

Empirical welfare analysis for discrete choice: Some general results

DEBOPAM BHATTACHARYA
Faculty of Economics, University of Cambridge

This paper develops nonparametric methods for welfare-analysis of economic changes in the common setting of multinomial choice. The results cover (a) simultaneous price-change of multiple alternatives, (b) introduction/elimination of an option, (c) changes in choice-characteristics, and (d) choice among nonexclusive alternatives. In these cases, Marshallian consumer surplus becomes path-dependent, but Hicksian welfare remains well-defined. We demonstrate that under completely unrestricted preference-heterogeneity and income-effects, the distributions of Hicksian welfare are point-identified from structural choice-probabilities in scenarios (a), (b), and only set-identified in (c), (d). In program-evaluation contexts, our results enable the calculation of compensated-effects, that is, the program's cash-equivalent and resulting deadweight-loss. They also facilitate a theoretically justified cost-benefit comparison of interventions targeting different outcomes, for example, a tuition-subsidy and a health-product subsidy. Welfare analyses under endogeneity is briefly discussed. An application to data on choice of fishing-mode illustrates the methods.

KEYWORDS. Multinomial choice, general heterogeneity, income effects, compensating variation, deadweight loss, multiple price change, elimination of alternative, change in characteristics, weak separability, nonexclusive choice, compensated program-effects.

JEL CLASSIFICATION. C14, C25, D12, D61, H22.

1. INTRODUCTION

Welfare calculations, based on the compensating and equivalent variation (CV and EV, henceforth), lie at the heart of economic policy analysis. Although theoretically well understood, these measures are rarely calculated or reported as part of econometric program evaluations. Indeed, empirical studies—both reduced-form and structural—usually adopt a paternalistic view and evaluate a policy intervention in terms of its *uncompensated* effects on individual outcomes. But they ignore the compensated effects which are required for cost-benefit analysis of the intervention. For example, in an educational context, researchers typically evaluate the impact of a tuition subsidy via its

Debopam Bhattacharya: debobhatta@gmail.com

I am grateful to Pascaline Dupas, Jerry Hausman, Whitney Newey, and seminar participants at the University of Wisconsin for feedback and suggestions. All errors are mine. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 681565).

net effect on college-enrollment (cf. [Ichimura and Taber \(2002\)](#), [Kane \(2003\)](#)). But they stop short of evaluating how much the subsidy is valued by the potential beneficiaries themselves, that is, the lump-sum income transfer that would result in the same individual utilities as the subsidy, and any resulting deadweight loss thereof.¹ The present paper develops methods for calculating such welfare effects as part of program evaluation studies, without imposing restrictive assumptions on the nature of preference-heterogeneity or income effects. The setting is where we observe individual level data on a cross section of consumers making choices among discrete alternatives, and the goal is to estimate the effects of hypothetical economic changes on consumer welfare.

In practical settings, allowing for general heterogeneity in consumer preferences makes empirical welfare analysis a challenging problem. Some advances based on Roy's identity have recently been made for the case of a continuous good, such as gasoline consumption (cf., [Hausman \(1981\)](#), [Hausman and Newey \(2016, 2017\)](#), [Lewbel \(2001\)](#), [Lewbel and Pendakur \(2017\)](#)). However, many important real-life decisions involve discrete choice, such as college attendance, choice of commuting method, school choice, retirement decisions, and so forth. The Roy's identity based methods used for continuous choice cannot be applied to these settings, owing to the nonsmoothness of individual demand in price and income.²

Until recently, available methods for welfare analysis in discrete choice settings were based on restrictive and arbitrary assumptions on preference heterogeneity, for example, quasilinear preferences implying absence of income effects (cf., [Domencich and McFadden \(1975\)](#), [Small and Rosen \(1981\)](#)), or parametrically specified utility functions and heterogeneity distributions ([McFadden \(1981\)](#), [Dagsvik and Karlstrom \(2005\)](#)). Indeed, two key concerns with parametric analyses are (a) model misspecification leading to erroneous substantive conclusions, and (b) identification of welfare distributions solely from functional form assumptions. Recently, [Bhattacharya \(2015\)](#) has shown that for heterogeneous consumers facing the choice between discrete alternatives, the marginal distributions of EV/CV resulting from ceteris paribus price change of a single alternative can be expressed as closed-form transformations of choice probabilities. Taking [Bhattacharya \(2015\)](#) as the point of departure, the present paper makes three new contributions, as follows.

Our first set of results, presented in Section 2, concern welfare effects of simultaneous price changes of several alternatives. Multiple price changes are common in real life. For example, a reduction in the ticket price of an airline serving a route is likely

¹[Stiglitz \(2000, p. 276\)](#) noted that empirical researchers typically ignore income effects owing to the perceived difficulty of calculating them, and [Goolsbee \(1999, p. 10\)](#) pointed out that whereas the economic theory of welfare concerns *compensated* elasticities, common program evaluation studies in public finance typically report uncompensated effects.

²The closest parallel to Roy's Identity for discrete choice is the so-called Daly-Williams-Zachary theorem ([McFadden, \(1981, Section 5.8\)](#)), which shows that in an *additive random utility model with scalar heterogeneity*, the choice probabilities equal certain derivatives of the *average* indirect utility. Even if we restrict preferences in this way, this result is not useful for analysis of *individual* welfare distributions, since the income compensation that maintains average utility is not the same as the average of the income compensations that maintain individual utilities, unless preferences are quasilinear. We are interested in the latter distribution, and hence the DWZ theorem is not relevant to our problem.

to lower the prices of competing airlines; a tuition subsidy in an area with prevalence of teenage labor is likely to raise teens' wages in response; a rise in gasoline prices is likely to raise prices of several transport alternatives, and so on. The impact of such simultaneous price changes on consumer welfare is *the* key consideration for policy-makers in regard to their decision on whether to implement a proposed policy change (cf., Willig, Salop, and Scherer (1991)). We show that under completely general heterogeneity, money-metric welfare effects of such price changes continue to remain well-defined, whereas the analog of the Marshallian Consumer Surplus becomes path dependent (Proposition 1). Furthermore, EV/CV distributions in this case are expressible as closed-form transformations of estimable choice probabilities (Theorem 1). These results cover situations where some price changes are negative, some zero and some positive. A key issue here, elaborated in Section 2.1 below, is that although welfare effects of multiple price changes are well-defined, their distributions *cannot* be obtained by iterating results for a single price change. This is because the income at which welfare distributions are to be evaluated varies in an unobservable way across individuals from the second iteration onward, due to unobservable preference heterogeneity, and thus cannot be conditioned on. Consequently, new results are required for welfare analysis in these situations.

For welfare evaluations in applications, it has been common practice to use the so-called log-sum formula (cf., Small and Rosen (1981), Train (2009)), which relies on strong parametric assumptions regarding both preferences and heterogeneity; see Carneiro, Das, and Reis (2016) and Knittel and Metaxoglou (2014) for some applications.³ Here, we show that under completely general preference heterogeneity and income effects, one can express welfare distributions resulting from multiple simultaneous price changes in terms of simple transformations of choice probability functions, thereby reducing welfare analysis to the problem of estimating choice probabilities. Crucially, our welfare expressions (a) hold when income effects are nonnegligible, as is likely for bigger purchases like children's education and consumer durables, and (b) apply to arbitrary patterns of price changes across alternatives, thereby making the results useful across a wide range of empirical situations.

The above results are also shown to imply the distribution of welfare loss resulting from elimination of an existing alternative, or equivalently, retrospective welfare gain from having added a new alternative to the choice set. Previous econometric studies of such problems either ignored consumer heterogeneity, and implicitly assumed a "representative consumer" model, and/or worked under restrictive parametric assumptions (Willig, Salop, and Scherer (1991), Hausman (1996), Hausman and Leonard (2002), Hausman (2003), Trajtenberg (1989)). In contrast, our set-up allows for (a) unrestricted consumer heterogeneity, and (b) arbitrary effects of eliminating the alternative on the

³Some of these restrictive assumptions have subsequently been replaced with less stringent parametric assumptions; cf. Herriges and Kling (1999), McFadden and Train (2000), Dagsvik and Karlstrom (2005), and Goolsbee and Klenow (2006). All of these still require specification of the dimension and distribution of unobserved heterogeneity and functional form of utility functions, about which no a priori information is available. Indeed, all welfare related measures are identified in these settings solely from assumed knowledge of these functional forms.

price of *existing* alternatives, thus greatly increasing the scope of the results.⁴ Importantly, these results are robust to failure of parametric assumptions, and clarify which features of welfare distributions can and which ones cannot be learned from demand data alone without functional form assumptions. Indeed, we show below that Hicksian welfare distributions in this case can again be expressed as closed-form transformations of choice probabilities without any assumption on heterogeneity. Nonetheless, calculating the tails of these distributions require observation of demand at prices where aggregate demand at an adjusted income level reaches zero. At high income levels, such prices may not be observed, since producers have no incentive to set a price where aggregate demand is zero. In that case, a lower bound on the welfare distributions can be obtained using the value of aggregate demand at the highest observed price.⁵ A heuristic empirical practice, popular in empirical research, for calculating welfare effects of new goods involves integrating the choice probability function from the current price to infinity. We provide a formal justification of this ad hoc practice, and show that this does indeed yield the average EV, if prices of substitutes remain unchanged; however, calculation of the average CV and/or allowing for prices of substitutes to change entail different expressions.

Section 3 of this paper considers two extensions, namely welfare analysis for (i) change in characteristics of an alternative, with or without an accompanying price change, and (ii) multinomial choice with nonexclusive options. An example of the former is the classic transport choice setting where, say, increasing the frequency of buses would raise utility from choosing the bus option, as opposed to driving or walking. Typically, characteristics of alternatives do not appear as additive components of net income in utilities, unlike prices (via the budget constraint), and hence the price change results are not useful for analyzing welfare effects of quality changes. We show that typically such welfare distributions are not identified from structural choice probabilities; nonetheless, bounds can be constructed on them, and hence on average welfare. We also show that further restrictions on preferences, such as weak separability of heterogeneity, restores point identification.

Our second extension pertains to choice among nonexclusive options. An example is where a cable-TV company offering a sports package and a movie package raises product prices; the resulting welfare calculations require new results because the packages are not exclusive alternatives for a potential consumer. We show that the welfare distribution resulting from a single price change can be directly expressed as a closed-form functional of choice probabilities but that for multiple price changes it cannot, unlike the case of multinomial choice among *exclusive* options. We then show how to construct nonparametric bounds for these distributions.

Taken together, these results provide new insights into welfare analysis in practical discrete choice situations, as well as providing practitioners with useful empirical tools

⁴Welfare calculations in the two scenarios described above correspond to situations where the price vectors before and after are given. The process through which the final price vector following the relevant change is calculated requires modeling the supply side; see Hausman and Leonard (2002, pp. 256–258), for an illustration of the methodology.

⁵For common *parametric* models, like mixed logit, these issues are assumed away, and welfare distributions are point-identified simply via functional form assumptions.

for evaluating policy changes in real life settings. In Section 4, we discuss some issues related to practical implementation of our results, and state a new and useful finding, namely that under income endogeneity and corresponding to a price increase, the EV but not the CV can be used for legitimate welfare analysis even in the absence of instruments or control functions.

Section 5 provides an empirical illustration of our methods in a setting of multinomial choice among modes of fishing. Using microdata, previously analyzed in [Herriges and Kling \(1999\)](#), we conduct private welfare analysis of price and quality changes, and the elimination of choice alternatives (e.g., for environmental protection). We show that logsum based estimates are substantively very different from those obtained using our methods. In particular, our estimates vary much less with income, and more importantly, have an income-gradient of the opposite sign, implying the exact opposite conclusion regarding progressivity of the economic changes, compared with the logsum estimates.

The proofs of all theoretical results are provided in the [Appendix](#).

2. MULTIPLE PRICE CHANGES

Set-up and notation: Consider a multinomial choice situation where alternatives are indexed by $j = 1, \dots, J$; individual income is denoted by Y , and price of alternative j by P_j . Individual utility from choosing alternative j is $U_j(Y - P_j, \eta)$, $j = 1, \dots, J$, where η denotes individual heterogeneity of unknown dimension; η is distributed in the population with unknown marginal CDF $F_\eta(\cdot)$. We have a cross-sectional random sample of consumers, and observe their characteristics including income, the prices they face, and the choice they make. The following analysis implicitly conditions on observable nonincome characteristics.

Define the *structural* choice probability for alternative j evaluated at price vector \mathbf{p} and income y , denoted $\{q_j(\mathbf{p}, y)\}$, $j = 1, \dots, J$, as

$$q_j(\mathbf{p}, y) = \int \mathbf{1}\left\{U_j(y - p_j, r) > \max_{k \neq j} \{U_k(y - p_k, r)\}\right\} dF_\eta(r). \quad (1)$$

In words, if we randomly sample individuals from the population, and offer the price vector \mathbf{p} and income y to each sampled individual, then a fraction $q_j(\mathbf{p}, y)$ will choose alternative j , in expectation.

ASSUMPTION 1. *Assume that for each η and for each $j = 1, \dots, J$, the utility function $U_j(\cdot, \eta)$ is continuous and strictly increasing.*

The first part of Assumption 1 simply says that utilities are nonsatiated with respect to the numeraire. Note that this assumption leaves the dimension of heterogeneity completely unspecified, and says nothing about how utility changes with unobserved heterogeneity.

Now consider a hypothetical change in the price vector from $\mathbf{p}_0 \equiv (p_{10}, p_{20}, \dots, p_{J0})$ to $\mathbf{p}_1 \equiv (p_{11}, p_{21}, \dots, p_{J1})$. Then the EV at income y for an η type consumer is the income reduction S in the initial situation that would lead to attainment of the eventual

indirect utility. Formally, the EV is the solution S to the equation:

$$\begin{aligned} & \max\{U_1(y - p_{11}, \eta), U_2(y - p_{21}, \eta), \dots, U_J(y - p_{J1}, \eta)\} \\ & = \max\{U_1(y - S - p_{10}, \eta), U_2(y - p_{20} - S, \eta), \dots, U_J(y - p_{J0} - S, \eta)\}. \end{aligned} \quad (2)$$

Similarly, the CV is the income compensation in the eventual situation necessary to restore the initial indirect utility; formally, the CV is the solution S to the equation:

$$\begin{aligned} & \max\{U_1(y + S - p_{11}, \eta), U_2(y + S - p_{21}, \eta), \dots, U_J(y + S - p_{J1}, \eta)\} \\ & = \max\{U_1(y - p_{10}, \eta), U_2(y - p_{20}, \eta), \dots, U_J(y - p_{J0}, \eta)\}. \end{aligned} \quad (3)$$

As η varies in the population, the CV and EV will have a distribution across consumers. Our goal is to estimate these distribution functions using the cross-sectional dataset.

Note that by analogy with the single price change case, one can attempt to define the change in average Marshallian consumer surplus in the multiple price change case via the line-integral

$$CS(L) = - \int_L \sum_{j=1}^J q_j(\mathbf{p}, y) dp_j, \quad (4)$$

where L denotes a path from \mathbf{p}_0 to \mathbf{p}_1 (cf. [Auerbach \(1985, equation \(2.2\)\)](#)). The negative sign stems from the fact that rise in price leads to a loss in consumer surplus.

In the set-up described above, we first show that for arbitrary price changes, the EV and CV are well-defined under [Assumption 1](#), but the Marshallian Consumer Surplus is not. We then show why one cannot get welfare distributions for multiple price changes by iterating a simpler result for a single price change, and then establish the first main result of this paper, namely that the marginal distributions of individual-level EV and CV for arbitrary changes in the price vector can be obtained as closed-form functionals of the structural choice probabilities.

We will assume without loss of generality that

$$p_{J1} - p_{J0} \geq p_{J-1,1} - p_{J-1,0} \geq \dots \geq p_{11} - p_{10}. \quad (5)$$

That is, label the alternative with the smallest price change (the smallest could be a negative number, representing a fall in price) alternative 1, the next smallest as alternative 2 and so on.

Now, it is well known that for continuous choice with multiple prices changing simultaneously, the Marshallian Consumer Surplus is generically undefined in the sense that the corresponding line-integral is path-dependent, but the Hicksian CV and EV continue to remain well-defined (cf., [Tirole \(1988, p. 11\)](#)). Our first result establishes that the same conclusion holds for discrete choice.

PROPOSITION 1. *Consider the multinomial choice set-up with J alternatives. Consider a price change from $\mathbf{p}_0 \equiv (p_{10}, p_{20}, \dots, p_{J0})$ to $\mathbf{p}_1 \equiv (p_{11}, p_{21}, \dots, p_{J1})$. Under [Assumption 1](#), the individual compensating and equivalent variations are uniquely defined.*

PROOF. Let $\mathbf{p}_0 = (p_{10}, p_{20}, \dots, p_{J0})$ denote the initial price vector and $\mathbf{p}_1 = (p_{11}, p_{21}, \dots, p_{J1})$ denote the final price vector. By definition,

$$\begin{aligned} & \max\{U_1(y - p_{11}, \eta), \dots, U_J(y - p_{J1}, \eta)\} \\ & = \max\{U_1(y - p_{10} - S, \eta), \dots, U_J(y - p_{J0} - S, \eta)\}. \end{aligned} \tag{6}$$

Since each $U_j(\cdot, \eta)$ is continuous, the RHS must be continuous in S . Now, if $S < p_{11} - p_{10}$, then the RHS will be larger than the LHS, and if $S > p_{J1} - p_{J0}$, then the RHS will be smaller than the LHS. By continuity and the intermediate value theorem, there must exist at least one S that solves (6). Now, since each $U_j(\cdot, \eta)$ is strictly increasing, the RHS of (6) is strictly decreasing in S . Therefore, the solution must be unique. An exactly analogous argument works for the CV where the corresponding equality is

$$\begin{aligned} & \max\{U_1(y - p_{11} + S, \eta), \dots, U_J(y - p_{J1} + S, \eta)\} \\ & = \max\{U_1(y - p_{10}, \eta), \dots, U_J(y - p_{J0}, \eta)\}. \end{aligned} \tag{7}$$

□

Marshallian consumer surplus: It can be shown that for discrete choice with *multiple* price changes, the integral (4) is path-dependent. In the [Appendix](#), we demonstrate this for the case with 3 alternatives.

2.1 Welfare distributions for multiple price changes

[Bhattacharya \(2015\)](#) showed that in the above setting, when the price of a single alternative changes *ceteris paribus*, the resulting CV and EV distributions can be expressed as closed-form functionals of choice probabilities. Interestingly, one cannot iterate this single price change result to obtain the welfare distribution for simultaneous changes in the prices of *multiple* alternatives. This fact is shown in the [Appendix](#). Consequently, an independent result is required for welfare analysis corresponding to *simultaneous* changes in multiple prices, and it is given by the following theorem.

THEOREM 1. *Consider the multinomial choice set up with J exclusive alternatives. Consider a price change from $\mathbf{p}_0 \equiv (p_{10}, p_{20}, \dots, p_{J0})$ to $\mathbf{p}_1 \equiv (p_{11}, p_{21}, \dots, p_{J1})$ satisfying (5). Denote $p_{j1} - p_{j0}$ by Δp_j for $j = 1, \dots, J$. Under Assumption 1, the marginal distribution of the individual EV evaluated at income y is given by*

$$\begin{aligned} & \Pr(\text{EV} \leq a) \\ & = \begin{cases} 0, & \text{if } a < \Delta p_1, \\ \sum_{k=1}^j q_k \left(\begin{matrix} p_{11}, \dots, p_{j1}, \\ p_{j+1,0} + a, \dots, p_{J0} + a, y \end{matrix} \right), & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, a \geq 0, 1 \leq j \leq J - 1, \\ \sum_{k=1}^j q_k \left(\begin{matrix} p_{11} - a, \dots, p_{j1} - a, \\ p_{j+1,0}, \dots, p_{J0}, y - a \end{matrix} \right), & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, a < 0, 1 \leq j \leq J - 1, \\ 1, & \text{if } a \geq \Delta p_J, \end{cases} \end{aligned} \tag{8}$$

while that of the individual CV evaluated at income y is given by

$$\Pr(\text{CV} \leq a) = \begin{cases} 0, & \text{if } a < \Delta p_1, \\ \sum_{k=1}^j q_k \left(\begin{array}{l} p_{11}, \dots, p_{j1}, \\ p_{j+1,0} + a, \dots, p_{J0} + a, y + a \end{array} \right), & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, a \geq 0, 1 \leq j \leq J-1, \\ \sum_{k=1}^j q_k \left(\begin{array}{l} p_{11} - a, \dots, p_{j1} - a, \\ p_{j+1,0}, \dots, p_{J0}, y \end{array} \right), & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, a < 0, 1 \leq j \leq J-1, \\ 1, & \text{if } a \geq \Delta p_J, \end{cases} \quad (9)$$

where q_k s are defined above in equation (1). (The separate entries for $a \geq 0$ and $a < 0$ in each line of (8) and (9) arise from accommodating rise and fall of prices, respectively).

These distributional results cover positive, zero, and negative price changes. For example, in the 3-alternative case, suppose alternative 2 is the outside option with price $p_{21} = p_{20} = 0$, and $\Delta p_1 < 0 < \Delta p_3$, then (8) becomes

$$\Pr(\text{EV} \leq a) = \begin{cases} 0, & \text{if } a < \Delta p_1, \\ q_1(p_{11} - a, 0, p_{30}, y - a), & \text{if } \Delta p_1 \leq a < 0, \\ q_1(p_{11}, 0, p_{30} + a, y) + q_1(p_{11}, 0, p_{30} + a, y), & \text{if } 0 \leq a < \Delta p_3, \\ 1, & \text{if } a \geq \Delta p_3. \end{cases} \quad (10)$$

Finally, note that in the above results, Assumption 1 guarantees that the CDFs are non-decreasing. For instance, in the 3-alternative case, for the CV we require

$$q_1(p_{11}, p_{20} + a, p_{30} + a, y + a) \leq q_1(p_{11}, p_{20} + a', p_{30} + a', y + a') \quad (11)$$

whenever $p_{11} - p_{10} \leq a < a'$. But this is true because the LHS is the probability of the event

$$\begin{aligned} U_1(y + a - p_{11}, \eta) &\geq \max\{U_2(y - p_{20}, \eta), U_3(y - p_{30}, \eta)\} \\ \Rightarrow U_1(y + a' - p_{11}, \eta) &\geq \max\{U_2(y - p_{20}, \eta), U_3(y - p_{30}, \eta)\}, \quad \text{since } a' > a \\ \Leftrightarrow U_1(y + a' - p_{11}, \eta) &\geq \max \left\{ \begin{array}{l} U_2(y + a' - (p_{20} + a'), \eta), \\ U_3(y + a' - (p_{30} + a'), \eta) \end{array} \right\}, \end{aligned}$$

whose probability is the RHS of the previous display.

REMARK 1. It is implicit throughout that the price changes do not alter the distribution of heterogeneity, for example, a large tuition subsidy in a district might attract outsiders with a strong preference for education to migrate in, altering the distribution of preferences relative to the status-quo. In other words, the price changes considered here are assumed to be modest enough to have no impact on the distribution of η , as is implicitly assumed in this literature.

The expected welfare change and deadweight loss follow from the CDF. For example, if (10) results from a subsidy $|\Delta p_1|$ on alternative 1, and a tax Δp_3 on alternative 3, then

$$E(\text{Deadwt_Loss}) = E(\text{EV}) - \Delta p_1 \times q_1(p_{11}, 0, p_{31}, y) \\ - \Delta p_3 \times q_3(p_{11}, 0, p_{31}, y + \Delta p_3).$$

2.2 Elimination of an alternative

Consider a setting of multinomial choice among exclusive alternatives $\{1, \dots, J + 1\}$. Suppose the alternative $J + 1$ is eliminated subsequently, which can potentially affect consumer welfare by both restricting the choice set and also by affecting the prices of other alternatives. Assume that we have data on a cross section of individual choices in the preelimination situation. We wish to calculate the distribution of Hicksian welfare effects that would result from eliminating the $J + 1$ th alternative. Applied researchers typically use the heuristic idea that eliminating an alternative is “like” raising its price to infinity, and thereafter use the welfare formulae for price change to evaluate welfare effects of eliminating an alternative. Our analysis below provides a formal justification for this heuristic approach, and shows that it is valid only when prices of other alternatives do not change and when one is interested in the EV. Furthermore, the incorporation of preference heterogeneity and income effects reveals interesting differences between CV and EV based formulae and identifiability of their distribution.

Toward that end, consider an individual at income y and unobserved heterogeneity η whose utility from consuming alternative j at price p_j is given by $U_j(y - p_j, \eta)$. The problem is to find the distribution of welfare effects across such individuals resulting from potentially eliminating alternative $J + 1$. Suppose from an initial price vector $(p_{10}, \dots, p_{J0}, p_{J+1})$, following elimination of the $J + 1$ th alternative, the eventual price vector becomes (p_{11}, \dots, p_{J1}) . The marginal distributions of EV and CV resulting from this change are given by the following corollary, whose proof appears in the [Appendix](#).

COROLLARY 1. *Suppose Assumption 1 holds for all alternatives. Also assume that for any bundle $(j, y - p_j)$, if $p_j \uparrow \infty$ with prices of other alternatives finite, then for each η at least one other bundle $(k, y - p_k)$ will be strictly preferred to $(j, y - p_j)$. Let $q_k(\cdot, \dots, \cdot, y)$ be as*

defined in (1), with $J + 1$ alternatives. Then

$$\Pr(\text{CV} \leq a) = \begin{cases} 0, & \text{if } a < \Delta p_1, \\ \sum_{k=1}^j q_k \begin{pmatrix} p_{11}, \dots, p_{j1}, \\ p_{j+1,0} + a, \dots, p_{J0} + a, p_{J+1} + a, \\ y + a \end{pmatrix}, & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, j = 1, \dots, J - 1, a \geq 0, \\ \sum_{k=1}^j q_k \begin{pmatrix} p_{11} - a, \dots, p_{j1} - a, \\ p_{j+1,0}, \dots, p_{J0}, p_{J+1}, y \end{pmatrix}, & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, 1 \leq j < J - 1, a < 0, \\ 1 - q_{J+1}(p_{11}, \dots, p_{J1}, p_{J+1} + a, y + a), & \text{if } \Delta p_J \leq a, a \geq 0, \\ 1 - q_{J+1}(p_{11} - a, \dots, p_{J1} - a, p_{J+1}, y), & \text{if } \Delta p_J \leq a < 0. \end{cases} \quad (12)$$

On the other hand,

$$\Pr(\text{EV} \leq a) = \begin{cases} 0, & \text{if } a < p_{11} - p_{10}, \\ \sum_{k=1}^j q_k \begin{pmatrix} p_{11}, \dots, p_{j1}, \\ p_{j+1,0} + a, \dots, p_{J0} + a, p_{J+1} + a, y \end{pmatrix}, & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, j = 1, \dots, J - 1, a \geq 0, \\ \sum_{k=1}^j q_k \begin{pmatrix} p_{11} - a, \dots, p_{j1} - a, \\ p_{j+1,0}, \dots, p_{J0}, p_{J+1}, y - a \end{pmatrix}, & \text{if } \Delta p_j \leq a < \Delta p_{j+1}, 1 \leq j < J - 1, a < 0, \\ 1 - q_{J+1}(p_{11}, \dots, p_{J1}, p_{J+1} + a, y), & \text{if } \Delta p_J \leq a, a \geq 0, \\ 1 - q_{J+1}(p_{11} - a, \dots, p_{J1} - a, p_{J+1}, y - a), & \text{if } \Delta p_J \leq a < 0. \end{cases} \quad (13)$$

As in the previous theorem, the pairs of results for $a < 0$ and $a \geq 0$ correspond to which existing alternatives have become more and less expensive, respectively, following the elimination of the $J + 1$ th alternative. Also, note that in order to calculate the probabilities appearing in Theorem 2, we need to observe adequate cross-sectional variation in the price of all $J + 1$ alternatives in the preelimination period.

COROLLARY 2. *If elimination of the alternative has no effect on prices of the other alternatives, then $p_{j1} = p_{j0}$ for all $j = 1, \dots, J$, and the above results simplify to*

$$\Pr(\text{CV} \leq a) = \begin{cases} 0, & \text{if } a < 0, \\ 1 - q_{J+1}(p_{10}, \dots, p_{J0}, p_{J+1} + a, y + a), & \text{if } 0 \leq a, \end{cases} \quad (14)$$

$$\Pr(\text{EV} \leq a) = \begin{cases} 0, & \text{if } a < 0, \\ 1 - q_{J+1}(p_{10}, \dots, p_{J0}, p_{J+1} + a, y), & \text{if } 0 \leq a. \end{cases} \quad (15)$$

The corresponding average values are given by

$$E(\text{CV}) = \int_{p_{J+1}}^{\infty} q_{J+1}(p_{10}, \dots, p_{J0}, r, y + r - p_{J+1}) dr, \quad (16)$$

$$E(\text{EV}) = \int_{p_{J+1}}^{\infty} q_{J+1}(p_{10}, \dots, p_{J0}, r, y) dr, \quad (17)$$

using the change of variable $r = p_{J+1} + a$.

The expression (17) is commonly used as an ad hoc measure of the welfare effect of introducing a new product. Thus it follows from the above discussion that if elimination of the alternative entails no price change for the other alternatives, then the commonly used expression happens to equal the mean EV, but ceases to be so if one is interested in the mean CV, or prices of substitutes also change.

In order to calculate the above expressions *nonparametrically*, a researcher needs to observe demand up to the price where the choice probability becomes zero. Typically, in a dataset, one is unlikely to observe such prices, since producers have no incentive to raise prices where revenue is zero. But then one can obtain a lower bound for $E(\text{CV})$ by integrating up to the highest price observed in the dataset where demand from consumers is nonzero. This is in contrast to the case of price changes for multiple alternatives, reported in equation (8) above, where welfare distributions are nonparametrically identified as long as the hypothetical price changes are within the range of the observed price data. Of course, for *parametric* choice probabilities, for example, random coefficient logit, expressions like (17) are identified directly from functional form assumptions.

Finally, note that the above expressions can also be used for retrospective calculation of welfare distributions corresponding to *introduction* of a new alternative, simply by interchanging the labels of EV and CV. Here, “retrospective” means that we must have consumption data after the product has been introduced.

3. TWO EXTENSIONS

In this section, we develop results for nonparametric welfare analysis in two practically important settings, namely (i) quality change in an alternative, and (ii) price changes when alternatives are nonexclusive. In these two cases, in general, one *cannot* express the distribution of welfare using functionals of structural choice probabilities; nonetheless, bounds can be constructed on the welfare distribution functions, and thus, on average welfare, as shown below.

3.1 Change in characteristics

The first setting we consider is where characteristics of a choice alternative changes. For example, consider a household in a poor country that is deciding whether to send its teenage daughter to school, or to let her stay at home. It is intuitive that the teacher student ratio in the local school would be an important determinant of the utility from

school attendance. Similarly, in transportation choice settings, increasing the frequency of buses would raise every individual's utility from choosing the bus option, as opposed to driving or walking. These characteristics, unlike prices, do not appear as additive components of net income inside utilities. The additivity of price, resulting from the budget constraint, was instrumental in deriving the results of the previous section. Therefore, welfare analysis of changes in characteristics would require a result that is different from those above. In this subsection, we develop such a result.

To keep the exposition transparent, we will focus on binary choice, and then outline how to generalize to the multinomial case. Accordingly, suppose an individual at income y and preference η is choosing between two options 0 and 1, with utilities given by $U_0(y, \eta)$ and $U_1(y - p, x, \eta)$, where x represents a vector of characteristics of option 1. For example, in the schooling example, option 0 is staying at home, option 1 is going to school, p is the tuition fee, and x is the teacher-student ratio in the local school. Define the structural choice probability of alternative 1 at price p , income y and quality x as

$$q_1(p, y, x) = \int 1\{U_1(y - p, x, r) > U_0(y, r)\} dF_\eta(r), \quad (18)$$

where $F_\eta(\cdot)$ represents the marginal distribution of unobserved heterogeneity.

We will make the following assumption.

ASSUMPTION 2. (i) For each x and η , $U_1(\cdot, x, \eta)$ is strictly increasing and continuous; additionally, for each η , $U_0(\cdot, \eta)$ is strictly increasing and continuous.

Now, suppose from an initial price p_0 and initial quality x_0 , the price and attribute change to p_1 and x_1 , respectively. It is reasonable to assume that improvement of quality will be funded, at least partially, by an increase in prices, and hence we consider $p_1 \geq p_0$. Setting $p_1 = p_0$, we can cover the case where there is no change in price, for example, the improvement is funded by external resources.

The private CV resulting from this change is defined by the solution S to the equation

$$\max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} = \max\{U_0(y + S, \eta), U_1(y + S - p_1, x_1, \eta)\}. \quad (19)$$

Our first result is that the distribution of S cannot be identified from knowledge of the structural choice probabilities. We demonstrate this in the [Appendix](#) via a counterexample, and state the conclusion as a theorem.

THEOREM 2. Under Assumption 2, the marginal distribution of S is not identified from the structural choice probabilities (the proof is in the [Appendix](#)).

One can, of course, construct bounds on these distributions, which are provided in Propositions 2 and 3 in the [Appendix](#). To see where these bounds come from, note from

(19) and Assumption 2(i), that for $a < 0$,

$$\begin{aligned} & \Pr(\text{CV} \leq a) \\ &= \Pr \left[\begin{array}{l} \max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} \\ \leq \max\{U_0(\underbrace{y+a}_{<y}, \eta), U_1(y+a - p_1, x_1, \eta)\} \end{array} \right] \\ &= \Pr[\max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} \leq U_1(y + a - p_1, x_1, \eta)]. \end{aligned}$$

The Frechet upper bound on this is given by

$$\begin{aligned} & \min \left\{ \begin{array}{l} \Pr[U_0(y, \eta) \leq U_1(y + a - p_1, x_1, \eta)], \\ \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] \end{array} \right\} \\ &= \min \left\{ \begin{array}{l} q_1(p_1 - a, y, x_1), \\ \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] \end{array} \right\}. \end{aligned}$$

The ultimate upper bound comes from a lower bound on $\Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)]$, obtained using

$$\begin{aligned} & \Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)] \\ & \geq \max_y \Pr[U_1(y - p_0, x_0, \eta) \geq U_0(y', \eta) \geq U_1(y + a - p_1, x_1, \eta)], \end{aligned}$$

and then applying Frechet bounds to this. In the application below, we provide explicit numerical expressions for these bounds.

It is straightforward to extend the above result to the case of multinomial choice. Formally, if we have alternatives $0, 1, \dots, J$ with the quality of alternative 1 changing from x_{10} to x_{11} , and price changing from p_0 to p_1 , then the CV solves

$$\begin{aligned} & \max\{U_0(y, \eta), U_1(y - p_0, x_{10}, \eta), U_2(y - p_2, x_2, \eta), \dots, U_J(y - p_J, x_J, \eta)\} \\ &= \max \left\{ \begin{array}{l} U_0(y + S, \eta), U_1(y + S - p_1, x_{11}, \eta), \\ U_2(y + S - p_2, x_2, \eta), \dots, U_J(y + S - p_J, x_J, \eta) \end{array} \right\}. \end{aligned}$$

Letting

$$U_{-1}(y, p, x, \eta) \equiv \max\{U_0(y, \eta), U_2(y - p_2, x_2, \eta), \dots, U_J(y - p_J, x_J, \eta)\},$$

the previous display reduces to

$$\begin{aligned} & \max\{U_{-1}(y, p, x, \eta), U_1(y - p_0, x_{10}, \eta)\} \\ &= \max\{U_{-1}(y + S, p, x, \eta), U_1(y + S - p_1, x_{11}, \eta)\}, \end{aligned}$$

which is the same as (19) with $U_0(y, \eta)$ replaced by $U_{-1}(y, p, x, \eta)$. The corresponding bounds on the CV distribution (see equation (44) in the [Appendix](#)) would involve replacing $q_0(\cdot)$ s with the probability of not buying alternative 1.

REMARK 2. Defining and proving “sharpness” of such bounds without any assumptions on the structure of preferences and dimension of heterogeneity (except Assumption 2) would be an interesting exercise. We believe that such results are best developed as corollaries of some fundamental revealed preference type results, related to [McFadden and Richter \(1990\)](#). As such, there is sufficient material here for an independent paper. We are currently exploring these and some related questions in a separate project.

We conclude this subsection with the observation that further restrictions on preferences, though short of fully parametric specifications, can yield point-identification of welfare distributions for quality change. One such restriction is that the utility from alternative 1 is weakly separable in unobserved heterogeneity, that is, $U_1(y - p, x, \eta) \equiv V_1(h_1(y - p, x), \eta)$, where $V_1(\cdot, \eta)$ is strictly increasing for each η , and $h_1(\cdot, \cdot)$ is continuous and strictly increasing in the first component. A special case of weak separability, where the functional form of $h_1(\cdot, \cdot)$ is assumed *known* up to estimable real parameters is the familiar “index restriction,” popular in the semiparametrics literature. Weak separability has been used previously in econometrics for identifying effects of endogenous dummy variables (cf. [Vytlacil and Yildiz \(2007\)](#)), while index restrictions have been used for achieving identification in general demand systems; cf. [Berry and Haile \(2015\)](#).

Consider a change from (p_0, x_0) to (p_1, x_1) satisfying the following condition.

CONDITION S. *Assume that the change from (p_0, x_0) to (p_1, x_1) is small enough that there exists a finite a satisfying*

$$h_1(y + a - p_1, x_1) = h_1(y - p_0, x_0).$$

This basically says that the quality and price change are not so large that no amount of compensation can restore the utility from consuming alternative 1 to the prechange level for any η .

PROPOSITION 2. *Suppose weak separability holds. Consider a change from (p_0, x_0) to (p_1, x_1) satisfying condition (S). Then there is a unique \bar{a} that solves $q_1(p_1 - \bar{a}, y, x_1) = q_1(p_0, y, x_0)$. Furthermore, the distribution of individual CV resulting from this change is point-identified from the structural choice probabilities, and it is given by the following expressions:*

If $\bar{a} \leq 0$, then

$$\Pr(S \leq a) = \begin{cases} 0, & \text{if } a < \bar{a}, \\ q_1(p_1 - a, y, x_1), & \text{if } \bar{a} \leq a < 0, \\ 1, & \text{if } a \geq 0. \end{cases}$$

If $\bar{a} > 0$, then

$$\Pr(S \leq a) = \begin{cases} 0, & \text{if } a < 0, \\ q_0(p_0 + a, y + a, x_0), & \text{if } 0 \leq a < \bar{a}, \\ 1, & \text{if } a \geq \bar{a}. \end{cases}$$

(The proof is in the [Appendix](#).)

3.2 Nonexclusive discrete choice

Assume that there are two binary choices which are nonexclusive among themselves. For example, suppose choice 1 for a household is whether to subscribe to a sports package offered by a cable TV network which costs P_1 and choice 2 is whether to subscribe to a movie package which costs P_2 . A household with income Y and preferences η then has four exclusive options— $\{1\}$, $\{2\}$, $\{1, 2\}$, $\{0\}$ (where $\{0\}$ denotes choosing none of the two packages) with respective utilities $U_1(Y - P_1, \eta)$, $U_2(Y - P_2, \eta)$, $U_{12}(Y - P_1 - P_2, \eta)$, and $U_0(Y, \eta)$, respectively.

Single price change: Consider the CV corresponding to a rise in the price of the sports package from p_{10} to p_{11} with the price of the movie package fixed at p_2 . The CV evaluated at income $Y = y$ is the solution to the equation

$$\begin{aligned} & \max \left\{ \begin{array}{l} U_0(y + CV, \eta), U_1(y + CV - p_{11}, \eta), \\ U_2(y + CV - p_2, \eta), U_{12}(y + CV - p_{11} - p_2, \eta) \end{array} \right\} \\ & = \max \left\{ \begin{array}{l} U_0(y, \eta), U_1(y - p_{10}, \eta), \\ U_2(y - p_2, \eta), U_{12}(y - p_{10} - p_2, \eta) \end{array} \right\}. \end{aligned} \tag{20}$$

Now, group option $\{1\}$ and $\{1, 2\}$ together (call it group A) and options $\{0\}$ and $\{2\}$ together and call it group B . Define

$$\begin{aligned} \varepsilon & \stackrel{\text{def}}{=} (p_2, \eta), \\ V_A(y - p_1, \varepsilon) & \stackrel{\text{def}}{=} \max\{U_1(y - p_1, \eta), U_{12}(y - p_1 - p_2, \eta)\}, \\ V_B(y, \varepsilon) & \stackrel{\text{def}}{=} \max\{U_0(y, \eta), U_2(y - p_2, \eta)\}. \end{aligned}$$

Then (20) becomes

$$\max\{V_A(y + CV - p_{11}, \varepsilon), V_B(y + CV, \varepsilon)\} = \max\{V_A(y - p_{10}, \varepsilon), V_B(y, \varepsilon)\}. \tag{21}$$

If the U functions are strictly increasing in the first argument for each η , then so are $V_A(\cdot, \varepsilon)$ and $V_B(\cdot, \varepsilon)$ for each ε . Now we can apply Theorem 1 of [Bhattacharya \(2015\)](#) for binary choice to this problem and get the marginal distribution of the compensating variation. For example, for $0 \leq a < p_{11} - p_{10}$, Theorem 1 of [Bhattacharya \(2015\)](#) gives

$$\begin{aligned} & \Pr(CV \leq a) \\ & = \Pr[V_B(y + a, \varepsilon) \geq V_A(y + a - (p_{10} + a), \varepsilon)] \\ & = \Pr \left[\begin{array}{l} \max\{U_0(y + a, \eta), U_2(y + a - p_2, \eta)\} \\ \geq \max \left\{ \begin{array}{l} U_1(y + a - (p_{10} + a), \eta), \\ U_{12}(y + a - (p_{10} + a) - p_2, \eta) \end{array} \right\} \end{array} \right] \\ & = q_0(p_{10} + a, p_2, y + a) + q_2(p_{10} + a, p_2, y + a). \end{aligned}$$

Thus

$$\Pr(\text{CV} \leq a) = \begin{cases} 0, & \text{if } a < 0, \\ q_0(p_{10} + a, p_2, y + a) + q_2(p_{10} + a, p_2, y + a), & \text{if } 0 \leq a < p_{11} - p_{10}, \\ 1, & \text{if } a \geq p_{11} - p_{10}. \end{cases}$$

Multiple price changes: The key fact enabling us to write (20) as the binary choice CV (21) is that P_2 is being held fixed at p_2 ; if P_2 also varied across individuals, then the distribution of ε would vary beyond the variation of η and the binary formulation would no longer be applicable. Indeed, for multiple price changes in the nonexclusive alternatives case, welfare distributions can no longer be written in terms of choice probabilities. To see this, consider a simultaneous rise in P_1 and P_2 from (p_{10}, p_{20}) to (p_{11}, p_{21}) . Assume that $0 < p_{11} - p_{10} < p_{21} - p_{20}$.

Then the CV is defined via

$$\begin{aligned} & \max \left\{ U_0(y + \text{CV}, \eta), U_1(y + \text{CV} - p_{11}, \eta), \right. \\ & \quad \left. U_2(y + \text{CV} - p_{21}, \eta), U_{12}(y + \text{CV} - p_{11} - p_{21}, \eta) \right\} \\ & = \max \{ U_0(y, \eta), U_1(y - p_{10}, \eta), U_2(y - p_{20}, \eta), U_{12}(y - p_{10} - p_{20}, \eta) \}. \end{aligned}$$

Now,

$$\begin{aligned} & \Pr[\text{CV} = p_{11} - p_{10}] \\ & = \Pr \left[U_1(y - p_{10}, \eta) \geq \max \left\{ \begin{array}{l} U_0(y, \eta), U_2(y - p_{20}, \eta), \\ U_{12}(y - p_{10} - p_{20}, \eta), \\ U_0(y + p_{11} - p_{10}, \eta), \\ U_2(y + p_{11} - p_{10} - p_{21}, \eta), \\ U_{12}(y - p_{10} - p_{21}, \eta) \end{array} \right\} \right] \\ & = \Pr \left[U_1(y - p_{10}, \eta) \geq \max \left\{ \begin{array}{l} U_2(y - p_{20}, \eta), U_{12}(y - p_{10} - p_{20}, \eta), \\ U_0(y + p_{11} - p_{10}, \eta) \end{array} \right\} \right]. \end{aligned} \tag{22}$$

This probability is not generically point-identified from the choice probabilities. To see this, consider the following counterexample.

Counterexample: Consider a classic McFadden (1973) type utility specification: $U_j(a, \eta) = \beta_j a + \eta_j$, for $j = 0, 1, 2, 12$, where the β_j are nonstochastic coefficients, and the η s are distributed standardized extreme valued. Then because the regressors $(y, y - p_1, y - p_2, y + p_1 - p_2)$ constitute a 3-dimensional subspace—the sum of the 1st and the 4th regressors equals the sum of the 2nd and the 3rd implying exact multicollinearity—the 4 coefficients $(\beta_0, \beta_1, \beta_2, \beta_{12})$ are not separately identified. A direct way to verify this is to note that the Hessian of the likelihood function is of the form $W \otimes XX'$, where $X \equiv (y, y - p_{10}, y - p_{20}, y - p_{10} - p_{20})'$ (cf. Böhning (1992, pp. 198–199)), and so its expectation is of the form $E(W \otimes XX')$. Given that the 4th component of X

is an exact linear combination of the first 3, in each block $E(W_{ij}XX')$ of the Kronecker product the 4th row is exactly the same linear combination of the first 3 rows. So it follows that $E(W \otimes XX')$ is globally singular. Therefore, by [Rothenberg \(1971, Theorem 1\)](#), the parameter vector β is not even locally identified.

Now the expression for $\Pr[CV = p_{11} - p_{10}]$ in this case equals

$$\begin{aligned}
 & q_1(y + p_{11} - p_{10}, y - p_{10}, y - p_{20}, y - p_{10} - p_{20}) \\
 &= \frac{\exp(\beta_1(y - p_{10}))}{\left(\begin{aligned} & \exp(\beta_0(y + p_{11} - p_{10})) + \exp(\beta_1(y - p_{10})) \\ & + \exp(\beta_2(y - p_{20})) + \exp(\beta_{12}(y - p_{10} - p_{20})) \end{aligned} \right)}. \tag{23}
 \end{aligned}$$

Since $x \equiv (y + p_{11} - p_{10}, y - p_{10}, y - p_{20}, y - p_{10} - p_{20})$ spans a 4-dimensional space (unless $p_{11} = p_{10}$), we cannot have that for some $\theta \neq \beta$,

$$\begin{aligned}
 & q_1(y + p_{11} - p_{10}, y - p_{10}, y - p_{20}, y - p_{10} - p_{20}; \beta) \\
 &= q_1(y + p_{11} - p_{10}, y - p_{10}, y - p_{20}, y - p_{10} - p_{20}; \theta)
 \end{aligned}$$

for all $p_{10}, p_{11}, p_{20}, y$, if the joint distribution of $(p_{10}, p_{11}, p_{20}, y)$ is nondegenerate. To see this explicitly, consider the following thought experiment. Suppose at a specific value x of the regressors, $q_1(x, \beta) = q_1(x, \theta)$, for some $\theta \neq \beta$. Now holding p_{10}, p_{20}, y fixed, if we increase p_{11} , then the denominator of $q_1(\cdot, \beta)$ will increase more (resp., less) than that of $q_1(\cdot, \theta)$ if $\beta_0 > \theta_0$ (resp., $\beta_0 < \theta_0$), so that $q_1(\cdot, \beta) = q_1(\cdot, \theta)$ cannot continue to hold. Therefore, we must have $\beta_0 = \theta_0$. Next, if we increase y, p_{10} and p_{20} by the same amount such that $y - p_{10}, y - p_{20}$ remain fixed, only the 4th term in the denominator of (23) will change, and by the same logic as above, maintaining $q_1(\cdot, \beta) = q_1(\cdot, \theta)$ would require $\beta_{12} = \theta_{12}$. Changing y and p_{10} by the same amount holding p_{11} and p_{20} fixed would deliver $\beta_2 = \theta_2$, and changing y while changing p_{20} by the same amount and changing p_{11} the same amount in the opposite direction while holding p_{10} fixed would deliver $\beta_1 = \theta_1$.

Thus although there exist $\theta \neq \beta$ which will produce the same choice probabilities for all potentially observable values of prices and income, the expression (22) is different with positive probability (w.r.t. the joint distribution of the regressors) if $\theta \neq \beta$. This implies that the CV distribution cannot be learned from the choice probabilities unless $p_{11} = p_{10}$, which is the single price change scenario.

We summarize the above discussion as a theorem.

THEOREM 3. *For multinomial choice with nonexclusive alternatives, the marginal distribution of EV and CV resulting from a change in the price of a single alternative can be expressed as a closed-form functional of the structural choice probabilities. For a simultaneous price change of multiple alternatives, the welfare distributions are not point-identified from the structural choice probabilities.*

Note that the conclusion of the above theorem is starkly different from the multinomial case with *exclusive* alternatives where welfare distributions are point-identified

from the structural choice probabilities for *both* single and multiple price changes. An intuition for the result is that in the nonexclusive case, there is a systematic relationship between the prices of the different ultimate options, so that in a dataset we can never observe independent price variations across the various (composite) options, no matter how many price combinations are observed across the constituent individual alternatives.

Bounds for the CV distribution: One can nonetheless bound the probability (22) using Assumption 1. In particular, let

$$A = \left\{ p_1^*, p_2^*, y^* : y^* - p_1^* \geq y - p_{10}, y^* - p_2^* \leq y - p_{20}, \right. \\ \left. y^* - p_1^* - p_2^* \leq y - p_{10} - p_{20}, y^* \leq y + p_{11} - p_{10} \right\},$$

$$B = \left\{ p_1^*, p_2^*, y^* : y^* - p_1^* \leq y - p_{10}, y^* - p_2^* \geq y - p_{20}, \right. \\ \left. y^* - p_1^* - p_2^* \geq y - p_{10} - p_{20}, y^* \geq y + p_{11} - p_{10} \right\}.$$

Then an upper bound consistent with Assumption 1 is given by

$$\Pr \left[U_1(y - p_{10}, \eta) \geq \max \left\{ U_2(y - p_{20}, \eta), U_{12}(y - p_{10} - p_{20}, \eta), \right. \right. \\ \left. \left. U_0(y + p_{11} - p_{10}, \eta) \right\} \right] \\ \leq \inf_{(p_1^*, p_2^*, y^*) \in A} \Pr \left[U_1(y^* - p_1^*, \eta) \geq \max \left\{ U_2(y^* - p_2^*, \eta), \right. \right. \\ \left. \left. U_{12}(y^* - p_1^* - p_2^*, \eta), \right. \right. \\ \left. \left. U_0(y^*, \eta) \right\} \right] \quad (24) \\ = \inf_{(p_1^*, p_2^*, y^*) \in A} q_1(p_1^*, p_2^*, y^*).$$

These bounds arise from the fact that if, for example, $y^* - p_1^* \geq y - p_{10}$, then $U_1(y^* - p_1^*, \eta) \geq U_1(y - p_{10}, \eta)$, by Assumption 1, so that the probability of $U_1(y - p_{10}, \eta)$ exceeding a specific number is smaller than that of $U_1(y^* - p_1^*, \eta)$ exceeding that same number. Similarly, $y^* - p_2^* \leq y - p_{20}$ implies that $U_2(y - p_{20}, \eta) \geq U_2(y^* - p_2^*, \eta)$ so that the probability of $U_2(y - p_{20}, \eta)$ being exceeded by a number is smaller than that of $U_2(y^* - p_2^*, \eta)$ being exceeded by it, etc.

By a similar logic, a lower bound on (22) is given by

$$\Pr \left[U_1(y - p_{10}, \eta) \geq \max \left\{ U_0(y + p_{11} - p_{10}, \eta), U_2(y - p_{20}, \eta), \right. \right. \\ \left. \left. U_{12}(y - p_{10} - p_{20}, \eta), \right\} \right] \\ \geq \sup_{(p_1^*, p_2^*, y^*) \in B} \Pr \left[U_1(y^* - p_1^*, \eta) \geq \max \left\{ U_2(y^* - p_2^*, \eta), \right. \right. \\ \left. \left. U_{12}(y^* - p_1^* - p_2^*, \eta), \right. \right. \\ \left. \left. U_0(y^*, \eta) \right\} \right] \quad (25) \\ = \sup_{(p_1^*, p_2^*, y^*) \in B} q_1(p_1^*, p_2^*, y^*).$$

REMARK 3. In some practical situations, suppliers may provide a discount if a household buys both options. If this discounted price varies in the cross section over and above that of the constituent options, then this setting becomes identical to a multinomial choice with four alternatives and independent price variation across them. One can then point-identify the EV/CV distributions for multiple price changes simply by using Theorem 1.

4. DISCUSSIONS ON IMPLEMENTATION

The results reported in Theorem 1, and the associated corollaries are fully nonparametric in that no functional form assumptions are required to derive them. When implementing these results in practical applications, one can therefore estimate conditional choice probabilities nonparametrically, for example, using kernel or series regressions, and then use those estimates to calculate welfare distributions. For instance, recall the 3-alternative case, leading to (10), where we have, say, $p_{11} - p_{10} < 0 = p_{21} - p_{20} < p_{31} - p_{30}$. For any random variable X with CDF $F(\cdot)$ and finite mean, the expectation satisfies

$$E(X) = - \int_{-\infty}^0 F(x) dx + \int_0^{\infty} (1 - F(x)) dx;$$

thus the mean EV from (10) is given by

$$- \int_{\Delta p_1}^0 q_1(p_{11} - a, 0, p_{30}, y - a) da + \int_0^{\Delta p_3} q_3(p_{11}, 0, p_{30} + a, y) da, \quad (26)$$

where $q_j(p_1, p_2, p_3, y)$ denotes the choice probability of alternative j when the price of the three alternatives are (p_1, p_2, p_3) and income is y . If the dataset is of modest size, so that kernel regressions are imprecise, then one can alternatively use a parametric approximation to the choice probabilities. Numerical integration routines are now available in popular software packages like STATA and MATLAB, and can be used to calculate the integrals of choice probabilities in the same way that consumer surplus was traditionally calculated by earlier researchers.

4.1 Endogeneity

In this subsection, we briefly discuss the issue of price/income “endogeneity,” that is, situations where price or income are correlated with unobserved preference heterogeneity η , conditional on all other covariates. Note that our theoretical results in Sections 2 and 3 establish the closed-form mapping between welfare distributions and *structural* choice probabilities. Once the structural choice probabilities (or nonparametric bounds on them) are identified, our results deliver identification of the corresponding welfare distributions *without requiring any additional modification due to endogeneity*. Nonetheless, in practical applications, estimation of the structural choice probabilities would be necessary before our results can be applied. In this subsection, we (a) briefly outline

how existing econometric methods can be used to estimate structural choice probabilities, and (b) more importantly, highlight that for price increase (resp., decrease) the estimation of EV (resp., CV) does not require any correction for income endogeneity.

Income endogeneity: In applications, observed income may be endogenous with respect to individual choice, for example, when omitted variables, such as unrecorded education level, can both determine individual choice and be correlated with income. Under such endogeneity, the *observed* choice probabilities would potentially differ from the *structural* choice probabilities, and one can define welfare distributions either unconditionally, or conditionally on income, analogous to the average treatment effect and the average effect of treatment on the treated, respectively, in the program evaluation literature. In this context, an important and useful insight, not previously noted, is that for a price rise, the distribution of the income-conditioned EV is not affected by income endogeneity, whereas that of the CV is; for a fall in price, the conclusion holds with CV and EV reversed.

To see why that is the case, recall the three alternative cases discussed above, and define the conditional-on-income structural choice probability at income y' as

$$q_j^c(p_1, p_2, p_3, y', y) = \int 1\{U_j(y' - p_j, \eta) \geq \max_{k \neq j} U_k(y' - p_k, \eta)\} dF(\eta|y),$$

where $F(\cdot|y)$ denotes the distribution of the unobserved heterogeneity η for individuals whose realized income is y , where y may or may not equal y' . Now, for a real number a , satisfying $p_{11} - p_{10} \leq a < p_{21} - p_{20}$, it is easy to see that similar to equation (10), the distributions of EV at a , evaluated at income y , conditional on realized income being y , are given by

$$\Pr(\text{EV} \leq a | \text{Inc} = y) = q_1^c(p_{11}, p_{20} + a, p_{30} + a, y, y), \quad (27)$$

while for CV it is given by

$$\Pr(\text{CV} \leq a | \text{Inc} = y) = q_1^c(p_{11}, p_{20} + a, p_{30} + a, y + a, y). \quad (28)$$

Now, $q_1^c(p_{11}, p_{20} + a, p_{30} + a, y, y)$, by definition, is the fraction of individuals currently at income y who would choose alternative 1 at prices $(p_{11}, p_{20} + a, p_{30} + a)$, had their income been y . Now if prices are exogenous in the sense that $\mathbf{P} \perp \eta|Y$, then the observable choice probability conditional on price \mathbf{p} and income y

$$\begin{aligned} & \bar{q}_1(\mathbf{p}, y) \\ &= \int 1\{U_1(y - p_1, \eta) \geq \max_{k \neq 1} U_k(y - p_k, \eta)\} f(\eta|\mathbf{p}, y) d\eta \\ & \stackrel{\text{by } \mathbf{P} \perp \eta|Y}{=} \int 1\{U_1(y - p_1, \eta) \geq \max_{k \neq 1} U_k(y - p_k, \eta)\} f(\eta|y) d\eta \\ &= q_1^c(\mathbf{p}, y, y). \end{aligned}$$

Therefore, (27) equals $\bar{q}_1(p_{11}, p_{20} + a, p_{30} + a, y, y)$, so no corrections are required owing to endogeneity. However, in (28), $q_1^c(p_{11}, p_{20} + a, p_{30} + a, y + a, y)$ is the fraction of individuals currently at income y who would choose alternative 1 at prices $(p_{11}, p_{20} + a, p_{30} + a)$, had their income been $y + a$. This fraction is counterfactual and not directly estimable because the distribution of η is likely to be different across people with income $y + a$ relative to those with income y , due to endogeneity. To summarize, if the objective of welfare analysis is to calculate the EV distribution resulting from price rise for individuals whose realized income equals the hypothetical income, then endogeneity of income is irrelevant to the analysis. This implies that if exogeneity of income is suspect and no obvious instrument or control function is available, then a researcher can still perform meaningful welfare analysis based on the EV distribution at realized income, provided price is exogenous conditional on income and other observed covariates. Furthermore, one can calculate aggregate welfare in the population by integrating $\bar{q}_1(\mathbf{p}, y) = q_1^c(\mathbf{p}, y, y)$ over the marginal distribution of income.

Price endogeneity: Price can be correlated with individual heterogeneity either because there are unobserved choice attributes which vary across individuals, or spatial concentration of individuals with strong preferences for the good could raise local prices. In such cases, control function methods (e.g., [Blundell and Powell \(2003\)](#)) can potentially be used to identify structural choice probabilities. For example, prices in neighboring markets (a la [Hausman \(1996\)](#)) can be used as instruments to generate residuals in a first stage, and then these residuals used as control functions in a second step. Such instruments are more likely to be valid if spatial correlation in prices result primarily from spatial correlation in supply costs and not from spatial correlation of consumer preferences. A complete resolution of these issues is left to future research.

5. EMPIRICAL ILLUSTRATION

In this section, we provide a brief empirical illustration of our methods using a dataset on choice of fishing mode in Southern California, previously analyzed in [Herriges and Kling \(1999\)](#). Our primary purpose is to see what sorts of numerical results one gets by applying our methods to a real dataset. As an aside, we also compare our results to a fully parametric, logsum-type approach, popular in applications.

The key features of our microdata, which are publicly available, are discussed in [Kling and Thomson \(1996\)](#). The dataset consists of a cross section of 1182 individuals. For each individual, there are 4 potential modes of fishing, namely from the beach, the pier, a charter boat, and a private boat, labeled 1 through 4, respectively. Individual income (\$ per month) is recorded, together with the chosen mode, the price of each alternative faced by the individual, and the exogenous catch rate, that is, the average number of fish per hour of fishing at the site (obtained from an external source) corresponding to each option. The price reflects primarily transportation costs to the fishing site; see [Kling and Thomson \(1996\)](#) for a more detailed description of the variables. Table 1 provides summary statistics.

We use these data to calculate distributions of individual compensating variation corresponding to two hypothetical policy changes. The first is a simultaneous change in

TABLE 1. Summary statistics.

Variable	Mean	Std. Dev.	Min	Max
beach	0.11	0.32	0.00	1.00
pier	0.15	0.36	0.00	1.00
private	0.35	0.48	0.00	1.00
charter	0.38	0.49	0.00	1.00
catch-rate beach	0.24	0.19	0.07	0.53
catch-rate pier	0.16	0.16	0.00	0.45
catch-rate private	0.17	0.21	0.00	0.74
catch-rate charter	0.63	0.71	0.00	2.31
income	5250	2958	500	10,000
price-beach	103.42	103.64	1.29	843.19
price-pier	103.42	103.64	1.29	843.19
price-private	55.26	62.71	2.29	666.11
price-charter	84.38	63.54	27.29	691.11

Note: Beach is a dummy for fishing from the beach, price-beach is its price, catch-rate beach is the number of fish caught per hour, and income is individual monthly income in dollars. Similarly for the other alternatives.

the prices of fishing from the beach, the pier, and from a private boat. The second is the private welfare effect of availability of the private boat option and, separately, that of the charter boat option at hypothetical prices. That is, we consider a thought experiment where we assume that the private option was initially available at a specific price, and was subsequently eliminated, for example, due to environmental concerns. The third exercise involves welfare effects of quality-change, namely a lowering of the catch rate of the beach option. For each exercise, the mean CV is reported across a range of incomes, while keeping all other variables (i.e., those not being targeted by the change) fixed at their mean value. We maintain the assumption of independence between budget sets and preferences, for this illustration.

In order to avoid curse of dimensionality problems during estimation, we model the choice probability of each alternative j in a partially linear fashion as

$$q_j(\mathbf{p}, y, x) = h_j(y - p_j) + x'\beta_j + (y\mathbf{1}_{-j} - \mathbf{p}_{-j})'\gamma_j, \quad (29)$$

where $h_j(\cdot, \cdot)$ denotes a cubic in income and own price, $y\mathbf{1}_{-j} - \mathbf{p}_{-j}$ denotes the vector of income minus price of other alternatives and x denotes catch rates. These specifications are to be interpreted as series/sieve approximations to the true nonparametric choice probabilities.

Multiple price change: For the first exercise, we consider an initial price vector $\mathbf{p}_0 = (140, 140, 60, 12)$ and final price vector $\mathbf{p}_1 = (40, 40, 60, 72)$. Note that the price of alternatives 1 and 2 have fallen by 100, while that of alternative 4 has risen by 60, with the price of 3 fixed at 60. These changes in prices correspond to a fall from approximately the 75th to the 25th percentile for alternatives 1 and 2, and a rise from the 25th to the 75th percentile for alternative 4. The marginal distribution of individual welfare losses at income y with catch rates set at their mean values x is given by (cf. equation

(10)

$$\Pr(CV \leq a) = \begin{cases} 0, & \text{if } a < -100, \\ (q_1 + q_2)(40 - a, 40 - a, 60, 12, y, x), & \text{if } -100 \leq a < 0, \\ 1 - q_4(40, 40, 60, 12 + a, y + a, x), & \text{if } 0 \leq a < 60, \\ 1, & \text{if } a \geq 60. \end{cases} \quad (30)$$

We will contrast the mean welfare estimates resulting from this exercise with the popular [Small and Rosen \(1981\)](#) approach using the so-called “log-sum” formula (cf., [Train \(2009, p. 56\)](#)). For the multiple price change case, for example, adding income and catch rate as controls, the log-sum values are given by

$$E(CV) = \frac{1}{|\alpha|} \left[\ln \left[\frac{\exp(b_{01} + \alpha p_{10} + \beta'_1 z) + \exp(b_{02} + \alpha p_2 + \beta'_2 z) + \exp(b_{03} + \alpha p_2 + \beta'_3 z) + \exp(b_{04} + \alpha p_{40} + \beta'_4 z)}{\exp(b_{01} + \alpha p_{11} + \beta'_1 z) + \exp(b_{02} + \alpha p_2 + \beta'_2 z) + \exp(b_{03} + \alpha p_2 + \beta'_3 z) + \exp(b_{04} + \alpha p_{41} + \beta'_4 z)} \right] \right], \quad (31)$$

where the *bs* represent the alternative specific intercepts, $|\alpha|$ is the constant marginal utility of income and $z = (x, y)$, the mean values of catch rates and income. The *b*’s and α are estimated via the constrained “mlogit” command in Stata after imposing constant marginal utility of income by requiring price coefficients to be identical across alternatives.

The substantive difference in functional forms between (29), (30), and (31) is in how price and income appear in them. The logsum formula is based on the constant marginal utility of income, and thus requires price coefficients to be identical across alternatives. In addition, income is arbitrarily incorporated as a covariate, as is usually done in applications. Our expressions, instead, include income *y* as appearing in the form of $y - p_j$, and are thus consistent with a coherent, theoretically justified specification.

In [Figure 1](#), we plot the average CV by income, evaluated at mean values of controls, based on the logsum formula of [Small and Rosen](#), adding income as an additional control,⁶ and then using our methods (30) with specification (29) for the choice probabilities. Evidently, the logsum estimates are small, and *decrease* with income, whereas our CV estimates imply larger (about 2 times) welfare losses which *increase* with income. The numerical results do vary a bit if a quadratic or a linear spline is used as opposed to a cubic, but these two substantive features continue to hold. It is also interesting to note that due to a higher proportion (3 times) of people choosing alternative 4 whose

⁶The logsum formula comes from an assumption of quasilinear utilities where there is no income effect. Taken literally, this would imply a flat welfare estimate against income. But in most empirical applications, income is added as a demographic variable (as opposed to coming from the budget constraint), which is what we use here.

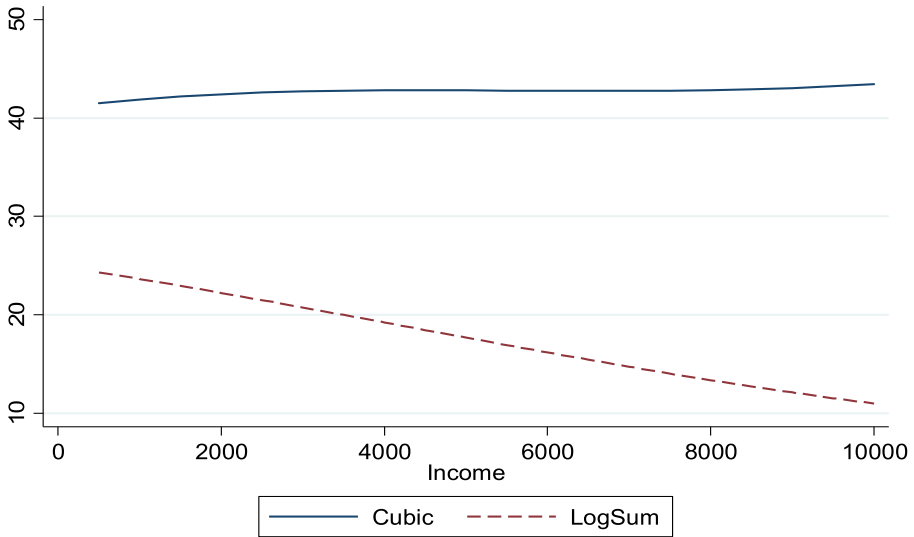


FIGURE 1. Average CV for multiple price change. Average welfare loss in \$ by monthly income from a price fall from the 75th to the 25th percentile for beach and pier fishing, and a rise from the 25th to the 75th percentile for private-boat fishing, evaluated at mean values of income and catch rate of all alternatives and mean price of charter-boat fishing. Solid line corresponds to our nonparametric estimates, dashed line to the logsum estimates.

price has risen, the overall welfare is reduced, that is, average CV is positive, although the prices of alternatives 1 and 2 have fallen by a large amount.

Eliminate alternative: Suppose the local government is considering means to reduce environmental damage by banning either the private boat or the charter boat option. For deciding which option to ban, it is useful to know the loss in private welfare from each. Toward that end, we first consider eliminating the private option from a situation where the price vector was $(75, 75, 60, 33)$, which are essentially the median values in the data, and the catch rates were at their mean value, denoted x , and assuming no changes in the prices of alternatives 1–3. In this case, average CV is given by

$$E(\text{CV}) = \int_0^{\infty} q_4(75, 75, 60, 33 + a, y + a, x). \quad (32)$$

These will be compared with the corresponding numbers resulting from eliminating the charter boat option.

In Figures 2 and 3, we plot the average welfare losses from elimination of respectively the private and the charter option. We compare our estimates using (29) and (32) with those obtained using a logsum approach. To calculate the integral up to infinity in (32), we simply use the highest observed price as the upper limit of the integral. As such, our estimates may be viewed as a lower bound on the average CV.

Evidently, our estimates are much flatter against income relative to the logsum estimates; they are much lower (less than 50%) for elimination of the private option, and

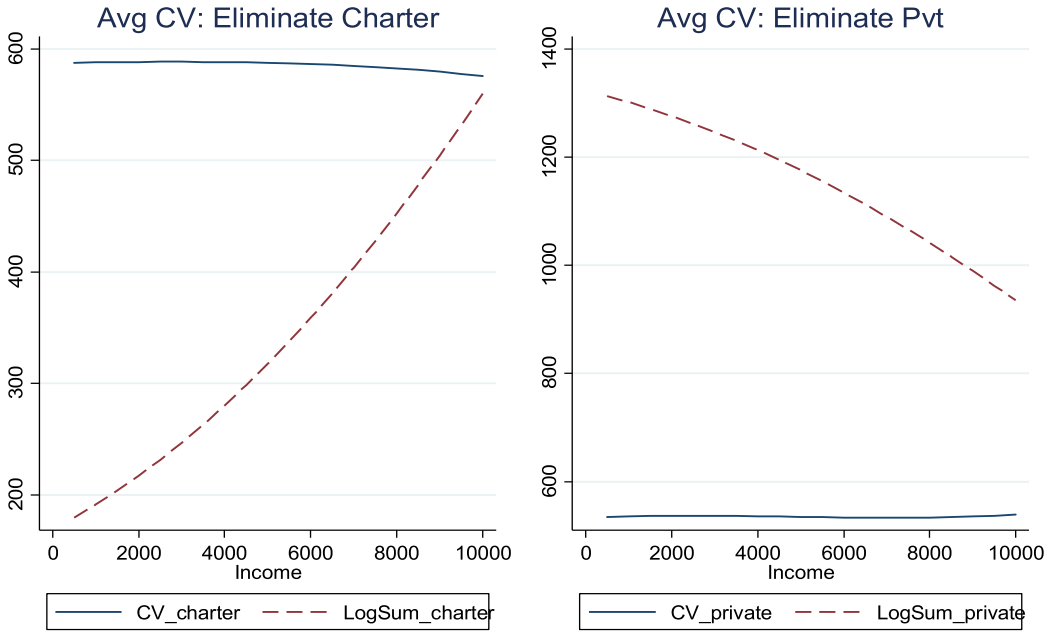


FIGURE 2. Elimination of option: LogSum vs cubic. Average welfare loss \$ by monthly income from elimination of charter-boat and of private boat fishing. Solid line corresponds to our non-parametric estimates, dashed line to the logsum estimates. All calculations are performed at mean values of other characteristics.

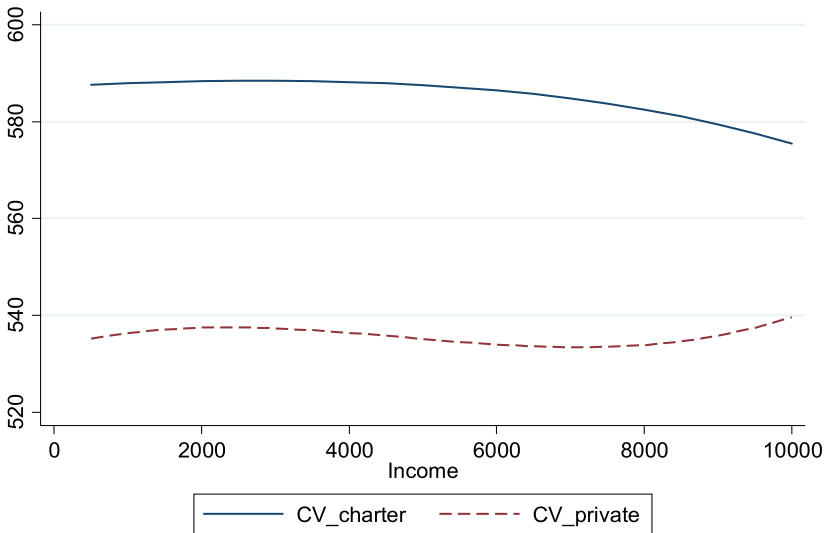


FIGURE 3. Eliminate charter vs pvt. Average welfare loss \$ by monthly income from elimination of charter-boat and of private boat fishing. Dashed line corresponds to charter-boat and solid line to private boat. All calculations are performed at mean values of other characteristics.

much higher (about twice on average for lower incomes) for the charter option. It is also important to note that the slope of average welfare w.r.t. income has different signs, for example, the logsum estimates imply that eliminating the charter boat option will hurt the rich more than the poor, whereas our CV estimates imply the exact opposite. Figure 3 shows, in a slightly blown up scale, that elimination of these options entail fairly significant losses in private welfare, amounting to about 10% of the average monthly income.

Putting Figures 2 and 3 together, it is also evident that the logsum estimates would imply a much larger *difference* in welfare losses between eliminating the private versus the charter option, compared with our CV estimates.

Quality change: Finally, the third exercise involves decreasing the catch rate of alternative 1, namely beach fishing from the 4th quartile 0.53 to the 1st quartile 0.06, keeping all other variables fixed at their mean value, denoted z .

Assuming that catch rate is a normal characteristic for all individuals, the bounds on the CDF of the CV are given by Proposition 2 in the Appendix, namely

$$\Pr(\text{CV} \leq a) \begin{cases} \in [L, U], & \text{if } a < 0, \\ = 1, & \text{if } a \geq 0, \end{cases}$$

where

$$L = \max \left\{ \begin{array}{l} \max_{y' < y} \left\{ \begin{array}{l} q_1(75 - a + y' - y, y', z, 0.53) \\ + q_0(75 + y' - y, y', z, 0.06) \end{array} \right\} \\ - 1 - q_0(75 - a, z, y, 0.53), 0 \end{array} \right\},$$

$$U = \min \left\{ \begin{array}{l} q_1(75 - a, y, z, 0.53), \\ \min_{y' \geq y} \left\{ \begin{array}{l} q_1(75 - a + y' - y, y', z, 0.53) \\ + q_0(75 + y' - y, y', z, 0.06) \end{array} \right\} \end{array} \right\}, \quad (33)$$

$$q_0(p, z, y, x) \equiv 1 - q_1(p, y, z, x).$$

To compare these bounds with the logsum formulae, which are based on quasilinear utility assumption, we use two alternative specifications of the logsum, one that incorporates income as an additional covariate and one that does not. The logsum expressions are essentially (31) with the same prices but different catch rates in the numerator and denominator. We finally obtain the following graph, where we plot the average welfare loss (i.e., negative of the CV) by income.

The bounds (dashed lines) are relatively tight. The lower bound remains zero, but the upper bound hovers around \$7.50 for most of the income range. Relative to the logsum estimates with income included as a covariate (the nearly straight downward sloping line), our bounds are much flatter, and higher than the logsum estimates at high incomes. Indeed, the logsum estimate yields negative welfare losses, that is, welfare *gains* from reduced catch rates, which is an artifact of the functional form assumption, and

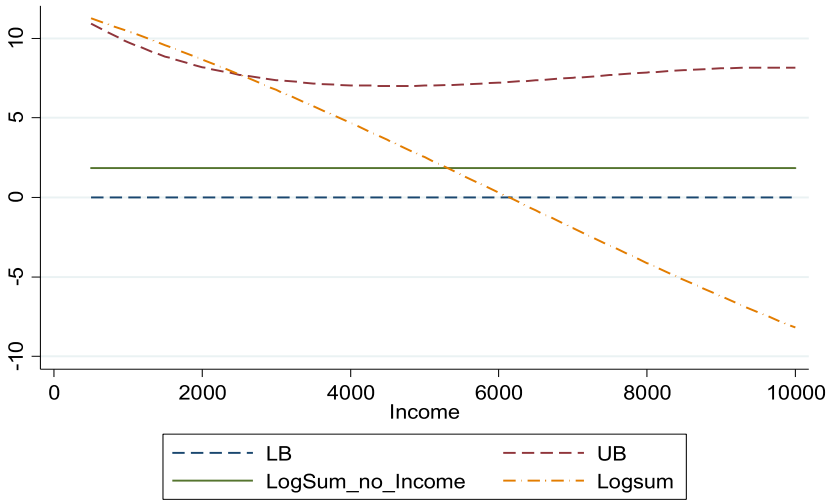


FIGURE 4. Avg. welfare loss from reduced catch rate. Average welfare loss \$ by monthly income from decreasing the catch rate of alternative 1, namely, beach fishing from the 4th quartile 0.53 to the 1st quartile 0.06, keeping all other variables fixed at their mean value. Lower and upper bounds from our calculations are shown via dotted lines and Logsum estimates with no income effect by the solid line, and with income as covariate by downward sloping, almost a straight line.

TABLE 2. Mean welfare loss & 95% C.I. in \$ at mean income.

	Cubic Approximation	Bootstrap C.I.	LogSum
Multiple Price Change	42.57	40.0, 43.1	19.21
Eliminate Pvt. Boat	536.37	292.31, 705.08	1212.4
Eliminate Charter	588.16	582.4, 591.8	280.29
Decrease Catch-rate of Beach	7.05	4.55, 11.91	4.68

Note: Mean CV and associated bootstrapped confidence interval for (i) simultaneous price change of beach and pier fishing from \$140 to \$40 and of private boat fishing from \$12 to \$72, (ii) elimination of private, and (iii) of charter boat, and (iv) a fall in catch rate of beach fishing from 0.53 to 0.06 per hour (upper bound of average welfare loss). All quantities are evaluated at mean income of \$4000 per month, and mean values of all other regressors.

is difficult to interpret. Of course, these differences in slopes are also important from a policy perspective, because the variation of welfare with income determines the progressivity of the economic change. As discussed above, the key difference in functional forms between our estimates and the logsum formulae lies in the treatment of income. One can see from Figure 4 that if one ignores income entirely, given the quasilinear basis of logsum estimates, then one gets the solid flat line at around \$1.84 which lies entirely within our bounds.

Our final set of results in Table 2 report the welfare changes discussed above, together with bootstrapped confidence intervals, evaluated at Income = \$4000 per month,

which is about the mean income. The confidence intervals are evaluated over 100 replications, and the top 1% and bottom 1% of values are discarded. The logsum estimates are reported in the last column for comparison. It is evident from this table that the cubic approximation based estimates of compensating variation are reasonably precise, and the corresponding confidence intervals do not contain the logsum estimates, except for the case of quality change. Finally, the sizes of the point estimates show that elimination of the private or the charter boat option entails very high welfare loss, amounting to nearly 13% of mean income; but the change in catch rate has a very small welfare impact, suggesting that it is the overall experience of fishing from one's chosen alternative, rather than the quantity of fish caught, that is primarily valued by individuals.

It is not obviously possible to conclude *empirically* whether our estimates are more accurate than the logsum ones. But since our estimates correspond to a theoretically consistent utility structure, do not impose constant marginal utility of income, and are not based on specific distributional assumptions, they should be viewed as more robust on theoretical grounds. In the above application, our estimates differ substantially from the logsum ones, which does therefore raise doubt about the strong assumptions underlying the logsum approach. Indeed, our theoretical results above show that specification of choice probabilities fully determines welfare estimates (or bounds thereof), without any reference to underlying utility structures. So any differences between our results and the logsum ones must necessarily arise from differences in functional forms for choice probabilities. One could also try using alternative and more flexible parametric models such as the mixed or nested logit, and compare them with our results. Since our primary goal here is to provide a "proof of concept" regarding empirical feasibility of our theoretical methods, we leave such comparisons to future research.

6. SUMMARY AND CONCLUSION

In this paper, we have developed expressions for welfare effects of economic changes in multinomial choice settings, allowing for completely general consumer heterogeneity and income effects. The paper considers four practical scenarios: (a) simultaneous change in prices of multiple alternatives, (b) the introduction or elimination of an alternative, possibly accompanied by price changes of other alternatives, (c) change in the quality of an alternative possibly accompanied by a price change, and (d) situations where choice alternatives are nonexclusive. The key results are: (1) Hicksian welfare changes are well-defined in all cases under the mild assumption that the numeraire is a normal good, (2) in cases (a) and (b) the marginal distributions of CV and EV can be expressed as simple closed-form functions of choice probabilities without requiring any assumption on the functional forms of utilities, preference heterogeneity or income effects and (3) this last conclusion fails for quality change or when alternatives are nonexclusive, but welfare distributions can still be bounded in these cases. Our approach delivers welfare analysis without requiring identification of preference distributions. As such, it strictly dominates random-coefficient based methods which require a researcher to either make arbitrary assumptions on the distribution of random coeffi-

cients (e.g., that they are normally distributed), or to require full support of regressors for identification and to solve difficult ill-posed inverse problems in estimation.

At a practical level, our methods can be used in program evaluation studies to calculate “compensated” program effects, that is, the program’s value to the subjects themselves, measured in terms of its cash equivalent, and the associated deadweight loss, while allowing for arbitrary preference distribution in the population. These money-metric welfare measures can also be compared across interventions with different *outcomes*. For example, a tuition subsidy for school attendance, and an adoption subsidy for take-up of a health-product cannot be directly compared in terms of their average outcomes which are in different units; but their cash equivalents and the associated deadweight loss can be directly compared, since both are expressible in monetary units. This comparison can then inform the decision of which public subsidy program should be adopted.

APPENDIX

PATH-DEPENDENCE OF LINE INTEGRAL DEFINING MARSHALLIAN CONSUMER SURPLUS FOR DISCRETE CHOICE

Consider a setting with three mutually exclusive alternatives with initial prices $\mathbf{p}_0 \equiv (p_{10}, p_{20}, p_{30})$ and final prices $\mathbf{p}_1 \equiv (p_{11}, p_{21}, p_{31})$. Let y denote income and $q_j(\mathbf{p}, y)$ denote the choice probability of alternative j when the price vector is \mathbf{p} and income is y . Then the change in average consumer surplus arising from the price change from \mathbf{p}_0 to \mathbf{p}_1 can be defined using (4) via the line integral

$$CS(L) = - \int_L q_1(\mathbf{p}, y) dp_1 + q_2(\mathbf{p}, y) dp_2 + q_3(\mathbf{p}, y) dp_3,$$

where L denotes a path $L(t)$ from $t = 0$ to $t = 1$ such that $L(0) \equiv \mathbf{p}_0 \equiv (p_{10}, p_{20}, p_{30})$ and $L(1) \equiv \mathbf{p}_1 \equiv (p_{11}, p_{21}, p_{31})$. Consider two different such paths

$$L_1(t) = (p_{10} + t(p_{11} - p_{10}), p_{20} + t(p_{21} - p_{20}), p_{30} + t(p_{31} - p_{30})),$$

$$L_2(t) = (p_{10} + t^2(p_{11} - p_{10}), p_{20} + t(p_{21} - p_{20}), p_{30} + t(p_{31} - p_{30})).$$

Then

$$CS(L_1) = - \int_0^1 \left[\begin{array}{l} (p_{11} - p_{10}) \times q_1(\mathbf{p}_0 + t(\mathbf{p}_1 - \mathbf{p}_0), y) \\ + (p_{21} - p_{20}) \times q_2(\mathbf{p}_0 + t(\mathbf{p}_1 - \mathbf{p}_0), y) \\ + (p_{31} - p_{30}) \times q_3(\mathbf{p}_0 + t(\mathbf{p}_1 - \mathbf{p}_0), y) \end{array} \right] dt.$$

But

$CS(L_2)$

$$= - \int_0^1 \left[\begin{aligned} & 2t(p_{11} - p_{10}) \times q_1 \left(\mathbf{p}_0 + t(\mathbf{p}_1 - \mathbf{p}_0) + (t^2 - t) \begin{pmatrix} p_{11} - p_{10} \\ 0 \\ 0 \end{pmatrix}, y \right) \\ & + (p_{21} - p_{20}) \times q_2 \left(\mathbf{p}_0 + t(\mathbf{p}_1 - \mathbf{p}_0) + (t^2 - t) \begin{pmatrix} p_{11} - p_{10} \\ 0 \\ 0 \end{pmatrix}, y \right) \\ & + (p_{31} - p_{30}) \times q_3 \left(\mathbf{p}_0 + t(\mathbf{p}_1 - \mathbf{p}_0) + (t^2 - t) \begin{pmatrix} p_{11} - p_{10} \\ 0 \\ 0 \end{pmatrix}, y \right) \end{aligned} \right] dt,$$

which would in general differ from $CS(L_1)$. Thus the CS is not well-defined for simultaneous change in multiple prices. Note that if only p_1 changes, then

$$CS(L_1) = -(p_{11} - p_{10}) \times \int_0^1 [q_1(p_{10} + t(p_{11} - p_{10}), y)] dt,$$

$$\begin{aligned} CS(L_2) &= -2 \int_0^1 [(p_{11} - p_{10}) \times q_1(p_{10} + t^2(p_{11} - p_{10}), y)] t dt \\ &= -(p_{11} - p_{10}) \times \int_0^1 q_1(p_{10} + r(p_{11} - p_{10}), y) dr, \quad \text{substituting } t^2 = r \\ &= CS(L_1), \end{aligned}$$

and we get back path independence. Thus the loss of path-independence arises only for multiple simultaneous price changes.

Iterating single price change: We show that welfare analysis of multiple price changes cannot be done by iterating the single price change result of [Bhattacharya \(2015\)](#). Consider a choice among three alternatives ($J = 3$), and suppose that price of alternative 1 changes from p_{10} to p_{11} , and that of 2 from p_{20} to p_{21} , and price of 3 is unchanged at p_3 . Suppose we try to calculate the overall CV, starting with, say, price change of alternative 1, followed by 2 (the order in which we do this does not matter, by [Proposition 1](#)) and applying [Theorem 2](#) of [Bhattacharya \(2015\)](#) at each stage. Suppose the CVs corresponding to the two price changes are denoted by S_1 and S_2 , respectively. Then, by definition,

$$\begin{aligned} & \max\{U_1(y - p_{10}, \eta), U_2(y - p_{20}, \eta), U_3(y - p_3, \eta)\} \\ &= \max\{U_1(y + S_1 - p_{11}, \eta), U_2(y + S_1 - p_{20}, \eta), U_3(y + S_1 - p_3, \eta)\} \\ &= \max \left\{ \begin{array}{l} U_1(y + \underbrace{S_1 + S_2}_{=S, \text{ overall CV}} - p_{11}, \eta), U_2(y + \underbrace{S_1 + S_2}_S - p_{21}, \eta), \\ U_3(y + \underbrace{S_1 + S_2}_S - p_3, \eta) \end{array} \right\}. \end{aligned}$$

Then using Theorem 2 of [Bhattacharya \(2015\)](#), we can get the marginal distribution of S_1 but we cannot get the marginal distribution of $S_1 + S_2$ because the price of both alternatives 1 and 2 have changed between lines 1 and 3 of the previous display, and so Theorem 2 of [Bhattacharya \(2015\)](#) does not apply. Second, because we cannot calculate the *value* of the CV S_1 for an individual (we can only calculate its *distribution* across all individuals), we cannot apply Theorem 2 of [Bhattacharya \(2015\)](#) to calculate the marginal distribution of S_2 , since the income at which we could potentially apply the theorem depends on S_1 which is unknown, unlike y , which is a fixed known constant.

PROOF OF THEOREM 1. First, consider EV.

Note that the EV is defined by

$$\begin{aligned} & \max\{U_1(y - p_{11}, \eta), U_2(y - p_{21}, \eta), \dots, U_J(y - p_{J1}, \eta)\} \\ & = \max\{U_1(y - S - p_{10}, \eta), U_2(y - p_{20} - S, \eta), \dots, U_J(y - p_{J0} - S, \eta)\}. \end{aligned} \tag{34}$$

We have already shown in Proposition 1 that under Assumption 1, the solution to (34) exists and is unique. Furthermore, because the RHS is strictly decreasing in S , we have that

$$EV \leq a \Leftrightarrow \left\{ \begin{array}{l} \max\{U_1(y - p_{11}, \eta), U_2(y - p_{21}, \eta), \dots, U_J(y - p_{J1}, \eta)\} \\ \geq \max\left\{ \begin{array}{l} U_1(y - a - p_{10}, \eta), U_2(y - p_{20} - a, \eta), \\ \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right\}. \tag{35}$$

Now, if a satisfies $p_{l1} - p_{l0} \leq a < p_{l+1,1} - p_{l+1,0}$, then

$\Pr(EV \leq a)$

$$\begin{aligned} & = \Pr \left[\begin{array}{l} U_1(y - p_{11}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_2(y - p_{21}, \eta), \dots, U_l(y - p_{l1}, \eta), \dots, U_J(y - p_{J1}, \eta), \\ U_1(y - a - p_{10}, \eta), U_2(y - p_{20} - a, \eta), \dots, \\ U_l(y - a - p_{l0}, \eta), \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right] \\ & + \Pr \left[\begin{array}{l} U_2(y - p_{21}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y - p_{11}, \eta), \dots, U_l(y - p_{l1}, \eta), \dots, U_J(y - p_{J1}, \eta) \\ U_1(y - a - p_{10}, \eta), U_2(y - p_{20} - a, \eta), \dots, \\ U_l(y - a - p_{l0}, \eta), \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right] \\ & + \dots \end{aligned}$$

$$\begin{aligned}
& + \Pr \left[\begin{array}{l} U_l(y - p_{l1}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y - p_{11}, \eta), U_2(y - p_{21}, \eta), \dots, U_{l-1}(y - p_{l-1,1}, \eta), \\ U_{l+1}(y - p_{l+1,1}, \eta), \dots, U_J(y - p_{J1}, \eta) \\ U_1(y - a - p_{10}, \eta), U_2(y - p_{20} - a, \eta), \dots, \\ U_l(y - a - p_{l0}, \eta), \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right] \\
\stackrel{(1)}{=} & \Pr \left[\begin{array}{l} U_1(y - p_{11}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_2(y - p_{21}, \eta), \dots, U_l(y - p_{l1}, \eta), \\ U_{l+1}(y - a - p_{l+1,0}), \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right] \\
& + \Pr \left[\begin{array}{l} U_2(y - p_{21}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y - p_{11}, \eta), \dots, U_l(y - p_{l1}, \eta), \\ U_{l+1}(y - a - p_{l+1,0}), \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right] \\
& + \dots \\
& + \Pr \left[\begin{array}{l} U_l(y - p_{l1}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y - p_{11}, \eta), \dots, U_{l-1}(y - p_{l-1,1}, \eta), \\ U_{l+1}(y - a - p_{l+1,0}, \eta), \dots, U_J(y - p_{J0} - a, \eta) \end{array} \right\} \end{array} \right],
\end{aligned}$$

where the equality marked (1) uses the fact that $y - p_{j1} \geq y - a - p_{j,0}$, for all $j = 1, \dots, l$, since $p_{l1} - p_{l0} \leq a$. The above probability equals

$$\begin{aligned}
& q_1(p_{11}, p_{21}, \dots, p_{l1}, p_{l+1,0} + a, \dots, p_{J0} + a, y) \\
& + q_2(p_{11}, p_{21}, \dots, p_{l1}, p_{l+1,0} + a, \dots, p_{J0} + a, y) \\
& + \dots + q_l(p_{11}, p_{21}, \dots, p_{l1}, p_{l+1,0} + a, \dots, p_{J0} + a, y),
\end{aligned}$$

if $a \geq 0$, and equals

$$\begin{aligned}
& = q_1(p_{11} - a, p_{21} - a, \dots, p_{l1} - a, p_{l+1,0}, \dots, p_{J0}, y - a) \\
& + q_2(p_{11} - a, p_{21} - a, \dots, p_{l1} - a, p_{l+1,0}, \dots, p_{J0}, y - a) \\
& + \dots + q_l(p_{11} - a, p_{21} - a, \dots, p_{l1} - a, p_{l+1,0}, \dots, p_{J0}, y - a),
\end{aligned}$$

if $a < 0$. This is precisely expression (8). The different expressions for $a \geq 0$ and $a < 0$ are needed to accommodate an outside good whose price remains 0 before and after. For example, if the l th alternative is the outside good, and thus $p_{l1} = p_{l0} = 0$ and $a \geq 0$, then one can evaluate the necessary probability by observing behavior at income y , for

example,

$$\begin{aligned} & \Pr \left[\begin{array}{l} U_1(y - p_{11}, \eta) \\ \geq \max \left\{ U_2(y - p_{21}, \eta), \dots, U_l(y - p_{l1}, \eta), \right. \\ \left. U_{l+1}(y - a - p_{l+1,0}), \dots, U_J(y - p_{J0} - a, \eta) \right\} \end{array} \right] \\ &= \Pr \left[\begin{array}{l} U_1(y - p_{11}, \eta) \\ \geq \max \left\{ U_2(y - p_{21}, \eta), \dots, U_l(y, \eta), \right. \\ \left. U_{l+1}(y - a - p_{l+1,0}), \dots, U_J(y - p_{J0} - a, \eta) \right\} \end{array} \right] \\ &= q_1(p_{11}, p_{21}, \dots, p_{l1}, p_{l+1,0} + a, \dots, p_{J0} + a, y). \end{aligned}$$

But if the outside alternative is $l + 1$, then $p_{l+1,1} = p_{l+1,0} = 0$ and $a < 0$, and so

$$\begin{aligned} & \Pr \left[\begin{array}{l} U_1(y - p_{11}, \eta) \\ \geq \max \left\{ U_2(y - p_{21}, \eta), \dots, U_l(y - p_{l1}, \eta), \right. \\ \left. U_{l+1}(y - a - p_{l+1,0}), \dots, U_J(y - p_{J0} - a, \eta) \right\} \end{array} \right] \\ &= \Pr \left[\begin{array}{l} U_1(y - p_{11}, \eta) \\ \geq \max \left\{ U_2(y - p_{21}, \eta), \dots, U_l(y - p_{l1}, \eta), \right. \\ \left. U_{l+1}(y - a), \dots, U_J(y - p_{J0} - a, \eta) \right\} \end{array} \right]. \end{aligned}$$

But we cannot evaluate this probability by observing behavior at income y , as no one in the data faces a price of $a < 0$ for the $l + 1$ th alternative, that is, the outside good. So the only way to learn this probability is to observe individuals at income $y - a$ facing a price of zero for the outside good.

Next, consider CV, defined as the solution S to the equation

$$\begin{aligned} & \max\{U_1(y + S - p_{11}, \eta), U_2(y + S - p_{21}, \eta), \dots, U_J(y + S - p_{J1}, \eta)\} \\ &= \max\{U_1(y - p_{10}, \eta), U_2(y - p_{20}, \eta), \dots, U_J(y - p_{J0}, \eta)\}. \end{aligned} \quad (36)$$

By exactly similar logic as above, $CV \leq a$ is equivalent to

$$\begin{aligned} & \max\{U_1(y - p_{10}, \eta), U_2(y - p_{20}, \eta), \dots, U_J(y - p_{J0}, \eta)\} \\ & \leq \max\{U_1(y + a - p_{11}, \eta), U_2(y + a - p_{21}, \eta), \dots, U_J(y + a - p_{J1}, \eta)\}. \end{aligned} \quad (37)$$

Accordingly, we have that

$\Pr(CV \leq a)$

$$\begin{aligned}
 &= \sum_{k=1}^j \Pr \left[\begin{array}{l} U_k(y + a - p_{k1}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y - p_{10}, \eta), \dots, U_{k-1}(y - p_{k-1,0}, \eta), U_k(y - p_{k0}, \eta), \\ U_{k+1}(y - p_{k+1,0}, \eta), \dots, U_J(y - p_{J0}, \eta) \\ U_1(y + a - p_{11}, \eta), \dots, U_{j-1}(y + a - p_{j-1,1}, \eta), \\ U_j(y + a - p_{j1}, \eta), U_{j+1}(y + a - p_{j+1,1}, \eta), \\ \dots, U_J(y + a - p_{J1}, \eta) \end{array} \right\} \end{array} \right] \\
 &= \sum_{k=1}^j \Pr \left[\begin{array}{l} U_k(y + a - p_{k1}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y - p_{10}, \eta), \dots, U_{k-1}(y - p_{k-1,0}, \eta), \\ U_{k+1}(y - p_{k+1,0}, \eta), \dots, U_J(y - p_{J0}, \eta) \\ U_1(y + a - p_{11}, \eta), \dots, U_{j-1}(y + a - p_{j-1,1}, \eta), \\ U_j(y + a - p_{j1}, \eta), U_{j+1}(y + a - p_{j+1,1}, \eta), \\ \dots, U_J(y + a - p_{J1}, \eta) \end{array} \right\} \end{array} \right] \\
 &= \sum_{k=1}^j \Pr \left[\begin{array}{l} U_k(y + a - p_{k1}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y + a - p_{11}, \eta), \dots, U_j(y + a - p_{j1}, \eta), \\ U_{j+1}(y - p_{j+1,0}, \eta), \dots, U_J(y - p_{J0}, \eta) \end{array} \right\} \end{array} \right].
 \end{aligned}$$

The second equality follows from the fact that $k \leq j$ and $p_{j1} - p_{j0} \leq a$, and so we have by Assumption 1 that $U_k(y + a - p_{k1}, \eta) \geq U_k(y - p_{k0}, \eta)$. The third equality follows from $p_{j1} - p_{j0} \leq a < p_{j+1,1} - p_{j+1,0}$, so by Assumption 1, $U_k(y - p_{k,0}, \eta) \leq U_k(y + a - p_{k1}, \eta)$ for all $k \leq j$ and $U_k(y - p_{k,0}, \eta) > U_k(y + a - p_{k1}, \eta)$ for all $k > j$. Finally,

$$\begin{aligned}
 &\sum_{k=1}^j \Pr \left[\begin{array}{l} U_k(y + a - p_{k1}, \eta) \\ \geq \max \left\{ \begin{array}{l} U_1(y + a - p_{11}, \eta), \dots, U_j(y + a - p_{j1}, \eta), \\ U_{j+1}(y - p_{j+1,0}, \eta), \dots, U_J(y - p_{J0}, \eta) \end{array} \right\} \end{array} \right] \\
 &\equiv \begin{cases} \sum_{k=1}^j q_k(p_{11}, \dots, p_{j1}, p_{j+1,0} + a, \dots, p_{J0} + a, y + a), & \text{if } a \geq 0, \\ \sum_{k=1}^j q_k(p_{11} - a, \dots, p_{j1} - a, p_{j+1,0}, \dots, p_{J0}, y), & \text{if } a < 0. \end{cases} \quad \square
 \end{aligned}$$

PROOF OF COROLLARY 1. The EV is the solution S to the equation:

$$\begin{aligned}
 &\max\{U_1(y - S - p_{10}, \eta), \dots, U_J(y - S - p_{J0}, \eta), U_{J+1}(y - S - p_{J+1}, \eta)\} \\
 &= \max\{U_1(y - p_{11}, \eta), \dots, U_J(y - p_{J1}, \eta)\}. \tag{38}
 \end{aligned}$$

We seek to find the marginal distribution of S . To do so, we first show that for each η , the $S(\eta)$ solving (38) is identical to the EV $T(\eta)$ for a different problem, and the distribution of $T(\eta)$ in that other problem follows as a corollary of our Theorem 1 above. The other problem is as follows: consider a multinomial choice problem with $J + 1$ alternatives, where the price vector changes from an initial value of $(p_{10}, \dots, p_{J0}, p_{J+1})$ to $(p_{11}, \dots, p_{J1}, p')$, where $p_{J1} - p_{J0} \geq p_{J-1,1} - p_{J-1,0} \geq \dots \geq p_{11} - p_{10}$. Then the resulting EV satisfies

$$\begin{aligned} &\max\{U_1(y - T - p_{10}, \eta), \dots, U_J(y - T - p_{J0}, \eta), U_{J+1}(y - T - p_{J+1}, \eta)\} \\ &= \max\{U_1(y - p_{11}, \eta), U_2(y - p_{21}, \eta), \dots, U_J(y - p_{J1}, \eta), U_{J+1}(y - p', \eta)\}. \end{aligned} \tag{39}$$

Now, suppose p' tends to infinity. Recall the assumption that for any bundle $(j, y - p_j)$, if $p_j \uparrow \infty$, then for each η , at least one other bundle $(k, y - p_k)$ will be strictly preferred to $(j, y - p_j)$. Therefore, the maximum on the RHS of (39) cannot be the $J + 1$ th term. Therefore, as p' tends to infinity, (39) reduces to (38), and we conclude that for each η , the EV $S(\eta)$ for elimination of the $J + 1$ th alternative is the same as the EV $T(\eta, p')$ when the new price p' of the $J + 1$ th alternative increases to infinity, that is, $\lim_{p' \uparrow \infty} T(\eta, p') = S(\eta)$ pointwise in η . To calculate the distribution of $T(\eta, p')$ for any fixed p' , we can apply Theorem 1 with price of the $J + 1$ th alternative changing to p' . Now, observe that the random variable $1\{T(\eta, p') \leq a\}$ is bounded uniformly in p' by 1 for each a , $\int 1 dF(\eta) = 1 < \infty$, and $T(\eta, p') \rightarrow T(\eta, \infty) \equiv S(\eta)$ pointwise in η as $p' \uparrow \infty$. Therefore, it follows by the dominated convergence theorem that for each a ,

$$\begin{aligned} \lim_{p' \uparrow \infty} \Pr(T(\eta, p') \leq a) &= \lim_{p' \uparrow \infty} \int 1\{T(\eta, p') \leq a\} dF(\eta) \\ &= \int 1\left\{\lim_{p' \uparrow \infty} T(\eta, p') \leq a\right\} dF(\eta) \\ &= \Pr(S(\eta) \leq a). \end{aligned}$$

(I am grateful to an anonymous referee for pointing out the need for the DCT argument.) The RHS of the previous display is the object we want, namely the CDF of the EV for problem (38), and the LHS is the limit of the CDF for EV in (39) as $p' \uparrow \infty$, which is obtained from Theorem 1 by taking the new price of the $J + 1$ th alternative to infinity. An exactly analogous conclusion holds for the CV, and this proves Corollary 1. \square

PROOF OF THEOREM 2. The proof works by constructing a counterexample. Suppose $\eta \equiv (\eta_1, \eta_0)$ is jointly independent of price, income, and quality (P, Y, X) and $\eta_1 \perp \eta_0$. Assume that the support of price distribution is $[p_L, p_H]$, support of income Y is $[y_L, y_H]$ and that x is a scalar with support $[x_L, x_H]$, with $p_H < y_L$. Let

$$U_1(Y - P, X, \eta) = Y - P + X\eta_1, U_0(Y, \eta) = (1 - \eta_0)Y,$$

where η_0 is distributed uniform $(0, 1)$ and the support of η_1 —denoted by $[L, U]$ —satisfies $0 < L < 1, U \leq \frac{p_H}{x_H}$. Assume η_1 is continuously distributed, and denote the CDF

of η_1 by $G(\cdot)$. An individual of type (y, η) and facing price p and quality x chooses alternative 1 if and only if $y - p + x\eta_1 > (1 - \eta_0)y$.

Now, consider the choice probability in this model. Since $0 \leq x_L \eta_1 \leq x_H \eta_1 \leq p_L$ w.p. 1, it follows that for any p, y, x in the support of the data, we must have that $p - y < x\eta_1 < p$, or

$$0 < \frac{p - x\eta_1}{y} < 1, \quad \text{w.p. 1.} \quad (40)$$

Therefore, the structural choice probability of alternative 1 at price p , quality x and income y is given by

$$\begin{aligned} q_1(p, y, x) &= \Pr\{y - p + x\eta_1 > (1 - \eta_0)y\} \\ &= \Pr\{\eta_0 y + x\eta_1 > p\} \\ &= \Pr\left\{\eta_0 > \frac{p - x\eta_1}{y}\right\} \quad \text{since } y > 0 \\ &= \int_L^U \left(1 - \frac{p - x\eta_1}{y}\right) dG(\eta_1), \\ &\quad \text{by } \eta_1 \perp \eta_0, \text{ inequality (40) and } \eta_0 \sim U(0, 1) \\ &= \left(1 - \frac{p}{y}\right) + \frac{x}{y} E(\eta_1). \end{aligned} \quad (41)$$

Now, consider a change from (p_0, x_0) to (p_1, x_1) satisfying

$$x_1 > x_0, \quad p_1 > p_0, \quad \frac{p_1 - p_0}{x_1 - x_0} \in (L, U), \quad \text{and} \quad \frac{p_1}{p_0} < \frac{x_1}{x_0}. \quad (42)$$

We wish to calculate the distribution of CV resulting from this change. The CV is defined by the solution S to the equation

$$\max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} = \max\{U_0(y + S, \eta), U_1(y + S - p_1, x_1, \eta)\}.$$

Note that

$$\begin{aligned} \Pr(S = 0) &= \Pr[\max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} = \max\{U_0(y, \eta), U_1(y - p_1, x_1, \eta)\}] \\ &= \Pr[U_0(y, \eta) \geq \max\{U_1(y - p_0, x_0, \eta), U_1(y - p_1, x_1, \eta)\}], \end{aligned}$$

since $U_1(y - p_1, x_1, \eta) - U_1(y - p_0, x_0, \eta)$ equals $(x_1 - x_0)\eta_1 + p_0 - p_1$, which is continuously distributed as η_1 is continuously distributed, so that

$$\Pr[U_1(y - p_1, x_1, \eta) = U_1(y - p_1, x_1, \eta)] = 0.$$

Therefore,

$$\begin{aligned}
& \Pr(S = 0) \\
&= \Pr[U_0(y, \eta) \geq \max\{U_1(y - p_0, x_0, \eta), U_1(y - p_1, x_1, \eta)\}] \\
&= \Pr[U_0(y, \eta) \geq U_1(y - p_0, x_0, \eta) \geq U_1(y - p_1, x_1, \eta)] \\
&\quad + \Pr[U_0(y, \eta) \geq U_1(y - p_1, x_1, \eta) \geq U_1(y - p_0, x_0, \eta)] \\
&= \Pr[(1 - \eta_0)y \geq y - p_0 + x_0\eta_1 \geq y - p_1 + x_1\eta_1] \\
&\quad + \Pr[(1 - \eta_0)y \geq y - p_1 + x_1\eta_1 \geq y - p_0 + x_0\eta_1] \\
&= \Pr\left[\eta_0 \leq \frac{p_0 - x_0\eta_1}{y}, \eta_1 \leq \frac{p_1 - p_0}{x_1 - x_0}\right] \\
&\quad + \Pr\left[\eta_0 \leq \frac{p_1 - x_1\eta_1}{y}, \eta_1 > \frac{p_1 - p_0}{x_1 - x_0}\right] \tag{43} \\
&= \int_L^{\frac{p_1 - p_0}{x_1 - x_0}} \left(\frac{p_0 - x_0\eta_1}{y}\right) dG(\eta_1) + \int_{\frac{p_1 - p_0}{x_1 - x_0}}^U \left(\frac{p_1 - x_1\eta_1}{y}\right) dG(\eta_1) \\
&= \int_L^U \left(\frac{p_0 - x_0\eta_1}{y}\right) dG(\eta_1) + \int_{\frac{p_1 - p_0}{x_1 - x_0}}^U \left(\frac{p_1 - x_1\eta_1}{y} - \frac{p_0 - x_0\eta_1}{y}\right) dG(\eta_1) \\
&= \frac{p_0}{y} - \frac{x_0}{y} E(\eta_1) + \int_{\frac{p_1 - p_0}{x_1 - x_0}}^U \left(\frac{p_1 - p_0 + (x_0 - x_1)\eta_1}{y}\right) dG(\eta_1) \\
&= 1 - q_1(p_0, y, x_0) + \int_{\frac{p_1 - p_0}{x_1 - x_0}}^U \left(\frac{p_1 - p_0 + (x_0 - x_1)\eta_1}{y}\right) dG(\eta_1).
\end{aligned}$$

Clearly, this expression is different for different choices of $G(\cdot)$ with the same expectation. In particular, if $\frac{p_1 - p_0}{x_1 - x_0}$ is smaller than $E(\eta_1)$, then for a degenerate distribution G_1 with point mass 1 at its expected value will yield a value of $1 - q_1(p_0, y, x_0)$ for (43), whereas a nondegenerate G_2 with the same expectation will yield a strictly smaller value than $1 - q_1(p_1, y, x_1)$, since $p_1 - p_0 + (x_0 - x_1)\eta_1 < 0$ if $\eta_1 > \frac{p_1 - p_0}{x_1 - x_0}$. However, since G_1 and G_2 have identical expectations, they will yield identical choice probabilities for all p, y, x , by (41). Thus we cannot identify $\Pr(S = 0)$ from the choice probabilities. \square

PROOF OF PROPOSITION 2. Assume $U_1(y - p, x, \eta) \equiv V_1(h_1(y - p, x), \eta)$. Define \bar{a} as the unique solution to $h_1(y + \bar{a} - p_1, x_1) = h_1(y - p_0, x_0)$. The existence and uniqueness of \bar{a} follow from the properties of $h_1(\cdot, \cdot)$ and condition S. Then \bar{a} can be obtained as the solution to $q_1(p_1 - \bar{a}, y, x_1) = q_1(p_0, y, x_0)$, since

$$\begin{aligned}
& \Pr[U_0(y, \eta) \leq V_1(h_1(y + \bar{a} - p_1, x_1), \eta)] \\
&= \Pr[U_0(y, \eta) \leq V_1(h_1(y - p_0, x_0), \eta)].
\end{aligned}$$

Now,

$$\begin{aligned}
 & \Pr(S \leq a) \\
 &= \Pr \left[\begin{array}{l} \max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} \\ \leq \max\{U_0(y + a, \eta), U_1(y + a - p_1, x_1, \eta)\} \end{array} \right] \\
 &= \Pr \left[\begin{array}{l} \max\{U_0(y, \eta), V_1(h_1(y - p_0, x_0), \eta)\} \\ \leq \max\{U_0(y + a, \eta), V_1(h_1(y + a - p_1, x_1), \eta)\} \end{array} \right] \\
 &= \begin{cases} \Pr[\max\{U_0(y, \eta), V_1(h_1(y - p_0, x_0), \eta)\} \leq U_0(y + a, \eta)], & \text{if } a < \bar{a}, \\ \Pr[U_0(y, \eta) \leq \max\{U_0(y + a, \eta), V_1(h_1(y + a - p_1, x_1), \eta)\}], & \text{if } a \geq \bar{a}. \end{cases}
 \end{aligned}$$

The first of these probabilities equals

$$\begin{aligned}
 & \begin{cases} 0, & \text{if } a < 0, \\ \Pr[V_1(h_1(y - p_0, x_0), \eta) \leq U_0(y + a, \eta)], & \text{if } a \geq 0 \end{cases} \\
 &= \begin{cases} 0, & \text{if } a < 0, \\ q_0(p_0 + a, y + a, x_0), & \text{if } a \geq 0, \end{cases}
 \end{aligned}$$

while the second equals

$$\begin{aligned}
 & \begin{cases} \Pr[U_0(y, \eta) \leq V_1(h_1(y + a - p_1, x_1), \eta)], & \text{if } a < 0, \\ 1, & \text{if } a \geq 0. \end{cases} \\
 &= \begin{cases} q_1(p_1 - a, y, x_1), & \text{if } a < 0, \\ 1, & \text{if } a \geq 0. \end{cases}
 \end{aligned}$$

Therefore,

$$\Pr(S \leq a) = \begin{cases} 0, & \text{if } a < \min\{\bar{a}, 0\}, \\ q_1(p_0 + a, y + a, x_0), & \text{if } 0 \leq a < \bar{a}, \\ q_1(p_1 - a, y, x_1), & \text{if } \bar{a} \leq a < 0, \\ 1, & \text{if } a \geq \max\{0, \bar{a}\}. \end{cases} \quad \square$$

BOUNDS FOR QUALITY CHANGE

PROPOSITION 3. *Suppose quality is a scalar. If higher values of x represent higher quality, that is, $U_1(a, \cdot, \eta)$ is strictly increasing for each a and η , and Assumption 2 holds, then the marginal distribution function of CV at income y resulting from a change of (p_0, x_0) to*

(p_1, x_1) satisfies the following bounds:

$$\Pr(S \leq a) \left\{ \begin{array}{l} \in \left[\begin{array}{l} \max \left\{ \max_{y' \leq y} \left\{ \begin{array}{l} q_1(p_1 - a + y' - y, y', x_1) \\ + q_0(p_0 + y' - y, y', x_0) \end{array} \right\} - 1 - q_0(p_1 - a, y, x_1), 0 \right\}, \\ \min \left\{ \begin{array}{l} q_1(p_1 - a, y, x_1), \\ \min_{y' \geq y} \left\{ \begin{array}{l} q_0(p_0 + y' - y, y', x_0) \\ + q_1(p_1 - a + y' - y, y', x_1) \end{array} \right\} \end{array} \right\} \end{array} \right], \\ \text{if } a < 0, \\ \in \left[\begin{array}{l} \max \left\{ \begin{array}{l} q_0(p_0 + a, y + a, x_0), \\ \max_{y' \geq y+a} \left\{ \begin{array}{l} q_1(p_1 - a + y' - y, y', x_1) \\ + q_0(p_0 + y' - y, y', x_0) \end{array} \right\} - 1 \end{array} \right\}, \\ \min \left\{ \left(\begin{array}{l} q_0(p_0 + a, y + a, x_0) \\ + \min_{y' \leq y+a} \left\{ \begin{array}{l} q_0(p_0 + y' - y, y', x_0) \\ + q_1(p_1 - a + y' - y, y', x_1) \end{array} \right\} \end{array} \right), 1 \end{array} \right\}, \\ \text{if } 0 \leq a \leq p_1 - p_0, \\ = 1, \quad \text{if } a \geq p_1 - p_0. \end{array} \right. \end{array} \right. \tag{44}$$

COROLLARY 3. *If we allow x to be multidimensional, or that utility is increasing in quality, the corresponding (wider) bounds would involve max/min over all y' and not, for example, $y' \leq y$ or $y' \geq y + a$, etc. For example, the first line of (44) would become*

$$\left[\begin{array}{l} \max \left\{ \max_{y'} \left\{ \begin{array}{l} q_1(p_1 - a + y' - y, y', x_1) \\ + q_0(p_0 + y' - y, y', x_0) \end{array} \right\} - 1 - q_0(p_1 - a, y, x_1), 0 \right\}, \\ \min \left\{ q_1(p_1 - a, y, x_1), \min_{y'} \left\{ \begin{array}{l} q_0(p_0 + y' - y, y', x_0) \\ + q_1(p_1 - a + y' - y, y', x_1) \end{array} \right\} \right\} \end{array} \right], \quad \text{if } a < 0,$$

etc.

PROOF OF PROPOSITION 3. Note that the CV is defined as the solution S to the equation

$$\max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} = \max\{U_0(y + S, \eta), U_1(y + S - p_1, x_1, \eta)\}.$$

Now, by Assumption 2(i),

$$\begin{aligned}
 \Pr(S \leq a) &= \Pr \left[\begin{array}{l} \max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} \\ \leq \max\{U_0(y + a, \eta), U_1(y + a - p_1, x_1, \eta)\} \end{array} \right]. \tag{45}
 \end{aligned}$$

Now, for $a < 0$, by Assumption 2(i). (45) reduces to

$$\begin{aligned} \Pr(S \leq a) &= \Pr \left[\begin{array}{l} \max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} \\ \leq U_1(y + a - p_1, x_1, \eta) \end{array} \right] \\ &= \Pr \left[\begin{array}{l} U_0(y, \eta) \leq U_1(y + a - p_1, x_1, \eta), \\ U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta) \end{array} \right]. \end{aligned} \quad (46)$$

This probability has the Frechet upper bound given by

$$\begin{aligned} \min \left\{ \begin{array}{l} \Pr[U_0(y, \eta) \leq U_1(y + a - p_1, x_1, \eta)], \\ \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] \end{array} \right\} \\ = \min\{q_1(p_1 - a, y, x_1), \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)]\}. \end{aligned} \quad (47)$$

Now,

$$\begin{aligned} \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] &= 1 - \Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)] \\ &\leq 1 - \min_{y'} \Pr[U_1(y - p_0, x_0, \eta) \geq U_0(y', \eta) \geq U_1(y + a - p_1, x_1, \eta)] \\ &\leq 1 - \min_{y'} \left\{ \max \left\{ \left(\begin{array}{l} \Pr[U_1(y - p_0, x_0, \eta) \geq U_0(y', \eta)] \\ + \Pr[U_0(y', \eta) \geq U_1(y + a - p_1, x_1, \eta)] \end{array} \right) - 1, 0 \right\} \right\} \\ &= 1 - \min_{y'} \left\{ \max \left\{ \left(\begin{array}{l} q_1(p_0 + y' - y, y', x_0) \\ + q_0(p_1 - a + y' - y, y', x_1) \end{array} \right) - 1, 0 \right\} \right\} \\ &= \min \left\{ \min_{y'} \left\{ \begin{array}{l} q_0(p_0 + y' - y, y', x_0) \\ + q_1(p_1 - a + y' - y, y', x_1) \end{array} \right\}, 1 \right\}. \end{aligned}$$

Replacing in (47), we get the upper bound as

$$\begin{aligned} \min \left\{ q_1(p_1 - a, y, x_1), \min_{y'} \left\{ \begin{array}{l} q_0(p_0 + y' - y, y', x_0) \\ + \underbrace{q_1(p_1 - a + y' - y, y', x_1)}_{\text{If } y' < y, \text{ then this term exceeds } q_1(p_1 - a, y, x_1)} \end{array} \right\} \right\} \\ = \min \left\{ q_1(p_1 - a, y, x_1), \min_{y' \geq y} \left\{ \begin{array}{l} q_0(p_0 + y' - y, y', x_0) \\ + q_1(p_1 - a + y' - y, y', x_1) \end{array} \right\} \right\}. \end{aligned}$$

The Frechet lower bound is given by

$$\begin{aligned}
& \max \left\{ \begin{array}{l} \Pr[U_0(y, \eta) \leq U_1(y + a - p_1, x_1, \eta)] \\ + \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] - 1, 0 \end{array} \right\} \\
&= \max \left\{ \begin{array}{l} q_1(p_1 - a, y, x_1) \\ + \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] - 1, 0 \end{array} \right\} \\
&\geq \max \left\{ \begin{array}{l} \max_{y'} \{ \Pr[U_1(y - p_0, x_0, \eta) \leq U_0(y', \eta) \leq U_1(y + a - p_1, x_1, \eta)] \} \\ - q_0(p_1 - a, y, x_1), 0 \end{array} \right\} \\
&\geq \max \left\{ \begin{array}{l} \max \left\{ \max_{y'} \left\{ \begin{array}{l} \Pr[U_0(y', \eta) \leq U_1(y + a - p_1, x_1, \eta)] \\ + \Pr[U_1(y - p_0, x_0, \eta) \leq U_0(y', \eta)] - 1 \end{array} \right\}, 0 \right\} \\ - q_0(p_1 - a, y, x_1), 0 \end{array} \right\} \\
&= \max \left\{ \begin{array}{l} \max_{y'} \left\{ \max_{y' < y} \{ \underbrace{q_1(p_1 - a + y' - y, y', x_1) + q_0(p_0 + y' - y, y', x_0)}_{\text{If } y \geq y', \text{ then sum is less than 1, as } a < 0} \} - 1, 0 \right\} \\ - q_0(p_1 - a, y, x_1), 0 \end{array} \right\} \\
&= \max \left\{ \begin{array}{l} \max \left\{ \max_{y' < y} \{ q_1(p_1 - a + y' - y, y', x_1) + q_0(p_0 + y' - y, y', x_0) \} - 1, 0 \right\} \\ - q_0(p_1 - a, y, x_1), 0 \end{array} \right\} \\
&= \max \left\{ \begin{array}{l} \max_{y' < y} \{ q_1(p_1 - a + y' - y, y', x_1) + q_0(p_0 + y' - y, y', x_0) \} \\ - 1 - q_0(p_1 - a, y, x_1), 0 \end{array} \right\}.
\end{aligned}$$

Finally, for $a \geq 0$, we get that

$$\begin{aligned}
& \Pr(S \leq a) \\
&= \Pr \left[\begin{array}{l} \max\{U_0(y, \eta), U_1(y - p_0, x_0, \eta)\} \\ \leq \max\{U_0(y + a, \eta), U_1(y + a - p_1, x_1, \eta)\} \end{array} \right] \\
&= \Pr \left[\begin{array}{l} U_1(y - p_0, x_0, \eta) \\ \leq \max\{U_0(y + a, \eta), U_1(y + a - p_1, x_1, \eta)\} \end{array} \right],
\end{aligned}$$

which equals 1 if $a \geq p_1 - p_0$. If $a < p_1 - p_0$, then

$$\begin{aligned}
& \Pr(S > a) \\
&= \Pr \left[\begin{array}{l} U_1(y - p_0, x_0, \eta) \\ > \max\{U_0(y + a, \eta), U_1(y + a - p_1, x_1, \eta)\} \end{array} \right] \\
&= \Pr \left[\begin{array}{l} U_1(y - p_0, x_0, \eta) > U_0(y + a, \eta), \\ U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta) \end{array} \right].
\end{aligned}$$

The Frechet lower bound on this is given by

$$\begin{aligned} & \max \left\{ \begin{array}{l} \Pr[U_1(y - p_0, x_0, \eta) > U_0(y + a, \eta)] \\ + \Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)] \\ - 1, 0 \end{array} \right\} \\ & \geq \max \left\{ \begin{array}{l} q_1(p_0 + a, y + a, x_0) \\ + \max_{y'} \Pr[U_1(y - p_0, x_0, \eta) > U_0(y', \eta) \geq U_1(y + a - p_1, x_1, \eta)] \\ - 1, 0 \end{array} \right\} \\ & \geq \max \left\{ \begin{array}{l} q_1(p_0 + a, y + a, x_0) \\ + \max_{y'} q_1(p_0 + y' - y, y', x_0) + q_0(p_1 - a + y' - y, y', x_1, \eta) \\ - 2, 0 \end{array} \right\}. \end{aligned}$$

The Frechet upper bound is given by

$$\begin{aligned} & \Pr \left[\begin{array}{l} U_1(y - p_0, x_0, \eta) > U_0(y + a, \eta), \\ U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta) \end{array} \right] \\ & \leq \min \left\{ \begin{array}{l} \Pr[U_1(y - p_0, x_0, \eta) > U_0(y + a, \eta)], \\ \Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)] \end{array} \right\} \\ & \leq \min \left\{ \begin{array}{l} \Pr[U_1(y - p_0, x_0, \eta) > U_0(y + a, \eta)], \\ \Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)] \end{array} \right\} \\ & = \min \left\{ \begin{array}{l} q_1(p_0 + a, y + a, x_0), \\ \Pr[U_1(y - p_0, x_0, \eta) > U_1(y + a - p_1, x_1, \eta)] \end{array} \right\}. \end{aligned}$$

The previous two displays imply that the upper bound on $\Pr(S \leq a)$ is given by

$$\begin{aligned} & \min \left\{ \begin{array}{l} q_0(p_0 + a, y + a, x_0) \\ + \min_{y'} \{q_0(p_0 + y' - y, y', x_0) + q_1(p_1 - a + y' - y, y', \eta)\}, 1 \end{array} \right\} \\ & = \min \left\{ 1, \left\{ \begin{array}{l} q_0(p_0 + a, y + a, x_0) \\ + \min_{y'} \{q_0(p_0 + y' - y, y', x_0) + q_1(p_1 - a + y' - y, y', x_1, \eta)\} \right\} \right\}, \end{array} \right\} \end{aligned}$$

and a lower bound by

$$\begin{aligned} & \max \left\{ q_0(p_0 + a, y + a, x_0), \right. \\ & \left. \Pr[U_1(y - p_0, x_0, \eta) \leq U_1(y + a - p_1, x_1, \eta)] \right\} \\ & \geq \max \left\{ q_0(p_0 + a, y + a, x_0), \right. \\ & \left. \max_{y'} \{ q_1(p_1 - a + y' - y, y', x_1) + q_0(p_0 + y' - y, y', x_0) \} - 1 \right\}. \end{aligned}$$

Putting all of this together, we get that

$$\Pr(S \leq a) \left\{ \begin{aligned} & \in \left[\begin{aligned} & \max \left\{ \max_{y' < y} \left\{ q_1(p_1 - a + y' - y, y', x_1) \right\} - 1 - q_0(p_1 - a, y, x_1), 0 \right\}, \right. \\ & \left. \min \left\{ q_1(p_1 - a, y, x_1), \min_{y' \geq y} \left\{ q_0(p_0 + y' - y, y', x_0) \right. \right. \right. \right. \\ & \left. \left. \left. \left. + q_1(p_1 - a + y' - y, y', x_1) \right\} \right\} \right\} \right], \\ & \text{if } a < 0, \\ & \in \left[\begin{aligned} & \max \left\{ q_0(p_0 + a, y + a, x_0), \right. \\ & \left. \max_{y' \geq y+a} \left\{ q_1(p_1 - a + y' - y, y', x_1) \right\} - 1 \right\}, \right. \\ & \left. \min \left\{ \left(q_0(p_0 + a, y + a, x_0) \right. \right. \right. \\ & \left. \left. \left. + \min_{y' \leq y+a} \left\{ q_0(p_0 + y' - y, y', x_0) \right. \right. \right. \right. \\ & \left. \left. \left. \left. + q_1(p_1 - a + y' - y, y', x_1) \right\} \right) \right\}, 1 \right\} \right], \\ & \text{if } 0 \leq a < p_1 - p_0, \\ & = 1, \quad \text{if } a \geq p_1 - p_0. \end{aligned} \right. \end{aligned}$$

For $p_0 = p_1$, we have

$$\Pr(S \leq a) \left\{ \begin{aligned} & \in \left[\begin{aligned} & \max \left\{ \max_{y' < y} \left\{ q_1(p_0 - a + y' - y, y', x_1) \right\} - 1 - q_0(p_0 - a, y, x_1), 0 \right\}, \right. \\ & \left. \min \left\{ q_1(p_0 - a, y, x_1), \min_{y' \geq y} \left\{ q_0(p_0 + y' - y, y', x_0) \right. \right. \right. \\ & \left. \left. \left. \left. + q_1(p_0 - a + y' - y, y', x_1) \right\} \right\} \right\} \right], \\ & \text{if } a < 0, \\ & = 1, \quad \text{if } a \geq 0. \end{aligned} \right. \quad \square$$

REFERENCES

Auerbach, A. J. (1985), “The theory of excess burden and optimal taxation.” In *Handbook of Public Economics*, Vol. 1, 61–127. [576]

Berry, S. and P. Haile (2015), “Identification in differentiated products markets.” NBER Working Paper No. 21500. [584]

- Bhattacharya, D. (2015), “Nonparametric welfare analysis for discrete choice.” *Econometrica*, 83 (2), 617–649. [572, 577, 585, 600, 601]
- Blundell, R. and J. Powell (2003), “Endogeneity in nonparametric and semiparametric regression models.” In *Advances in Economics and Econometrics*, Cambridge University Press, Cambridge, UK. [591]
- Böhning, D. (1992), “Multinomial logistic regression algorithm.” *Annals of the Institute of Statistical Mathematics*, 44 (1), 197–200. [586]
- Carneiro, P., J. Das, and H. Reis (2016), “The value of private schools: Evidence from Pakistan.” Working paper, University College London. [573]
- Dagsvik, J. and A. Karlstrom (2005), “Compensating variation and Hicksian choice probabilities in random utility models that are nonlinear in income.” *Review of Economic Studies*, 72 (1), 57–76. [572, 573]
- Domencich, T. and D. McFadden (1975), *Urban Travel Demand—A Behavioral Analysis*. North-Holland, Oxford. [572]
- Goolsbee, A. (1999), “Evidence on the high-income Laffer curve from six decades of tax reform.” *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution*, 30 (2), 1–64. [572]
- Goolsbee, A. and P. Klenow (2006), “Valuing consumer products by the time spent using them: An application to the Internet.” *American Economic Review*, 96, 108–113. [573]
- Hausman, J. (1981), “Exact consumer’s surplus and deadweight loss.” *The American Economic Review*, 71 (4), 662–676. [572]
- Hausman, J. (1996), “Valuation of new goods under perfect and imperfect competition.” In *The Economics of New Goods*, 207–248, University of Chicago Press, Chicago. [573, 591]
- Hausman, J. (2003), “Sources of bias and solutions to bias in the consumer price index.” *The Journal of Economic Perspectives*, 17 (1), 23–44. [573]
- Hausman, J. A. and G. K. Leonard (2002), “The competitive effects of a new product introduction: A case study.” *The Journal of Industrial Economics*, 50 (3), 237–263. [573, 574]
- Hausman, J. and W. Newey (2016), “Individual heterogeneity and average welfare.” *Econometrica*, 84 (3), 1225–1248. [572]
- Hausman, J. and W. Newey (2017), “Nonparametric welfare analysis.” *Annual Review of Economics*, 9, 521–546. [572]
- Herriges, J. and C. Kling (1999), “Nonlinear income effects in random utility models.” *The Review of Economics and Statistics*, 81 (1), 62–72. [573, 575, 591]
- Ichimura, H. and C. Taber (2002), “Semiparametric reduced-form estimation of tuition subsidies.” *American Economic Review*, 2002, 286–292. [572]
- Kane, T. J. (2003), “A quasi-experimental estimate of the impact of financial aid on college-going.” No. w9703, National Bureau of Economic Research. [572]

- Kling, C. L. and C. J. Thomson (1996), “The implications of model specification for welfare estimation in nested logit models.” *American Journal of Agricultural Economics*, 78 (1), 103–114. [591]
- Knittel, C. R. and K. Metaxoglou (2014), “Estimation of random-coefficient demand models: Two empiricists’ perspective.” *Review of Economics and Statistics*, 96 (1), 34–59. [573]
- Lewbel, A. (2001), “Demand systems with and without errors.” *The American Economic Review*, 91 (3), 611–618. [572]
- Lewbel, A. and K. Pendakur (2017), “Unobserved preference heterogeneity in demand using generalized random coefficients.” *Journal of Political Economy*, 125 (4), 1100–1148. [572]
- McFadden, D. (1973), “Conditional logit analysis of qualitative choice behavior.” [586]
- McFadden, D. (1981), “Econometric models of probabilistic choice.” In *Structural Analysis of Discrete Data With Econometric Applications* (C. F. Manski and D. McFadden, eds.), Chapter 5. MIT Press, Cambridge, MA. [572]
- McFadden, D. and M. K. Richter (1990), “Stochastic rationality and revealed stochastic preference.” In *Preferences, Uncertainty, and Optimality: Essays in Honor of Leo Hurwicz*, 161–186, Westview Press, Boulder, CO. [584]
- McFadden, D. and K. Train (2000), “Mixed MNL models for discrete response.” *Journal of Applied Econometrics*, 15 (5), 447–470. [573]
- Rothenberg, T. J. (1971), “Identification in parametric models.” *Econometrica: Journal of the Econometric Society*, 39 (3), 577–591. [587]
- Small, K. and H. Rosen (1981), “Applied welfare economics with discrete choice models.” *Econometrica*, 49 (1), 105–130. [572, 573, 593]
- Stiglitz, J. (2000), *Economics of the Public Sector*. W. W. Norton, New York. [572]
- Tirole, J. (1988), *The Theory of Industrial Organization*. MIT Press, Cambridge. [576]
- Train, K. (2009), *Discrete Choice Methods With Simulation*. Cambridge University Press, Cambridge, UK. [573, 593]
- Trajtenberg, M. (1989), “The welfare analysis of product innovations, with an application to computed tomography scanners.” *Journal of Political Economy*, 97 (2), 444–479. [573]
- Vytlačil, E. and N. Yildiz (2007), “Dummy endogenous variables in weakly separable models.” *Econometrica*, 75 (3), 757–779. [584]
- Willig, R., S. Salop, and F. Scherer (1991), “Merger analysis, industrial organization theory, and merger guidelines.” *Brookings Papers on Economic Activity. Microeconomics*, 22, 281–332. [573]

Co-editor Christopher Taber handled this manuscript.

Manuscript received 22 July, 2017; final version accepted 19 October, 2017; available online 23 October, 2017.