldots,M_1+q_1-1$, are base polynomials of degree $q_1-1$ constructed on the system of knots $\{z_k^{(1)}\}$, $k=2,\ldots,M_1+2q_1$ (the multiplicity of each boundary knot is now reduced by 1), and $\mathcal{R}_{2;\tilde{m}_2,M_2+q_1-1}(\cdot;q_2-1)$, $\tilde{m}_2=1,\ldots,M_2+q_2-1$, are base polynomials of degree $q_2-1$ constructed on the system of knots $\{z_k^{(2)}\}$, $k=2,\ldots,M_2+2q_2$ (once again, the multiplicity of each boundary knot is now reduced by 1). Using the formulas from the recursive definition of B-splines, we rewrite these partial derivatives as

$$\frac{\partial\mathcal{T}_{M_1,M_2}(p,y)}{\partial p}=q_1\sum_{m_1=2}^{M_1+q_1}\mathcal{R}_{1;m_1-1,M_1+q_1-1}(p;q_1-1)\sum_{m_2=2}^{M_2+q_2}\mathcal{R}_{2;m_2-1,M_2+q_2-1}(y;q_2-1)$$
$$\left(\omega_{m_2,q_2}^{(2)}(y)\frac{h_{m_1m_2}-h_{m_1-1,m_2}}{z_{m_1+q_1}^{(1)}-z_{m_1}^{(1)}}+(1-\omega_{m_2,q_2}^{(2)}(y))\frac{h_{m_1,m_2-1}-h_{m_1-1,m_2-1}}{z_{m_1+q_1}^{(1)}-z_{m_1}^{(1)}}\right),$$

$$\frac{\partial \mathcal{T}_{M_1,M_2}(p,y)}{\partial y} = q_2 \sum_{m_2=2}^{M_2+q_2} \mathcal{R}_{2;m_2-1,M_2+q_2-1}(y;q_2-1) \sum_{m_1=2}^{M_1+q_1} \mathcal{R}_{1;m_1-1,M_1+q_1-1}(p;q_1-1)$$
$$\left(\omega^{(1)}_{m_1,q_1}(p)\frac{h_{m_1m_2}-h_{m_1,m_2-1}}{z^{(2)}_{m_2+q_1}-z^{(1)}_{m_2}} + (1-\omega^{(1)}_{m_1,q_1}(p))\frac{h_{m_1-1,m_2}-h_{m_1-1,m_2-1}}{z^{(2)}_{m_2+q_1}-z^{(1)}_{m_2}}\right),$$

$$\text{where} \qquad \omega^{(1)}_{m_1,q_1}(p) = \frac{p-z^{(1)}_{m_1}}{z^{(1)}_{m_1+q_1}-z^{(1)}_{m_1}}, \quad \omega^{(2)}_{m_2,q_2}(y) = \frac{y-z^{(2)}_{m_2}}{z^{(2)}_{m_2+q_2}-z^{(2)}_{m_2}}.$$

Since $\mathcal{R}_{j;m_j-1,M_j+q_j-1}(\cdot;q_j-1)$ is non-zero only on the interval $[z^{(j)}_{m_j}, z^{(j)}_{m_j}+q_j]$, $m_j = 1,\ldots,M_j+q_j$, $j=1,2$ (as follows from the properties of B-splines), then for all the relevant (non-zero) terms $\omega^{(j)}_{m_j,q_j}(\cdot)$ is between 0 and 1. Also notice that we can represent

$$\begin{aligned}
&q_1\left(\omega^{(2)}_{m_2,q_2}(y)\frac{h_{m_1m_2}-h_{m_1-1,m_2}}{z^{(1)}_{m_1+q_1}-z^{(1)}_{m_1}} + (1-\omega^{(2)}_{m_2,q_2}(y))\frac{h_{m_1,m_2-1}-h_{m_1-1,m_2-1}}{z^{(1)}_{m_1+q_1}-z^{(1)}_{m_1}}\right)\\
&= q_1\left(\omega^{(2)}_{m_2,q_2}(y)\frac{h_{m_1m_2}-h_{m_1-1,m_2}}{z^{(1)}_{m_1+q_1}-z^{(1)}_{m_1}} + (1-\omega^{(2)}_{m_2,q_2}(y))\frac{h_{m_1,m_2-1}-h_{m_1-1,m_2-1}}{z^{(1)}_{m_1+q_1}-z^{(1)}_{m_1}}\right)(\omega^{(1)}_{m_1,q_1}(p)+1-\omega^{(1)}_{m_1,q_1}(p))\\
&= \sum_{r_1=0}^{1}\sum_{r_2=0}^{1}\left((1-r_1)\omega^{(1)}_{m_1,q_1}(p)+r_1(\omega^{(1)}_{m_1,q_1}(p))\right)\left((1-r_2)\omega^{(2)}_{m_2,q_2}(y)+r_2(\omega^{(2)}_{m_2,q_2}(y))\right)c^{(1)}_{m_1,m_2-r_2},
\end{aligned}$$

$$\begin{aligned}
&q_2\left(\omega^{(1)}_{m_1,q_1}(p)\frac{h_{m_1m_2}-h_{m_1,m_2-1}}{z^{(2)}_{m_2+q_1}-z^{(1)}_{m_2}} + (1-\omega^{(1)}_{m_1,q_1}(p))\frac{h_{m_1-1,m_2}-h_{m_1-1,m_2-1}}{z^{(2)}_{m_2+q_1}-z^{(1)}_{m_2}}\right)\\
&= q_2\left(\omega^{(1)}_{m_1,q_1}(p)\frac{h_{m_1m_2}-h_{m_1,m_2-1}}{z^{(2)}_{m_2+q_1}-z^{(1)}_{m_2}} + (1-\omega^{(1)}_{m_1,q_1}(p))\frac{h_{m_1-1,m_2}-h_{m_1-1,m_2-1}}{z^{(2)}_{m_2+q_1}-z^{(1)}_{m_2}}\right)(\omega^{(2)}_{m_2,q_2}(y)+1-\omega^{(2)}_{m_2,q_2}(y))\\
&= \sum_{r_1=0}^{1}\sum_{r_2=0}^{1}\left((1-r_1)\omega^{(1)}_{m_1,q_1}(p)+r_1(\omega^{(1)}_{m_1,q_1}(p))\right)\left((1-r_2)\omega^{(2)}_{m_2,q_2}(y)+r_2(\omega^{(2)}_{m_2,q_2}(y))\right)c^{(2)}_{m_1-r_1,m_2},
\end{aligned}$$

where

$$c^{(1)}_{m_1,m_2} \equiv q_1\left(h_{m_1,m_2}-h_{m_1-1,m_2}\right), \quad c^{(2)}_{m_1,m_2} \equiv q_2\left(h_{m_1,m_2}-h_{m_1,m_2-1}\right).$$

The derivative of $\mathcal{T}_{M_1,M_2}$ in the direction of $(d_1, d_2)$ is

$$\nabla_{(d_1,d_2)}\mathcal{T}_{M_1,M_2}(p,y) = \sum_{m_2=1}^{M_1+q_1}\sum_{m_2=2}^{M_2+q_2}\sum_{r_1=0}^{1}\sum_{r_2=0}^{1}\mathcal{R}_{1;m_1-1,M_1+q_1-1}(p;q_1-1)\mathcal{R}_{2;m_2-1,M_2+q_2-1}(y;q_2-1)\times$$
$$\times\left((1-r_1)\omega^{(1)}_{m_1,q_1}(p)+r_1(\omega^{(1)}_{m_1,q_1}(p))\right)\left((1-r_2)\omega^{(2)}_{m_2,q_2}(y)+r_2(\omega^{(2)}_{m_2,q_2}(y))\right)\left(d_1c^{(1)}_{m_1,m_2-r_2}+d_2c^{(2)}_{m_1-r_1,m_2}\right).$$

Taking into account that or all the relevant (non-zero) terms $\omega^{(j)}_{m_j,q_j}(\cdot)$ is between 0 and 1, it is easy to see that the conditions

$$d_1c^{(1)}_{m_1,m_2-r_2}+d_2c^{(2)}_{m_1-r_1,m_2}\geq 0,\quad m_j=2,\ldots,M_j+q_j,\ r_j\in\{0,1\},\quad j=1,2,$$

guarantee that $\mathcal{T}_{M_1,M_2}(\cdot,\cdot)$ is increasing in the direction $(d_1, d_2)$. It can also bee shown that these conditions become necessary for monotonicity in the direction $(d_1, d_2)$ as $M_1, M_2 \to \infty$.

Without any endogeneity issues, with a sample $\{(q_i, p_i, y_i, x_i)\}_{i=1}^{N_1}$, where $q_i$ are binary choices, we can estimate both $\{h_{m_1m_2}\}$, $m_j = 1, \ldots, M_j + q_j$, $j = 1, 2$, and $\beta$ by solving the constrained optimization problem

$$\max_{b,\{h_{m_1m_2}\}}\sum_{i=1}^{N_1}\left(q_i\log\Phi\left(\mathcal{T}_{M_1,M_2}(p_i,y_i)+x'b\right)+(1-q_i)\log\left(1-\Phi\left(\mathcal{T}_{M_1,M_2}(p_i,y_i)+x'b\right)\right)\right) \tag{36}$$

subject to shape constraints on the coefficients.

$$c^{(1)}_{m_1,m_2-r_2}+c^{(2)}_{m_1-r_1,m_2}\leq 0,\quad m_j=2,\ldots,M_j+q_j,\ r_j\in\{0,1\},\quad j=1,2, \tag{37}$$
$$c^{(1)}_{m_1,m_2-r_2}\leq 0,\quad m_2=2,\ldots,M_2+q_2,\ r_2\in\{0,1\}. \tag{38}$$

The use of tensor-product B-spline leads to a large number of coefficients $h_{m_1,m_2}$ that need to be estimated. Indeed, even choosing $M_1 = M_2 = 5$ – so each domain is split into 5 intervals of equal length – and $q_1 = q_2 = 3$ (cubic splines), gives 64 coefficients $h_{m_1,m_2}$. A large number of parameters often results in overfitting. To avoid overfitting, the approximation literature often uses a "wiggliness" penalty matrix (cf. Eilers and Marx 1996).

For univariate B-splines penalty term usually has the form $\lambda\gamma' D' D\gamma$, where $\gamma$ is the vector of all B-spline coefficients, $\lambda$ is the penalty constant and $D$ is the matrix describing $k$-th order differences for components in $\gamma$ (usually $k = 2$). In the case of the tensor-product B-spline, we use two penalty terms – let us denote them as $\lambda_1 h' D_1' D_1 h$ and $\lambda_2 h' D_2' D_2 h$, where $h$ is the $(M_1 + q_1)(M_2 + q_2)$ vector of all coefficients $h_{m_1 m_2}$ arranged in the lexicographic order, $D_1$ is the matrix describing 2-nd order differences in the first dimension and $D_2$ is the matrix describing 2-nd order differences in the second dimension for the components of $h$.

The objective function with these penalty terms is

$$\sum_{i=1}^{N_1} \left(q_i \log \Phi\left(\mathcal{T}_{M_1,M_2}(p_i, y_i) + x'b\right) + (1 - q_i) \log\left(1 - \Phi\left(\mathcal{T}_{M_1,M_2}(p_i, y_i) + x'b\right)\right)\right) + \lambda_1 h' D_1' D_1 h + \lambda_2 h' D_2' D_2 h, \quad (39)$$

where $\lambda_1, \lambda_2 \leq 0$ since we are maximizing the objective function. Thus, we can estimate the demand function by maximizing (39) with respect to $h$ and $b$ subject to (37)-(38).

An IV approach shall be used if the model has endogeneity. By the partition of unity property, all the $\mathcal{R}_{1;m_1,M_1+q_1}(p; q_1)\mathcal{R}_{2;m_2,M_2+q_2}(y; q_2)$ sum up to 1. Hence, the problem can be reparametrized and equivalently rewritten as the one with a constant term and $(M_1+q_1)(M_2+q_2)-1$ endogenous explanatory variables expressed through $\mathcal{R}_{1;m_1,M_1+q_1}(p; q_1)\mathcal{R}_{2;m_2,M_2+q_2}(y; q_2)$. Then we will need to have at least $(M_1 + q_1)(M_2 + q_2) - 1$ valid instrumental variables.

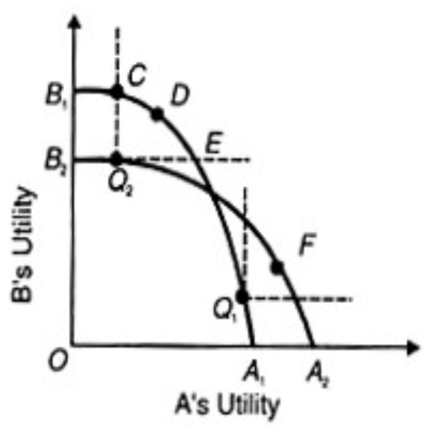


Figure 1: Comparison of ASW with CV: An illustration.

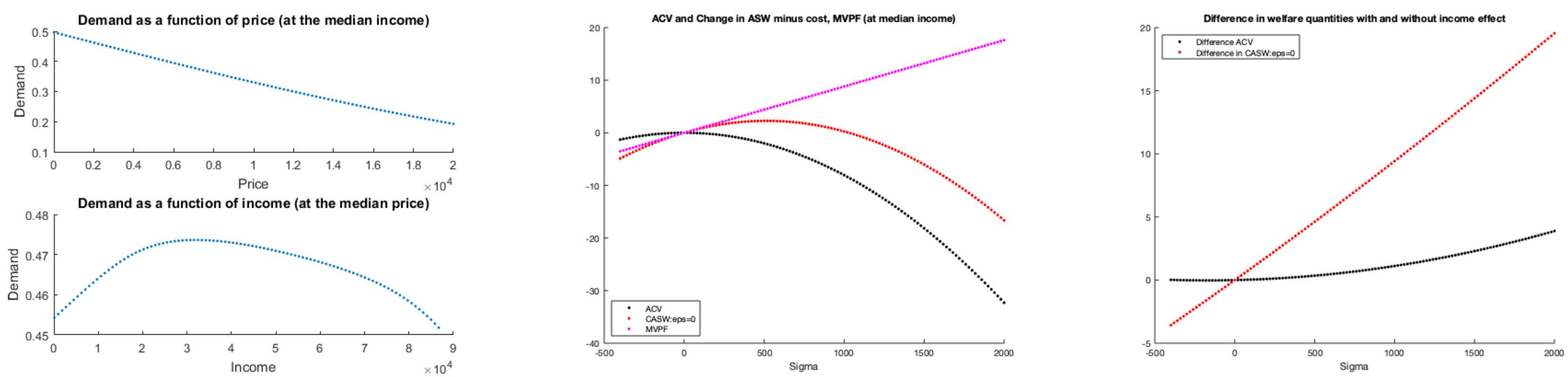


Figure 2: Left: Illustration of demand estimation. Middle: ACV, change in ASW ($\varepsilon = 0$) net of average cost, and the first-order effect based on MVPF ($\bar{p}$ and $y$ are at their median levels). Covariate values are at their median levels. Right: Difference in ACV, change in ASW ($\varepsilon = 0$) with and without income effect.

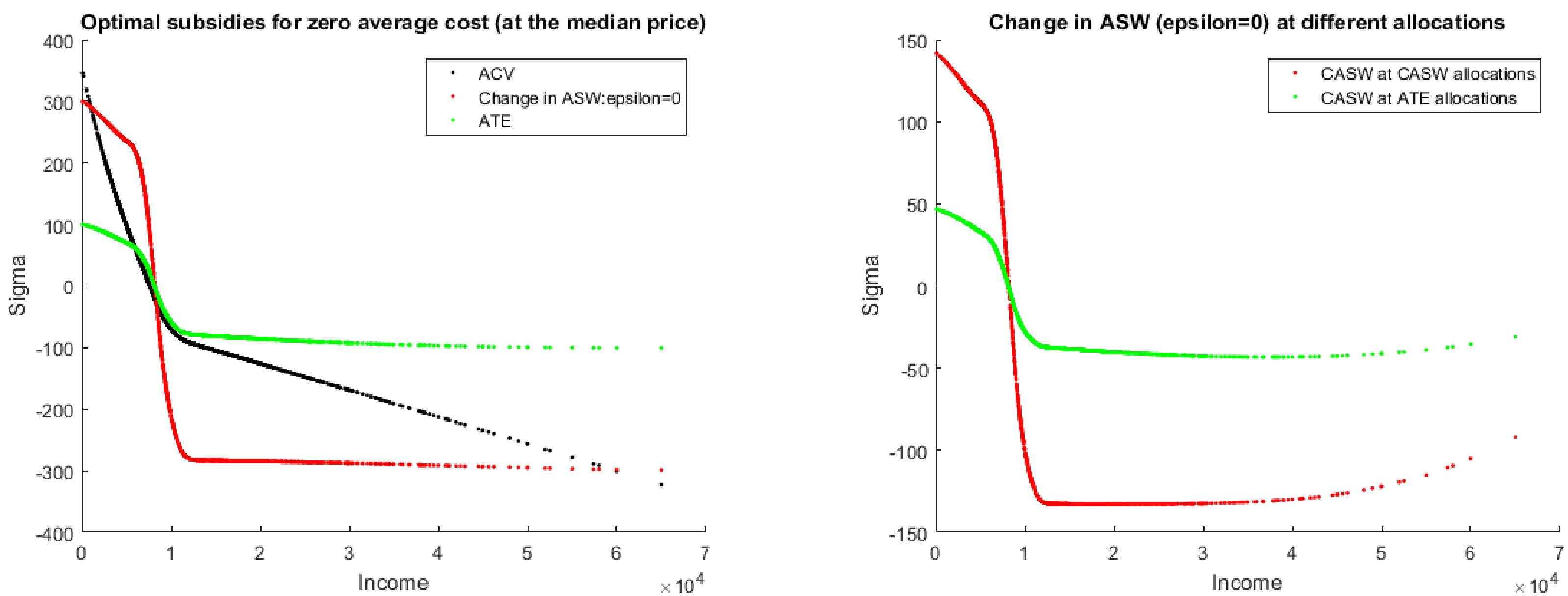


Figure 3: Left: Optimal subsidies for three different welfare criteria with the budget constraint giving the zero average cost. Right: Change in ASW ($\epsilon = 0$) calculated at the CASW ($\epsilon = 0$) optimal allocation path and at the ATE optimal allocation path.

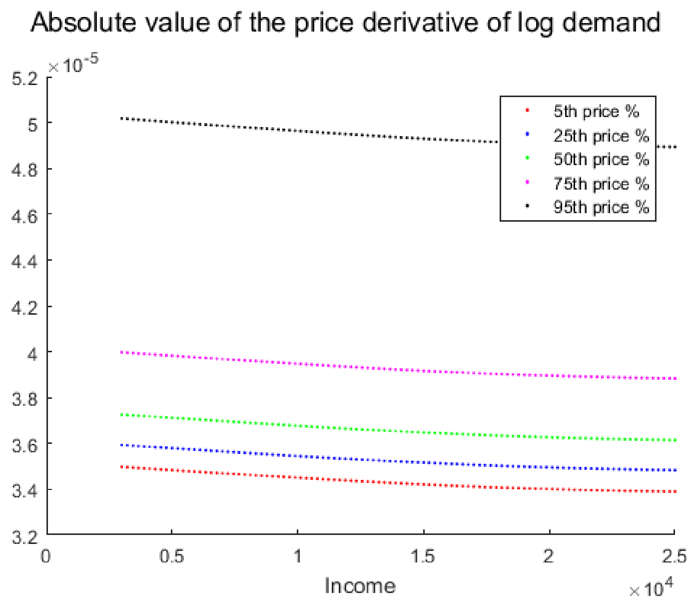


Figure 4: Private tuition application. Absolute values of the derivatives $\left|\frac{\partial}{\partial p}\ln q\left(p,y\right)\right|$ as functions of incomes (for incomes from the 3rd till the 97th percentile) at different price percentiles.