WORKING PAPER · NO. 2022-75

Discrete-Choice Models and Representative Consumer Theory

Jean-Pierre H. Dubé, Joonhwi Joo, and Kyeongbae Kim

JUNE 2022

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

DISCRETE-CHOICE MODELS AND REPRESENTATIVE CONSUMER THEORY

Jean-Pierre H. Dubé
Joonhwi Joo
Kyeongbae Kim

June 2022

Dube acknowledges research support from the Kilts Center for Marketing and the Charles E. Merrill faculty research fund.

Discrete-Choice Models and Representative Consumer Theory
Jean-Pierre H. Dubé, Joonhwi Joo, and Kyeongbae Kim
June 2022
JEL No. C43,D01,D11,D60,E1,L00,M3

ABSTRACT

We establish the Hurwicz-Uzawa integrability of the broad class of discrete-choice additive random-utility models of individual consumer behavior with perfect substitutes preferences and divisible goods. We derive the corresponding indirect uility function and then establish a representative consumer formulation for this entire class of models. The representative consumer is always normative, facilitating aggregate welfare analysis. These findings should be of interest to the literatures in macro, trade, industrial organization, labor and ideal price index measurement that use representative consumer models, such as CES and its variants. Our results generalize such representative consumer formulations to the broad, empirically-relevant class of models of behavior that are routinely used in the discrete-choice analysis of micro data, including specifications that do not suffer from the IIA property and that allow for heterogeneous consumer preferences and incomes. When products are indivisible, we show that Hurwicz-Uzawa integrability fails; although some model variants might satisfy a stronger version of quasi-linear integrability.

Jean-Pierre H. Dubé
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
jdube@chicagobooth.edu

Joonhwi Joo
University of Texas at Dallas
800 W. Campbell Rd.
JSOM 13-326
Richardson, TX 75080
joonhwi.joo@utdallas.edu

Kyeongbae Kim
Sejong University
209 Neungdong-ro, Gwangjin-gu
Seoul 05006
South Korea
kimkb@sejong.ac.kr

1 Introduction

We study the link between a broad class of additive random utility "discrete-choice" models of demand, that are well-known to fit individual-level choice behavior with micro data, and a representative consumer theory of aggregate demand. At least since Archibald et al. (1986), the literature on differentiated products markets has sought a connection between the convenient representative consumer framework (e.g., Spence, 1976; Dixit and Stiglitz, 1977), in which the representative consumer consumes all the product variety, and the empirical reality whereby most individual consumers' consumption bundles exhibit mostly corner solutions. A closely related concern is whether the representative consumer's utility function is also normative, so it can be treated as a social welfare function for the underlying consumer population (e.g., Dow and da Costa Werlang, 1988).

The extant literature has only been able to obtain a closed-form solution for the special cases of simple Logit and Nested Logit discrete-choice models, both of which can be rationalized by the endogenous consumption of a single representative consumer with variations of CES preferences (e.g., Anderson et al., 1987, 1992; Verboven, 1996). This link has been relied on extensively to justify the use of the CES representative consumer functional form to derive practical macroeconomic models with imperfect competition in the literatures on industrial organization (e.g., Judd, 1985; B. Curtis Eaton, 1989; Zhelobodko et al., 2012; Dhingra and Morrow, 2019), macroeconomics (e.g., Blanchard and Kiyotaki, 1987), monetary policy (e.g., Beck and Lein, 2020), economic growth (e.g., Romer, 1987), trade and economic geography (e.g., Krugman, 1995; Melitz, 2003, 2008; Fajgelbaum et al., 2011; Khandelwal, 2010; Bernini and Tomasi, 2015; Crin'o and Ogliari, 2017), labor (e.g., Berger et al., Forthcoming; Card, 2022) and ideal price index measurement (e.g., Redding and Weinstein, 2020). In their list of research priorities, Anderson et al. (1992, p.91) specifically indicate: "...it would be interesting to find the form of the representative consumer's utility function for other discrete choice models, such as the Probit, the nested MNL, and the GEV." We generalize the link between the representative consumer model and a broad class of discrete-choice, additive random utility models of consumer behavior that have been widely studied in the empirical marketing literature.

We study the Hurwicz and Uzawa (1971) integrability of the expected discrete-continuous demand function corresponding to a consumer with perfect substitutes preferences (i.e., linear indifference curves) and an absolutely-continuous, additive random utility component, hereafter the *additive random utility model* (ARUM). In the ARUM, individual demand exhibits corner solutions for all but at most one of the market goods (e.g., Deaton and Muellbauer, 1980; Hanemann, 1984). We derive and characterize the indirect utility function for the ARUM for *any* absolutely continuous, full-support distribution, of the random-utility component. We also derive necessary conditions on the functional form of each product's mean utility in order for the ARUM discrete-continuous demand model to satisfy integrability.

To establish the link to a representative consumer theory, we then study the aggregate demand from an underlying population of consumers making discrete choices. From the Gorman form of the indirect utility, it follows that the expected discrete-continuous aggregate demand system for the ARUM always has a positive and normative representative consumer formulation with a class of preferences that nests the popular CES model depending on the assumed distribution of the random utility. We therefore generalize Anderson et al. (1987, 1992); Verboven (1996) beyond simple Logit or one-level Nested Logit. We present a number of examples, including the Becker-Lancaster "characteristics" models, GEV, random-coefficients Logit and multinomial Probit models that allow for flexible substitution patterns between products and for which the representative consumer's demand and indirect utility can be expressed in closed form in most cases. This characterization of the normative representative consumer's indirect utility function facilitates a practical means for welfare analysis utilizing a broad class of ARUM discrete-choice models of demand. Many of these formulations do not exhibit the IIA property or the corresponding unrealistic consumer substitution patterns associated with the popular logit model and its CES formulation.[1] We also explore ARUM formulations that allow for both heterogeneous preferences and incomes at the individual level, that should lead to more robust ideal price index and cost-of-living measurements (e.g., Redding and Weinstein, 2020).

In practice, many empirical applications of discrete-choice analysis to individual-level demand impose indivisibility, requiring consumers to form totally-inelastic demands over the chosen alternative (e.g., Berry et al., 1995; Goldberg, 1995). We show that the Hurwicz and Uzawa-integrability of the expected pure-discrete-choice aggregate demand function corresponding to indivisible goods fails, even for the simple Logit model. Indivisibility can lead to well-known problems with the budget-balancedness condition (Nocke and Schutz, 2017), with either a non-binding budget constraint or negative consumption of the outside good. In some cases, a stronger version of quasi-linear integrability Nocke and Schutz might hold.

The representative consumer formulation for a broad class of underlying populations of consumers making discrete choices provides additional empirical justification for such popular representative consumer models as the CES and its variants used for macro modeling and welfare analysis. In a broad study of consumer packaged goods shopping behavior in the U.S., Dubé (2019) finds that consumers select a single brand alternative in a category for over 90% of the observed transactions. However, our results herein are limited to discrete-choice models in which individual consumers purchase at most a single product. Our results do not extend to the more general context of multiple-discrete-choice whereby a consumer purchases an assortment of different products in the category, with corner solutions arising for a subset of the available products (e.g., Wales and Woodland, 1983; Lee and Pitt, 1986; Hendel, 1999; Kim et al., 2002; Dubé, 2004; Bhat, 2005, 2008).

Our work adds to the extant literature on expected discrete-continuous demand systems with

[1] Helpman (2011) discusses the adverse implications of the CES representative consumer model's IIA property for the estimated gains from variety in a model of trade. Usual solutions, like the nested logit, only partially offset these problems between nests and require the researcher to pre-classify products into nests.

single-product choices (e.g., Dubin and McFadden, 1984; Hanemann, 1984; Chiang, 1991; Chintagunta, 1993). Typically, that literature derives demand from the "dual" problem and applyies Roy's Identity to a pre-specified indirect utility function. In contrast, we derive the demand and indirect utility functions from the "primal" problem with ARUM form. From the primal problem, we find that the necessary conditions for integrability exclude the use of arbitrary flexible functional forms for the ARUM indirect utility function. In addition, we establish that this class of models has a positive and normative representative consumer formulation with a known class of preferences.

Our results on the integrability of aggregate discrete-choice demand systems also relate to the recent literature studying non-binding budget constraints and the potential for strictly negative consumption of the numeraire good in the context of indivisibility (Hosoya, 2017; Nocke and Schutz, 2017, 2018). Our results add to this literature by demonstrating the (non)-integrability of a similar class of aggregate demand systems with indivisibility when we assume the representative consumer faces a binding budget constraint.

The remainder of the article is organized as follows. Section 2 presents the setup of ARUM individual demand with perfect substitutes preferences under divisibility and establishes the Hurwicz and Uzawa (1971) integrability of the individual expected demand. Section 3 studies the aggregation problem of the ARUM individual expected demand, establishes the existence of the positive and normative representative consumer associated with the demand system, and presents several examples. Section 4 studies the integrability problem of the ARUM individual demand under indivisibility. Section 5 concludes.

2 Individual Demand with Perfect Substitutes Preferences and Divisible Goods

In this section, we start with the general formulation of the ARUM for an individual consumer with perfect substitutes preferences. We first derive the corresponding expected discrete-continuous demand system. We then prove the Hurwicz and Uzawa (1971) integrability of the expected demand system and characterize the analytic form of the corresponding indirect utility function.

2.1 Perfect Substitutes and Discrete-Continuous Choice

We focus on a market that supplies J variants of a differentiated product – *brands*. An individual consumer consumes at most one of the brands. To obtain this *discrete-choice* behavior, we use a random-utility formulation in which consumers perceive the J brands as perfect substitutes as in the *simple re-packaging model with varieties* (e.g., Deaton and Muellbauer, 1980; Hanemann, 1984).[2]

[2] We do not consider the alternative "mutual exclusivity" formulation of discrete-choice herein (e.g., Hanemann, 1984).

Formally, a consumer with total income y chooses a consumption bundle $\mathbf{q} = (q_1, ..., q_J)'$ to solve the following utility-maximization problem:

$$\mathbf{q}^* = \underset{\mathbf{q}\in\mathbb{R}_+^J}{\text{argmax}} \sum_{j=1}^{J} q_j \psi_j(\epsilon_j) \text{ s.t. } \mathbf{q}'\mathbf{p} \leq y \tag{2.1}$$

where $\boldsymbol{\psi} = (\psi_1, ..., \psi_J)'$, is the vector of constant marginal utilities for each product, or "brand qualities," with $\psi_j \geq 0$, $\mathbf{p} = (p_1, ..., p_J)'$ are the strictly positive brand prices, and $\boldsymbol{\epsilon} = (\epsilon_1, ..., \epsilon_J)' \sim F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon})$ is a vector of random-utility disturbances with absolutely-continuous distribution function $F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon})$ with full support on $\mathbb{R}^J$.

Due to the linearity of the indifference curves associated with the perfect substitutes preferences, the utility maximization problem (2.1) leads to a corner solution in which only a single brand is chosen.[3] The consumer chooses brand j (WLOG) if

$$\frac{p_j}{\psi_j(\epsilon_j)} = \min_{k\in\{1,2,...,K\}} \left\{ \frac{p_k}{\psi_k(\epsilon_k)} \right\}.$$

Assumption 1. $\log(\psi_j(\epsilon_j)) = \theta_j + \epsilon_j, \forall j$, *where* $\theta_j \in \mathbb{R}$ *is the deterministic component of product* j'*s quality.*

Under Assumption 1, the consumer chooses brand j if,

$$\begin{aligned} \epsilon_j - \epsilon_k &\geq [\theta_k - \log(p_k)] - [\theta_j - \log(p_j)], \forall k \\ &= -(\delta_j - \delta_k), \end{aligned} \tag{2.2}$$

where $\delta_j := \theta_j - \log(p_j)$. The corresponding probability that the consumer chooses brand j is

$$\begin{aligned} \pi_j(\boldsymbol{\delta}(\mathbf{p})) &\equiv \Pr(\epsilon_j - \epsilon_k \geq \delta_k - \delta_j, \forall k) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\delta_j - \delta_1 + \tilde{\epsilon}} \cdots \int_{-\infty}^{\delta_j - \delta_{j-1} + \tilde{\epsilon}} \int_{-\infty}^{\delta_j - \delta_{j+1} + \tilde{\epsilon}} \cdots \int_{-\infty}^{\delta_j - \delta_J + \tilde{\epsilon}} \\ &\quad f_{\boldsymbol{\epsilon}}(\epsilon_1, ..., \epsilon_{j-1}, \tilde{\epsilon}, \epsilon_{j+1}, ..., \epsilon_J) \, d\epsilon_J ... d\epsilon_{j+1} d\epsilon_{j-1} ... d\epsilon_1 d\tilde{\epsilon} \end{aligned} \tag{2.3}$$

which does not depend on the consumer's income.

Therefore, the consumer has the following *expected* discrete-continuous demand for brand j:

$$q_j(\mathbf{p}, y) \equiv \mathbb{E}_{\boldsymbol{\epsilon}}\left[q_j^* | \mathbf{p}\right] = \pi_j(\boldsymbol{\delta}(\mathbf{p})) \frac{y}{p_j}. \tag{2.4}$$

4

[3] Note that the probability of ties is zero because $F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon})$ is absolutely continuous on $\mathbb{R}^J$.

[4] There is no loss of generality in the fact that $\delta_j = \theta_j - \log(p_j)$ has a coefficient equal to -1 on log-price. This property is a result, not an assumption, and will be true for any random utility distribution, $F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon})$. From an empirical perspective, this result also does not imply that log-price has a coefficient equal to -1 in the demand system $q_j(\mathbf{p}, y)$. As we show in Section 3.3 below, the scale of the random utility distribution enters the demand model $q_j(\mathbf{p}, y)$ as the price coefficient.

The following analysis will help with the integrability results derived in the sections below. The expected indirect brand utility per unit purchased with expectations taken against the distribution of $\boldsymbol{\epsilon}$ is

$$\mathcal{G}(\boldsymbol{\delta}) := \mathbb{E}_{\boldsymbol{\epsilon}}\left[\max_{j\in\{1,2,\ldots,J\}}\{\delta_j+\epsilon_j\}\right]. \tag{2.5}$$

In a pure discrete-choice model with totally-inelastic quantity demanded, $\mathcal{G}(\boldsymbol{\delta})$ is the social surplus function (e.g., McFadden, 1981; Small and Rosen, 1981). From the Williams-Daly-Zachary (WDZ) theorem (Williams, 1977; Daly and Zachary, 1978; Fosgerau et al., 2013, WDZ thoeorem henceforth), we know that:

$$\nabla_{\boldsymbol{\delta}}\mathcal{G}(\boldsymbol{\delta}) = \boldsymbol{\pi}' \tag{2.6}$$

where $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_J)'$ is the vector of choice probabilities. Therefore, we can re-write the consumer's choice probability for brand j (2.3) as

$$\pi_j(\boldsymbol{\delta}(\mathbf{p})) = \frac{H_j(\boldsymbol{\delta})}{\sum_{k=1}^{J} H_k(\boldsymbol{\delta})} \tag{2.7}$$

where $H(\boldsymbol{\delta}) := \exp(\mathcal{G}(\boldsymbol{\delta}))$ is a *choice probability generating function* (Fosgerau et al., 2013). In the remainder of the paper, we let $H_j(\cdot)$ denote the partial derivative of $H(\cdot)$ with respect to its jth element δ_j: $H_j(\boldsymbol{\delta}) := \{\nabla_{\boldsymbol{\delta}} H(\boldsymbol{\delta})\}_j$. Similarly, we let $H_{jk}(\boldsymbol{\delta})$ denote the cross partial derivative with respect to δ_j and δ_k, which we assume exist and are continuous on $\mathbb{R}^J$ as in McFadden (1981). As is well-known in the literature, the formulation of the choice probability in (2.7) is consistent with any absolutely-continuous distribution of the random utilities, $F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon})$, including popular empirical implementations like the simple Logit, GEV, Probit and random coefficients models.

It follows immediately that since $\mathcal{G}(\boldsymbol{\delta})$ is convex in $\boldsymbol{\delta}$ (Rust, 1994; Chiong et al., 2016; Chiong and Shum, 2019), $H(\boldsymbol{\delta})$ is also convex in $\boldsymbol{\delta}$ because $\exp(\cdot)$ is a nondecreasing convex function on $\mathbb{R}$. $H_j(\boldsymbol{\delta})$ is non-negative for any $\boldsymbol{\delta} \in \mathbb{R}^J$ and for any j because $H(\boldsymbol{\delta})$ and π_j are all strictly positive.

2.2 Integrability of the Individual's Expected Demand System

We now state and prove the main results of the paper. We first establish the integrability of the *expected* individual discrete-continuous demand function (2.4). We then derive the analytic form of the corresponding expenditure function and indirect utility function.

The following Theorem 1 establishes the integrability of the demand function (2.4) by showing that it satisfies the necessary and sufficient conditions from Hurwicz and Uzawa (1971).

Theorem 1. *The demand function $\mathbf{q}(\mathbf{p}, y)$ of the form (2.4) is integrable in the Hurwicz and Uzawa (1971) sense as, for any $\mathbf{p} \in \mathbb{R}_+^J$ and for any income level $y \in \mathbb{R}_+$, it satisfies the following conditions:*

(i) (Differentiability of Demand) $\mathbf{q}(\mathbf{p}, y)$ is a continuously differentiable function with bounded partial derivatives,

(ii) (Budget Balancedness) $\mathbf{p}'\mathbf{q} = y$,

(iii) (Slutsky Symmetry) The associated Slutsky matrix $\mathbf{S} := \left\{ \frac{\partial q_k}{\partial p_j} + q_k \frac{\partial q_k}{\partial y} \right\}_{j,k \in \{1,2,\ldots,J\}}$ *is symmetric, and*

(iv) (Negative Semidefiniteness of Slutsky Matrix) $\mathbf{S}$ *is negative semidefinite.*

Proof. See Appendix A. □

An interesting component of the proof is that a necessary and sufficient condition for the Slutsky Symmetry and Negative Semidefiniteness of the ARUM demand system is that the mean utility per unit consumed of a brand has the form: $\delta_j = c_j - \eta \log(p_j)$ where $c_j \in \mathbb{R}$ and $\eta \geq 0$. This condition holds mechanically for our ARUM, where $\eta = 1$ and $c_j = \theta_j$. However, it also shows why alternative, ad hoc formulations of the discrete-continuous demand system might not be integrable. For instance, some researchers have used the dual approach to derive the discrete-continuous model of demand from a flexible functional form assumption for the indirect utility (e.g., Hanemann, 1984; Chiang, 1991; Chintagunta, 1993).

The following corollary characterizes the analytic forms of the expenditure function and indirect utility functions, respectively, corresponding to the expected demand system (2.4).

Corollary 1. *The expenditure function corresponding to the demand function (2.4) is*

$$e(\mathbf{p}, u) = \frac{u}{H(\boldsymbol{\delta}(\mathbf{p}))}, \; u \geq 0,$$

and the corresponding indirect utility function is

$$v(\mathbf{p}, y) = yH(\boldsymbol{\delta}(\mathbf{p})). \tag{2.8}$$

Proof. See Appendix A. □

Theorem 1 and Corollary 1 characterize the preferences of a broad class of ARUM models with choice probabilities, $\{\pi_j(\mathbf{p})\}_{j=1}^{J}$, spanning many well-known discrete-choice models from the empirical literature including Logit, random coefficients ("mixed") Logit, GEV and Probit.

An immediate consequence of Corollary 1 is that the indirect utility function for our expected aggregate discrete-continuous-choice demand system (2.4) has the Gorman form. We will use this property extensively in the aggregation results below.

3 Aggregate-Discrete-Continuous-Choice Demand and the Representative Consumer

We now derive the aggregate demand and representative consumer formulation corresponding to the ARUM from Section 2. The Gorman form of the indirect utility function leads to the well-known result that aggregate demand depends only on aggregate income and is independent of the distribution of income across consumers. Furthermore, from the Gorman form, it follows that a positive and normative representative consumer exists and can be used for welfare analysis. We now derive the analytic form of aggregate demand and characterize the class of preferences for the representative consumer corresponding to the ARUM.

3.1 Aggregate Demand

Consider a population of N consumers indexed $i = 1, ..., N$ each with ARUM preferences as in Section 2.1 and incomes $\left\{y^i\right\}_{i=1}^N$. The corresponding expected aggregate discrete-continuous-choice demand system is:

$$Q_j(\mathbf{p}, Y) = \sum_{i=1}^{N} q_j^i\left(\mathbf{p}, y^i\right) = \frac{Y}{p_j}\pi_j(\boldsymbol{\delta}(\mathbf{p})) \qquad \text{for } j = 1, ..., J \tag{3.1}$$

where $Y = \sum_{i=1}^N y^i$ is the aggregate income.

The demand system (3.1) can also be derived as the aggregation over a continuum of consumers making deterministic choices, as in Anderson et al. (1992) for the case of pure discrete choice with indivisible goods. Suppose the population consists of a continuum of consumers with mass N and income distribution $F_y(\cdot)$. The corresponding aggregate demand system is:

$$\begin{aligned} Q_j(\mathbf{p}, Y) &= N \iint \mathbf{1}\left(\delta_j + \epsilon_j \geq \delta_k + \epsilon_k \, \forall k \neq j\right) \tfrac{y^i}{p_j} dF_y\left(y^i\right) dF_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}) \\ &= \tfrac{Y}{p_j}\pi_j(\boldsymbol{\delta}(\mathbf{p})) \end{aligned} \tag{3.2}$$

where $\pi_j(\boldsymbol{\delta}(\mathbf{p})) = \frac{H_j(\boldsymbol{\delta})}{\sum_{k=1}^J H_k(\boldsymbol{\delta})}$ as before in (2.7) and $Y = N \int y^i dF_y\left(y^i\right)$ is aggregate income.

3.2 The Representative Consumer

We now characterize the representative consumer formulation of the aggregation of the ARUM. Of particular interest is the ability to use the representative consumer for welfare analysis.

Consider a Bergson-Samuelson social welfare function (Samuelson, 1956) $\mathcal{W} : \mathbb{R}^N \rightarrow \mathbb{R}$ that maps vectors of individual utilities, $\left(u^1, u^2, ..., u^N\right)$, into a "social utility" and is increasing, continuously differentiable, and concave. We can define the social indirect utility function as the solution to the following income allocation problem:

$$v^{\mathcal{W}}(\mathbf{p}, Y) = \max_{\{y^i\}_{i=1}^N} \quad \mathcal{W}\left(v^1\left(\mathbf{p}, y^1\right), ..., v^N\left(\mathbf{p}, y^N\right)\right) \quad \text{s.t.} \sum_{i=1}^N y^i = Y, \tag{3.3}$$

where $v^i\left(\mathbf{p}, y^i\right)$ is consumer i's indirect utility function. A positive representative consumer with demand $\mathbf{Q}(\mathbf{p}, Y) = \sum_{i=1}^N \mathbf{q}^i\left(\mathbf{p}, y^{i,\mathcal{W}}\right)$ is also a *normative representative consumer* with respect to the social welfare function $\mathcal{W}(\cdot)$ when her indirect utility is $v(\mathbf{p}, Y) = v^{\mathcal{W}}(\mathbf{p}, Y)$.

The following proposition establishes that a positive and normative representative consumer for our ARUM exists and has an analytic form for her indirect utility.

Proposition 1. *(Integrability of the Aggregate Discrete-Continuous-Choice Demand System) The demand system $\mathbf{Q}(\mathbf{p}, Y)$ characterized by (3.1) corresponds to a representative consumer with indirect utility*

$$v(\mathbf{p}, Y) = Y \cdot H(\boldsymbol{\delta}(\mathbf{p})) \tag{3.4}$$

where

$$v(\mathbf{p}, Y) = v^{\mathcal{W}}(\mathbf{p}, Y)$$

for any Bergson-Samuelson social welfare function, $\mathcal{W}(\cdot)$.

Proof. See Appendix B. □

The proof of Proposition 1 relies on the Gorman form of the indirect utility in the ARUM (2.8). The proposition establishes that a population of consumers with perfect substitutes preferences, divisible consumption and an absolutely-continuous distribution of random utility always has a positive and normative representative consumer. This result applies in both the case of a population of N consumers and a population with a continuum of consumers with total mass N.

This characterization of the representative consumer formulation spans many popular empirical ARUM models such as random coefficients models (e.g., mixed-Logit), multinomial Probit and GEV. Below we provide some examples. However, for many popular empirical models like mixed-Logit and multivariate Probit, the function $H(\cdot)$ may not have a closed-form representation and would need to be computed numerically.

Proposition 1 generalizes Anderson et al. (1987) and Anderson et al. (1992, Section 3.7), who study the CES representative consumer with aggregate Logit discrete-choice demand, as well as Verboven (1996, Section 4) who studies the extension to a Nested Logit model. Our results are also related to Anderson et al. (1992, Section 3.4) who prove the existence of a representative consumer for the pure discrete-choice model under indivisible (not divisible) goods and quasi-linear preferences. However, Nocke and Schutz (2017) show that the quasi-linear pure discrete-choice model can generate negative consumption, leading to a non-binding budget constraint. This problem does not arise in our model where the quantities consumed will be strictly positive due to divisibility. In Section 4 below, we show that integrability does not hold in the Hurwicz and Uzawa (1971) sense for a pure discrete-choice aggregate demand system with indivisible quantities.

The following corollary indicates how to conduct welfare analysis using the normative representative consumer formulation in 1.

Corollary 2. *(Welfare Analysis with the Aggregate Discrete-Continuous-Choice Demand System) Consider the representative consumer characterized by (3.4). The represenative consumer's compen-*

sating variation for a change in price from $\mathbf{p}^0$ *to* $\mathbf{p}^1$*is*

$$CV = Y^0 \frac{H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right) - H\left(\boldsymbol{\delta}\left(\mathbf{p}^0\right)\right)}{H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right)}.$$

Proof. See Appendix B for derivation. □

Below we explore several commonly-used example of ARUMs from the empirical literature for which $H(\cdot)$ can be derived in closed form in most cases.

Proposition 1 can be extended in several ways to connect various assumptions about the representative consumer to other popular implementations of the discrete-choice model.

3.2.1 Adding a Numeraire

Many empirical applications include an additional, divisible numeraire good to allow for the case that some consumers do not purchase any of the brands in the commodity group. Consider the following extension of our model in Section 2 to include the consumption $q_z \in \mathbb{R}_+$ of a numeraire good (all other goods) with price normalized to one. We redefine the consumer's utility-maximization problem (2.1) as follows

$$(\mathbf{q}^*, q_z^*) = \underset{(\mathbf{q}, q_z) \in \mathbb{R}_+^{J+1}}{\operatorname{argmax}} \left(\sum_{j=1}^{J} q_j \psi_j(\epsilon_j) \right) (q_z)^{\alpha} \text{ s.t. } q_z + \mathbf{q}'\mathbf{p} \leq y. \tag{3.5}$$

The consumer's expected demand is as follows

$$q_j(\mathbf{p}, y) = \frac{1}{1+\alpha} \cdot \frac{y}{p_j} \cdot \pi_j \tag{3.6}$$

$$q_z(\mathbf{p}, y) = \frac{\alpha}{1+\alpha} \cdot y \tag{3.7}$$

where the choice probabilities $\{\pi_j\}_{j=1}^{J}$ are identical to (2.2) and the aggregate demand system is

$$Q_j(\mathbf{p}, Y) = \sum_{i=1}^{I} q_j^i = \frac{Y}{p_j} \pi_j(\mathbf{p}) \qquad \text{for } j = 1, ..., J. \tag{3.8}$$

Appendix B.3 shows that Proposition 1 still holds for the aggregate demand system (3.8) and that the corresponding representative consumer still has preferences with indirect utility: $v(\mathbf{p}, Y) = \frac{Y}{1+\alpha} \{H(\boldsymbol{\delta}(\mathbf{p}))\} \left(\frac{\alpha Y}{1+\alpha}\right)^{\alpha}$. Once again, this result generalizes for any absolutely-continuous distribution of random utility.

3.2.2 Characteristics Models

Proposition 1 also nests characteristics models in the spirit of Lancaster (1966). We can re-write the deterministic component of consumer's quality as $\theta_j = \mathbf{x}_j'\boldsymbol{\beta}$ where $\mathbf{x}_j$ is a vector of product characteristics and $\boldsymbol{\beta}$ is a vector of tastes. We then obtain

$$\delta_j(p_j, \mathbf{x}_j) = -\log p_j + \mathbf{x}_j'\boldsymbol{\beta} \tag{3.9}$$

and the proof of Proposition 1 remains the same.

3.3 Examples

We now work out a number of representative consumer formulations that correspond to popular specifications of discrete-choice demand at the individual consumer level. In several cases, we show that models with flexible substitution patterns can still generate a tractable and analytic form for the representative consumer. However, our results generalize to a much broader class of discrete-choice models, such as the multinomial Probit, which does not exhibit the IIA property. While a positive and normative representative consumer exists for the multinomial Probit, $H(\boldsymbol{\delta})$ and $v(\mathbf{p}, Y)$ can only be derived numerically due to the non-analytic form of the integration over a multivariate Normal distribution.

Example 1. (Simple Logit and One-level Nested Logit Demand) As an illustrative example, we re-work the example of Verboven (1996, Section 4) to determine the representative consumer's preferences for a population of underlying consumers whose product choices are drawn from the Nested Logit model. Suppose aggregate demand (3.1) has the the one-level aggregate discrete-continuous-choice Nested Logit demand functional form:

$$Q_j(\mathbf{p}, Y) = \frac{Y}{p_j} \frac{\left(\exp(\sigma\delta_j)\right)^{\frac{1}{\lambda_l}} \left(\sum_{k\in B_l} \left(\exp(\sigma\delta_k)\right)^{\frac{1}{\lambda_l}}\right)^{\lambda_l - 1}}{\sum_{1\leq K\leq \mathcal{K}} \left(\sum_{k\in B_K} \left(\exp(\sigma\delta_k)\right)^{\frac{1}{\lambda_K}}\right)^{\lambda_K}} \qquad \text{for } j = 1, ..., J,$$

where B_l denotes the group to which product j belongs and $\sigma = 1/\mu > 0$, where μ is the scale parameter of the GEV random utility distribution. The corresponding function $H(\boldsymbol{\delta})$ is then as follows:

$$H(\boldsymbol{\delta}) = \left\{ \sum_{1\leq K\leq \mathcal{K}} \left(\sum_{k\in B_K} \left(\exp(\sigma\delta_k)\right)^{\frac{1}{\lambda_K}} \right)^{\lambda_K} \right\}^{\frac{1}{\sigma}} \tag{3.10}$$

From Proposition 1, the representative consumer's indirect utility function is

$$v(\mathbf{p}, Y) \;=\; Y \left\{ \sum_{1 \le K \le \mathcal{K}} \left(\sum_{k \in B_K} \left(\exp\left(\sigma \delta_k\right)\right)^{\frac{1}{\lambda_K}} \right)^{\lambda_K} \right\}^{\frac{1}{\sigma}} \tag{3.11}$$

and the corresponding direct utility function is

$$u(\mathbf{Q}) = \sum_{1 \le K \le \mathcal{K}} \left(\sum_{k \in B_K} C_k^{\frac{\sigma}{\sigma+\lambda_K}} Q_k^{\frac{\sigma}{\sigma+\lambda_K}} \right)^{\frac{\sigma+\lambda_K}{\sigma+1}}, \tag{3.12}$$

where $C_k = \exp(\theta_k)$. We can easily conduct welfare analysis by substituting the formulation of $H(\boldsymbol{\delta})$ from (3.10) into (2) to obtain the compensating differential for a price change.

If we let $C_k = 1 \; \forall k$, $\rho = \frac{\sigma}{\sigma+1}$, and $\rho_K = \frac{\sigma}{\sigma+\lambda_K}$, the utility function (3.12) reduces to the Nested Logit model of demand in Verboven (1996, Section 4). If we also let λ_K =1, the model further reduces to to the simple Logit model of demand studied by Anderson et al. (1987, 1992) with standard CES direct utility function.

Example 2. ("Product Differentiation" Logit (GEV)) We now derive the indirect utility function for a representative consumer with demand equivalent to the "Product Differentiation" version of the aggregate discrete-continuous GEV Logit demand system (e.g., Bresnahan et al., 1997). Suppose aggregate demand (3.1) has the the aggregate discrete-continuous-choice "Product Differentiation Logit" demand functional form:

$$Q_j(\mathbf{p}, Y) = \frac{Y}{p_j} \frac{\left(\exp\left(\sigma \delta_j\right)\right)^{\frac{1}{\lambda_g}} \sum_g a_g \left(\sum_{k \in B_{gl}(j)} \left(\exp\left(\sigma \delta_k\right)\right)^{\frac{1}{\lambda_g}} \right)^{\lambda_g - 1}}{\sum_g a_g \left(\sum_{B_{gl} \in g} \left(\sum_{k \in B_{gl}} \left(\exp\left(\sigma \delta_k\right)\right)^{\frac{1}{\lambda_g}} \right)^{\lambda_g} \right)} \qquad \text{for } j = 1, ..., J,$$

where $B_{gl}(j)$ denotes the group to which product j belongs. The corresponding probability generating function $H(\boldsymbol{\delta})$ is then as follows:

$$H(\boldsymbol{\delta}) = \left\{ \sum_g a_g \left(H^g(\boldsymbol{\delta})\right)^{\sigma} \right\}^{\frac{1}{\sigma}} = \left\{ \sum_g a_g \left(\sum_{B_{gl} \in g} \left(\sum_{k \in B_{gl}} \left(\exp\left(\sigma \delta_k\right)\right)^{\frac{1}{\lambda_g}} \right)^{\lambda_g} \right) \right\}^{\frac{1}{\sigma}}, \tag{3.13}$$

where $H^g(\boldsymbol{\delta})$ is the same as in the Nested Logit and for each g, $a_g \in (0, 1]$ such that $\sum_g a_g = 1$, and $\lambda_g \in (0,1)$. From Proposition 1, the representative consumer's indirect utility function is

$$v(\mathbf{p}, Y) = Y \left\{ \sum_g a_g \left(\sum_{B_{gl} \in g} \left(\sum_{k \in B_{gl}} \left(\exp(\sigma \delta_k) \right)^{\frac{1}{\lambda_g}} \right)^{\lambda_g} \right) \right\}^{\frac{1}{\sigma}}. \tag{3.14}$$

This formulation allows for more flexible underlying substitution patterns at the individual consumer level when the nests are overlapping. At the same time, this formulation still generates an analytic formulation of the representative consumer's demand and indirect utility. We can easily conduct welfare analysis by substituting the formulation of $H(\boldsymbol{\delta})$ from (3.13) into (2) to obtain the compensating differential for a price change.

Example 3. (Random Coefficients Logit) As another illustrative example, we derive the indirect utility function for a representative consumer with demand equivalent to the aggregate discrete-continuous random coefficients Logit demand system. In the trade literature, allowing for random coefficients to offset the implications of the IIA property in a Logit have been shown to lead to markedly different measurements of the gains from product variety (e.g., Sheu, 2014). Unlike the Nested Logit and GEV specifications above, the random coefficients "mixed logit" does not require pre-classifying products into nests.

For the case of a discrete distribution of heterogeneity, we have:

$$Q_j(\mathbf{p}, Y) = \frac{Y}{p_j} \sum_{i=1}^{K} \frac{\exp\left(\delta_j + \nu_j^i\right)}{\sum_k \exp\left(\delta_k + \nu_k^i\right)} \lambda^i \qquad \text{for } j = 1, ..., J \tag{3.15}$$

where $\delta_j = \bar{\gamma}_j - \ln p_j$ and

$$\boldsymbol{\nu} := \begin{cases} \boldsymbol{\gamma}^1 - \bar{\boldsymbol{\gamma}}, & \text{with probability } \lambda^1 \\ \vdots & \\ \boldsymbol{\gamma}^K - \bar{\boldsymbol{\gamma}}, & \text{with probability } \lambda^K \end{cases}.$$

From the WDZ theorem, we know that:

$$\nabla_{\boldsymbol{\delta}} \mathcal{G}(\boldsymbol{\delta}) = (\pi_1, \pi_2, ..., \pi_J)$$

where $\pi_j = \sum_{i=1}^{K} \frac{\exp\left(\delta_j + \nu_j^i\right)}{\sum_k \exp\left(\delta_k + \nu_k^i\right)} \lambda^i, \ \forall j$ and $\mathcal{G}(\boldsymbol{\delta}) := \ln(H(\boldsymbol{\delta})) = \sum_i \ln\left(\sum_k \exp\left(\delta_k + \nu_k^i\right)\right) \lambda^i$. Hence, the formulation of $H(\boldsymbol{\delta})$ is as follows:

$$H(\boldsymbol{\delta}) = \exp\left(\sum_i \ln\left(\sum_k \exp\left(\delta_k + \nu_k^i\right) \right) \lambda^i \right). \tag{3.16}$$

From Proposition 1, the representative consumer has the following indirect utility function:

$$v(\mathbf{p}, Y) \quad = Y\left\{\prod_i \left(\sum_k \exp\left(\delta_k + \nu_k^i\right)\right)\lambda^i\right\}. \tag{3.17}$$

Equations (3.15) and (3.17) indicates that the richer underlying behavior of a heterogeneous population of consumers making discrete choices can be captured using a representative consumer model without losing analytic tractability. Once again, we can easily conduct welfare analysis by substituting the formulation of $H(\boldsymbol{\delta})$ from (3.15) into (2) to obtain the compensating differential for a price change.

For the case of an absolutely continuous distribution of heterogeneity, $F_{\boldsymbol{\nu}}(\boldsymbol{\nu})$, we have:

$$Q_j(\mathbf{p}, Y) = \frac{Y}{p_j}\int \frac{\exp(\delta_j + \nu_j)}{\sum_k \exp(\delta_k + \nu_k)} dF_{\boldsymbol{\nu}}(\boldsymbol{\nu}) \qquad \text{for } j = 1, ..., J$$

where $H(\boldsymbol{\delta})$ is as follows:

$$H(\boldsymbol{\delta}) = \exp\left(\int \ln\left(\sum_k \exp(\delta_k + \nu_k)\right) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu})\right).$$

The representative consumer's indirect utility function is:

$$v(\mathbf{p}, Y) \quad = \quad Y\left\{\exp\left[\int \ln \sum_k \exp(\delta_k + \nu_k)\, dF_{\boldsymbol{\nu}}(\boldsymbol{\nu})\right]\right\}. \tag{3.18}$$

However, this specification requires numerical integration and does not yield closed-form expressions for demand, indirect utility and $H(\boldsymbol{\delta})$.

Example 4. (Simple Logit with a Continuum of Products) We now consider the typical differentiated products model used in the macro and trade literatures which assumes a continuum of products distributed uniformly over the choice set Ω (or the "continuous Logit" (Ben-Akiva et al., 1985)). Although this example is technically not nested in Proposition 1, which only considers the case with a finite number of products, we show that the corresponding indirect utility function still has the form of (3.4). As always, a consumer chooses brand j if

$$\epsilon(j) - \epsilon(\omega) \geq \delta(\omega) - \delta(j), \quad \forall \omega \in \Omega.$$

The expected demand for brand j across the population of consumers with measure N is

$$Q_j(\mathbf{p}, Y) = \frac{Y}{p(j)} \frac{\exp(\sigma\theta(j))p(j)^{-\sigma}}{\int_{\omega\in\Omega} \exp(\sigma\theta(\omega))p(\omega)^{-\sigma} d\omega} \qquad \text{for } j \in \Omega. \tag{3.19}$$

Now consider the usual representative consumer model with CES utility defined over a continuum of brands and with income Ny. The utility function is given by

$$u=\left(\int_{\omega\in\Omega}\exp(\sigma\theta(\omega))^{\frac{\sigma}{\sigma+1}}Q(\omega)^{\frac{\sigma}{\sigma+1}}d\omega\right)^{\frac{\sigma+1}{\sigma}}.$$

It is straightforward to show that the corresponding demand function is

$$\begin{aligned}Q_j\left(\mathbf{p},Y\right) &= Y\exp(\theta(j))p(j)^{-\sigma-1}P^{-1}\\ &= \frac{Y}{p(j)}\frac{\exp(\sigma\theta(j))p(j)^{-\sigma}}{\int_{\omega\in\Omega}\exp(\sigma\theta(\omega))p(\omega)^{-\sigma}d\omega},\end{aligned} \tag{3.20}$$

where $P=\int_{\omega\in\Omega}\exp(\theta(\omega))p(\omega)^{-\sigma}d\omega$ is the Dixit-Stiglitz price index. The equivalence of the aggregate demand (3.19) and the representative consumer's demand (3.20) establishes the representative consumer formulation for a population of consumers making discrete choices from a continuum of product alternatives. It is also straightforward to show that the representative consumer's indirect utility function is:

$$v\left(\mathbf{p},Y\right) = Y\left(\int_{\omega\in\Omega}\exp\left(\delta(\omega)\right)d\omega\right)^{\frac{1}{\sigma}}$$

which has the same form as (3.4).

Example 5. (Multinomial Probit) We now explore the implementation of our results for a population of consumers making *Probit* discrete choices with correlated random utilities. Since a multinomial Probit demand system cannot be derived in closed form, our analysis herein will be numeric. We use parameter values for the trinomial Probit demand system for competing Chinese movie theaters in Dubé et al. (2017):[5]

$$\begin{array}{ll}u_1=\delta_1+\epsilon_1= & -0.0486-0.5474\log p_1+\epsilon_1\\ u_2=\delta_2+\epsilon_2= & -0.1460-0.5474\log p_2+\epsilon_2\\ u_0=\delta_0+\epsilon_0= & \epsilon_0\end{array},\qquad \Delta\boldsymbol{\epsilon}\equiv\begin{bmatrix}\epsilon_1-\epsilon_0\\ \epsilon_2-\epsilon_0\end{bmatrix}\sim\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1 & -0.1739\\ -0.1739 & 1.1745\end{bmatrix}\right) \tag{3.21}$$

where alternative $j=0$ is a no-purchase alternative.

We now illustrate our findings by computing the aggregate welfare change associated with a change in the movie theater price level from $\mathbf{p}^0=(\$75,\$75)$ to $\mathbf{p}^1=(\$20,\$20)$. In Dubé et al. (2017), the latter price represented the equilibrium prices at an off-peak hour of the day whereas the former prices represented the typical "regular" box office prices. We assume that the aggregate income level of the representative consumer is $Y^0=1,000$. We compute $G\left(\boldsymbol{\delta}\left(\mathbf{p}\right)\right)$ using Monte

[5] Dubé et al. (2017) look at demand between two competing movie theaters along with a "no purchase" outside alternative.

Carlo simulation to evaluate the integral over the Gaussian random utility terms:

$$\begin{aligned} H(\boldsymbol{\delta}(\mathbf{p})) &\equiv \exp(\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))) \\ &= \exp\left(\mathbb{E}_{\boldsymbol{\epsilon}}\left[\max_{j\in\{1,2,0\}}\{\delta_j + \Delta\epsilon_j\}\right]\right) \\ &\approx \exp\left(\frac{1}{N_s}\sum_{s=1}^{N_s}\left\{\max_{j\in\{1,2,0\}}\{\delta_j + \Delta\epsilon_j^s\}\right\}\right), \end{aligned}$$

where $\Delta\epsilon_0 = \epsilon_0 - \epsilon_0 = 0$ and $N_s = 5{,}000$ simulation draws. The compensating differential, CV, can then be calculated using Corollary 2. The simulation gives $H\left(\boldsymbol{\delta}\left(\mathbf{p}^0\right)\right) = 1.005$, $H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right) = 1.046$, and $CV = \$39$ as the compensating variation associated with the price decrease.

4 (Non)-Integrability of the Aggregate Pure-Discrete-Choice Demand System with Perfect Substitutes and Indivisibility

We now turn to another popular implementation of the discrete-choice model in empirical research that assumes indivisibility of the products. We show that this *pure discrete-choice model* is not integrable in the sense of Hurwicz and Uzawa (1971) as the budget balancedness condition no longer holds for the aggregate demand system over the entire domain of prices and income levels. Therefore, the representative consumer formulation from Proposition 1 no longer holds. Even under Nocke and Schutz (2017)'s stronger form of *quasi-linear* integrability, the pure discrete-choice model of demand is only integrable in the special case of a linear marginal utility of expenditure on other goods and services.

To obtain the usual pure discrete-choice model, we retain the assumption of perfect substitutes preferences, as in Section 2. Unlike the case of divisible goods, we now need to include a numeraire good comprising expenditure on other goods and services to satisfy the *adding-up* condition with indivisibility of the quantity of the products consumed.

The consumer's choice-specific values have the well-known additive random utility form

$$\begin{aligned} v_j &= \theta_j + g(y - p_j) + \epsilon_j \qquad j = 1, 2, ..., J \\ &= \delta_j + \epsilon_j \end{aligned} \tag{4.1}$$

where we normalize $\delta_J \equiv 0$. The consumer's probability of choosing alternative j is

$$\begin{aligned} \pi_j(\boldsymbol{\delta}(\mathbf{p}), y) &\equiv \Pr(\epsilon_j - \epsilon_k \geq \delta_k - \delta_j, \forall k) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\delta_j-\delta_1+\tilde{\epsilon}}\cdots\int_{-\infty}^{\delta_j-\delta_{j-1}+\tilde{\epsilon}}\int_{-\infty}^{\delta_j-\delta_{j+1}+\tilde{\epsilon}}\cdots\int_{-\infty}^{\delta_j-\delta_J+\tilde{\epsilon}} \\ &\quad f_{\boldsymbol{\epsilon}}(\epsilon_1, ..., \epsilon_{j-1}, \tilde{\epsilon}, \epsilon_{j+1}, ..., \epsilon_J)\, d\epsilon_J ... d\epsilon_{j+1} d\epsilon_{j-1} ... d\epsilon_1 d\tilde{\epsilon}. \end{aligned} \tag{4.2}$$

The consumer has the following *expected* demand for the numeraire good:

$$q_z(\mathbf{p}, y) \equiv \mathbb{E}_{\boldsymbol{\epsilon}}[q_z^*|\mathbf{p}] = y - \sum_{j=1}^{J} \pi_j(\boldsymbol{\delta}(\mathbf{p}), y) p_j. \tag{4.3}$$

If there are N consumers in the market with aggregate income $Y = Ny$, then the expected aggregate demand system is given by:

$$\begin{aligned} Q_j(\mathbf{p}, Y) &= N\pi_j \qquad \text{for } j = 1, ..., J \\ Q_z(\mathbf{p}, Y) &= N\left(y - \sum_{j=1}^{J} \pi_j p_j\right) \end{aligned} \tag{4.4}$$

where the brand-choice probabilities $\{\pi_j\}_{j=1}^{J}$ are given by (4.2).

The Hurwicz and Uzawa integrability of the demand system (4.4) requires $Q_j(\mathbf{p}, Y) \geq 0$ and $Q_z(\mathbf{p}, Y) \geq 0$ over the entire price-income space, $\mathbb{R}_{++}^{J} \times \mathbb{R}_{+}$. To see that demand is not defined over the entire domain, consider the case where $Y < \min\{p_j, j = 1, ..., J\}$ and, hence, $Q_z(\mathbf{p}, Y) < 0$. Indeed, empirical applications of the pure discrete-choice model must restrict the domain to ensure $Y \geq \max\{p_j, j = 1, ..., J\}$ to ensure the empirical demand model is well defined. We state the non-integrability in the Hurwicz and Uzawa (1971) sense of the demand system (4.4) in the following proposition.[6]

Proposition 2. *(Non-Hurwicz and Uzawa-integrability of Aggregate Pure Discrete-Choice Demand Model) The aggregate pure discrete-choice demand system (4.4) does not satisfy the budget balancedness condition and, therefore, is not integrable in the sense of Hurwicz and Uzawa (1971).*

Nocke and Schutz (2017) propose a stronger form of "quasi-linear integrability" that is defined only on the restricted domain of $(\mathbf{p}, Y)$ where $Q_j(\mathbf{p}, Y) \geq 0$ and $Q_z(\mathbf{p}, Y) \geq 0$. The following proposition establishes that the demand system (4.4) is quasi-linear integrable only in the special case of linear utility for the numeraire: $g(Y - p_j) = \alpha(Y - p_j)$ where $\alpha > 0$.

Proposition 3. *(Quasi-linear Integrability of Aggregate Pure Discrete-Choice Demand System When the Marginal Utility of Income is Linear in the Numeraire) Assume $g(Y - p_j) = \alpha(Y - p_j)$ for some constant $\alpha > 0$ in the aggregate pure discrete-choice demand system given by (4.4). Then, the demand system is quasi-linearly integrable.*

Proof. See Appendix C.1. □

[6] The non-integrability result also holds if we add a $J+1$ "no purchase" (e.g., home production) option in the choice set with $p_{J+1} = 0$ and $\mathbb{E}(v_{J+1}) = 0$, as is often the case in empirical settings. Demand for the numeraire $Q_z(\mathbf{p}, y) = N\left(y - \sum_{j=1}^{J+1} \pi_j p_j\right) = N\left(y - \sum_{j=1}^{J} \pi_j p_j\right)$ can be negative for a small y whenever $\pi_{J+1} \neq 1$, leading to the failure of Hurwicz-Uzawa integrability. For example, $\pi_{J+1} \neq 1$ whenever $g(y - p_j)$ is finite and the random utility is absolutely-continuously distributed with full-support.

The proof relies on the fact that linearity ensures that income, Y, "differences out" of the brand choice probabilities (4.4), thereby satisfying Theorem 1 (iv) for quasi-linear integrability in Nocke and Schutz (2017). It is unlikely that quasi-linear integrability would hold under more general, non-linear forms of $g(Y - p_j)$ since income would no longer difference out of the brand choice probabilities.

5 Conclusions

We have shown that the expected aggregate demand corresponding to a broad class of ARUMs is integrable and has a normative representative consumer formulation with a specific form of preferences. We provide several examples of representative consumer formulations that can be rationalized as populations of consumers with discrete-continuous demands generating flexible substitution patterns between products. However, we also show that in the case of indivisible quantities, or "pure discrete choice," integrability fails. Even a stronger form of quasi-linear integrability only holds under the special case of quasi-linearity in a numeraire good.

References

Anderson, S. P., A. de Palma, and J.-F. Thisse (1987): "The CES is a discrete choice model?" *Economics Letters*, 24, 139–140.

——— (1992): *Discrete choice theory of product differentiation*, Cambridge, Massachusetts: The MIT Press.

Archibald, G., B. Eaton, and R. Lipsey (1986): *New Developments in the Analysis of Market Structure*, Published by Macmillan for the International Economics Association, chap. Address Models of Value, 3–47.

B. Curtis Eaton, R. G. L. (1989): *Handbook of Industrial Organization*, North-Holland, vol. 1, chap. Product differentiation, 723–768.

Beck, G. W. and S. M. Lein (2020): "Price elasticities and demand-side real rigidities in micro data and in macro models," *Journal of Monatery Economics*, 115, 200–212.

Ben-Akiva, M., N. Litinas, and K. Tsunokawa (1985): "Continuous spatial choice: The continuous logit model and distributions of trips and urban densities," *Transportation Research A*, 19A, 119–154.

Berger, D., K. Herkenhoff, and S. Mongey (Forthcoming): "Labor market power," *American Economic Review*.

Bernini, M. and C. Tomasi (2015): "Exchange rate pass-through and product heterogeneity: Does quality matter on the import side?" *European Economic Review*, 77, 117–138.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, 63, 841–890.

Bhat, C. R. (2005): "A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions," *Transportation Research Part B*, 39, 679–707.

——— (2008): "The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions," *Transportation Research Part B*, 42, 274–303.

Blanchard, O. J. and N. Kiyotaki (1987): "Monopolistic Competition and the Effects of Aggregate Demand," *The American Economic Review*, 77, 647–666.

Bresnahan, T. F., S. Stern, and M. Trajtenberg (1997): "Market segmentation and the sources of rents from innovation: Personal computers in the late 1980s," *RAND Journal of Economics*, 28, S17–S44.

Card, D. (2022): "Who Set Your Wage," *UC Berkeley Working Paper.*

Chiang, J. (1991): "A Simultaneous Approach to the Whether, What and How Much to Buy Questions," *Marketing Science*, 10, 297–315.

Chintagunta, P. K. (1993): "Investigating Purchase Incidence, Brand Choice and Purchase Quantiy Decisions of Households," *Marketing Science*, 12, 184–208.

Chiong, K. X., A. Galichon, and M. Shum (2016): "Duality in dynamic discrete-choice models," *Quantitative Economics*, 7, 83–115.

Chiong, K. X. and M. Shum (2019): "Random projection estimation of discrete-choice models with large choice sets," *Management Science*, 65, 256–271.

Crin'o, R. and L. Ogliari (2017): "Financial imperfections, product quality, and international trade," *Journal of International Economics*, 104, 63–84.

Daly, A. and S. Zachary (1978): "Improved multiple choice models," *Determinants of Travel Choice*, Identifying and Measuring the Determinants of Mode Choice.

Deaton, A. and J. Muellbauer (1980): *Economics and Consumer Behavior*, Cambridge University Press.

Dhingra, S. and J. Morrow (2019): "Monopolistic competition and optimum product diversity under firm heterogeneity," *Journal of Political Economy*, 127, 196–232.

Dixit, A. K. and J. E. Stiglitz (1977): "Monopolistic competition and optimum product diversity," *American Economic Review*, 67, 297–308.

Dow, J. and S. R. da Costa Werlang (1988): "The Consistency of Welfare Judgments with a Representative Consumer," *Journal of Economic Theory*, 44, 269–280.

Dubé, J.-P. (2004): "Multiple discreteness and product differentiation: Demand for carbonated soft drinks," *Marketing Science*, 23, 66–81.

——— (2019): *Handbook of the Economics of Marketing, Vol. 1*, Elsevier, chap. Microeconometric modelsof consumer demand, 1–68.

Dubé, J.-P., Z. Fang, N. Fong, and X. Luo (2017): "Competitive price targeting with smartphone coupons," *Marketing Science*, 36, 944–975.

Dubin, J. A. and D. McFadden (1984): "An econometric analysis of residential electric appliance holdings and consumption," *Econometrica*, 52, 345–362.

Fajgelbaum, P., G. M. Grossman, and E. Helpman (2011): "Income distribution, product quality, and international trade," *Journal of Political Economy*, 119, 721–765.

Fosgerau, M., D. McFadden, and M. Bierlaire (2013): "Choice probability generating functions," *Journal of Choice Modelling*, 8, 1–18.

Goldberg, P. K. (1995): "Product differentiation and oligopoly in international markets: The case of the U.S. automobile industry," *Econometrica*, 63, 891–951.

Hanemann, W. M. (1984): "Discrete/continuous models of consumer demand," *Econometrica*, 52, 541–561.

Helpman, E. (2011): *Understanding Global Trade*, Harvard University Press.

Hendel, I. (1999): "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *Review of Economic Studies*, 66, 423–446.

Hosoya, Y. (2017): "The relationship between revealed preference and the Slutsky matrix," *Journal of Mathematical Economics*, 70, 127–146.

Hurwicz, L. and H. Uzawa (1971): *On the integrability of demand functions*, Harcourt Brace Jovanovich, chap. 6, 114–148, Preferences, Utility, and Demand.

Jackson, M. O. (1986): "Integration of demand and continuous utility functions," *Journal of Economic Theory*, 38, 298–312.

Judd, K. L. (1985): "Marginal Excess Burden in a Dynamic Economy," *Economic Letters*, 18, 213–216.

Khandelwal, A. (2010): "The long and short (of) quality ladders," *Review of Economic Studies*, 77, 1450–1476.

Kim, J., G. M. Allenby, and P. E. Rossi (2002): "Modeling consumer demand for variety," *Marketing Science*, 21, 229–250.

Krugman, P. (1995): *Handbook of International Economics*, North-Holland, vol. 3, chap. Increasing returns, imperfect competition and the positive theory of international trade, 1243–1277.

Lancaster, K. J. (1966): "A new approach to consumer theory," *Journal of Political Economy*, 74, 132–157.

Lee, L.-F. and M. M. Pitt (1986): "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach," *Econometrica*, 5, 123–1242.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory*, Oxford University Press.

McFadden, D. (1981): *Econometric models of probabilistic choice*, The MIT Press, chap. 5, 198–272, Structural analysis of discrete data with econometric applications.

Melitz, M. (2008): *International trade and heterogeneous firms*, Palgrave Macmillan.

Melitz, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71, 1695–1725.

Nocke, V. and N. Schutz (2017): "Quasi-linear integrability," *Journal of Economic Theory*, 169, 603–628.

——— (2018): "Multiproduct-firm oligopoly: An aggregative games approach," *Econometrica*, 86, 523–557.

Redding, S. J. and D. E. Weinstein (2020): "Measuring aggregate price indices with taste shocks: Theory and evidence for CES preferences," *Quarterly Journal of Economics*, 135, 503–560.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton University Press.

Romer, P. M. (1987): "Growth Based on Increasing Returns Due to Specialization," *American Economic Association*, 77, 56–62.

Rust, J. (1994): *Structural estimation of Markov decision process*, North-Holland Publishing Company, vol. 4 of *Handbook of Econometrics*, chap. 51.

Samuelson, P. A. (1956): "Social indifference curves," *Quarterly Journal of Economics*, 70, 1–22.

Sheu, G. (2014): "Price, Quality, andVariety: Measuring the Gains from Trade inDierentiated Products," *American Economic Journal: Applied Economics*, 6, 66–89.

Small, K. A. and H. S. Rosen (1981): "Applied welfare economics with discrete choice models," *Econometrica*, 49, 105–130.

Spence, M. (1976): "Product selection, fixed costs, and monopolistic competition," *Review of Economic Studies*, 43, 217–235.

Varian, H. R. (1992): *Microeconomic Analysis*, Norton, 3rd ed. ed.

Verboven, F. (1996): "The nested logit model and representative consumer theory," *Economics Letters*, 50, 57–73.

Wales, T. and A. Woodland (1983): "Estimation of consumer demand systems with binding non-negativity constraints," *Journal of Econometrics*, 21, 263–285.

Williams, H. (1977): "On the formation of travel demand models and economic evaluation measures of user benefit," *Environment and Planning A: Economy and Space*, 9, 285–344.

Zhelobodko, E., S. Kokovin, M. Parenti, and J.-F. Thisse (2012): "Monopolistic competition: Beyond the constant elasticity of substitution," *Econometrica*, 80, 2765–2784.

Appendix

A Appendix of Section 2: Proof of Theorem 1 and Corollary 1

This Appendix proves that the demand system (2.4) satisfies the four necessary and sufficient conditions for Hurwicz and Uzawa (1971) integrability and derives the exact form of the corresponding indirect utility function. We break this proof into three sub-sections. We prove the budget balancedness and differentiability conditions in Section A.1, and the Slutsky Symmetry condition in Section A.2. In Section A.3, we first characterize the expenditure function and indirect utility functions corresponding to the demand system (2.4) and use them to establish the negative semi-definiteness condition, which completes the proof of integrability.

Throughout the proof, we omit the argument of $H(\cdot), H_j(\cdot), H_{jk}(\cdot)$, and $\boldsymbol{\delta}(\cdot)$ when it is clear from the context for simplicity in the notation.

A.1 Proof of the Differentiability and Budget Balancedness

To establish differentiability, we substitute the discrete-choice probabilities (2.7) into (2.4):

$$q_j(\mathbf{p}, y) = \frac{y}{p_j} \frac{H_j}{H} \qquad \text{for } j \in \{1, 2, ..., J\}. \tag{A.1}$$

Differentiability follows from the smoothness assumptions for $H(\boldsymbol{\delta})$ and $\delta_j(p_j, \cdot, \cdot)$. The boundedness of the partial derivatives of $\mathbf{q}(\mathbf{p}, Y)$ follows from the assumption that $H(\cdot) > 0$, the mixed partial derivatives $H_{jk}(\cdot)$ exist, and $H_{jk}(\cdot)$ is continuous on $\mathbb{R}^J \; \forall \; j, k$.

To show the budget balancedness, we show that for any $y \geq 0$,

$$y = \sum_{k=1}^{J} p_k q_k(\mathbf{p}, y).$$

The right-hand side can be written as

$$\sum_{k=1}^{J} p_k q_k(\mathbf{p}, y) = y \sum_{k=1}^{J} \pi_k(\boldsymbol{\delta}(\mathbf{p})) = y,$$

where equality to y follows from the fact that $\sum_{k=1}^{J} \pi_k(\mathbf{p}) = 1$.

A.2 Proof of the Slutsky Symmetry Condition

We now establish the Slutsky symmetry of the individual demand system (2.4). Recall $\delta_j = \theta_j - \ln(p_j)$, which is continuously differentiable in p_j. For $k \neq j$,

$$\begin{aligned}\frac{\partial q_k(\mathbf{p}, y)}{\partial p_j} =& \frac{y}{p_k}\left\{\frac{\delta_j'(p_j) H_{jk}}{H} - \frac{\delta_j'(p_j) H_j H_k}{H^2}\right\} \\ \frac{\partial q_k(\mathbf{p}, y)}{\partial y} =& \frac{1}{p_k}\frac{H_k}{H}\end{aligned}$$

The Slutsky symmetry condition is:

$$\frac{\partial q_k(\mathbf{p}, y)}{\partial p_j} + q_j(\mathbf{p}, y)\frac{\partial q_k(\mathbf{p}, y)}{\partial y} = \frac{\partial q_j(\mathbf{p}, y)}{\partial p_k} + q_k(\mathbf{p}, y)\frac{\partial q_j(\mathbf{p}, y)}{\partial y} \tag{A.2}$$

The left-hand side can be re-written as follows:

$$\begin{aligned}&\frac{y}{p_k}\left\{\frac{\delta_j'(p_j) H_{jk}}{H} - \frac{\delta_j'(p_j) H_j H_k}{H^2}\right\} + \frac{y}{p_j p_k}\left\{\frac{H_j H_k}{H^2}\right\} \\ =& \frac{1}{H^2}\left\{\frac{y}{p_k}\delta_j'(p_j) H H_{jk} - \frac{y}{p_k}\delta_j'(p_j) H_j H_k + \frac{y}{p_j p_k} H_j H_k\right\}.\end{aligned}$$

The symmetry condition becomes:

$$\begin{aligned}&\frac{1}{p_k}\left[\delta_j'(p_j) H H_{jk} - \delta_j'(p_j) H_j H_k\right] \\ =& \frac{1}{p_j}\left[\delta_k'(p_k) H H_{jk} - \delta_k'(p_k) H_j H_k\right]\end{aligned}$$

which further simplifies to:

$$\frac{\delta_j'(p_j)}{p_k}\left[H H_{jk} - H_j H_k\right] = \frac{\delta_k'(p_k)}{p_j}\left[H H_{jk} - H_j H_k\right].$$

The Slutsky symmetry condition (A.2) can therefore be characterized as a system of ordinary differential equations for each j, k pair:

$$p_j \delta_j'(p_j) = p_k \delta_k'(p_k).$$

This yields the form of the solution:

$$\delta_j(p_j) = -\eta \log p_j + c_j, \forall j \tag{A.3}$$

for $\eta \in \mathbb{R}$ and some $c_j \in \mathbb{R}$, which is sufficient for Slutsky symmetry. Condition (A.3) is also necessary for Slutsky symmetry due to the uniqueness of the solution of an ordinary differential equation (up to initial conditions determined by c_j). This condition holds mechanically for the ARUM where $\eta = 1$ and $c_j = \theta_j$.

A.3 Proof of the Negative Semidefiniteness of the Slutsky Matrix and the Integrability of the Expected Discrete-Continuous-Choice Demand System

In this section, we first characterize the expenditure function and indirect utility function corresponding to the individual demand system (2.4). We then prove the negative semidefiniteness of the Slutsky matrix which both completes our proof of integrability and characterizes the corresponding preferences. Our proof consists of showing that $\eta > 0$ in (A.3) is sufficient for the negative semidefiniteness of the associated Slutsky matrix, which again holds mechanically for the ARUM where $\eta = 1$.

The strategy for our proof is *guess-and-verify*. We start with the *guess* that the indirect utility function corresponding to the ARUM individual demand system (2.4) has the form:

$$v(\mathbf{p}, y) = yH(\boldsymbol{\delta}(\boldsymbol{p})) \tag{A.4}$$

with corresponding expenditure function

$$e(\mathbf{p}, u) = \frac{u}{H(\boldsymbol{\delta}(\mathbf{p}))}, \tag{A.5}$$

where the function $H(\cdot)$ is defined as in Section 2.1 and $u \geq 0$ WLOG.[7] Our intuition for this guess comes from the following observation. The log-indirect utility function for the consumer problem (2.1) where $\eta = 1$ is:

$$\begin{aligned}\log\{\tilde{v}(\mathbf{p}, y, \boldsymbol{\epsilon})\} &= \log\left\{\max_{\mathbf{q}\in\mathbb{R}_+^J} \sum_{j=1}^{J} q_j \psi_j(\epsilon_j) \text{ s.t. } \mathbf{q}'\mathbf{p} \leq y\right\} \\ &= \max_{j\in\{1,2,\ldots,J\}} \{\log y + \delta_j + \epsilon_j\} \end{aligned} \tag{A.6}$$

because of the corner solution in which the chosen brand k has demand: $q_k = \frac{y}{p_k}$. Taking the exponentiated form of the expectation of (A.6) with respect to the random utility, $\boldsymbol{\epsilon}$, gives:

$$\exp(\mathbb{E}_{\boldsymbol{\epsilon}}[\log \tilde{v}(\mathbf{p}, y, \boldsymbol{\epsilon})]) = \exp(\log y + \mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))) = yH(\boldsymbol{\delta}(\mathbf{p})). \tag{A.7}$$

We verify that the candidate $e(\mathbf{p}, u)$ is a valid expenditure function associated with the demand system (2.4) by showing that it satisfies the two following sufficient conditions (for sufficiency, see, e.g., Mas-Colell et al., 1995, p.80): (i) $e(\mathbf{p}, u)$ is concave, and, (ii) $\nabla_{\mathbf{p}} e(\mathbf{p}, u)$ evaluated at $e(\mathbf{p}, u) = y$ generates the individual demand system (2.4), and therefore, the Hessian of $e(\mathbf{p}, u)$ lines up with

[7] See, e.g., Jackson (1986) which imposed similar restrictions.

the Slutsky matrix we examined in Section A.2.

Lemma 1. *The function (A.5) is concave in* $\mathbf{p}$ *and therefore, the associated Hessian is negative semidefinite.*

Proof. We can re-write the candidate expenditure function (A.5) as:

$$e(\mathbf{p}, u) = u \exp\left(-\mathcal{G}\left(\boldsymbol{\delta}\left(\mathbf{p}\right)\right)\right),$$

which is nonnegative since $u \geq 0$.

We first establish the convexity of $H(\boldsymbol{\delta}(\mathbf{p})) = \exp(\mathcal{G}(\boldsymbol{\delta}(\mathbf{p})))$ in $\mathbf{p}$. The function $\delta_j(p_j) = -\eta \log p_j + \theta_j$ is convex in p_j if and only if $\eta \geq 0$. The surplus function, $\mathcal{G}(\boldsymbol{\delta})$, is convex in $\boldsymbol{\delta}$ because the max function preserves convexity (Rockafellar, 1970, Theorem 5.5), and a linear combination of convex functions, $\delta_j(p_j) + \epsilon_j$ (indexed by the realization of $\boldsymbol{\epsilon}$) with nonnegative weights associated with the operator $\mathbb{E}_{\boldsymbol{\epsilon}}$, is convex (Rockafellar, 1970, p.33). Therefore, $\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))$ is convex in $\mathbf{p}$ and the composite function $\exp(\mathcal{G}(\boldsymbol{\delta}(\mathbf{p})))$ is also convex in $\mathbf{p}$ because $\exp(\cdot) : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing (Rockafellar, 1970, Theorem 5.1).

Since $H(\boldsymbol{\delta}(\mathbf{p}))$ is non-negative and convex in $\mathbf{p}$, $\{H(\boldsymbol{\delta}(\mathbf{p}))\}^{-1} = \exp(-\mathcal{G}(\boldsymbol{\delta}(\mathbf{p})))$ is concave (Rockafellar, 1970, p.32). It follows that $e(\mathbf{p}, u)$ is concave in $\mathbf{p}$. The concavity also establishes the negative semidefiniteness of the Slutsky matrix generated from (A.5). □

We now verify that $e(\mathbf{p}, u)$ is the expenditure function associated with the individual demand system (2.4).

Lemma 2. *Let* $\delta_j = -\log p_j + \theta_j$*. For the expenditure function (A.5),*

$$\nabla_{\mathbf{p}} e(\mathbf{p}, u) \big|_{e(\mathbf{p},u)=y} = \left(\frac{y}{p_1}\pi_1, ..., \frac{y}{p_J}\pi_J\right),$$

where u *is the utility level that satisfies* $e(\mathbf{p}, u) = y$*.*

Proof. By Shephard's lemma and the duality of demand, we obtain:

$$\begin{aligned} \frac{\partial \left\{ \frac{u}{H(\boldsymbol{\delta}(\mathbf{p}))} \right\}}{\partial p_j} &= \frac{u H_j(\boldsymbol{\delta}(\mathbf{p}))}{p_j H(\boldsymbol{\delta}(\mathbf{p}))^2}, \; j = 1, ..., J \\ &= \frac{u \pi_j}{p_j H(\boldsymbol{\delta}(\mathbf{p}))}. \end{aligned} \tag{A.8}$$

Evaluating (A.8) at $y = e(\mathbf{p}, u) \equiv \frac{u}{H(\boldsymbol{\delta}(\mathbf{p}))}$ gives

$$\frac{\partial \left\{ \frac{u}{H(\boldsymbol{\delta}(\mathbf{p}))} \right\}}{\partial p_j} \Bigg|_{u = yH(\boldsymbol{\delta}(\mathbf{p}))} = \frac{y}{p_j}\pi_j,$$

which is the ARUM discrete-continuous-choice demand system (2.4). This result also confirms that the Slutsky matrix in Section A.2 is the Hessian of $e(\mathbf{p},u)$. $\square$

The non-negativity of $H(\boldsymbol{\delta}(\mathbf{p}))$ and the unboundedness of income, y, ensures that the corresponding indirect utility function, $v(\mathbf{p},y)=yH(\boldsymbol{\delta}(\mathbf{p}))$, has the range $\mathbb{R}_+$.

B Proofs of Section 3

B.1 Proof of Proposition 1

The individual's indirect utility function

$$v^i\left(\mathbf{p},y^i\right)=y^iH(\boldsymbol{\delta}(\mathbf{p})),$$

has the Gorman form. Therefore, the positive representative consumer with income level $Y=\sum_{i=1}^N y^i$ has the indirect utility function

$$v(\mathbf{p},Y)=YH(\boldsymbol{\delta}(\mathbf{p})),$$

and demand $\mathbf{q}(\mathbf{p},Y)=\sum_{i=1}^N \mathbf{q}^i\left(\mathbf{p},y^i\right)$ when Roy's identity is applied, for any $\mathbf{p}\in\mathbb{R}_+^J$ and for any allocation of income $\left\{y^1,y^2,...,y^N\right\}$ (see, e.g., Varian, 1992, Section 9.4). The Gorman form also ensures that the representative consumer is normative robust to any increasing, continuously differentiable, and concave social welfare function $\mathcal{W}(\cdot)$ (see, e.g., Mas-Colell et al., 1995, pp.119-120.).

B.2 Proof of Corollary 2

The representative consumer's compensating variation is characterized by $v\left(\mathbf{p}^0,Y^0\right)=v\left(\mathbf{p}^1,Y^0-CV\right)$. Substituting $v\left(\mathbf{p}^0,Y^0\right)=Y^0H\left(\boldsymbol{\delta}\left(\mathbf{p}^0\right)\right)$ and $v\left(\mathbf{p}^1,Y^0-CV\right)=\left(Y^0-CV\right)H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right)$ back and equating gives

$$Y^0H\left(\boldsymbol{\delta}\left(\mathbf{p}^0\right)\right)=\left(Y^0-CV\right)H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right).$$

Rearranging gives

$$CV=Y^0\frac{H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right)-H\left(\boldsymbol{\delta}\left(\mathbf{p}^0\right)\right)}{H\left(\boldsymbol{\delta}\left(\mathbf{p}^1\right)\right)}.$$

B.3 Extension to Including the Numeraire

The remainder of the this subsection consists of extending the results above to include the numeraire in a manner similar to Anderson et al. (1992, Section 3.7) and Verboven (1996, Section 4). Specifically, we use the Cobb-Douglas bivariate utility over the numeraire and the products group.

Let $Y = \sum_{i=1}^{N} y^i$ be the representative consumer's total income. We have shown the existence of the direct utility function $u(Q_1, ..., Q_J)$ such that the observed demand system $\mathbf{Q}(\mathbf{p}, Y)$ solves the usual budget-constrained utility maximization problem:

$$\max_{(Q_1,...,Q_J)} u(Q_1, ..., Q_J) \qquad \text{s.t.} \qquad \sum_{1 \le j \le J} p_j Q_j = Y.$$

We now modify the utility function $u(Q_1, ..., Q_J)$ as $u(Q_1, ..., Q_J)Q_z^{\alpha}$. The separability of the Cobb-Douglas utility $u(Q_1, ..., Q_J)Q_z^{\alpha}$ between the consumption of the products, $(Q_1, ..., Q_J)$, and of the numeraire, Q_z, ensures that the representative consumer always spends $Y/(1+\alpha)$ and $\alpha Y/(1+\alpha)$ on the inside goods and the numeraire, respectively. Therefore,

$$(Q_1, ..., Q_J; Q_z) := \left(Q_1, ..., Q_J; Y - \sum_{j=1}^{J} p_j Q_j \right)$$

solves

$$\max_{(Q_1,...,Q_J,Q_z)} u(Q_1, ..., Q_J) Q_z^{\alpha} \qquad \text{s.t.} \qquad Q_z + \sum_{1 \le j \le J} p_j Q_j = Y.$$

It is straightforward to verify that the resulting demand system has the following form:

$$\begin{aligned} Q_j(\mathbf{p}, Y) &= \left(\frac{1}{1+\alpha} \right) \frac{Y}{p_j} \frac{H_j}{H} \qquad \text{for } j \in \{1, 2, ..., J\} \\ Q_z(\mathbf{p}, Y) &= \left(\frac{\alpha}{1+\alpha} \right) Y. \end{aligned} \tag{B.1}$$

C Proofs of Section 4

C.1 Proof of Proposition 3

Nocke and Schutz (2017) define quasi-linear integrability as follows:

Definition. (Quasi-linear Integrability) A demand system $\mathbf{Q}(\mathbf{p}, Y)$ is quasi-linearly integrable if there exist a set $\mathcal{X} \subset \mathbb{R}^J$ and a function $u : \mathcal{X} \to \mathbb{R}$ such that for every $(\mathbf{p}, Y) \in \mathbb{R}_{++}^J \times \mathbb{R}_+$ where $Q_z(\mathbf{p}, Y) \ge 0$, $\mathbf{Q}(\mathbf{p}, Y)$ is the unique solution of

$$\max_{(\mathbf{Q}, Q_z)} (Q_z + u(\mathbf{Q})) \qquad s.t. \qquad Q_z + \mathbf{p}'\mathbf{Q} \le Y, Q_z \ge 0, \mathbf{q} \in \mathcal{X}.$$

Theorem 1 of Nocke and Schutz (2017) shows that a demand system is quasi-linear integrable if and only if the substitution matrix $\left(\frac{\partial Q_j(\mathbf{p}, Y)}{\partial p_k} \right)_{1 \le j,k \le J}$ is symmetric and negative semidefinite for every $\mathbf{p} \in \mathbb{R}_{++}^J$ and $Y \in \mathbb{R}_+$ for which the demand is well-defined. In our case, we need non-negative consumption of the numeraire: $Q_z(\mathbf{p}, Y) \equiv Y - \sum_{j=1}^{J} p_j Q_j(\mathbf{p}, Y) \ge 0$. We now show that the expected aggregate pure discrete-choice demand system (4.4) satisfies this condition when $g(Y - p_j) = \alpha (Y - p_j)$.

Consider the function $N\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))$. The substitution matrix is the Hessian of $N\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))$ with respect to $\mathbf{p}$. The symmetry of the substitution matrix follows from the assumed smoothness of $H(\cdot)$ function.

The negative semidefiniteness follows from the following argument. We claim the concavity of $\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))$ in $\mathbf{p}$, which would imply the negative semidefiniteness of the substitution matrix $\left(\frac{\partial Q_j(\mathbf{p},Y)}{\partial p_k}\right)_{1\leq j,k\leq J}$. Recall the definition of $\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))$ under the assumed form of δ_j with $\alpha > 0$:

$$\mathcal{G}(\boldsymbol{\delta}(\mathbf{p})) = \mathbb{E}\left[\max_{1\leq j\leq J}\{-\alpha p_j + \theta_j + \epsilon_j\}\right].$$

Dividing the above equality by $-\alpha$ gives

$$-\frac{1}{\alpha}\mathcal{G}(\boldsymbol{\delta}(\mathbf{p})) = \mathbb{E}\left[\max_{1\leq j\leq J}\left\{p_j - \frac{1}{\alpha}(\theta_j + \epsilon_j)\right\}\right].$$

Because we can consider the term $-\frac{1}{\alpha}(\theta_j + \epsilon_{i,j})$ as a location and scale transform of the random utility term $\boldsymbol{\epsilon}$ supported on the entire $\mathbb{R}^J$, $-\frac{1}{\alpha}\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))$ is convex in $\mathbf{p}$. In turn, $\mathcal{G}(\boldsymbol{\delta}(\mathbf{p})) = (-\alpha)\cdot\left(-\frac{1}{\alpha}\mathcal{G}(\boldsymbol{\delta}(\mathbf{p}))\right)$, is concave in $\mathbf{p}$ because $-\alpha < 0$.