IFN Working Paper No. 1430, 2022

Nonparametric Analysis of the Mixed-Demand Model

Per Hjertstrand

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Nonparametric Analysis of the Mixed-Demand Model

Per Hjertstrand*

May 10, 2022

Abstract

The mixed-demand model allows for very flexible specification of what should be considered endogenous and exogenous in demand system estimation. This paper introduces a revealed preference framework to analyze the mixed-demand model. The proposed methods can be used to test whether observed data (with measurement errors) are consistent with the mixed-demand model and calculate goodness-of-fit measures. The framework is purely non-parametric in the sense that it does not require any functional form assumptions on the direct or indirect utility functions. The framework is applied to demand data for food and provides the first non-parametric empirical analysis of the mixed-demand model.

Key words: demand systems, measurement errors, mixed-demand, non-parametric, revealed preference.

1 Introduction

The conventional specification of empirical demand models treats quantities as response variables and prices and income as predetermined variables (or predictors) in the system of demand equations. It is usually stipulated that the behavioral implications of these so-called direct demand models hold at the aggregate or market-level, i.e., over a set of collective consumers (Deaton and Muellbauer 1980b). An implication of predetermined prices and income is that, if the data are assumed to be the outcome of a market equilibrium model, then supply functions are perfectly elastic, meaning that demand adjust to clear the market. Of course, this assumption is valid in some markets (e.g., for tradeable goods in small open economies), but it is equally as easy to find examples where it is an inaccurate assumption to describe market characteristics.

On the other hand, inverse (or indirect) demand models where (expenditure-normalized) prices act as response variables are particularly useful in markets where supply functions are perfectly inelastic (e.g., for agricultural and natural resource commodities; See McLaren and Wong (2005) and Barten and Bettendorf (1989)). In such markets, (expenditure-normalized) prices adjust to clear the market.

*Research Institute of Industrial Economics (IFN). P.O. Box 55665, SE-102 15 Stockholm, Sweden. E-mail: Per.Hjertstrand@ifn.se. I acknowledge financial support from Torsten Söderbergs stiftelse. I thank Henrik Hällerfors for research assistance.

However, in practice, the demand of goods from different sub-markets are often modelled simultaneously, and for some of these goods it may be reasonable to treat prices and income as predetermined, while for the remaining goods it may be more reasonable to treat quantities as predetermined (at the market level). Hence, some quantities and some (expenditure-normalized) prices adjust to clear the market. Samuelson (1965) introduced a rich class of demand models, called mixed-demand models, that is able to exactly account for this type of market behavior; See also Chavas (1984) for a detailed theoretical analysis of this class. Mixed-demand models have obvious econometric appeal for the purpose of estimating demand behavior at the market level because they are able to cover a spectrum of possibilities between the two opposite cases of direct and inverse (indirect) demand estimation. As such, empirical applications of mixed-demand models have been used in a wide variety of fields such as agricultural economics (e.g., Moschini and Rizzi 2006, 2007), environmental economics (e.g., Cunha-e-Sá and Ducla-Soares 1999), development economics (e.g., Krishnan *et al.* 2019) and welfare economics (e.g., Ramadan and Thomas 2011).

This paper: (1) provides a nonparametric revealed preference analysis of the mixed-demand model, and (2) provides the first purely nonparametric empirical application of the mixed-demand model to the demand for food at the market level.

The first contribution consists of three parts. First, I derive a revealed characterization of the mixed-demand model. This characterization gives a simple condition to test if observed price-quantity data can be rationalized by the mixed-demand model, i.e., that the data are consistent this model. The condition can be implemented using simple linear programming (LP) techniques, and is therefore applicable to medium- and large-scaled data sets.

In the second part, I derive a measure of goodness-of-fit for the mixed-demand model. In case observed data violates the testable condition this index measures the severity of the violation. I relate the index to conventional measures of goodness-of-fit for the direct and indirect utility maximization models, and show that the index can be interpreted as a measure of "wasted" expenditure, i.e., it gives the minimal expenditure adjustment necessary to render the observed data consistent with the mixed-demand model.

In the third part, I derive a testable condition under the assumption that observed data are contaminated with measurement errors. It is well-recognized that aggregated market-level data contains measurement errors in various forms, e.g., due to aggregation issues and the use of price indices (See, for example, Carroll *et al.* (2015)). Although I focus on measurement errors in prices here, the condition can be straightforwardly generalized to instead account for measurement errors in quantities. The condition is based on solving a sequence of LP problem, and is therefore easy to implement in practice.

As a second contribution, I apply the new methods to annual, quarterly and monthly U.S. data over total food expenditures for 16 food commodities. I use measures of goodness-of-fit, power and predictive success to compare and contrast the empirical performance of the mixed-demand model with the direct utility maximization model (fully characterized by the generalized axiom of revealed preference, GARP) and a necessary condition for the mixed-demand model, which I call M-WARP. Because of the higher power of the mixed-demand model against (partial) uniform random behavior, I find that this model seem to perform better than the other models for at least annual data (with relatively few observations), and perhaps also for quarterly data. Using data sampled on higher frequencies, such as monthly data, this difference vanishes since the power of GARP increases to the same levels as the power for the mixed-demand model. Overall, I find that the mixed-demand model fits the data well.

Parametric vs. nonparametric modelling of mixed-demands. The parametric approach requires postulating a particular parametric functional form for the demand system and then estimating it using statistical techniques. The estimated demand functions can then be used to, for examples, calculate elasticities, perform welfare analysis and forecast demand behavior. However, the parametric approach will be satisfactory only when the postulated functional form is a good approximation to the "true" form. In fact, one may argue that any statistical inference based on parametric estimations of a mixed-demand model is a joint test of the null hypothesis and the hypothesis that correct functional forms and error structures have been employed.

In contrast to conventional demand analysis, many flexible direct demand systems, such as the translog (Christensen *et al.* 1975), almost ideal demand system (Deaton and Muellbauer 1980a) and quadratic almost ideal demand system (Banks *et al.* 1997), cannot be used in parametric mixed-demand modelling because there do not exist closed-form dual expressions of both the direct and indirect utility functions. Hence, parametric analysis may be rather limited since the most common flexible functional forms used in conventional (direct) demand analysis are unapplicable.[1]

Nevertheless, some parametric functional forms possess closed-form expressions for the direct and indirect utility functions. For examples, Barten (1992), Moschini and Vissa (1993), Matsuda (2004), Brown and Lee (2006) and Tabarestani *et al.* (2017) use different versions of the Rotterdam model to specify and estimate mixed-demand systems. More recently, Moschini and Rizzi (2006, 2007), Ramadan and Thomas (2011) and Krishnan *et al.* (2019) formulate and estimate mixed-demand systems for either the Stone–Geary model or normalized quadratic model.

The theory developed in this paper can be seen as an alternative to parametric demand analysis of the mixed-demand model. The framework is based on the classical revealed preference theory originally developed by Afriat (1967), Diewert (1973) and Varian (1982), and can be used to test a finite body of data for consistency with the mixed-demand model, recover preferences, perform counterfactual analysis and forecast demand behavior.[2] A distinct feature of the revealed preference framework is that it does not require any ad hoc assumptions regarding functional form and there is no parameter estimation. In other words, any test-procedure based on this approach is purely nonparametric, and as such, circumvent the problem of having to choose a suitable functional form as in the parametric approach.

---

[1] Wong and Park (2007) instead suggest to use conditional cost functions to generate empirical mixed-demand models.

[2] Revealed preference methods has sometimes been used as pre-tests to check whether observed market level data can be rationalized by a well-behaved utility function prior to conducting parametric demand analysis. If the data is "close" to being rationalizable, then this may serve as motivation to perform parametric demand estimation (Conversely, if the data is "far" from being rationalizable, then this might serve as motivation to refrain from such analysis.) Of course, the revealed preference methods proposed here can also be used prior to parametric demand analysis of the mixed-demand model for exactly the same purpose.

The nonparametric revealed preference approach appears particularly appropriate for empirical analysis of the mixed-demand model. In particular, the direct and indirect utility maximization models have been characterized in Varian (1982), Brown and Shannon (2000) and Hjertstrand and Swofford (2012), and these results are used to construct dual expressions of the direct and indirect utility functions. Hence, the revealed preference approach is very flexible in that it allows for closed-form expressions of the direct and indirect utility functions. More generally, each of the issues of concern to parametric demand analysis mentioned above is amenable to the nonparametric revealed preference approach.

This paper proceeds as follows: The next section gives a brief introduction to the mixed-demand model. Section 3 recalls the basic characterizations of direct and indirect utility maximization. Section 4 contains the theoretical results and provides a revealed preference characterization of the mixed-demand model, analogous to the characterizations of the direct and indirect utility maximization models in Section 3. Section 5 contains the application and Section 6 concludes. Proofs of the theoretical results and some additional empirical results are given in the appendix.

2 Mixed demands

In this section, I briefly recapitulate the mixed-demand model originally proposed by Samuelson (1965) and analyzed in detail by Chavas (1984).

Notation. We consider a market with K goods. Let $\mathbf{x} \in \mathbb{R}_+^K$ denote the quantity-vector and $\mathbf{p} \in \mathbb{R}_{++}^K$ denote the price-vector of the goods.[3] I assume that total expenditure, m, is exhaustive, and therefore given by $m = \mathbf{px}$. The vector of expenditure-normalized prices is defined as $\mathbf{r} = \frac{\mathbf{p}}{m} = \frac{\mathbf{p}}{\mathbf{px}}$.

Suppose that the goods are partitioned into two mutually exclusive and collectively exhaustive blocks. The two blocks of quantities are denoted $\mathbf{x} = (\mathbf{y}, \mathbf{w})$, where the expenditure-normalized prices of $\mathbf{y}$ and $\mathbf{w}$ are denoted $\mathbf{q}$ and $\mathbf{z}$, respectively, with $\mathbf{r} = (\mathbf{q}, \mathbf{z})$.

The model. In the mixed-demand model, consumers are price takers for all goods. However, at the market level, prices (and income) for some subset of goods are predetermined, while for the remaining goods it is the quantities that are predetermined. Suppose that $\mathbf{y}$ is the block of quantities that is chosen optimally, and that $\mathbf{z}$ is the block of expenditure-normalized prices that is optimally determined. Thus, $\mathbf{w}$ and $\mathbf{q}$ are the predetermined blocks of quantities and expenditure-normalized prices.

In the mixed-demand model, the consumer "chooses" the optimal values of the decision values $(\mathbf{y}, \mathbf{z})$ by solving the following maximization problem (Samuelson, 1965,

[3] I use the following notation: The inner product of two vectors $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^K$ is defined as $\mathbf{xy} = \sum_{k=1}^K x_k y_k$. For all $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^K$, $\mathbf{x} \geqq \mathbf{y}$ if $x_i \geq y_i$ for all $i = 1, \ldots, K$; $\mathbf{x} \geq \mathbf{y}$ if $\mathbf{x} \geqq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$; and $\mathbf{x} > \mathbf{y}$ if $x_i > y_i$ for all $i = 1, \ldots, K$. We denote $\mathbb{R}_+^K = \{\mathbf{x} \in \mathbb{R}^K : \mathbf{x} \geq (0, \ldots, 0)\}$ and $\mathbb{R}_{++}^K = \{\mathbf{x} \in \mathbb{R}^K : \mathbf{x} > (0, \ldots, 0)\}$. $\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})$ denotes the gradient of the function f with respect to the argument $\mathbf{x}$ (I write $\nabla = \nabla_{\mathbf{x}}$ when it is obvious from the context).

p.791):

$$0 = \max_{(\mathbf{y},\mathbf{z})}\{u(\mathbf{y},\mathbf{w}) - v(\mathbf{q},\mathbf{z})\} \qquad \text{subject to} \qquad \mathbf{qy} + \mathbf{zw} \leq 1, \tag{1}$$

where u is a continuous, strictly increasing and concave direct utility function, and v is a continuous, strictly decreasing and convex indirect utility function. The first-order conditions for an optimal (interior) solution of the problem (1) are:[4]

$$\begin{aligned} 0 &= \nabla_{\mathbf{y}} u\left(\mathbf{y},\mathbf{w}\right) - \psi\mathbf{q}, \\ 0 &= -\nabla_{\mathbf{z}} v\left(\mathbf{q},\mathbf{z}\right) - \psi\mathbf{w}, \end{aligned} \tag{2}$$

where ψ is the Lagrange multiplier. The solutions to the first-order conditions in (2) give the Marshallian mixed-demand vectors $\mathbf{y}^\star = \mathbf{y}(\mathbf{q},\mathbf{w},1)$ and $\mathbf{z}^\star = \mathbf{z}(\mathbf{q},\mathbf{w},1)$. Given the assumptions on u and v, the functions $\mathbf{y}(\mathbf{q},\mathbf{w},1)$ and $\mathbf{z}(\mathbf{q},\mathbf{w},1)$ satisfies Walras' law and are homogeneous of degree zero and one in $(\mathbf{p},\mathbf{px})$ (See Chavas (1984) and Moschini and Vissa (1993)).[5]

3 Characterizations of utility maximization

This section recalls the nonparametric revealed preference conditions that characterize direct and indirect utility maximization from Varian (1982), Brown and Shannon (2000) and Hjertstrand and Swofford (2012).

Notation. I assume that prices and quantities are observed in a *finite* number of time periods, denoted by T. I index observations by the set $\mathbb{T} = \{1, ..., T\}$, and let subscripts denote observations, that is, $\mathbf{x}_t = (x_{1t}, ..., x_{Kt})$, $\mathbf{p}_t = (p_{1t}, ..., p_{Kt})$ and $\mathbf{r}_t = (r_{1t}, ..., r_{Kt}) = (\frac{p_{1t}}{\mathbf{p}_t\mathbf{x}_t}, ..., \frac{p_{Kt}}{\mathbf{p}_t\mathbf{x}_t})$ denote quantities, prices and expenditure-normalized prices at observation $t \in \mathbb{T}$, respectively. I write $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ to signify all price-quantity observations.

3.1 Direct utility maximization

We begin with the definition of rationalization in terms of the direct utility function.

Definition 1 (Direct utility rationalization) *Consider a data set $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ and a direct utility function $u : \mathbb{R}^K_+ \mapsto \mathbb{R}$. For all $\mathbf{x} \in \mathbb{R}^K_+$ and all $t \in \mathbb{T}$ such that $\mathbf{r}_t\mathbf{x} \leq 1$, the data $\mathbb{D}$ is rationalized by u if $u(\mathbf{x}) \leq u(\mathbf{x}_t)$.*

This means that, for each observation $t \in \mathbb{T}$ it must be the case that the bundle $\mathbf{x}_t$ maximizes the utility function u over the set of all affordable bundles. The well-known Afriat's theorem gives necessary and sufficient *nonparametric* revealed preference conditions for when a data set $\mathbb{D}$ can be rationalized by a well-behaved (i.e., continuous, strictly increasing and concave) direct utility function:

[4] See Eqs. (28) and (29) in Chavas (1984).

[5] Since I do not characterize the mixed-demand model in terms of the cost function, I omit any discussion of duality etc.; See Chavas (1984) and Moschini and Vissa (1993) for a detailed discussion of such results.

Theorem 1 (Afriat's theorem, Varian 1982) *The data set $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ can be rationalized by a continuous, strictly increasing and concave direct utility function $u(\mathbf{x})$ if and only if there exist numbers U_t and $\lambda_t > 0$ such that the Afriat inequalities hold:*

$$U_s \leq U_t + \lambda_t \mathbf{r}_t \left(\mathbf{x}_s - \mathbf{x}_t\right),$$

for all observations $s, t \in \mathbb{T}$.

This is a limited version of Afriat's theorem. It usually contains additional procedures to test the direct utility maximization model such as the generalized axiom of revealed preference (GARP), but since I only use the Afriat inequalities in the analysis of mixed-demands, I omit these other conditions.[6] The Afriat inequalities are convenient for empirical analysis since they are linear in the unknown variables U_t and λ_t. In particular, U_t and λ_t can be interpreted as the utility and marginal utility of income values at the observation $t \in \mathbb{T}$. To see this, suppose that the data $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ were generated by a differentiable and concave utility function $u(\mathbf{x})$, in which case, by the properties of concavity, it must be that for all observations $s, t \in \mathbb{T}$:

$$u(\mathbf{x}_s) \leq u(\mathbf{x}_t) + \nabla u(\mathbf{x}_t) \left(\mathbf{x}_s - \mathbf{x}_t\right).$$

Additionally, the first-order conditions for utility maximization subject to a linear budget constraint (excluding boundary conditions) yield:

$$\nabla u(\mathbf{x}_t) = \lambda_t \mathbf{r}_t,$$

where λ_t is the marginal utility of income (i.e., the Lagrange multiplier). If we substitute this into the concavity inequalities and define $U_t = u(\mathbf{x}_t)$, we obtain the Afriat inequalities. Diewert (1973) and Fleissig and Whitney (2005) propose linear programing (LP) procedures to check whether there exists a feasible solution to the Afriat inequalities.

3.2 Indirect utility maximization

Brown and Shannon (2000) and Hjertstrand and Swofford (2012) give nonparametric revealed preference conditions for when a data set $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ can be rationalized by a well-behaved (i.e., continuous, strictly decreasing and convex) indirect utility function:

Definition 2 (Indirect utility rationalization) *Consider a data set $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ and an indirect utility function $v : \mathbb{R}^K_{++} \mapsto \mathbb{R}$. For all $\mathbf{r} \in \mathbb{R}^K_{++}$ and all $t \in \mathbb{T}$ such that $\mathbf{r}\mathbf{x}_t \leq 1$, the data $\mathbb{D}$ is rationalized by v if $v(\mathbf{r}_t) \leq v(\mathbf{r})$.*

Theorem 2 (Brown and Shannon 2000; Hjertstrand and Swofford 2012) *The data set $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ can be rationalized by a continuous, strictly decreasing and convex indirect utility function $v(\mathbf{r})$ if and only if there exist numbers V_t and $\mu_t > 0$ such that the indirect Afriat inequalities hold:*

$$-V_s \leq -V_t + \mu_t \mathbf{x}_t \left(\mathbf{r}_s - \mathbf{r}_t\right),$$

for all observations $s, t \in \mathbb{T}$.

[6] See Demuynck and Hjertstrand (2019) for a recent survey of Afriat's theorem that also includes other testable conditions.

This theorem mirrors Afriat's theorem. The indirect utility function in Theorem 2 is continuous, strictly decreasing and convex in $\mathbf{r}$. By defining $\mathbf{r} = \frac{\mathbf{p}}{\mathbf{px}}$ it follows that the indirect utility function $v(\mathbf{r}) = v(\frac{\mathbf{p}}{\mathbf{px}})$ is strictly decreasing in prices $\mathbf{p}$, strictly increasing in expenditure $\mathbf{px}$ and homogeneous of degree zero in $(\mathbf{p}, \mathbf{px})$.

A heuristic derivation of the indirect Afriat inequalities follows along the same lines as for the Afriat inequalities by first noticing that the indirect utility function is convex in $\mathbf{r}$ and then applying Roy's identity which states that $\nabla v(\mathbf{r}) = \mathbf{x} \sum_{j=1}^{K} r_j \frac{\partial v(\mathbf{r})}{\partial r_j}$ (Roy, 1947, p.219). Evaluating Roy's identity in every observation $t \in \mathbb{T}$ and defining $V_t = v(\mathbf{r}_t)$ and $\mu_t = -\sum_{j=1}^{K} r_{jt} \frac{\partial v(\mathbf{r}_t)}{\partial r_{jt}}$ gives the indirect Afriat inequalities. Diewert (1973)'s and Fleissig and Whitney (2005)'s LP procedures can be straightforwardly modified to also check whether there exist numbers V_t and $\mu_t > 0$ satisfying the indirect Afriat inequalities.

Notice that the Afriat inequalities and the indirect Afriat inequalities are formally equivalent with expenditure-normalized prices and quantities interchanged. This implies that the characterizations in Theorems 1 and 2 are equivalent.

Proposition 1 (Chavas and Cox 1987) *A data set $\mathbb{D} = \{\mathbf{r}_t, \mathbf{x}_t\}_{t \in \mathbb{T}}$ can be rationalized by a well-behaved direct utility function if and only if $\mathbb{D}$ can be rationalized by a well-behaved indirect utility function.*

4 Characterization of mixed-demands

In this section, I give nonparametric revealed preference characterizations for the mixed-demand model. I begin by deriving a characterization result for mixed-demands that is analogous to Theorems 1 and 2. I then show how the framework can be extended to test if the mixed-demand model can rationalize a data set with measurement errors.

Notation. As in the previous section, it is assumed that prices and quantities are observed in a *finite* number of time periods. Denote the two blocks of quantities and expenditure-normalized prices at observation $t \in \mathbb{T}$ as $\mathbf{x}_t = (\mathbf{y}_t, \mathbf{w}_t)$ and $\mathbf{r}_t = (\mathbf{q}_t, \mathbf{z}_t)$, respectively. I write $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}}$ to signify all price-quantity observations.

4.1 Revealed preference conditions

I begin with the definition of rationalization in the context of the mixed-demand model.

Definition 3 (Mixed demand rationalization) *Consider a data set $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}}$, a direct utility function $u : \mathbb{R}^K_+ \mapsto \mathbb{R}$ and an indirect utility function $v : \mathbb{R}^K_{++} \mapsto \mathbb{R}$. For all $(\mathbf{y}, \mathbf{z}) \in \mathbb{R}^K_+ \times \mathbb{R}^K_{++}$ and all $t \in \mathbb{T}$ such that $\mathbf{q}_t\mathbf{y} + \mathbf{z}\mathbf{w}_t \leq 1$, the data $\mathbb{D}$ is rationalized by the mixed-demand model if $u(\mathbf{y}, \mathbf{w}_t) - v(\mathbf{q}_t, \mathbf{z}) \leq 0$.*

This means that, for each observation $t \in \mathbb{T}$ it must be the case that $(\mathbf{y}_t, \mathbf{z}_t)$ maximizes the function $0 = \{u(\mathbf{y}, \mathbf{w}_t) - v(\mathbf{q}_t, \mathbf{z})\}$ over the set of all non-predetermined expenditure-normalized prices and bundles $(\mathbf{y}, \mathbf{z})$ given the budget constraint $\mathbf{q}_t\mathbf{y} + \mathbf{z}\mathbf{w}_t \leq 1$. Thus, Definition 3 is equivalent to that $(\mathbf{y}_t, \mathbf{z}_t)$ are the optimal solutions to the problem (1) at every observation.

The next theorem gives a nonparametric revealed preference characterization of the mixed-demand model, and is analogous to the characterizations of direct and indirect utility maximization in Theorems 1 and 2.

Theorem 3 *Consider the data set* $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}}$. *The following two conditions are equivalent:*

- *There exists a continuous, strictly increasing and concave direct utility function* $u(\mathbf{y}, \mathbf{w})$*, and a continuous, strictly decreasing and convex indirect utility function* $v(\mathbf{q}, \mathbf{z})$*, such that the data* $\mathbb{D}$ *can be rationalized by the mixed-demand model.*

- *For every* $t \in \mathbb{T}$*, there exist numbers* $U_t, V_t, \lambda_t > 0$ *and* $\mu_t > 0$ *such that the following (in)equalities hold:*

$$\begin{aligned} U_s &\leq U_t + \lambda_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right), \\ -V_s &\leq -V_t + \mu_t \mathbf{w}_t \left(\mathbf{z}_s - \mathbf{z}_t\right), \\ U_t &= V_t, \\ \lambda_t &= \mu_t, \end{aligned} \tag{3}$$

  for all observations $s, t \in \mathbb{T}$.

This theorem gives necessary and sufficient conditions for when the mixed-demand model can be rationalized by well-behaved direct and indirect utility functions. The inequalities in (3) are the Afriat inequalities and indirect Afriat inequalities for the non-predetermined bundles and expenditure-normalized prices, respectively. The equality restrictions in (3) ensures that v is the indirect utility function corresponding to the direct utility function u, and therefore guarantees that the difference of the function values at the optimal solution is effectively zero. By defining $F_t = U_t = V_t$ and $\psi_t = \lambda_t = \mu_t > 0$, the condition (3) can be equivalently stated as:

$$\begin{aligned} F_s &\leq F_t + \psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right), \\ -F_s &\leq -F_t + \psi_t \mathbf{w}_t \left(\mathbf{z}_s - \mathbf{z}_t\right), \end{aligned} \tag{4}$$

for all observations $s, t \in \mathbb{T}$.

4.2 Necessary conditions and non-nestedness

In this section, I derive some necessary conditions for the inequalities (3) to hold and show by example that the mixed-demand model is neither nested within nor nests the direct (or indirect) utility maximization model.

The inequalities (4) are linear in the unknowns F_t and ψ_t, and checking whether a solution exists to these inequalities can be achieved in polynomial time by solving a LP problem in the $2T$ parameters $\{F_t, \psi_t\}_{t \in \mathbb{T}}$ subject to the $2T(T-1)$ linear restrictions given by (4). However, for very large data sets (i.e., large T) this LP problem may become computationally difficult to implement since the number of restrictions grows quadratically in T. In this case, it is possible to test a set of necessary conditions, which if violated, implies that there cannot exist numbers F_t and $\psi_t > 0$ solving the inequalities (4).

One necessary condition is that the blocks $\{\mathbf{q}_t, \mathbf{y}_t\}_{t\in\mathbb{T}}$ and $\{\mathbf{z}_t, \mathbf{w}_t\}_{t\in\mathbb{T}}$ individually satisfies the Afriat inequalities (or equivalently, by Proposition 1, the indirect Afriat inequalities).[7] Another necessary condition follows directly by combining the two sets of inequalities in (4):

$$-\psi_t \mathbf{w}_t (\mathbf{z}_s - \mathbf{z}_t) \leq F_s - F_t \leq \psi_t \mathbf{q}_t (\mathbf{y}_s - \mathbf{y}_t).$$

Thus, since $\psi_t > 0$, for every pair of observations $s, t \in \mathbb{T}$ it must hold that:

$$\mathbf{q}_t \mathbf{y}_s + \mathbf{z}_s \mathbf{w}_t \geq \mathbf{q}_t \mathbf{y}_t + \mathbf{z}_t \mathbf{w}_t = 1. \tag{5}$$

I call this condition the mixed weak axiom of revealed preference (M-WARP):

Definition 4 (M-WARP) *The data $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t\in\mathbb{T}}$ satisfies the mixed weak axiom of revealed preference (M-WARP) if, for every pair of observations $s, t \in \mathbb{T}$, it holds that $\mathbf{q}_t \mathbf{y}_s + \mathbf{z}_s \mathbf{w}_t \geq 1$.*

The following proposition formally states that M-WARP is a necessary condition for the inequalities (4), or equivalently the inequalities (3) to hold:

Proposition 2 *If the data $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t\in\mathbb{T}}$ satisfies the necessary and sufficient conditions for the mixed-demand model given by the inequalities* (3) *then $\mathbb{D}$ satisfies M-WARP.*

In order to show that the mixed-demand model is empirically distinguishable from the utility maximization model, I provide a data set that satisfies the rationalizability condition for the direct utility maximization model but not the mixed-demand model, and also a data set showing the opposite.

Example 1 *Suppose $T = 2$ and $K = 2$. The expenditure-normalized prices are given by:*

$$\mathbf{q}_1 = 1 \qquad \mathbf{z}_1 = 1$$
$$\mathbf{q}_2 = \frac{1}{2} \qquad \mathbf{z}_2 = 1$$

and the quantities are given by:

$$\mathbf{y}_1 = \frac{1}{2} \qquad \mathbf{w}_1 = \frac{1}{2}$$
$$\mathbf{y}_2 = 1 \qquad \mathbf{w}_2 = \frac{1}{2}$$

It is easy to verify that the Afriat inequalities are satisfied by choosing $U_1 = 1, U_2 = \frac{1}{2}, \lambda_1 = 1$ and $\lambda_2 = 2$. However, it is also easy to verify from (4) *that $F_1 = F_2$, which implies $\frac{\psi_1}{2} \geq 0$ but $\frac{\psi_2}{4} \leq 0$. Thus, $\psi_2 \ngtr 0$, which violates condition* (3).

[7] Note that when implementing this necessary condition it is computationally simpler to apply other equivalent test procedures, such as the generalized axiom of revealed preference (GARP).

Example 2 *Suppose $T = 3$ and $K = 4$. The expenditure-normalized prices are given by:*

$$\mathbf{q}_1 = \left(\frac{2}{302}, \frac{22}{302}\right) \qquad \mathbf{z}_1 = \left(\frac{3}{302}, \frac{16}{302}\right)$$
$$\mathbf{q}_2 = \left(\frac{5}{96}, \frac{2}{96}\right) \qquad \mathbf{z}_2 = \left(\frac{4}{96}, \frac{4}{96}\right)$$
$$\mathbf{q}_3 = \left(\frac{2}{169}, \frac{7}{169}\right) \qquad \mathbf{z}_3 = \left(\frac{11}{169}, \frac{5}{169}\right)$$

and the quantities are given by:

$$\mathbf{y}_1 = (17, 8) \qquad \mathbf{w}_1 = (4, 5)$$
$$\mathbf{y}_2 = (4, 10) \qquad \mathbf{w}_2 = (12, 2)$$
$$\mathbf{y}_3 = (3, 9) \qquad \mathbf{w}_3 = (5, 9)$$

It is easy to verify that there cannot exist a feasible solution to the Afriat inequalities, since $\mathbf{r}_1 (\mathbf{x}_2 - \mathbf{x}_1) = -\frac{6}{302}$, $\mathbf{r}_2 (\mathbf{x}_3 - \mathbf{x}_2) = -\frac{7}{96}$ and $\mathbf{r}_3 (\mathbf{x}_1 - \mathbf{x}_3) = -\frac{10}{169}$. On the other hand, the following numbers (rounded to the fourth decimal) solve the necessary and sufficient condition for the mixed-demand model in (3)*:*

$$F_1 = 0 \qquad \psi_1 = 36.2146$$
$$F_2 = -0.6354 \qquad \psi_2 = 1$$
$$F_3 = -0.7195 \qquad \psi_3 = 5.7902$$

Example 2 gives a data set that can be rationalized by the mixed-demand model, but not the utility maximization model. Examples 1 and 2 taken together shows that the utility maximization model and the mixed-demand model are distinct and may yield different answers regarding whether or not a data set can be rationalized. In other words, the models have identifying power against each other (i.e., they may generate different predictive success). Finally, Example 1 gives a data set which violates the rationalizability condition for the mixed-demand model, which shows that this model has substantial empirical bite even when the data set consists of two observations.

4.3 Goodness-of-fit

The test given in condition (4) is "sharp" in the sense that a single violation of the inequalities rejects the hypothesis that the mixed-demand model rationalizes the data. However, observed data may violate the condition even though underlying preferences are consistent with the mixed-demand model. This may arise because of optimization errors, minor measurement errors etc. Afriat (1972) proposed to measure the severity of violations of the direct utility maximization model by calculating the minimal expenditure adjustment necessary to render the observed data consistent with the Afriat inequalities. This index, usually called the Afriat efficiency index (AEI) or the critical cost efficiency index (CCEI), is a measure of wasted expenditure due to inconsistency with the Afriat inequalities: If the consumer has an AEI less than one, then he could have obtained the same level of utility by only spending a fraction 1−AEI of expenditure (at every observation). In this regard, Varian (1990) interprets the AEI as a measure of goodness-of-fit of the direct utility maximization model.

In practice, the AEI is calculated by multiplying total expenditure (normalized to unity here) at every observation with a scalar $e \in [0,1]$ as $e\mathbf{r}_t\mathbf{x}_t$, which gives the following set of "relaxed" Afriat inequalities:[8]

$$U_s \leq U_t + \lambda_t \mathbf{r}_t \left(\mathbf{x}_s - e\mathbf{x}_t\right), \tag{6}$$

for all observations $s, t \in \mathbb{T}$. The AEI is defined as the largest e such that the observed data satisfies condition (6). Analogously, the AEI for the indirect utility maximization model is defined as the largest e such that the observed data satisfies the "relaxed" indirect Afriat inequalities:

$$-V_s \leq -V_t + \mu_t \mathbf{x}_t \left(\mathbf{r}_s - e\mathbf{r}_t\right), \tag{7}$$

for all observations $s, t \in \mathbb{T}$.

Given condition (3) in Theorem 3 and conditions (6) and (7), the same idea can be straightforwardly applied to the mixed-demand model: I define the AEI for the mixed-demand model as the largest value of e such that the observed data satisfies the following "relaxed" version of condition (4):

$$\begin{aligned} F_s &\leq F_t + \psi_t \mathbf{q}_t \left(\mathbf{y}_s - e\mathbf{y}_t\right), \\ -F_s &\leq -F_t + \psi_t \mathbf{w}_t \left(\mathbf{z}_s - e\mathbf{z}_t\right), \end{aligned} \tag{8}$$

for all observations $s, t \in \mathbb{T}$. Hence, the AEI in this context calculates the minimal expenditure adjustment necessary to render the observed data consistent with condition (8). As such, analogous to the direct and indirect utility maximization models, the AEI can be interpreted as a measure of goodness-of-fit of the mixed-demand model.

4.4 Measurement errors

In this section, I derive a testable condition for the mixed-demand model under the assumption that prices are measured with errors. As explained in the Introduction, I focus on errors in prices because, in aggregated data, prices are usually constructed from index number theory, and are therefore likely to be measured with errors due to aggregation bias etc.

Following Varian (1985), it is assumed that the measurement errors, $\mathbf{e}_t = \{e_{1t}, ..., e_{Kt}\} \in \mathbb{R}^K$, enter prices multiplicatively via the Berkson multiplicative measurement error model:

$$\overline{\mathbf{p}}_t = \mathbf{p}_t \odot \mathbf{e}_t, \tag{9}$$

for all $t \in \mathbb{T}$, where $\overline{\mathbf{p}}_t$ denotes the "true" prices, and $\odot$ denotes Hadamard (elementwise) multiplication. Notice that, as a consequence of this assumption, total expenditure is

[8] See Theorem 2 in Heufer and Hjertstrand (2017) for a justification and proof that the "relaxed" Afriat inequalities can be used to calculate the minimal expenditure adjustment.

also measured with errors, and the 'true' amount of total expenditure, denoted $\overline{m}_t$, is given by:

$$\overline{m}_t = \overline{\mathbf{p}}_t \mathbf{x}_t = (\mathbf{p}_t \odot \mathbf{e}_t)\mathbf{x}_t. \tag{10}$$

I define the "true" expenditure-normalized prices as $\overline{\mathbf{r}}_t = \frac{\overline{\mathbf{p}}_t}{\overline{m}_t}$, and let $\overline{\mathbf{r}}_t$ be split into two blocks $(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t)$ corresponding to $(\mathbf{q}_t, \mathbf{z}_t)$.

The purpose is to test the following hypothesis:

$$\text{H}_0 : \text{The "true" data } \overline{\mathbb{D}} = \{(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}} \text{ satisfies condition (3).} \tag{11}$$

The null, H_0, corresponds to that there exists a continuous, strictly increasing and concave direct utility function $u(\mathbf{y}, \mathbf{w})$, and a continuous, strictly decreasing and convex indirect utility function $v(\overline{\mathbf{q}}, \overline{\mathbf{z}})$, such that the "true" data $\overline{\mathbb{D}} = \{(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}}$ can be rationalized by the mixed-demand model. The alternative hypothesis is that $\overline{\mathbb{D}}$ cannot be rationalized by the mixed-demand model, i.e., $\overline{\mathbb{D}}$ does not satisfy condition (3).

Consider the following optimization problem:

$$\begin{aligned} \Phi = \min_{\{F_t, \psi_t, \phi\}_{t \in \mathbb{T}}} \phi \qquad & \text{subject to} \\ F_s - F_t - \psi_t \mathbf{q}_t (\mathbf{y}_s - \mathbf{y}_t) &\leq \psi_t \phi, \\ -F_s + F_t - \psi_t \mathbf{w}_t (\mathbf{z}_s - \mathbf{z}_t) &\leq \psi_t \phi, \\ 1 &\leq \psi_t, \\ 0 \leq \phi &\leq 1. \end{aligned} \tag{12}$$

This problem computes the minimal slacks at each observation, $\psi_t \phi$, such that the inequalities (4) hold. If $\phi = 0$ then the first two restrictions in the problem (12) are equivalent to the inequalities (4). In this case, the observed data $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}}$ satisfies the rationalizability conditions in Theorem 3. On the other hand, if $\phi > 0$ then ϕ can be interpreted as the minimal (additive) adjustment of expenditure such that the inequalities hold, which can be seen by rewriting the restrictions in the problem (12) as:

$$\begin{aligned} F_s - F_t - \psi_t (\mathbf{q}_t \mathbf{y}_s - (\mathbf{q}_t \mathbf{y}_t - \phi)) &\leq 0, \\ -F_s + F_t - \psi_t (\mathbf{w}_t \mathbf{z}_s - (\mathbf{w}_t \mathbf{z}_t - \phi)) &\leq 0. \end{aligned}$$

The third restriction, $1 \leq \psi_t$, can be imposed without loss of generality because the inequalities are homogeneous of degree one in the parameters F_t, F_s and ψ_t. Moreover, since total expenditure sums to unity, we have $\mathbf{q}_t \mathbf{y}_t \leq 1$ and $\mathbf{w}_t \mathbf{z}_t \leq 1$, which implies that the inequalities always have a solution for $\phi = 1$. Thus, the fourth restriction, $0 \leq \phi \leq 1$, follows without loss of generality.

The problem (12) is nonlinear in the parameters $\psi_t \phi$ and may therefore be difficult to implement in practice. However, notice that if it has a feasible solution for a particular value of ϕ, then it also has a feasible solution for all values of $\phi' \geq \phi$. This monotonicity condition implies that we can solve the problem using a binary search algorithm, and treat ϕ as fixed in each iteration of the procedure.[9] As such, the problem (12) is reduced

[9]Consider the infeasible bound $\phi_l = 0$ and the feasible bound $\phi_u = 1$. For each step of the binary search algorithm, the procedure evaluates whether the midpoint $\frac{\phi_l + \phi_u}{2}$ is feasible. If it is, then in the next iteration the upper bound ϕ_u is replaced by this midpoint. However, if the midpoint is not feasible, then the midpoint replaces the lower bound ϕ_l. At each iteration of the algorithm, the range $[\phi_l, \phi_u]$, which contains the solution of the problem, is halved. Thus, the width of the interval decreases exponentially in the number of iterations.

to a standard LP problem, which is to be solved in every step. Since a LP problem can be solved efficiently, i.e., in polynomial time, the procedure is applicable for medium- and large scaled data sets.

The optimization problem (12) can be used to derive a statistical decision rule to test the hypothesis (11) under the assumption of the multiplicative Berkson measurement error model (9). This decision rule is based on the following theorem.

Theorem 4 *Consider the optimization problem* (12)*. Under* H_0 *in* (11)*, it holds that:*

$$\Phi \leq C,$$

where:

$$C = \max\{C_1, C_2\}, \tag{13}$$

with:

$$C_1 = \max_{s,t \in \mathbb{T}} \{(\overline{\mathbf{q}}_t - \mathbf{q}_t)(\mathbf{y}_s - \mathbf{y}_t)\}, \tag{14}$$

$$C_2 = \max_{s,t \in \mathbb{T}} \{\mathbf{w}_t((\overline{\mathbf{z}}_s - \mathbf{z}_s) - (\overline{\mathbf{z}}_t - \mathbf{z}_t))\}. \tag{15}$$

This theorem suggests the statistical decision rule that the null, H_0, in the hypothesis (11) should be rejected whenever $\Phi > C_{1-\alpha}$, where α denotes the significance level and $C_{1-\alpha}$ denotes the $(1-\alpha) \times 100$th percentile of the distribution of the random variable C in (13). However, C is function of the "true" prices $(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t)$, which are unobserved. Nevertheless, for a given distributional assumption on the measurement errors $\mathbf{e}_t$, we can simulate $(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t)$ from the Berkson measurement error models (9) and (10). Suppose that the measurement errors, $\mathbf{e}_t$, follow a known distribution, and consider D draws (realizations) from this distribution denoted by $\boldsymbol{\varepsilon}_t^d$, for $d = 1, ..., D$. Expenditure-normalized prices, $\overline{\mathbf{r}}_t^d$, are simulated from (9) and (10) as:

$$\overline{\mathbf{p}}_t^d = \mathbf{p}_t \odot \boldsymbol{\varepsilon}_t^d,$$
$$\overline{m}_t^d = \overline{\mathbf{p}}_t^d \mathbf{x}_t,$$
$$\overline{\mathbf{r}}_t^d = \frac{\overline{\mathbf{p}}_t^d}{\overline{m}_t^d},$$

for all $t \in \mathbb{T}$ and $d = 1, ..., D$. If we split $\overline{\mathbf{r}}_t^d$ into two blocks $(\overline{\mathbf{q}}_t^d, \overline{\mathbf{z}}_t^d)$ corresponding to $(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t)$, then we can calculate the empirical distribution of C by drawing D copies of C as $C^d = \max\{C_1^d, C_2^d\}$, where:

$$C_1^d = \max_{s,t \in \mathbb{T}} \{(\overline{\mathbf{q}}_t^d - \mathbf{q}_t)(\mathbf{y}_s - \mathbf{y}_t)\},$$
$$C_2^d = \max_{s,t \in \mathbb{T}} \{\mathbf{w}_t((\overline{\mathbf{z}}_s^d - \mathbf{z}_s) - (\overline{\mathbf{z}}_t^d - \mathbf{z}_t))\},$$

for $d = 1, ..., D$. At a given nominal significance level, α, let $c_{1-\alpha}$ be the $(1-\alpha) \times 100$th percentile of the simulated sample $\left(C^1, ..., C^D\right)$, and let $\hat{\Psi}$ be the optimal solution from the optimization problem (12). The null H_0 in the hypothesis (11) is rejected if $\hat{\Psi} > c_{1-\alpha}$.

A few remarks. First, deriving the analytical distribution of C in order to find an analytical expression for $C_{1-\alpha}$ is difficult since C does not follow any standard distribution. However, as we have shown, it is simple to calculate the empirical distribution of C by Monte Carlo simulations, which makes the problem easy to implement in practice. Note that the number of draws D should be set relatively large so that the empirical percentile $c_{1-\alpha}$ closely approximates the "true" percentile $C_{1-\alpha}$.[10]

Second, in order to make draws from the distribution of the measurement errors we must know the parameters characterizing the distribution. For example, assuming that $\boldsymbol{\varepsilon}_t^d$ are draws from the normal distribution (as in the empirical application below), require assuming a value of the variance of the distribution. In practice, I implement the procedure for a grid of values of the variance. The smallest variance such that H_0 is rejected is a lower bound of the possible "amount" of measurement errors (measured by the variance). More precisely, if $\overline{\sigma}^2$ denotes the lower bound, then the procedure is also unable to reject H_0 at the given significance level for any $\sigma^2 > \overline{\sigma}^2$. Thus, the lower bound $\overline{\sigma}^2$ is a measure of what the unknown variance of the measurement errors would have to be in order to reject H_0 that the mixed-demand model can rationalize the "true" data $\overline{\mathbb{D}} = \{(\overline{\mathbf{q}}_t, \overline{\mathbf{z}}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t\in\mathbb{T}}$.

Third, the procedure is conservative in the sense that the probability of rejecting the null, H_0, when it is true is at most α.[11] Unfortunately, calculating the power of the procedure provides very limited information, since it would require specifying a model of irrationality. Thus, any such analysis would only show the power against one specific alternative, and not against any general notion of irrational consumption behavior.

Finally, a brief history of the origin of this procedure. Varian (1985) originally proposed a procedure to account and test for measurement errors in revealed preference. However, for the mixed-demand model, Varian's procedure would be computationally very difficult to implement in practice since it would be based on solving a highly nonlinear optimization problem. The framework proposed in this section is based on Fleissig and Whitney (2005), and in particular, on the modification of Fleissig and Whitney's (2005) procedure proposed by Jones and Edgerton (2009). The idea of calculating a lower bound for the measurement errors was proposed by Varian (1985) in the context of his model, while Jones and Edgerton (2009) showed how to calculate a lower bound for their modified version of Fleissig and Whitney's (2005) procedure.

5 Application

In this section, I apply the methods proposed in previous sections to aggregated market data on prices and quantities for food.

5.1 Data and classification

Data. I use data over total food expenditures in the U.S. These data consists of 16 different aggregated food commodities which are listed in Table 1 (Tables A1 and A2 in

[10] In the empirical application, I set $D = 1,000$.

[11] See Jones and Edgerton (2009) for a more detailed discussion of this issue in the context of testing for measurement errors in the utility maximization model.

Table 1: Commodities and classification.

Commodity		Classification utility function
A.	Cereals	Direct
B.	Bakery products	Direct
C.	Beef and veal	Direct
D.	Pork	Direct
E.	Other meats	Direct
F.	Poultry	Direct
G.	Fish and seafood	Indirect
H.	Fresh milk	Indirect
I.	Processed diary products	Direct
J.	Eggs	Indirect
K.	Fats and oils	Direct
L.	Fruit (fresh)	Indirect
M.	Vegetables (fresh)	Indirect
N.	Processed fruits & vegetables	Direct
O.	Sugar and sweets	Direct
P.	Unclassified food	Direct

Appendix C provides some descriptive statistics of the prices and budget shares). For each commodity, I collect data on total real personal consumption expenditures from the Bureau of Economic Analysis (BEA), which are sampled on three different frequencies: (i) annual data ranging from 1999 to 2017 ($T = 19$ observations); (ii) quarterly data ranging from 1999Q1 to 2017Q4 ($T = 76$ observations), and (iii) monthly data ranging from 1999M1 to 2017M12 ($T = 228$ observations). The quarterly and monthly data are seasonally adjusted at annual rates and all series are given in millions of chained (2009) dollars.[12]

To obtain per capita expenditures, I collect U.S. annual, quarterly and monthly population series (in thousands) from the FRED data base at the St. Louis FED. Per capita expenditures for every commodity are calculated by dividing the total real personal consumption expenditures with the corresponding population series.

I also collect price indices for each food commodity at the three sample frequencies from the BEA. In all these price series, the base year is 2009 and the quarterly and monthly series are seasonally adjusted.[13]

Classification of commodities into direct or indirect utility function. I classify all 16 commodities listed in Table 1 as arguments in either the direct or indirect utility functions. Following standard practice, all commodities that are easy to store are classified as arguments in the direct utility function (commodities A, B, I, K, N, O and P). Following Moschini and Vissa (1993) I also list beef and veal [C], pork [D], other meats [E]

[12]The data are collected from Table 2.4.6U downloadable from the BEA data archive, and listed in excel sheets 20405-A (annual data), 20405-Q (quarterly data) and 20405-M (monthly data).

[13]The price data are collected from Tables 2.4.4U-A (annual data), 2.4.4U-Q (quarterly data) and 2.4.4U-M (monthly data) downloadable from the BEA data archive.

Table 2: Goodness-of-fit (AEI).

Sample	Mixed-demand	M-WARP	GARP
Annual	0.813	0.813	1.000
Quarterly	0.806	0.806	1.000
Monthly	0.801	0.801	1.000

and poultry [F] as arguments in the direct utility function. Moreover, following Moschini and Rizzi (2006,2007), I list commodities that can be considered as fresh food such as fish and seafood [G], fresh milk [H], eggs [J], fresh fruits [L] and fresh vegetables [M] as arguments in the indirect utility function. These classifications are also listed in the last column of Table 1.

5.2 Goodness-of-fit

I begin by calculating the Afriat efficiency index (AEI) for the mixed-demand model, the necessary condition M-WARP and the direct utility maximization model (GARP).[14] Table 2 gives the results. As seen from this table, the direct utility maximization model rationalizes the data at full efficiency, that is, the consumer does not need to waste any of her income for the data to satisfy GARP. This holds irrespectively of the sample. The AEI for the mixed-demand model and M-WARP are lower, meaning that these models seem to provide a worse fit to the data. However, what should be considered a "too low" measure of goodness-of-fit? In several recent studies, a lower threshold is set to 0.8, implying that data reflecting an AEI below this threshold is deemed to far from utility maximization behavior.[15]

Following these recent studies, in this paper, I will consider a "waste" of 20% reasonable. However, I strongly emphasize that this level is not universal, but should depend on the problem at hand, that is, the number of observations, the power of the test, and the model under consideration. Therefore, I encourage readers to make up their own mind of what is a reasonable amount of "waste" such that the data should be considered "close" enough to being rationalizable by the model in question.

Finally, I note that although M-WARP is only a necessary condition for the mixed-demand model, the AEI for M-WARP and the mixed-demand model are equivalent, suggesting that M-WARP is a very accurate approximation of the testable conditions for the mixed-demand model.

[14]When measuring the goodness-of-fit for the direct demand model I treat all commodities as arguments in the direct utility function. The AEI is calculated using the binary search algorithm for the generalized axiom of revealed preference, GARP, as described in Varian (1990). Another equivalent but computationally more expensive way to calculate the AEI for the direct demand model is to use the Afriat inequalities as described in Section 4.3. Appendix A describes in more detail how the AEI for M-WARP is calculated.

[15]See e.g., Choi et al. (2007) and Fisman et al. (2007).

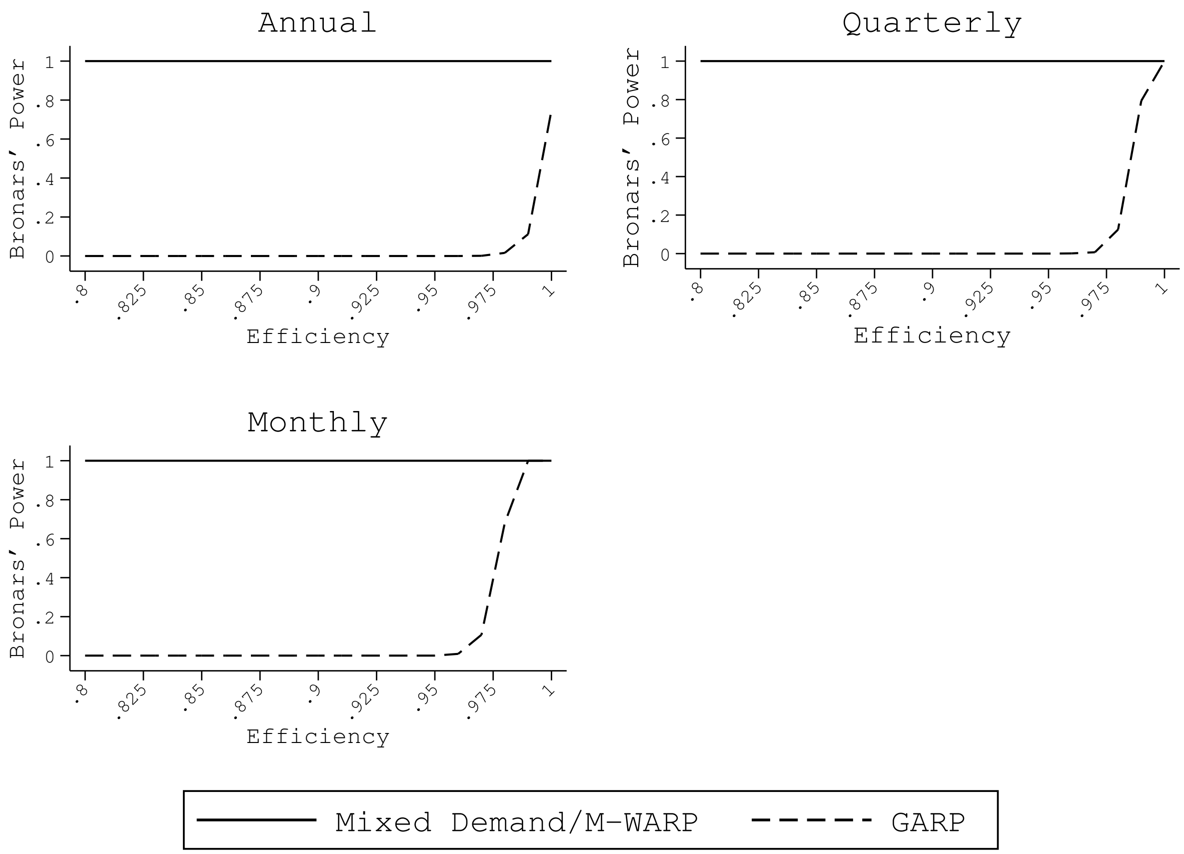

Figure 1: Power against uniformly random behavior (Bronars' power index) for the mixed-demand model, M-WARP and GARP at different levels of efficiency.

5.3 Power and predictive success

To meaningfully compare different models, it is important also to consider other diagnostic measures than goodness-of-fit. One important such measure is power, which is measured as the probability of rejecting the revealed preference test given that the model does not hold. A high goodness-of-fit have little value if the test has little discriminatory power (i.e., the conditions are hard to reject for the data at hand). In this section, I calculate and compare the power for the mixed-demand model, M-WARP and GARP using two different measures of power. I also compare the models using a measure of predictive success.

Power against uniformly random behavior. The first measure I consider is the power against uniformly random behavior. This index (often called Bronars power index) quantifies discriminatory power in terms of the probability to detect uniformly random behavior, and is based on Bronars (1987). This measure is implemented by first simulating 1,000 random series of 16 consumption choices (one for each commodity). At every observation, this is achieved by drawing a random quantity bundle from a uniform distribution on the budget hyperplane given the observed prices and total expenditure. At a given efficiency level, the power against uniformly random behavior is then calculated as one minus the proportion of these randomly generated consumption series that are consistent with the rationalizability conditions being evaluated.

Figure 1 presents plots of the power of the mixed-demand model, M-WARP and GARP at different levels of efficiency ranging from 0.8 to 1 for each of the three data samples.[16] As seen from these plots, both the mixed-demand model and M-WARP has optimal power (equal to 1) for any efficiency level equal to 0.8 or higher, which holds irrespectively of the sample. The results for the direct demand model (GARP) is quite different: the power is equal to zero up to an efficiency level of about 0.975 after which it increases rapidly. For the quarterly and monthly data, the power of GARP is 1 at full efficiency ($e = 1$), while it is 0.75 for the annual data at full efficiency.

These results suggest that GARP is a much less stringent condition than the testable conditions for the mixed-demand model and M-WARP. One way to interpret this is that, while GARP is associated with a higher degree of goodness-of-fit, it seems that this better fit may simply be due to a lower discriminatory power rather than a better model per se. To analyze this in more detail, we also consider another measure of power, which nests Bronars power index as a special case.

Power against partial uniformly random behavior. The second measure I consider is the power against *partial* uniformly random behavior. This index (called the partial uniformly random power (PURP) index) quantifies discriminatory power in terms of the probability to detect a model of irrationality, where only a certain fraction of expenditure is random, and was recently introduced by Hjertstrand (2021). In this model of irrationality, the vector of budget shares, $\mathbf{w}^I$, corresponds to the weighted average between the observed budget shares, $\mathbf{w}$, and the uniformly random budget shares, $\mathbf{w}^U$, i.e.,

$$\mathbf{w}^I = (1 - \lambda)\,\mathbf{w} + \lambda \mathbf{w}^U. \tag{16}$$

The parameter $\lambda \in [0, 1]$ is interpreted as the fraction of expenditure that is randomly allocated. This model encompasses a continuum of sub-models. At the one extreme $\lambda = 0$, it conform to when no part of the expenditure is being randomly allocated, in which case the budget shares are equal to the observed budget shares, i.e., $\mathbf{w}^I = \mathbf{w}$. At the other extreme $\lambda = 1$, the model reduces to Bronars' pure model of uniformly random behavior, implying that the budget shares are equal to the uniformly random budget shares, i.e., $\mathbf{w}^I = \mathbf{w}^U$.[17] An appealing property of the PURP index is that it can trace out the entire power curve against uniformly random expenditure allocation by implementing the model at each node in an equally-spaced grid for $\lambda \in [0, 1]$.[18]

I implement the PURP index for the mixed-demand model, M-WARP and GARP at each node using a grid of 0.05 for λ (starting at 0 and ending at 1).[19] Figure 2 plots the

---

[16] Thus, the power is calculated sequentially for the three models at different values of e as outlined in Section 4.3.

[17] Thus, in the case $\lambda = 1$, the PURP index and Bronars power index coincide.

[18] In this regard, note that Bronars power index only gives a measure of power at a specific point on the power curve.

[19] When implementing the testable condition for each model, I add efficiency corresponding to the AEI for the corresponding model. Thus, in the testable condition for the mixed-demand model, I add

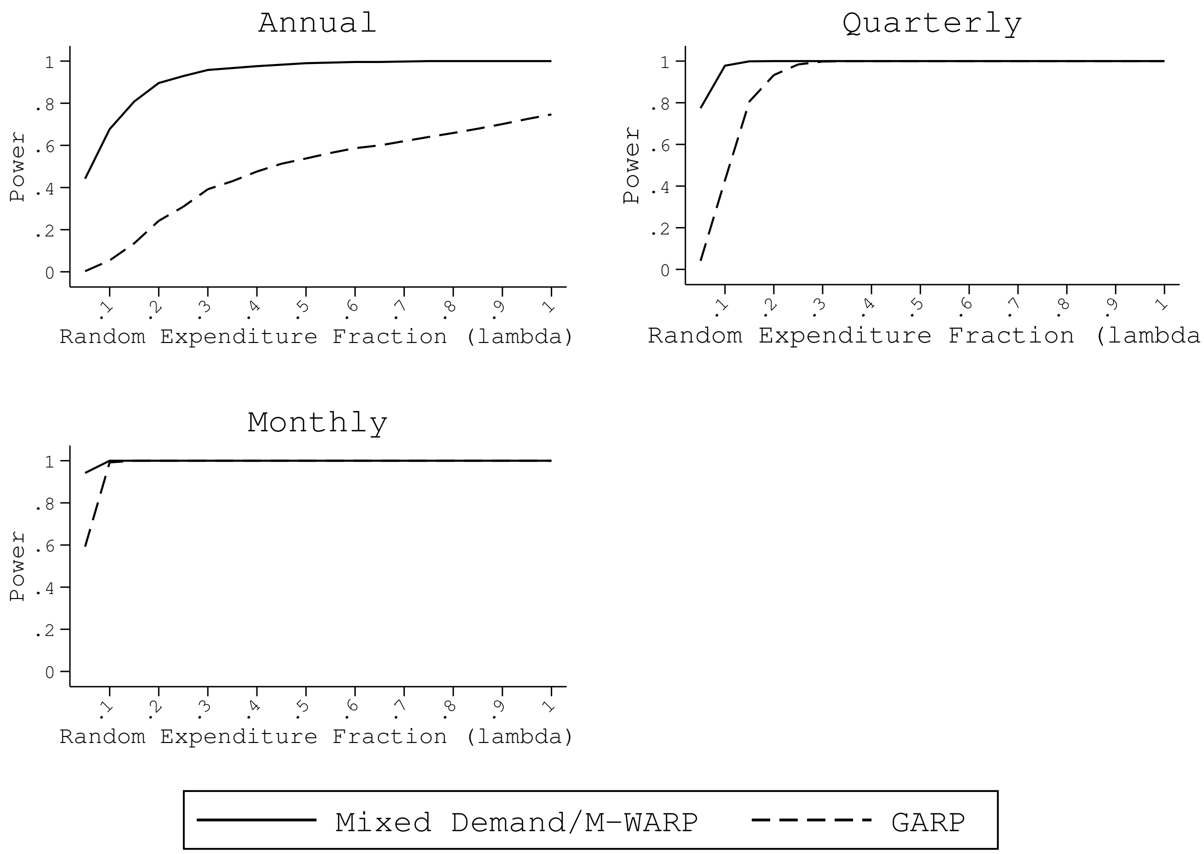

Figure 2: Power against partial uniformly random behavior (PURP index) for the mixed-demand model, M-WARP and GARP.

entire power curves against random expenditure allocation for the three models. First, I note that the PURP indices for the mixed-demand model and M-WARP coincide for each sample. For the annual data with relatively few observations, the power is considerably higher for the mixed-demand model and M-WARP than it is for GARP at lower levels of random expenditure allocations (i.e., for low values of λ). However, this difference becomes smaller as the sample size increases, and for the monthly data the difference is almost negligible. This suggests that the mixed-demand model and M-WARP has higher discriminatory power in smaller samples, which is perhaps as expected since the power should increase with sample size.

Predictive success. So far, we have compared the mixed-demand model, M-WARP and GARP in terms of goodness-of-fit and discriminatory power. Beatty and Crawford (2011) suggests to combine these two (often inversely related) performance measures into a single metric. Building on an original idea of Selten (1991), Beatty and Crawford (2011) suggest to assess the empirical performance of a model using a notion of predictive success (PS). This measure is computed as difference between the pass rate (either 1 or 0) and

efficiency values of 0.813, 0.806 and 0.801 for the annual, quarterly and monthly data, respectively.

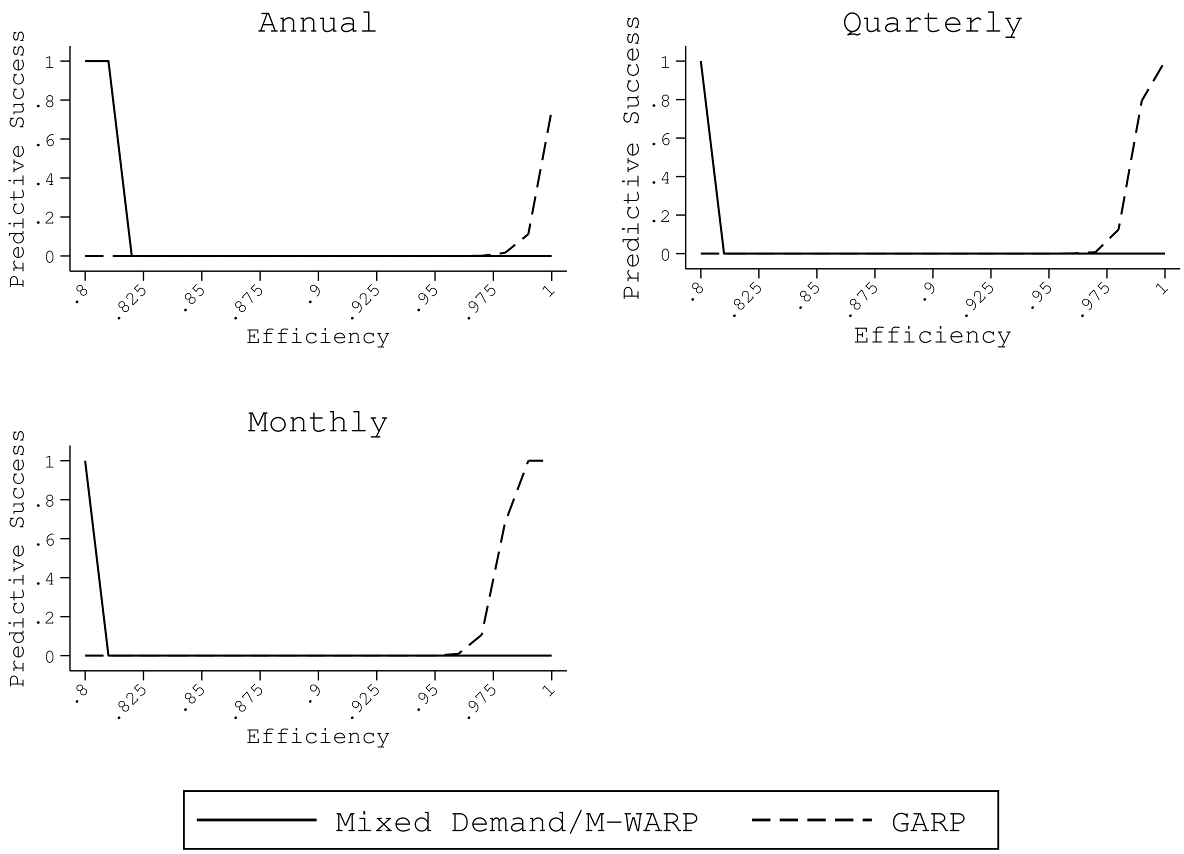

Figure 3: Predictive success for the mixed-demand model, M-WARP and GARP at different levels of efficiency.

1 minus the power, in which case $-1 \leq \text{PS} \leq 1$. Negative values ($\text{PS} < 0$) suggest that the model inadequately describes the data, that is, the model provides a poor fit (pass rate is zero) even though the model's power is low (i.e., the model is difficult to reject empirically). Conversely, a high and positive predictive success value points to a potentially useful model: it is able to explain the observed consumption behavior (i.e., pass rate equals 1) while its power is high (i.e., the model would rapidly be rejected in case of random behavior). Finally, a predictive success of zero, or approximately zero, indicates that the theory in question performs about as well as a theory that explains consumer behavior as purely uniformly random.

For every sample, I calculate the predictive success of the mixed-demand model, M-WARP and GARP at different levels of efficiency (starting at 0.8 and ending at 1, with increments 0.01). Figure 3 presents these results by plotting the predictive success against efficiency. Since the goodness-of-fit and power are the same for the mixed-demand model and M-WARP, these conditions will also have the same predictive success. The predictive success for these two models are high and close to 1 at efficiency levels close to 0.8 but rapidly drops to zero. For any sample, however, it is equal to 1 for some efficiency level equal to 0.8 or higher. In contrast, the predictive success for GARP is zero up to about 0.98, after which it rapidly increases. For the annual, quarterly and monthly samples, the

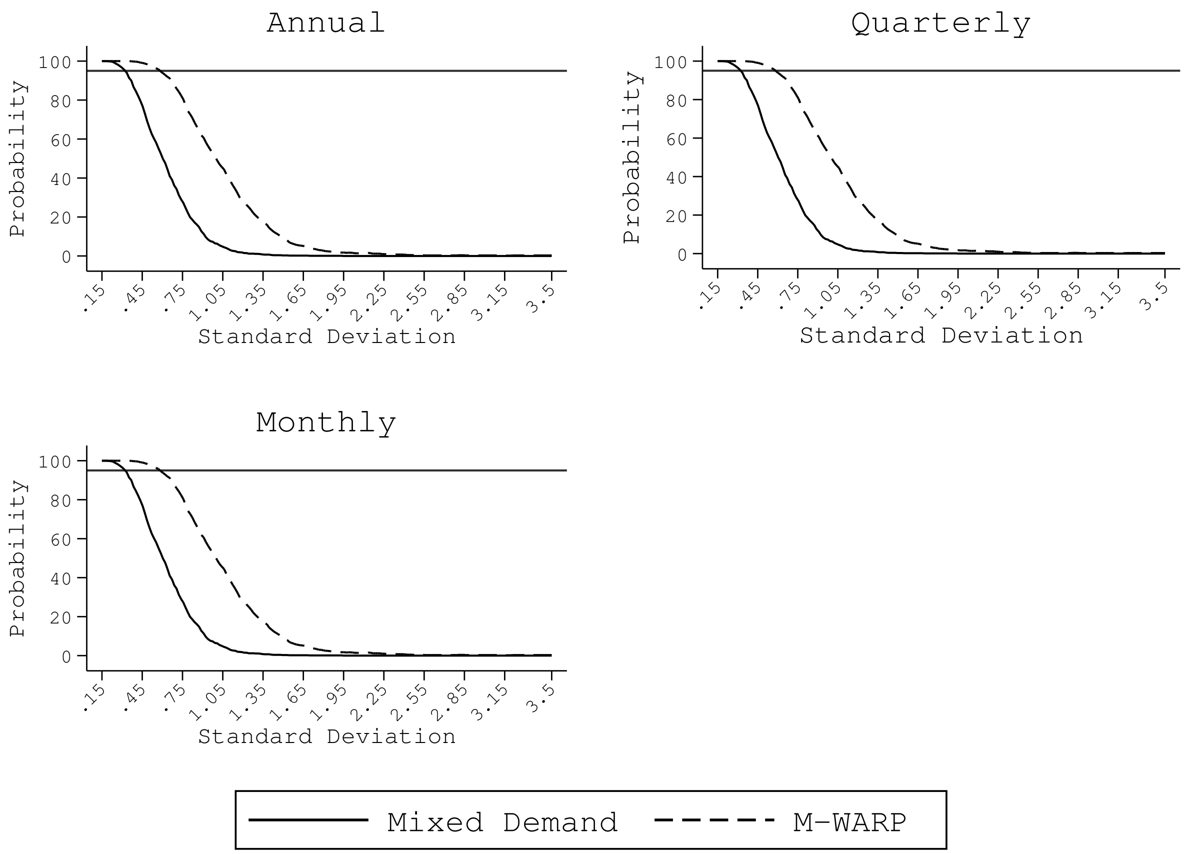

Figure 4: Rejection rates in percent of the null hypothesis (11) for the mixed-demand model and M-WARP. The horizontal solid line represents the 95$^{\text{th}}$ percentile, i.e., the 5% nominal significance level.

predictive success for GARP is equal to 0.75, 1.00 and 1.00 at full efficiency, respectively.

Taking an efficiency level of 0.8 as reasonable, it appear from the analyses of power and predictive success that the mixed-demand model is the preferred model for the annual data with a relatively short sample length. For the quarterly data and given that the Bronars power index is considerably higher for the mixed-demand model than GARP, one may also argue that the mixed-demand model perform slightly better than GARP. For the monthly data, however, with a relatively large number of observations, the power is very similar for the mixed-demand model and GARP, which would perhaps make the direct utility maximization model the preferred model in terms of predictive success.

5.4 Measurement errors

As a final exercise, I implement the measurement error procedure outlined in Section 4.4. Since the data in every sample satisfies GARP, I focus on the mixed-demand model and M-WARP.[20] I implement the procedure for an equally spaced grid of standard deviations

[20] Appendix A adapts the measurement error procedure in Section 4.4 to test M-WARP.

starting at 0.15 and ending at 3.5. Figure 4 plots the percentage rejection rates of the null hypothesis (11) at every point in the grid. The horizontal solid line represents the 95$^{\text{th}}$ percentile. Thus, a rejection rate above this line implies a rejection of the null hypothesis (11) at the 5% nominal significance level.

The lowest standard deviation at which the null cannot be rejected at the 5% nominal significance level for the mixed-demand model is about 0.32% for the annual data, 0.26% for the quarterly data and 0.24% for the monthly data. If one's prior belief is that the standard deviation of measurement errors in prices is lower, then the null is rejected, implying that the "true" data cannot be rationalized by a well-behaved mixed-demand model. If, however, one's belief is that the standard deviation of the errors in prices is higher, then the null of rationalizability is not rejected. In these type of data, given the possibility of aggregation errors and the potential bias in the construction of price indices, I would argue that an amount of error of about 0.3% appear low. Thus, for these data, the null that the "true" data can be rationalized by a well-behaved mixed-demand model should not be rejected.

As seen from Figure 4, the lowest standard deviation at which the null cannot be rejected at the 5% nominal significance level for M-WARP is slightly higher than for the mixed-demand model, and about 0.56% for the annual data, 0.46% for the quarterly data and 0.42% for the monthly data. Again, give the type of data, these numbers seem low enough not to reject the null of rationalizability.

6 Conclusions

This paper has: (1) provided a revealed preference characterization of Samuelson (1965) mixed-demand model, and (2) provided the first nonparametric empirical application of the mixed-demand model (Nonparametric here means that the direct and indirect utility functions are free from any functional assumptions). A simple testable condition was derived that can be implemented using computationally effective linear programming techniques. These results were also generalized to provide a simple testable condition under the assumption that observed data contains measurement errors.

I see a wide variety of potential follow-up studies based on the results in this paper. For example, the methods introduced here can be straightforwardly extended to also recover preferences following the ideas in Varian (1982). In particular, being based on revealed preference methods, there are usually many types of preferences that will rationalize data consistent with the rationalizability condition in Theorem 3. This exercise would aim at constructing inner and outer bounds for the indifference curves passing an arbitrary (not necessarily observed) block of optimally determined quantities and expenditure-normalized prices.

The methods proposed here can also be used to perform counterfactual analysis. Given some predetermined income and price- and quantity-vectors, the methods in Varian (1982) can be used to find optimally determined quantities and expenditure-normalized prices under the counterfactual scenario.

Given that some very popular parametric demand models does not have a mixed-demand representation (as discussed in the Introduction), it seem that the theory and methods developed in this paper could be a good alternative to researchers wanting to apply the mixed-demand model in empirical demand analysis.

References

Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International Economic Review*, **8**, 67–77.

— (1972). Efficiency estimation of production functions. *International Economic Review*, **13**, 568–598.

Banks, J., Blundell, R. and Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics*, **79**, 527–539.

Barten, A. P. (1992). The estimation of mixed demand systems. In R. Bewley and T. V. Hoa (eds.), *Contributions to Consumer Demand and Econometrics: Essays in Honour of Henri Theil*, London: Palgrave Macmillan.

— and Bettendorf, L. (1989). Price formation of fish: An application of an inverse demand system. *European Economic Review*, **33**, 1509–1525.

Beatty, T. K. M. and Crawford, I. A. (2011). How demanding is the revealed preference approach to demand. *American Economic Review*, **101**, 2782–2795.

Bronars, S. G. (1987). The power of nonparametric tests of preference maximization. *Econometrica*, **55**, 693–698.

Brown, D. J. and Shannon, C. (2000). Uniqueness, stability, and comparative statics in rationalizable Walrasian markets. *Econometrica*, **68**, 1529–1540.

Brown, M. G. and Lee, J.-Y. (2006). A mixed rotterdam demand system. In M. T. Holt and J.-P. Chavas (eds.), *Essays in Honor of Stanley R. Johnson*, Berkley: Berkley Electronic Press.

Carroll, C. D., Crossley, T. F. and Sabelhaus, J. (2015). Introduction to "Improving the measurement of consumer expenditures". In C. D. Carroll, T. F. Crossley and J. Sabelhaus (eds.), *Improving the Measurement of Consumer Expenditures*, Chicago: University of Chicago Press.

Chavas, J.-P. (1984). The theory of mixed demand functions. *European Economic Review*, **24**, 321–344.

— and Cox, T. L. (1987). *Non-parametric demand analysis: A duality approach*. Staff paper no. 265, Department of Agricultual Economics, University of Wisconsin-Madison.

Christensen, L. R., Jorgenson, D. W. and Lau, L. J. (1975). Transcendental logarithmic utility functions. *The American Economic Review*, **65**, 367–383.

Cunha-e-Sá, M. A. and Ducla-Soares, M. M. (1999). Specification tests for mixed demand systems with an emphasis on combining contingent valuation and revealed data. *Journal of Environmental Economics and Management*, **38**, 215–233.

Deaton, A. and Muellbauer, J. (1980a). An almost ideal demand system. *The American Economic Review*, **70**, 312–326.

— and — (1980b). *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

Demuynck, T. and Hjertstrand, P. (2019). Samuelson's approach to revealed preference theory: Some recent advances. In R. Anderson, W. Barnett and R. Cord (eds.), *Paul Samuelson: Master of Modern Economics*, London: Palgrave MacMillan, pp. 193–227.

Diewert, E. W. (1973). Afriat and revealed preference theory. *Review of Economic Studies*, **40**, 419–425.

Fleissig, A. and Whitney, G. A. (2005). Testing the significance of violations of Afriat's inequalities. *Journal of Business and Economic Statistics*, **23**, 355–362.

Heufer, J. and Hjertstrand, P. (2017). Homothetic efficiency: Theory and applications. *Journal of Business and Economics Statistics*, **37** (2), 235–247.

Hjertstrand, P. (2021). Power against random expenditure allocation for revealed preference tests. *Journal of Economic Behavior and Organization*, **188**, 36–45.

— and Swofford, J. L. (2012). Revealed preference tests for consistency with weakly separable indirect utility. *Theory and Decision*, **72**, 245–256.

Jones, B. and Edgerton, D. (2009). Testing utility maximization with measurement errors in the data. *Advances in Econometrics*, **24**, 199–236.

Krishnan, N., Olivieri, S. and Ramadan, R. (2019). Estimating the welfare costs of reforming the Iraq public distribution system: A mixed demand approach. *The Journal of Development Studies*, **55**, 91–106.

Matsuda, T. (2004). Incorporating generalized marginal budget shares in a mixed demand system. *American Journal of Agricultural Economics*, **86**, 1117–1126.

McLaren, K. and Wong, K. (2005). Specification and estimation of regular inverse demand systems: A distance function approach. *American Journal of Agricultural Economics*, **87**, 823–834.

Moschini, G. and Rizzi, P. L. (2006). Coherent specification of a mixed demand system: The Stone-Geary model. In M. T. Holt and J.-P. Chavas (eds.), *Essays in Honor of Stanley R. Johnson*, Berkley: Berkley Electronic Press.

— and — (2007). The normalized quadratic mixed demand system. *American Journal of Agricultural Economics*, **89**, 1034–1045.

— and Vissa, A. (1993). Flexible specification of mixed demand systems. *American Journal of Agricultural Economics*, **75**, 1–9.

Ramadan, R. and Thomas, A. (2011). Evaluating the impact of reforming the food subsidy program in Egypt: A mixed demand approach. *Food Policy*, **36**, 638–646.

Roy, R. (1947). La distribution du revenu entre les divers biens. *Econometrica*, **15**, 205–225.

Samuelson, P. A. (1965). Using full duality to show that simultaneously additive direct and indirect utilities implies unitary price elasticity of demand. *Econometrica*, **33**, 781–796.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences*, **21**, 153–167.

Tabarestani, M., Jr., W. R. K. and Marzoughi-Ardakani, H. (2017). An analysis of the US shrimp market: A mixed demand approach. *Marine Resource Economics*, **32**, 411–429.

Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica*, **50** (4), 945–973.

— (1985). Non-parametric analysis of optimizing behaviour with measurement error. *Journal of Econometrics*, **30**, 445–458.

— (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics*, **46**, 125–140.

Wong, K. K. G. and Park, H. (2007). The use of conditional cost functions to generate estimable mixed demand systems. *American Journal of Agricultural Economics*, **89**, 273–286.

Appendix

A AEI and measurement errors for M-WARP

AEI. The AEI for M-WARP is constructed in the same way as for the mixed-demand inequalities. Note that the right-hand side of the M-WARP condition equals total expenditure (normalized to unity). Consider the following "relaxed" version of M-WARP, where total expenditure, $\mathbf{q}_t\mathbf{y}_t + \mathbf{z}_t\mathbf{w}_t = 1$, at every observation is multiplied with a scalar $e \in [0, 1]$:

$$\mathbf{q}_t\mathbf{y}_s + \mathbf{z}_s\mathbf{w}_t \geq e. \tag{17}$$

The AEI is defined as the largest e such that the observed data satisfies this condition. Analogous to the AEI for the mixed-demand model, the AEI for M-WARP is calculated by implementing a binary search algorithm.

Measurement errors. Consider the following optimization problem:

$$\begin{gathered}\Phi = \min_{\phi} \phi \qquad \text{subject to} \\ \mathbf{q}_t\left(\mathbf{y}_t - \mathbf{y}_s\right) + \mathbf{w}_t\left(\mathbf{z}_t - \mathbf{z}_s\right) \leq \phi\end{gathered} \tag{18}$$

Analogous to the mixed-demand model, this optimization problem can be used to derive a statistical decision rule to test the hypothesis (11) under the assumption of the multiplicative Berkson measurement error model (9). This decision rule is based on the following theorem.

Theorem 5 *Consider the optimization problem* (18)*. Under H_0 in* (11)*, it holds that:*

$$\Phi \leq C,$$

where:

$$C = \max_{s,t\in\mathbb{T}}\{(\overline{\mathbf{q}}_t - \mathbf{q}_t)(\mathbf{y}_s - \mathbf{y}_t) + \mathbf{w}_t((\overline{\mathbf{z}}_s - \mathbf{z}_s) - (\overline{\mathbf{z}}_t - \mathbf{z}_t))\}. \tag{19}$$

As in Theorem 4, the null is rejected whenever $\Phi > C_{1-\alpha}$, where α denotes the significance level and $C_{1-\alpha}$ denotes the $(1-\alpha)\times 100$th percentile of the distribution of the random variable C in (19). The empirical distribution of C can be simulated as described in Section 4.4.

Finally, a word on solving the problem (18). Note that the constraints in (18) can be written as:

$$\mathbf{q}_t\mathbf{y}_s + \mathbf{z}_s\mathbf{w}_t \geq \mathbf{q}_t\mathbf{y}_t + \mathbf{z}_t\mathbf{w}_t - \phi = 1 - \phi$$

Thus, by the AEI in (17), we have that $\phi = 1 - e$, which implies that the decision rule in Theorem 5 can be written as $1 - \text{AEI} \leq C$, where C is defined in (19). Hence, if one calculates the AEI for M-WARP in empirical applications, then it is not necessary to also calculate ϕ.

B Proofs

Proof of Theorem 3

Necessity. By Afriat's Theorem, the existence of a continuous, strictly increasing and concave (and rationalizing) direct utility function implies that there exist numbers U_t and λ_t for all $t \in \mathbb{T}$ such that the Afriat inequalities (in the first row) hold. Analogously, from Hjertstrand and Swofford (2012), the existence of a continuous, strictly decreasing and convex (and rationalizing) indirect utility function implies that there exist numbers V_t and μ_t for all $t \in \mathbb{T}$ such that the indirect Afriat inequalities (in the second row) hold. By Definition 3, the optimum occur at $0 = u(\mathbf{y}^\star, \mathbf{w}) - v(\mathbf{q}, \mathbf{z}^\star)$. Hence, $u(\mathbf{y}_t^\star, \mathbf{w}_t) = v(\mathbf{q}_t, \mathbf{z}_t^\star)$ for all $t \in \mathbb{T}$. Defining $U_t = u(\mathbf{y}_t^\star, \mathbf{w}_t)$ and $V_t = v(\mathbf{q}_t, \mathbf{z}_t^\star)$, we obtain the equality in the third row. Finally, the first-order conditions of the individual utility and indirect utility maximization problems are:

$$\begin{aligned} 0 &= \nabla_{\mathbf{y}} u(\mathbf{y}, \mathbf{w}) - \lambda \mathbf{q}, \\ 0 &= -\nabla_{\mathbf{z}} v(\mathbf{q}, \mathbf{z}) - \mu \mathbf{w}, \end{aligned}$$

But by the first-order conditions from the mixed-demand model (2), we must have $\lambda = \mu$. Evaluating this in every time period $t \in \mathbb{T}$ yields the equality in the fourth row.

Sufficiency. Suppose that the second condition holds (i.e., condition (3) holds). For all $(\mathbf{y}, \mathbf{z})$ and any fixed $(\mathbf{w}, \mathbf{q})$ define the functions:

$$\begin{aligned} u(\mathbf{y}, \mathbf{w}) &= \min_{\{s \in \mathbb{T}\}} \{U_s + \lambda_s \mathbf{q}_s (\mathbf{y} - \mathbf{y}_s)\}, \\ -v(\mathbf{q}, \mathbf{z}) &= \min_{\{s \in \mathbb{T}\}} \{-V_s + \mu_s \mathbf{w}_s (\mathbf{z} - \mathbf{z}_s)\}. \end{aligned}$$

By Afriat's theorem, the function $u(\mathbf{y}, \mathbf{w})$ is continuous, strictly increasing and concave in $\mathbf{y}$. Rewrite $v(\mathbf{q}, \mathbf{z})$ as:

$$\begin{aligned} v(\mathbf{q}, \mathbf{z}) &= -\min_{\{s \in \mathbb{T}\}} \{-V_s + \mu_s \mathbf{w}_s (\mathbf{z} - \mathbf{z}_s)\} \\ &= \max_{\{s \in \mathbb{T}\}} \{V_s + \mu_s \mathbf{w}_s (\mathbf{z}_s - \mathbf{z})\}. \end{aligned}$$

Hjertstrand and Swofford (2012) showed that $v(\mathbf{q}, \mathbf{z})$ is continuous, strictly decreasing and convex in $\mathbf{z}$. Hence, it suffices to show that the mixed-demand model rationalizes the data $\mathbb{D} = \{(\mathbf{q}_t, \mathbf{z}_t); (\mathbf{y}_t, \mathbf{w}_t)\}_{t \in \mathbb{T}}$. For all $(\mathbf{y}, \mathbf{z})$ and any observation $(\mathbf{w}_t, \mathbf{q}_t)$ such that $\mathbf{q}_t \mathbf{y} + \mathbf{z} \mathbf{w}_t \leq 1$:

$$\begin{aligned} u(\mathbf{y}, \mathbf{w}_t) - v(\mathbf{q}_t, \mathbf{z}) &= u(\mathbf{y}, \mathbf{w}_t) + (-v(\mathbf{q}_t, \mathbf{z})) \\ &\leq (U_t + \lambda_t \mathbf{q}_t (\mathbf{y} - \mathbf{y}_t)) + (-V_t + \mu_t \mathbf{w}_t (\mathbf{z} - \mathbf{z}_t)) \\ &= (U_t + (-V_t)) + \lambda_t \mathbf{q}_t (\mathbf{y} - \mathbf{y}_t) + \mu_t \mathbf{w}_t (\mathbf{z} - \mathbf{z}_t) \\ &= (U_t - V_t) + \psi_t (\mathbf{q}_t (\mathbf{y} - \mathbf{y}_t) + \mathbf{w}_t (\mathbf{z} - \mathbf{z}_t)) \\ &= (U_t - V_t) + \psi_t ((\mathbf{q}_t \mathbf{y} + \mathbf{z} \mathbf{w}_t) - (\mathbf{q}_t \mathbf{y}_t + \mathbf{z}_t \mathbf{w}_t)) \\ &\leq U_t - V_t \\ &= 0, \end{aligned}$$

where $\psi_t = \lambda_t = \mu_t$ and $U_t = V_t$ holds for all $t \in \mathbb{T}$ from condition (3). Thus, the mixed-demand model rationalizes the data $\mathbb{D}$. ■

Proof of Theorem 4

Under H_0 in (11) and by Theorem 3, there exist numbers F_t and $\psi_t > 0$ such that the following inequalities hold:

$$\begin{aligned} F_s - F_t &\leq \psi_t \overline{\mathbf{q}}_t \left(\mathbf{y}_s - \mathbf{y}_t\right), \\ -F_s + F_t &\leq \psi_t \mathbf{w}_t \left(\overline{\mathbf{z}}_s - \overline{\mathbf{z}}_t\right), \end{aligned}$$

for all observations $s, t \in \mathbb{T}$. For every $s, t \in \mathbb{T}$, adding and subtracting $\psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right)$ from the right-hand side of the first set of inequalities gives:

$$\begin{aligned} F_s - F_t &\leq \psi_t \overline{\mathbf{q}}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) \\ &= \psi_t \overline{\mathbf{q}}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) + \psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) - \psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) \\ &= \psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) + \psi_t (\overline{\mathbf{q}}_t - \mathbf{q}_t)(\mathbf{y}_s - \mathbf{y}_t). \end{aligned}$$

Thus,

$$\begin{aligned} F_s - F_t - \psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) &\leq \psi_t (\overline{\mathbf{q}}_t - \mathbf{q}_t)(\mathbf{y}_s - \mathbf{y}_t) \\ &\leq \psi_t C_1, \end{aligned}$$

where C_1 is defined in (14). Analogously, for every $s, t \in \mathbb{T}$, adding and subtracting $\psi_t \mathbf{w}_t \left(\mathbf{z}_s - \mathbf{z}_t\right)$ from the right-hand side of the second set of the inequalities gives:

$$-F_s + F_t - \psi_t \mathbf{w}_t \left(\mathbf{z}_s - \mathbf{z}_t\right) \leq \psi_t C_2,$$

where C_2 is defined in (15). Since $C = \max\{C_1, C_2\}$, we have:

$$\begin{aligned} F_s - F_t - \psi_t \mathbf{q}_t \left(\mathbf{y}_s - \mathbf{y}_t\right) &\leq \psi_t C, \\ -F_s + F_t - \psi_t \mathbf{w}_t \left(\mathbf{z}_s - \mathbf{z}_t\right) &\leq \psi_t C. \end{aligned}$$

Because Ψ is the value that solves the problem (12) it holds that $\Psi \leq C$. ■

C Additional results

Table A1: Descriptive statistics of prices.

	Annual ($T = 19$)			Quarterly ($T = 76$)			Monthly ($T = 228$)		
Commodity	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Cereals	93.13	78.90	105.61	93.14	78.66	105.92	93.14	78.37	106.16
Bakery products	92.90	70.45	110.91	92.91	69.99	111.28	92.91	69.76	111.69
Beef and veal	102.40	63.81	146.63	102.44	62.75	148.28	102.45	62.55	148.59
Pork	102.87	80.49	125.10	102.90	79.55	128.06	102.90	78.95	128.61
Other meats	97.43	76.03	117.13	97.43	75.75	117.92	97.44	75.18	118.11
Poultry	97.17	77.29	116.66	97.17	77.20	117.73	97.18	76.75	118.06
Fish and seafood	96.92	77.03	120.03	96.93	76.12	121.18	96.93	75.86	121.39
Fresh milk	102.29	83.38	121.51	102.33	80.32	122.25	102.34	79.93	122.88
Processed diary products	97.57	80.08	112.94	97.59	79.64	114.45	97.59	79.48	114.82
Eggs	98.66	67.50	149.88	98.93	63.64	168.09	98.98	61.57	173.28
Fats and oils	95.12	73.27	115.94	95.14	72.66	116.14	95.14	71.97	116.35
Fruit (fresh)	97.27	79.59	110.92	97.28	76.98	112.27	97.29	75.14	112.85
Vegetables (fresh)	95.08	69.95	107.29	95.10	68.76	110.56	95.12	68.14	113.74
Processed fruits & vegetables	90.61	70.32	106.48	90.62	69.79	107.01	90.62	69.74	107.56
Sugar and sweets	94.43	77.34	109.75	94.44	76.54	110.42	94.44	76.42	110.80
Unclassified food	95.90	81.22	109.83	95.90	80.75	110.16	95.90	80.74	110.48

Table A2: Descriptive statistics of budget shares.

	Annual ($T = 19$)			Quarterly ($T = 76$)			Monthly ($T = 228$)		
Commodity	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Cereals	0.07	0.06	0.07	0.07	0.06	0.07	0.07	0.06	0.07
Bakery products	0.14	0.13	0.15	0.14	0.13	0.15	0.14	0.13	0.15
Beef and veal	0.06	0.04	0.07	0.06	0.04	0.08	0.06	0.04	0.08
Pork	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.05
Other meats	0.04	0.04	0.05	0.04	0.04	0.05	0.04	0.04	0.05
Poultry	0.07	0.07	0.08	0.07	0.07	0.08	0.07	0.07	0.08
Fish and seafood	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Fresh milk	0.03	0.03	0.04	0.03	0.03	0.04	0.03	0.03	0.04
Processed diary products	0.06	0.06	0.07	0.06	0.06	0.07	0.06	0.06	0.07
Eggs	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.01	0.02
Fats and oils	0.03	0.02	0.02	0.03	0.02	0.03	0.03	0.02	0.03
Fruit (fresh)	0.05	0.02	0.06	0.05	0.04	0.06	0.05	0.04	0.06
Vegetables (fresh)	0.07	0.06	0.07	0.07	0.06	0.08	0.07	0.06	0.08
Processed fruits & vegetables	0.04	0.04	0.05	0.04	0.04	0.05	0.04	0.04	0.05
Sugar and sweets	0.07	0.06	0.08	0.07	0.06	0.08	0.07	0.06	0.08
Unclassified food	0.20	0.18	0.20	0.20	0.18	0.20	0.20	0.18	0.20