This article has been published as: Boysen, Ole (2019). When does specification or aggregation across consumers matter for economic impact analysis models? An investigation into demand systems. *Empirical Economics* **56**(1), 137–172. DOI: 10.1007/s00181-017-1353-z

When does specification or aggregation across consumers matter for economic impact analysis models? An investigation into demand systems*

Ole Boysen[‡]

Abstract

Economic impact analysis simulation models frequently rely on some kind of representation of consumption behavior. However, the sensitivity of such results with respect to the choices of the specification and the level of aggregation across consumers has not yet been thoroughly examined. We exploit a unique dataset to simulate various stereotypical scenarios and investigate the influence of the choice between six demand system specifications and household-level versus national-level models on several outcome measures. We find that both choices have a large influence on simulation results and thus on policies deduced therefrom. Our results point to pragmatic recommendations for various settings.

Keywords: Demand systems, specification, aggregation, calibration, economic impact analysis
JEL classification: C69, D12

1 Introduction

Economic impact analysis simulation models frequently rely on some kind of representation of consumption demand behavior. This applies, for instance, to studies in the domains of public and development economics or international trade, such as assessments of the first-order impacts of food price changes on poverty (e.g., de Janvry and Sadoulet 2010; Simler 2010; Wodon and Zaman 2010), the partial equilibrium welfare effects of tax reforms (e.g., Banks et al. 1996) or the general equilibrium effects of agricultural policy reforms in individual countries (e.g., Boysen et al. 2016) or of multilateral trade agreements (e.g., Anderson et al. 2006). However, little is known about the sensitivity of the analyses' results to the specific demand representations chosen and, in turn, of the policies designed on that basis.

This article investigates this issue with respect to two important dimensions of choices for demand representations, namely the model specification – such as the Linear Expenditure System (LES) or the Almost Ideal Demand System (AIDS) – and the level of aggregation across

[‡]Ole Boysen, e-mail: ole.boysen@ucd.ie, School of Agriculture and Food Science, University College Dublin, Ireland.

*I would like to thank three anonymous referees for their constructive comments which greatly helped to improve this article. I am indebted to Sherman Robinson for many insightful discussions.

consumers ranging from representing individual consumers to representing a single, aggregate national consumer.

Both dimensions have been researched extensively. Banks et al. (1997), Cranfield et al. (2003), Fisher et al. (2001), Klevmarken (1979), Parks (1969) or Pollak and Wales (1995), for instance, find that the specification choice matters with respect to the theoretical properties and the empirical fit of econometrically estimated demand systems to observed data. Another strand of literature highlights the importance of the specification choice when estimating elasticities for the purpose of analyzing optimal commodity tax design or commodity tax reform, particularly if reforms are non-marginal (see, e.g., Decoster and Schokkaert 1990; Madden 1996; Ray 1986, 1999). Likewise, the aggregation choice and the bias it creates has been argued to matter from theoretical and empirical perspectives as is comprehensively reviewed in Blundell and Stoker (2007).

Nevertheless, it remains unclear what practical relevance these two choices and the corresponding theoretical and econometric insights have for economic impact analysis simulations, as exemplified by the studies previously mentioned. This is surprising since the application of models for both *ex-ante* and *ex-post* economic impact analysis has become widespread in both academic research and policy arenas in support of policy decision-making. The economic models applied might employ demand systems either as stand-alone models or as integral parts of larger, more complex economic model systems, such as Computable General Equilibrium (CGE) or microsimulation models.

The economist developing an impact analysis model is confronted with the choices of the demand model specification and the level of aggregation across consumers and correspondingly with the trade-off between the need to mimic observed real-world behavior and consistency with demand theory on the one hand and manageable parameterization and numerical solvability of the model on the other.

The less restrictive the specification and the higher the number of its parameters, the better it can be fitted to observed behavior. But it also requires more data and it is more challenging to determine a set of parameter values which is consistent with demand theory. Greater complexity frequently implies higher non-linearity of the equation system and thereby increases the difficulty for numerical algorithms to find the (optimal) solution to the impact analysis model. Similarly, the lower the level of aggregation is – i.e., the higher the number of consumers that are individually represented by demand systems – the better the representation of observed behavior in case of differentially affected households. But the more individual consumers are included, the more data and parameters are needed and the larger the model in terms of variables and equations. The memory and computer processing power required by a numerical optimization algorithm often depends not linearly but exponentially on the size of the model. Thus, data availability considerations and the need to numerically solve the model might lead the model developer to resort to pragmatic choices of a specification and an aggregation level.

The detail and complexity of simulation models has been steadily increasing with the rising availability of more and higher quality data, computing power, and advances in solution methods. Nevertheless, as far as demand systems are concerned, economic simulation models continue to predominantly adopt early popular specifications, such as the Cobb-Douglas (CD), the Constant Elasticity of Substitution (CES), and in particular, the LES, with only few exceptions. Given the range of specification options, there is surprisingly little research on the implications of these choices and the practical relevance of the theoretical and econometric properties of the various specifications for scenario simulation outcomes in general.

There are a few studies which consider the effects of specification choices in the context of CGE models. Yu et al. (2004) compare four specifications of aggregate demand systems (LES, CD, Constant Difference of Elasticities [CDE], and An Implicit Direct Additive Demand System [AIDADS]), integrated into a global CGE model named GTAP, to analyze their suitability for projecting food demand, in particular for developing countries, when populations become wealthier. Gohin (2005) also compares three aggregate specifications (CD, LES, and a specific "flexible model") with respect to some aggregate outcome variables within a CGE framework and indicates the relevance of the choice. Finally, de Boer (2010) finds the outcomes of a national CGE model to be sensitive with respect to the choice between two specifications (LES and Indirect Addilog System [IAS]). All of these consider only aggregate national demand systems within proprietary and complex models under specific scenarios and find the specification choice to be more or less relevant for particular outcome measures.

The second issue analyzed in this article, and which has also not been rigorously examined in the literature, is what bias the use of aggregated demand system models introduces to outcome measures in a practical simulation setting, in particular, when incomes change differentially across households. Other measures, such as poverty indices, are naturally only measurable with a high degree of household detail.

Hence, in contrast to the studies mentioned above which deal with highly aggregated, representative households, here we focus on individual households from a household survey which are subject to much larger variation in prices, incomes, and elasticities and investigate the influence of household aggregation.

This paper aims to provide a generalizable contribution to the issues of choosing a demand system specification and the aggregation level. More specifically, we are interested in the effect of these choices on a multitude of outcome measures in simulations of stereotypical scenarios. To this end, we conduct scenario simulations based solely on demand system models without potential contamination through other model components as is the case, for example, in a CGE framework. We examine under which circumstances biases occur for which outcome measure and in what magnitude and then draw generalizable conclusions.

The paper proceeds as follows. The background and general methodology is explained in section 2. Section 3 presents the mathematical models of the demand system specifications and their properties. Section 4 describes the household demand data which forms the basis for the experiments. In section 5, the calibrated models are utilized for the simulation of various price and income scenarios to assess their performances. Finally, section 6 summarizes and concludes.

2 Background and Methodology

Consumption demand models, as typically used for econometric estimation and impact analysis models, are built on the presumption of utility-maximizing consumers. Given some utility function that complies with microeconomic theory, i.e., it is consistent with the axioms of choice and monotonically increasing in quantities and quasi-concave (Caves and Christensen 1980), maximizing the consumer's utility subject to the budget constraint allows the derivation of the demand functions which determine the optimal consumption bundle for a specific combination of prices and income.

Econometric estimation of a demand model requires specifying some functional form and applying appropriate parameter restrictions to estimate the conditional mean demand behavior across the sample observation combinations of prices and incomes. Depending on the chosen specification and the parameter restrictions, the estimated model then may or may not comply with several or all of the underlying theoretical assumptions for some or all price-income combinations of the sample, for a certain range of combinations or for the full range of the price-income space. However, non-compliance puts the very existence of the demand functions into question and renders all results calculated theoretically invalid so that, depending on the type of economic simulation model, the scenarios to be simulated, and the outcome measure, theory-consistent behavior is crucial even beyond the price-income region covered by the observations. Even though a small degree of violation of these assumptions could actually be tolerable for some types of economic simulation models if the influence on the results is small, this magnitude is, unfortunately, not known in advance.

A system of demand functions consistent with theory is called *regular* and fulfills the properties of adding up, homogeneity, symmetry, and negativity (Deaton and Muellbauer 1980a, pp. 43–46). Even if functions are regular in the initial prices-income point, they might lose this property when simulating a scenario and, as a consequence, might exhibit inconsistent demand behavior, such as irrational preferences or negative income budget shares for individual consumption items, and might cause numerical solution methods to fail (Perroni and Rutherford 1998). The combinations of prices and incomes for which a demand function is regular is called the *regular region* and functions which are regular over the entire price-income space are called *globally regular* (Caves and Christensen 1980).

Regularity comes at the expense of *flexibility*, i.e., the capability to represent an arbitrary combination of price and income elasticities at a given point of the price-income space (Caves and Christensen 1980). The flexibility of a demand function is determined by its number of independent parameters and the restrictions on the parameters' domains which, in turn, are necessary to ensure regularity over some region. Let n be the number of items in the demand system, then there are n income plus n own-price plus $(n^2 - n)$ cross-price elasticities, and thus $n + n^2$ elasticities in total. These are reduced by 1 if adding-up, by n if homogeneity, and by $(n^2 - n)/2$ if symmetry properties are enforced, resulting in $(n^2 + n)/2 - 1$ independent elasticities (Deaton and Muellbauer 1980a). This is the minimum number of independent parameters a demand system must feature to be flexible. Here, we define the *number of independent parameters* as the number of pieces of information required in addition to the budget shares, prices, and income to make the system of equations exactly determined when the demand system parameters and only these are treated as unknowns of the demand system of equations.[1] However, the definition of flexibility only considers elasticities locally, i.e., at one point in the price-income space, and thus neglects the higher-order curvature properties which determine how elasticities change with the levels of prices and income.

A characterization of demand functions with respect to their capability to represent the curvature of Engel curves, that is, the relationship between the share of income allocated to a particular item and the income level, is provided by the *rank* of a demand system, first defined by Gorman (1981) and further generalized by Lewbel (1991).[2] Rank one demand systems define budget shares independently of the income level so that the income is spent on the different consumption items in constant shares. Rank two includes budget share functions depending linearly on the income or the logarithm of income. Finally, rank three comprises, amongst others, functions which are quadratic in the (logarithm) of income. The necessity of rank three demand systems was found, for example, by Lewbel (1991) using non-parametric rank tests on his samples of US and UK datasets.[3]

More parsimonious demand functions in terms of parameters are preferred for econometric estimations as these yield stronger results given a limited number of observations. Likewise, there is a preference for parsimonious demand functions in simulation models where information, i.e., elasticities, is typically scarce. On the downside, such functions are limited in their capability to capture observed demand behavior realistically, which calls for more flexible and higher rank functional forms. To find the right trade-off in this spectrum is a context-dependent matter and cannot be resolved universally.

When selecting a functional form, a modeler of scenario simulations needs to consider, in addition to the above tensions, the parameterization of the functional forms, the computational limitations, and the scenarios and outcome measures. The parameterization is commonly limited through data availability (e.g., availability of elasticity estimates) and requires some method of calibration[4], that is, the fitting of the functional form to the observed demand point and the behavioral information (often demand elasticities) such that the parameterized demand function replicates the observed demand quantities at given prices and income.

The computational limitations are determined by the computer memory, processing power, and the numerical solutions methods at hand. The memory requirements increase with the numbers of variables and equations which, depending on the numerical solution method applied,

[1] More precisely, we define the *number of independent parameters* as the number of parameters minus the number of independent equations of the demand model specification. In the context of systems of equations, this is usually termed *degrees of freedom*, see, e.g., Sydsæter et al. (2012, p.449f.).

[2] Lewbel defined the rank "...to be the maximum dimension of the function space spanned by the Engel curves of the demand system." (Lewbel 1991, p.711).

[3] He also found that the required rank decreases to two if the observations at the tails of the income distribution are excluded.

[4] For an extensive discussion of the term calibration, see Dawkins et al. (2001).

often also increase the required computing time more than proportionally. Thus, the computational requirements increase with the number of consumers in the simulation model, each represented by an individual demand system. This argument favors aggregated, representative consumers. However, how adequately a representative consumer can represent the demand behavior of heterogeneous individual consumers, particularly if the income distribution changes, is questionable. Aggregate demand is a function of prices and individual consumers' incomes. Only under strong assumptions (identical and quasi-homothetic utility functions for all individuals) is aggregate demand a function of aggregate income (see, e.g., Mas-Colell et al. 1995). Moreover, greater flexibility and the higher rank of a demand function result in an increased non-linearity of the functions, thereby increasing the difficulty for numerical solution methods and hence the computing time. Finally, the numerical solution algorithm might evaluate the demand function for price-income vectors which fall outside the regular region, possibly causing the results to exhibit perverse behavior or even leading the solution process to fail (Perroni and Rutherford 1998). The likelihood of the latter can be decreased by ensuring a "sufficiently large" regular region with respect to the scenarios simulated. For these reasons, regularity might be not only a desirable but also a necessary property of the demand function for, at least, an extended region of the price-income space.

Early popular functional forms in simulation models are the CES function and its special case, the CD function, which are both globally regular. However, the CES is highly inflexible and of rank one (Barnett and Serletis 2008). IAS and CDE models are more flexible and are also globally regular. To further distinguish the remaining functional forms with respect to their regularity properties, Cooper and McLaren (1996) defined *effectively global regularity* as regularity for all combinations of prices and incomes that result in price index-deflated real incomes exceeding some chosen minimum value.[5] The class of effectively globally regular demand systems includes, e.g., the LES and AIDADS models. Nevertheless, where households' real incomes can be negatively affected, effectively globally regular systems need to be deliberately calibrated to create a sufficiently large regular region.[6] Most models found in the literature that apply economic simulation models are of rank two, including the inflexible LES, IAS, and CDE models and the flexible AIDS and Translog models. Finally, examples of flexible, rank three demand models are the Quadratic Expenditure System (QES), Quadratic Almost Ideal Demand System (QUAIDS), and AIDADS.

The area of tension created between the need to mimic real-world demand behavior on one hand and the computational and data limitations on the other is the object of this study. The choices of model specification and aggregation level might have a major influence on the simulation results for some scenarios and outcome measures while they are negligible for others. For instance, the level of aggregation might not matter for the total, national quantity demanded

[5]More formally, a demand system is effectively globally regular if there exists a linearly homogeneous price index $P(p)$ non-decreasing in prices p and the system satisfies the regularity properties for all real incomes $\frac{M}{P(p)}$ greater than some minimum level κ (Cooper and McLaren 1996).

[6]Wolff et al. (2010) have introduced an approach that ensures regularity over a freely defined region during the econometric estimation of a demand system. However, nothing guarantees that this property will survive the calibration of a demand function to the parameters estimated.

when household incomes uniformly increase by some percentage whereas it is crucial for assessing the redistribution of income between households.

The following setup is used to elucidate the trade-offs between specification and aggregation options for a number of stereotypical scenarios and important outcome measures. We start from individual household-level (HL) demand data and estimates for a flexible functional form, rank three demand system which has been estimated separately for the subsets of rural and urban households. Furthermore, we select for comparison six demand model specifications that are more or less popular in impact analysis models and range over the spectrum from inflexible to flexible and ranks one and two.

The idea is the following. An estimated flexible, rank three QUAIDS[7] captures the mean household demand behavior conditional on various household characteristics and expenditure levels and thus heterogeneity in household demand. The detailed household data from the QUAIDS estimation, including household characteristics, is used to calculate elasticities for individual households. Then, each demand model specification is calibrated to the household-specific demand elasticities, for each individual household from the survey data set separately, such that the parameterized model precisely replicates the household's observed budget shares at observed prices and income while also enforcing regularity properties. Note that this modification not only adapts the parameters to a different functional form but also incorporates the estimation residuals, i.e., the unexplained household-specific variation, thereby adding additional information beyond the parameter estimates.

The sets of demand system parameters utilized in this study have been calibrated using the information-theoretic, generalized cross-entropy (GCE) approach which fits a set of household-specific prior demand elasticities as closely as possible while incorporating information in the form of the imposed model specification, theoretical regularity conditions of consumer demand, and the observed household demand data. The resulting set of parameters for an individual household's demand system will reflect the estimated behavior, modified as little as possible to account for the information available.

Those calibrated *household-level* (HL) demand systems are adopted to simulate a number of stereotypical income and price shock scenarios and compare the influence of the specification choices on outcomes in terms of national quantity demanded, equivalent income, poverty, inequality, and cost-of-living indices to the assumed "true" household demand behavior. Contrary to what might seem intuitive at first thought, we have chosen to not use calibrated QUAIDS models as the "true" behavior reference. Indeed, a QUAIDS specification was used to derive the prior elasticity data for this study. However, the calibration of QUAIDS models would require data in addition to what is usually available in the assumed simulation modeling context. More specifically, the QUAIDS includes a term quadratic in income and any meaningful calibration of the associated parameter would necessitate, in addition to a single consumption data point and a set of price and income elasticities, information about how the slope of demand changes with increasing income. Instead, as measures of fit have shown that the calibrated AIDS mod-

[7] See Appendix A for a formal description.

els match the prior elasticities very closely, we assume that the HL AIDS models represent the "true" household demand behavior and we adopt these as a benchmark to which the outcomes of the other specifications are compared.

What effects the aggregation of the HL demand systems to a single, aggregate national demand system has on outcomes under various settings is investigated for each specification separately. We calculate aggregate *national-level* (NL) elasticities from the corresponding calibrated HL demand systems separately for each specification and then calibrate a model of the same specification to the NL elasticities and aggregated demand data in order to generate parameters for a nationally representative household. Then, the NL demand systems are used to simulate the same scenarios as above and the results are compared to the corresponding HL results.

3 Demand System Specifications

Six demand system specifications are compared: the LES, CD, CDE, IAS, AIDS, and the Leontief demand system (LEO). The differences most relevant in the current context are their flexibility as determined by the number of independent parameters and their restrictions, rank, and regular region. The CD, with its easily traceable behavior, serves as a naïve reference. The CES is not included in the comparison since it has only a single independent parameter and a universally constant income elasticity of one so that it offers no noteworthy advantage over the CD in the context of multi-item demand systems.[8] The LEO embodies the extreme case of no substitutability and is thus a useful benchmark as the worst-case impact on welfare from price changes.

3.1 Linear Expenditure System (LES)

The LES model (Stone 1954) is probably the most popular demand system in simulation modeling and is defined by the budget share equations (1), the adding-up condition (2) plus the parameter domain definition of β_i (3). The income, own-, and cross-price elasticities are given by eqs. (4), (5) and eq. (6). Note that in these and all following elasticity equations w_i and q_i are placeholders for the formula of the budget share equation and its quantity demand transfor-

[8]The CES also needs some sort of calibration while the CD does not. In the literature, the CES mainly finds use as a demand specification in highly aggregated models, such as macroeconomic models.

mation, respectively.[9]

$$
\begin{aligned}
w_i &= \frac{p_i\gamma_i}{M} + \beta_i(1 - \frac{\sum_{j\in I} p_j\gamma_j}{M}) && \forall i \in I && (1)\\
\sum_{i\in I} \beta_i &= 1 && && (2)\\
0 \leq \beta_i &\leq 1 && \forall i \in I && (3)\\
\eta_i &= \frac{\beta_i}{w_i} && \forall i \in I && (4)\\
\varepsilon_{ii} &= -1 + (1-\beta_i)\frac{\gamma_i}{q_i} && \forall i \in I && (5)\\
\varepsilon_{ij} &= -\frac{\beta_i\gamma_j p_j}{p_i q_i} && \forall i,j \in \{I|i \neq j\} && (6)
\end{aligned}
$$

where I: set of consumption items. The β_i represent marginal budget shares and the γ_i can be interpreted as subsistence quantities as long as these are non-negative. However, the theory does not require the latter to be non-negative (see Pollak 1971) and our calibration model allows them to become negative.

The LES is an inflexible (n independent parameters), rank two (quasi-homothetic) demand system. Its regular region is larger the closer the γ parameters are to zero but restricting γ_i would decrease the flexibility, up to the point where all $\gamma_i = 0$ so that the LES turns into a CD which is completely inflexible, see below. More specifically, the LES is regular for all price-income combinations where both hold $M \geq -\frac{p_i\gamma_i}{\beta_i} + \sum_{j\in I} p_j\gamma_j$ for all $i \in I$ and $M \geq \sum_{j\in I} p_j\gamma_j$.

In the LES, the quantity demanded depends linearly on income but the budget share Engel curves are non-linear. The LES only allows non-negative income elasticities ($\eta_i \geq 0$) due to $\beta_i > 0$ and thus no inferior goods, but its own- and cross-price elasticities are unbounded *a priori*.

3.2 Cobb-Douglas (CD) Demand System

If all $\gamma_i = 0$, the LES turns into a CD demand system. The CD is globally regular, of rank one, and completely inflexible (zero independent parameters) and thus features universally constant income, own-price, and cross-price elasticities of $+1$, -1, and 0, respectively.

3.3 Constant Difference of Elasticities (CDE)

The CDE demand system (Hanoch 1975) is applied in the popular global CGE model GTAP (Hertel 1997) and its derivatives. The definition of the CDE requires the indirect utility function (or implicit expenditure function) (7) and the budget share eq. (8) together with the constraints on the distribution (9), expansion (10), and substitution (11) parameters. The income and price

[9] For example, the complete income elasticity equation for the LES results if w_i in eq. (4) is substituted by eq. (1).

elasticities are given by eqs. (12) and (13).

$$
\begin{aligned}
&1 \equiv \sum_{i \in I} \psi_i u^{\tau_i \beta_i} \left(\frac{p_i}{M}\right)^{\beta_i} && && (7)\\
&w_i = \frac{\psi_i \beta_i u^{\tau_i \beta_i} \left(\frac{p_i}{M}\right)^{\beta_i}}{\sum_{j \in I} \psi_j \beta_j u^{\tau_j \beta_j} \left(\frac{p_j}{M}\right)^{\beta_j}} && \forall i \in I && (8)\\
&\psi_i > 0 && \forall i \in I && (9)\\
&\tau_i > 0 && \forall i \in I && (10)\\
&\beta_i \leq 0 \quad \text{or} \quad 0 < \beta_i < 1 && \forall i \in I && (11)\\
&\eta_i = \frac{\tau_i \beta_i - \sum_{k \in I} w_k \tau_k (1 - \beta_k)}{\sum_{k \in I} w_k \tau_k} + 1 - \beta_i - \sum_{k \in I} w_k (1 - \beta_k) && \forall i \in I && (12)\\
&\varepsilon_{ij} = -w_j \left(\frac{\tau_i \beta_i - \sum_{k \in I} w_k \tau_k (1 - \beta_k)}{\sum_{k \in I} w_k \tau_k}\right) \\
&\qquad + (w_j - \delta_{ij})(1 - \beta_j) && \forall i, j \in I && (13)
\end{aligned}
$$

where the Kronecker delta $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise and u denotes utility. Note that exactly one of the two constraints in (11) has to be applied and then satisfied for all $i \in I$.

The CDE is globally regular, of rank two and, with $2n$ independent parameters, it is inflexible but still substantially more flexible than the LES. It is non-homothetic so that budget shares allocated to each item may vary with income.

A strong limitation for the own-price elasticities is due to eq. (11) which requires an *a priori* decision on the domain of all β_i, restricting *all* own-price elasticities to be either (a) inelastic if $0 < \beta_i < 1$ or (b) elastic if $\beta_i \leq 0$. The CDE's parameter domains imply no bounds on income and cross-price elasticity value ranges. However, a further limitation and name-giving feature of the CDE is that the difference between the Allen elasticities of substitution[10] of two particular items j and k ($\frac{\varepsilon^c_{ik}}{w_k} - \frac{\varepsilon^c_{il}}{w_l}$) are identical for all items $i \in I$. This reduces the number of substitution effects from $(n^2 - n)/2$ to n (Hanoch 1975, p.416).[11]

3.4 Indirect Addilog System (IAS)

The IAS was first introduced by Houthakker (1960) and was shown to be a special case of the CDE function by Hanoch (1975).[12] In contrast to the CDE, it is defined by an explicit indirect utility function and has fewer parameters. Here, we adopt the notation of Jensen et al. (2011, therein called Constant Differences of Elasticities of Substitution, CDES) who also prove that the parameter domains suggested by Hanoch (1975) can be relaxed, thereby crucially extending the range of own-price elasticities to also depict price-elastic demand whereas the common

[10]The Allen elasticity of substitution is defined as ε^c_{ij}/w_j with ε^c_{ij} denoting the compensated cross-price elasticity.

[11]See Hertel (2013) or Yu et al. (2004) for discussions on limitations of the CDE.

[12]Hanoch (1975) derives the IAS from the CDE by adding the restriction $\tau_i \beta_i = g$ with $g = -1$, which cancels out the utility term, and by substituting $\psi_i \beta_i$ with α_i.

parameter restrictions only allow price-inelastic demand.

The demand function is given by eq. (14) and the income, own-price, and cross-price elasticities are given by eqs. (18), (19), and (20), respectively. Eq. (15) is a normalization of α and (16) and (17) define parameter domains.

$$w_i = \frac{\alpha_i (M/p_i)^{\beta_i}}{\sum_{j \in I} \alpha_j (M/p_j)^{\beta_j}} \qquad \forall i \in I \tag{14}$$

$$\sum_{i \in I} \alpha_i = 1 \tag{15}$$

$$\alpha_i > 0 \qquad \forall i \in I \tag{16}$$

$$\beta_i \geq -1 \quad (\text{at most one } \beta_i = -1) \qquad \forall i \in I \tag{17}$$

$$\eta_i = 1 + \beta_i - \sum_{j \in I} w_j \beta_j \qquad \forall i \in I \tag{18}$$

$$\varepsilon_{ii} = -1 - \beta_i \cdot (1 - w_i) \qquad \forall i \in I \tag{19}$$

$$\varepsilon_{ij} = \beta_j w_j \qquad \forall i, j \in \{I | i \neq j\} \tag{20}$$

The IAS is globally regular and of rank two. Similar to the LES, it has n independent parameters and is thus inflexible.

The parameter restrictions impose the following bounds on the elasticity value ranges: $\epsilon_{ii} < 0$ and $-1 \leq \epsilon_{ij}$. There are no bounds on η_i. Nevertheless, the IAS severely restricts substitution effects. As eq. (20) reveals, the cross-price elasticities associated with an increase in the price of item j are all identical. de Boer and Paap (2009) argue that the IAS is just as easy to implement and has the same data needs as the LES but is richer in its representation of household demand behavior because it defines quantity demanded to depend non-linearly on income.

3.5 Almost Ideal Demand System (AIDS)

The AIDS (Deaton and Muellbauer 1980b) model is defined by the budget share eq. (21), the price index $a(p)$ (22), and the parameter constraints which impose theoretical properties on the demand system. These are adding-up (23), homogeneity of degree zero in prices and income (24),[13] and symmetry (25). The negativity property is imposed by adding the constraints (26) and (27) adopted from Ryan and Wales (1998) to the standard AIDS model where the matrix K denotes a lower triangular matrix with elements k_{ij}. However, the negativity property is only guaranteed to hold locally, i.e., at the point of calibration, and might be violated for other price-income combinations. The Ryan and Wales (1998) approach requires the demand data to be normalized (scaled) at the reference (or calibration) point such that $M = 1$ and $p_i = 1$ for all

[13] This equation is redundant if the constraint on γ from eq. (23) and symmetry (eq. 25) are enforced.

items $i \in I$. The income and price elasticities are defined by eqs. (28) and (29), respectively.

$$w_i = \alpha_i + \sum_{j \in I} \gamma_{ij} \ln p_j + \beta_i \ln \left[\frac{M}{a(p)} \right] \qquad \forall i \in I \tag{21}$$

$$\ln a(p) = \alpha_0 + \sum_{i \in I} \alpha_i \ln p_i + \frac{1}{2} \sum_{i \in I} \sum_{j \in I} \gamma_{ij} \ln p_i \ln p_j \tag{22}$$

$$\sum_{i \in I} \alpha_i = 1, \quad \sum_{i \in I} \beta_i = 0, \quad \sum_{i \in I} \gamma_{ij} = 0 \quad \forall j \in I \tag{23}$$

$$\sum_{j \in I} \gamma_{ij} = 0 \qquad \forall i \in I \tag{24}$$

$$\gamma_{ij} = \gamma_{ji} \qquad \forall i, j \in \{I | i < j\} \tag{25}$$

$$\begin{aligned} \gamma_{ij} = {} & -(KK')_{ij} + (\alpha_i - \beta_i \alpha_0)\delta_{ij} \\ & - (\alpha_j - \beta_j \alpha_0)(\alpha_i - \beta_i \alpha_0) + \beta_i \beta_j \alpha_0 \end{aligned} \qquad \forall i, j \in I \tag{26}$$

$$k_{ij} = 0 \qquad \forall i, j \in \{I | i < j\} \tag{27}$$

$$\eta_i = 1 + \frac{\beta_i}{w_i} \qquad \forall i \in I \tag{28}$$

$$\varepsilon_{ij} = -\delta_{ij} + \frac{1}{w_i} \left[\gamma_{ij} - \beta_i \left(\alpha_j + \sum_{k \in I} \gamma_{jk} \ln p_k \right) \right] \qquad \forall i, j \in I \tag{29}$$

with Kronecker δ_{ij} (see section 3.3). The standard AIDS (eqs. 21 to 25) is flexible ($(n^2 + n)/2$ independent parameters) and of rank two. While the aforementioned demand systems are all at least effectively globally regular by construction of their (indirect) utility functions and restrictions on their parameters, the standard AIDS might violate the theoretical properties of negativity and monotonicity. To establish negativity for the AIDS, we additionally impose eqs. (26) and (27). The predicted budget shares are identical to the observed positive ones due to the constraints of the GCE calibration procedure, ensuring that also monotonicity holds (Caves and Christensen 1980). Thus, the calibrated AIDS models are regular at the respective price-income point of calibration. It is hoped that regularity then also extends to price-income combinations as implied by the scenarios to be simulated, but whether the regular region is indeed large enough for the simulations is a matter of empirical testing. The AIDS imposes no bounds on the elasticity value ranges. The Engel curves are linear in the log of income.

3.6 Leontief Demand System (LEO)

The LEO assumes that all items are consumed in constant quantity ratios. It is not a usual specification for larger impact analysis models but its inherent assumptions are often hidden within models for the analysis of first-order welfare impacts (e.g., Wodon and Zaman 2010) or what is occasionally referred to as “non-behavioral microsimulation” or “microaccounting” (e.g., Bourguignon et al. 2008). The model is defined by the budget share equation (30) and the parameter restriction (31) (see, e.g., Pollak and Wales 1995).

$$w_i = \frac{\alpha_i p_i}{\sum_{j \in I} \alpha_j p_j} \qquad \forall i \in I \tag{30}$$

$$\alpha_i > 0 \qquad \forall i \in I \tag{31}$$

The LEO represents the extreme case where all items are perfect complements. It is globally regular but completely inflexible (zero independent parameters) and has income, own-price, and cross-price elasticities of 1, $-w_i$, and $-w_j$, respectively.

4 Data

We exploit a unique dataset that was created in several steps. The starting point is given by the results from a demand system estimation based on cross-sectional household survey data described in Boysen (2016). The author extracted the demand data from the Uganda National Household Survey 2012/2013 (UNHS) which includes a sample of 6,887 households and is nationally representative (see UBOS 2014).

The processed data included prices and budget shares of 15 consumption items as well as income and various characteristics for each household and was used to estimate a 15-item, two-stage budgeting demand system model for rural and urban household subsamples separately. In the first stage of the model, households allocate their budgets to two aggregate consumption item groups: food and non-food. This stage is represented by a Working-Leser-type demand function. The second budgeting stage adds detail to the food item group: A QUAIDS (Banks et al. 1997) is estimated which allocates the food item group's budget to 14 food items. The estimations account for various household-specific characteristics and for censoring, that is, the problem associated with the fact that most of the 14 items, albeit nationally important consumption items, are not part of the consumption of a substantial share of the households. The QUAIDS – an extension of the AIDS – is popular with household demand studies as it is a flexible, rank three demand system which is consistent with consumer demand theory while remaining parsimonious in the number of parameters to be estimated (Banks et al. 1997). However, due to the censoring approach used in those estimations, only homogeneity and symmetry properties were enforced so that frequent regularity violations are expected in the derived elasticity dataset. Indeed, elasticity datasets not conform with theory are a common reality for impact analysis model developers.

The parameter estimates together with the prices and incomes from the household dataset are utilized to calculate an individual set of demand elasticities for each of the 6,887 households in the survey. These serve as priors for the calibration of four demand system specifications and their distributions are shown in the form of boxplots[14] in Figure 1. The prior elasticities clearly

[14]The bottom, middle, and top lines of each box represent the first, second (median), and third quartiles, respectively, and the whiskers extend to 1.5 times the interquartile range. For the sake of an improved scale of the graph, outliers are not shown.

show a large variability across households. While demand for cassava, vegetables, legumes, and fats is predominantly price and income inelastic, it is predominantly price and income elastic for items such as meat, fish, and milk. Both types of elasticities reach levels of beyond absolute three. Moreover, cereals and vegetables are inferior items for some households, and meat and other food ('other') feature some own-price elasticities in the positive range.

Each of the four specifications, the LES, IAS, CDE, and AIDS, has been calibrated for each household such that they replicate the observed budget shares subject to the household's income and price vector and are as close as possible to the calculated prior elasticities. To this end, "closeness" is measured using the GCE objective function based on the Kullback-Leibler measure of information divergence (see Golan et al. 1996).[15] In this particular application, the GCE objective function value is a measure of the distance of the calibrated from the prior elasticities and is minimized during the process of calibration. Because the GCE approach makes use of all information provided and allows the estimation of overdetermined equation systems, all specifications have been calibrated to the full sets of income and own-price elasticities. Only in the case of the AIDS, which is flexible, have the cross-price elasticities been included in addition. Due to their specific restrictions on the parameters and differing degrees of flexibility, the four specifications differ in the extent of deviation from the prior elasticities. But all calibrated household models of all specifications are regular at least at the point of calibration because they have been calibrated using the full sets of equations and restrictions as given in their definitions in section 3. These calibrated demand systems, together with the household data of incomes and prices, constitute the dataset for our experiments. The demand elasticities implicit in the calibrated demand systems from the dataset are contrasted with the priors in Figure 1. The plots clearly show that the distributions of the prior and calibrated AIDS elasticities match closely. It also illustrates the limitation of the CDE to predominantly price-inelastic demand.

Because the GCE objective function values measure the distance from the priors they are also directly suitable for evaluating the goodness of fit of the calibrated demand systems. The lower the GCE measure, the better the fit. The GCE of 26.4 attests the AIDS is a very good fit to the priors in comparison to the other specifications. As it is the only flexible specification among the four which does not impose *a priori* bounds on the elasticity values, it is not surprising that its fit is much better than those of the other specifications.[16] Although the number of independent parameters of the CDE equals the number of income and own-price elasticities, its severe restrictions on the own-price elasticities result in a GCE value of 126.7. This is worse than that of the LES (69.5) and the IAS (71.4) which both have only half the number of independent parameters.

[15] For further details on the generalized cross-entropy estimation model, see Appendix B.

[16] The GCE value of the AIDS is indeed further penalized compared to the other specifications as it also includes the deviations in the cross-price elasticities.

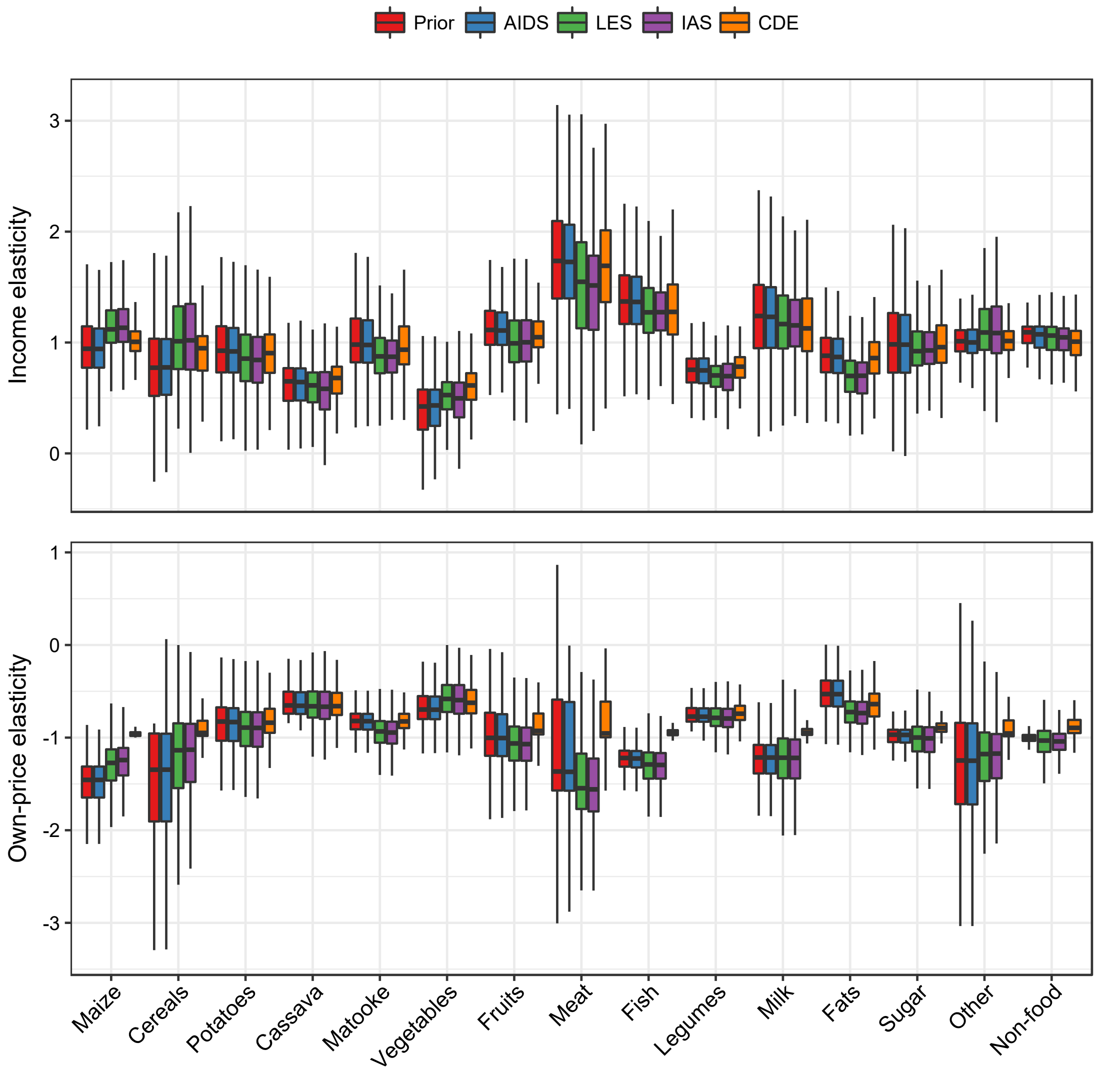

Figure 1: Boxplots comparing the prior elasticities calculated from the estimated demand system to the calibrated HL income and own-price elasticities by item. Note that there are no priors for the own-price elasticity of non-food.

5 Simulations

What influence the choice of the model specification and aggregation level has on the simulation results is case-specific and thus an empirical matter. We evaluate the sensitivity of several outcome measures with respect to these choices in simulations of stereotypical scenarios. In each scenario, either one or several prices are increased or incomes are cut or raised, either uniformly across the population or only for selected households. Depending on the objective of the analysis, i.e., the outcome measure, our choices might or might not matter. For example, the effect of a population-wide uniform income increase on aggregate national quantity demanded for a good should not be influenced by the choice of the aggregation across households. By contrast, the choice of the specification does matter in this case because the shape of the Engel curve determines how income elasticities change depending on the level of income. The level of aggregation across households is important, for example, if a shock affects households differently.

The analysis applies two broad types of outcome measures, national quantity demanded and equivalent income-based measures. In this section, first all outcome measures are defined and then the simulations and their results are presented separately for the two broad types of measures as their analysis requires different simulation strategies.

5.1 Outcome Measures

The influence of the choices of model specification and aggregation level on simulation results are evaluated in terms of five types of outcome measures: quantity demanded, equivalent income, poverty, inequality, and cost-of-living indices. All measures are aggregated from individual households' results to the national level using the survey's household weights.

The first measure is aggregate *national quantity demanded* which reflects the pure quantity effects on the economy from the demand side.

Equivalent income (see, e.g., King 1983) named in analogy to the "equivalent variation" welfare measure, is the second measure, defined using the cost function $C(u, p)$ as the income required to attain the post-shock utility level u' at base price levels p and thus as $C(u', p)$. The cost function is specific to the respective demand system specification. The equivalent income equations are shown below for the LES (eq. 32), the AIDS (eq. 33), and the LEO (eq. 34). For the CD, eq. (32) is modified by setting all γ_i to zero. The equivalent income for the IAS and CDE is calculated implicitly by solving the indirect utility function (eq. (35) for the IAS and eq. (7) for the CDE) for the income $C(u', p)$ such that it attains the post-shock utility level u' at the base price levels p.

$$C_{\text{LES}}(u', p) = \sum_{i \in I} \gamma_i p_i + (M' - \sum_{i \in I} \gamma_i p'_i) \prod_{i \in I} \left(\frac{p_i}{p'_i} \right)^{\beta_i} \tag{32}$$

$$\ln C_{\text{AIDS}}(u', p) = \ln a(p) + \big(\ln M' - \ln a(p')\big) \prod_{i \in I} \left(\frac{p_i}{p'_i} \right)^{\beta_i} \tag{33}$$

$$C_{\text{LEO}}(u', p) = \frac{M'}{\sum_{i \in I} \alpha_i p'_i} \sum_{i \in I} \alpha_i p_i \tag{34}$$

$$V_{\text{IAS}}(p, M) = \sum_{i \in I} \alpha_i \frac{(M/p_i)^{\beta_i} - 1}{\beta_i} \tag{35}$$

To measure poverty, equivalent income can be directly compared to an existing absolute poverty line since it represents the post-shock income's value at the vector of base prices. As a poverty measure, we calculate the Foster-Greer-Thorbecke (FGT) poverty indices P_α (Foster et al. 1984) directly on the basis of the equivalent income $C(u', p)$. The FGT indices are defined by eq. (36).

$$P_\alpha = \frac{1}{N} \cdot \sum_{h \in H} \left(\frac{z - M_h}{z} \right)^\alpha \cdot S(h) \tag{36}$$

where N: population size, H: set of household, z: poverty line, M_h: income of household h, and $S(h) = 1$ if $M_h < z$ and $S(h) = 0$ otherwise.

Setting the parameter α to 0 or 1 yields the *poverty headcount* and *gap* indices, respectively. The poverty headcount index P_0 measures the share of people whose income is below the poverty line. The poverty gap P_1 represents the proportion of the poverty line income level that the average poor person would need in addition to the current income to reach the poverty line. The poverty line has been set so that the official 2012/13 Ugandan poverty headcount of 19.7% (UBOS 2014), corresponding to the dataset underlying this study, is replicated.

Inequality is measured by the *Gini* index which ranges from zero for perfect equality to 100 for perfect inequality.

As a final measure, we consider the Konüs *true cost-of-living* index (see Deaton and Muellbauer 1980a), defined as $\frac{C(u,p')}{C(u,p)}$ and aggregated to a national index using democratic household weighting. In contrast to the usual consumer price index measures, this price index does not use a fixed basket of quantities consumed for weighting but instead accounts for changes in the quantities caused by substitution.

5.2 National Quantity Demanded

To investigate the influence of the choices of model specification and level of aggregation across households on simulation results measured in terms of aggregate national quantity demanded, the following six sets of price and income shock scenarios are simulated.

Scenario	**Description**
Own-price +20%	Each of these 15 scenarios increases the price of the individual named product by 20% while keeping all other prices constant.
All -20%	The incomes of all households are uniformly decreased by 20%.
Poor -20%	The incomes of poor households are uniformly decreased by 20%.
20% richest	The incomes of the 20% richest households are decreased at a uniform rate such that the total amount of income loss corresponds to the sum of loss in the 'poor -20%' scenario.
Firing poor	Randomly selected poor households experience a 50% decrease in their incomes due to job losses. The total amount of income loss is equal to that of the 'poor -20%' scenario.
Redistribution	Income is redistributed from the 20% richest households to the poor households (uniform percentage changes) such that the poverty headcount is halved.

These scenarios are designed to expose how the specifications in their two levels of aggregation differ with respect to their impacts on national quantity demanded through their own-price, cross-price, and income effects.

But before discussing the actual simulation results, it is worth reconsidering the regularity properties of the specifications at the price-income points of the various scenarios. While the CD, IAS, CDE, and LEO are globally regular – as ensured by their parameter domains – this does not apply to the AIDS and LES models. If indeed inconsistencies with demand theory occur, how relevant is this in an applied simulation setting? And by how much does it bias the results?

The AIDS might violate the negativity property for any deviation from the price-income vector of the base data. Every AIDS simulation result was checked for a violation of the negativity property by testing the Slutsky matrix for semi-definiteness.[17] Positive eigenvalues of the matrix indicate a violation of negativity.[18] While the representative NL AIDS model showed no violation in any scenario, between 4% and 56% of the HL AIDS models, varying by scenario, displayed at least one violation, albeit eigenvalues remain tiny in the majority of cases. A violation causes theory-inconsistent behavior, such as positive compensated own-price elasticities.

Moreover, the AIDS might also violate monotonicity and hence predict negative budget shares. On the level of individual demand systems in the HL AIDS simulations, negative budget shares occur when the incomes of the poorer households are affected either through income cuts or price increases. This occurs most severely in the non-food 'own-price +20%' and 'firing poor' scenarios where 0.16% and 0.22% of households have negative budget shares of less than -10%. However, on average across all scenarios, more than 99.2% of households display no budget shares of less than -0.1%. To examine what this means for the simulation results, we define the *income bias* as the sum of the household income-weighted deviations of the sum of individual households' absolute budget shares from one in terms of a percentage of total income.[19] The income bias across all NL AIDS simulations amounts to 0.03% on average and to 0.2% in the worst case. Note that this does not correspond to additional disposable income but rather indicates the extent of "irregular substitution," which should be much larger than the corresponding irregular welfare gain. Thus, the influence on household-level outcomes might be notable in some cases but in national-level aggregate outcome measures, such as the quantity demanded or equivalent income, it is most likely very small and should not invalidate the HL AIDS results as a benchmark.

Also, the LES might violate regularity if the income falls below some level related to the γ_i parameters (see section 3.1). This results in negative budget shares and occurs in the scenarios where the incomes of the poorest incur severe hits. Only in the 'all -20%' and the 'firing poor' scenarios do households have budget shares of less than -10%, more specifically, 0.01% and 0.59% of all households, respectively. On average, across all scenarios, 99.5% of households

[17] More specifically, we constructed a matrix according to eq. (14) from Deaton and Muellbauer (1980b).

[18] Here, eigenvalues are counted as being positive if they are greater than 10^{-6}.

[19] Mathematically, income bias is defined as $100 \cdot \frac{\sum_{h \in H} \Omega_h M_h (-1 + \sum_{i \in I} |w_{hi}|)}{\sum_{h \in H} \Omega_h M_h}$.

do not have budget shares of less than -0.1%. The average income bias across all scenarios is 0.02% and 0.12% at maximum. Consequently, the aggregate results of the HL LES simulations should only be marginally biased by these irregularities.

Even though these regularity violations in some of the HL LES and AIDS simulations are of little significance for most results presented in the following, they could be prohibitive in other settings, for instance, if these demand systems were components of larger economic model systems which have strong requirements with respect to theoretical properties. In a mathematical optimization model, negative budget shares could lead to an explosive process and could result in non-convergence of the numerical optimization algorithm or, directly, to the infeasibility of the model.

The actual results of the simulations are discussed in turn for own-price, cross-price, and income effects. The results of the CD serve as a benchmark of a naïve demand system as it does not require estimated parameters and assumes unit elastic demands and no cross-price effects. The LEO illustrates the extreme case when households do not explicitly adjust their consumption to price changes but only scale all quantities consumed by the same proportion to meet their budget constraints. Because the AIDS is flexible and fits the prior elasticities very well, its simulation results are taken as an approximation of the "true" consumption behavior.

Own-price effects. This first set of results explicitly looks at the performance with respect to own-price demand reactions. Starting with the HL simulations shown in 'own-price +20%' row of the left-hand panels of Figure 2, the percentage reactions in the aggregate, national quantity demanded of items caused by a 20% increase in the respective item's own price vary substantially between the specifications. The CD universally features an own-price elasticity of -1 so that the effect of a 20% increase in the own price is identical for all items and corresponds to a 16.7% decrease in the quantity demanded. In absolute terms, the percentage changes simulated by the CDE are smaller than those of the CD without exception, which is due to the strong restrictions it imposes on the own-price elasticities as discussed in section 3. In contrast to the CD and CDE, the LES and IAS also allow price elastic demand and the corresponding changes can be smaller or larger than those of the CD. The LES and IAS results are generally close together but the LES tends to be a little bit closer to the AIDS in more cases than vice versa. Both are closer to the AIDS than the CDE in all except one scenario. The simulated reaction of the CDE is larger than that of the AIDS in only one scenario, showing that it rather consistently underestimates the own price-induced demand changes. The LEO allows no substitution at all and scales down all consumption quantities proportionately in reaction to the own-price increase. Using the terms of the Slutsky equation, the LEO's quantity responses to own-price changes are exclusively due to the income effect and are thus the smallest possible in a theory-consistent demand system.

For further comparison, the differences between the specifications are summarized by calculating the absolute deviations (ADs) from our benchmark, the HL AIDS results. The MAD row of the 'own-price +20%' scenario in the HL statistics columns in Table 1 shows the mean

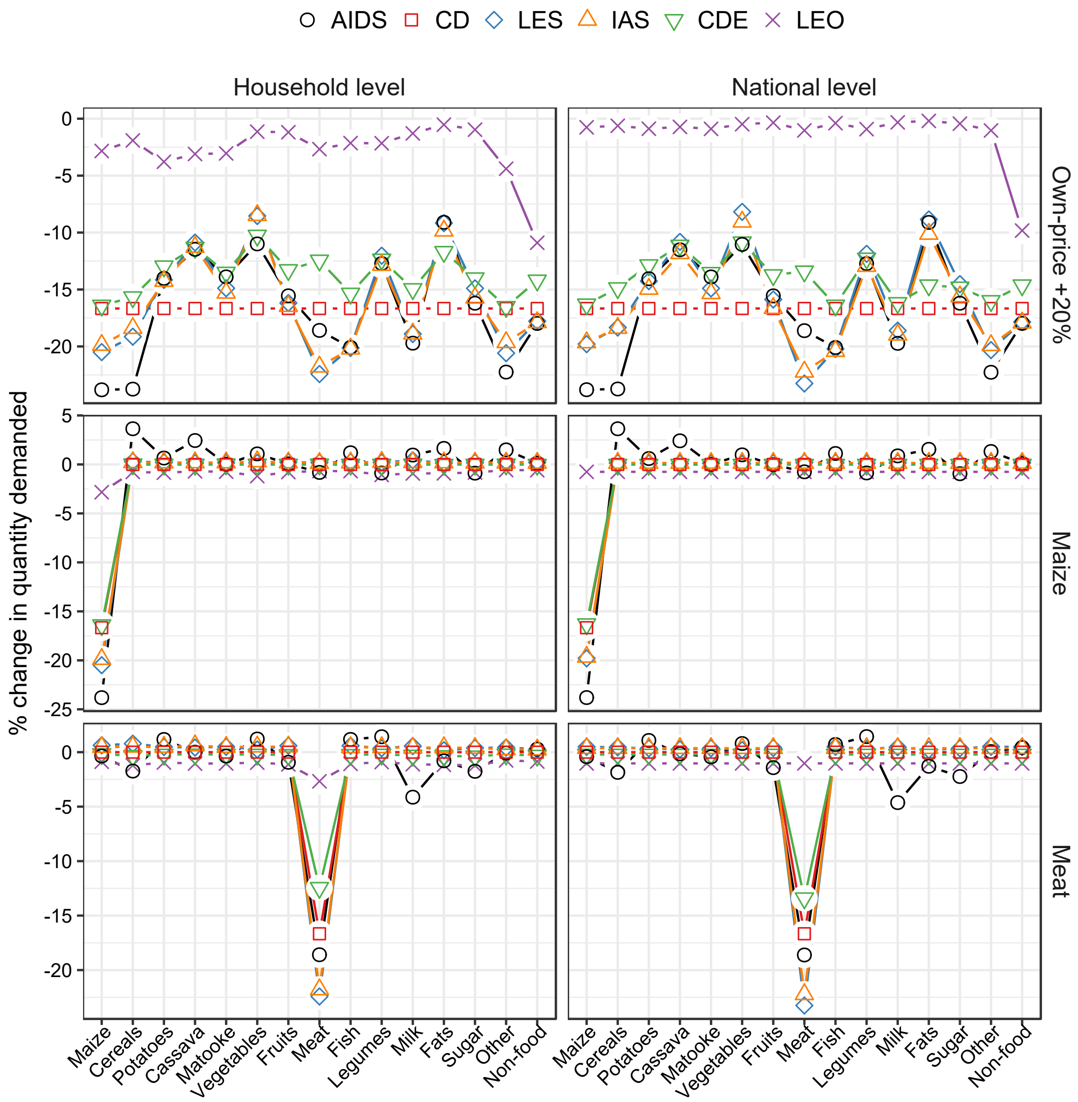

Figure 2: 'Own-price +20%' shows the percentage changes in national quantity demanded by item where each item's price is increased by 20% in an individual simulation. 'Maize' and 'meat' show the percentage changes in national quantity demanded for all items resulting from an increase in the maize and meat prices, respectively, of 20%.

absolute deviation (MAD) from the AIDS results. It indicates that the LES performs best with an MAD of 1.4 percentage points (pp) over all scenarios, closely followed by the IAS (1.5pp). In comparison, the CDE's MAD is more than twice as large and only a little better than that of the CD. Thus, given our elasticity data, which features many price-elastic items, the CD's constant own-price elasticity of -1 turns out to be an assumption that is not much worse than the CDE. The maximum absolute deviation (max. AD) represents the worst-case mismatch across all scenarios with respect to the AIDS results and yields a similar ranking to the MAD. However, here the CD outperforms the CDE and might thus be a more reliable choice. The MAD and maximum AD statistics highlight that the choice of the model specification can result in large differences in the quantity demanded for items which are subject to a price change, in this example of up to 8.6pp between the CDE and the AIDS. The statistics for the LEO emphasizes the extreme difference to the other specifications.

Table 1: Performance statistics in simulating national quantity demanded by demand model specification and aggregation

		HL						NL					
Scenario	Statistic	AIDS	CD	LES	IAS	CDE	LEO	AIDS	CD	LES	IAS	CDE	LEO
Own-price +20%	max. AD		7.6	4.6	5.4	8.1	21.8	0.0	0.0	0.8	0.7	2.9	3.4
	MAD		3.9	1.4	1.5	3.4	13.9	0.0	0.0	0.3	0.3	0.7	1.5
Maize	max. AD		3.6	3.3	3.4	3.6	4.4	0.2	0.0	0.3	0.2	0.0	0.4
	MAD		1.1	1.1	1.1	1.1	1.6	0.1	0.0	0.1	0.1	0.0	0.1
Meat	max. AD		4.1	4.7	4.5	3.8	3.0	0.5	0.0	0.3	0.2	0.2	0.3
	MAD		1.1	1.3	1.2	1.1	1.3	0.3	0.0	0.1	0.1	0.1	0.1
All -20%	max. AD		15.4	6.4	6.8	12.6	15.4	0.1	0.0	1.0	1.0	2.1	0.0
	MAD		4.8	2.0	1.9	5.4	4.8	0.0	0.0	0.3	0.3	0.6	0.0
20% Richest	max. AD		0.5	0.3	0.3	0.9	0.5	0.7	0.8	0.8	0.8	0.7	0.8
	MAD		0.3	0.1	0.1	0.4	0.3	0.4	0.4	0.4	0.4	0.3	0.4
Poor -20%	max. AD		2.2	0.5	0.5	1.5	2.2	2.3	2.6	2.0	2.2	2.4	2.6
	MAD		0.5	0.2	0.2	0.4	0.5	0.9	0.9	0.9	0.9	0.8	0.9
Firing poor	max. AD		1.6	0.5	0.5	1.2	1.6	2.4	2.6	2.2	2.4	2.5	2.6
	MAD		0.4	0.2	0.2	0.3	0.4	0.9	0.8	0.9	0.9	0.8	0.8
Redistribution	max. AD		3.7	1.5	1.1	2.4	3.7	4.3	5.6	4.1	4.0	5.0	5.6
	MAD		1.3	0.4	0.4	1.2	1.3	2.2	1.8	2.1	2.1	1.8	1.8

Source: Own computation from simulation results. HL statistics are calculated with reference to the HL AIDS results, each NL statistic with reference to the corresponding HL result. The statistics for the cross-price effect simulations 'maize' and 'meat' are calculated excluding the own-price effects.

In these scenarios, how well does a single, representative NL demand model of the same specification approximate the HL demand systems? This question is answered by comparing the right- to the left-hand side of the 'own-price +20%' panel of Figure 2. The NL models appear to closely mimic the effects on national quantity demanded overall but with some differences in the details. This first impression is confirmed by the AD statistics in the NL columns of Table 1. The MAD is far below 1pp for all specifications, suggesting that there is only a limited loss in accuracy when using a representative national household instead of all individual households if own-price shocks and aggregate demand are the focus of interest. The highest differences are found for the CDE with a MAD of 0.6pp and a maximum AD of 3.0pp in the fats scenario. All other specifications fare much better and the AIDS and CD even show no noticeable differences. Overall, it appears that the superiority of the AIDS, LES, and IAS also translates to the NL specifications in the given setting of measuring the effects of own-price shocks on national quantity demanded.

Cross-price effects. Next, the maize and meat scenarios are used as examples to examine cross-price reactions across the range of all items. Note that only the AIDS possesses sufficient independent parameters to allow the calibration of cross-price elasticities. By contrast, the CD assumes no cross-price effects at all and the LES, IAS, CDE, and LEO assume very specific reactions as discussed in section 3 and reflected in the quantity changes of the items not shocked.

Correspondingly, the cross-price effects of all specifications apart from the AIDS can be more or less regarded as artifacts. The left-hand panels of Figure 2 illustrate the effects of increasing the maize and meat prices, respectively, by 20% on the national demand for the full range of items. The results from the AIDS simulations feature both negative and positive demand changes of the non-shocked items, i.e., complements and substitutes. While the cross-price effects of the CD are zero, those of the others vary slightly in close proximity to zero. This means that the other specifications might sometimes miss sizable cross-price effects in comparison to the AIDS. The 'maize' and 'meat' rows of Table 1 quantify the ADs from the AIDS results. The LEO aside, these are in the order of 1pp on average but amount to up to 3.6pp in the maize and up to 4.7pp in the meat scenario.

Comparing the left- and right-hand panels of Figure 2 indicates a strong similarity of the NL to the HL cross-price effects for all specifications. In numbers, the NL results of the LES, IAS, and CDE specifications deviate by 0.1pp on average from their HL counterparts but, given the limited size of their HL cross-price effects, this is still noteworthy in relative terms. Given the larger size of the HL AIDS effects, it is not surprising that the AIDS NL results feature the highest deviations from the corresponding HL specifications. As expected, the aggregation choice is of minor importance next to choosing a flexible demand system specification when cross-price effects are of interest.

Income effects. The effects of income changes on national quantity demanded, subject to the specification choice, are illustrated in Figure 3. Focusing first on the HL panels on the left-hand side, the uniform 20% decrease of all incomes in the 'all -20%' scenario shows that the overall pattern of quantity changes is roughly the same for all specifications apart from the CD and LEO. However, a closer inspection reveals that even under a uniform income shock the differences in the size of the simulated reactions can differ substantially between specifications. The CD and LEO are identical with respect to income shocks and are inadequate specifications for this type of shock and outcome measure as they universally assume income elasticities of one, resulting in a uniform percentage decrease in all items' quantities. But the LES, IAS, and CDE results also deviate from those of the AIDS by between 1.6pp and 2pp, on average, and by 6.4pp to

7.2pp in the extremes as reported in Table 1. Thus, these three specifications present themselves as similarly inferior alternatives to the AIDS where the IAS marks the middle ground.

Comparing the HL results displayed in the left-hand panel to the NL results in the right-hand panels of Figure 3, the graphs of the 'all -20%' scenario show the change profiles of the NL specifications tracking their corresponding HL profiles closely. As measured by the AD statistics in Table 1, however, the CDE in particular deviates by a maximum of 3.1pp and an average of 0.6pp. The LES and IAS perform equally with a maximum AD of 1pp and 0.3pp on average. The AIDS deviates the least, by a maximum of 0.1pp.

The '20% richest,' 'poor -20%,' and 'firing poor' scenarios are constructed such that they all reduce national income by the same percentage of 1.2%, but individual households are affected differently. Reducing the incomes of the 20% richest households by this amount constitutes only a small relative income shock and their income elasticities are typically smaller so that the reactions turn out smaller than those in the next two scenarios where only poor households are affected. Although the decreases in national income are much smaller than in the 'all -20%' scenario, the HL AD statistics are no less substantial in relation to the AIDS reactions, considering the magnitude of the shocks. In all of these scenarios, the CD and LEO perform by far the worst and the LES, IAS, and CDE perform rather similarly with MADs of 0.2pp and below and a maximum AD below 1pp. Among the three, the CDE tends to feature the highest maximum AD but also the lowest MAD.

Given that the change in national income is identical in the '20% richest,' 'poor -20%,' and 'firing poor' scenarios, they are effectively identical in the NL simulations with a single national demand system as illustrated by the graphs. While the NL patterns of quantity changes in the '20% richest' scenario exhibit at least some similarity to their HL counterparts, this is not the case for the 'poor -20%' and the 'firing poor' scenarios where only poor households with high income elasticities are affected.

The AD statistics of the NL from the HL results are large, more specifically, the NL MAD frequently exceed factor four of the HL MAD. It is not surprising that all NL specifications perform similarly badly with respect to their HL counterparts, including the CD, LEO, and AIDS. This highlights how inappropriately even national quantity reactions are simulated once a scenario affects household incomes' differentially.

The final scenario 'redistribution' purely redistributes income of the same amount as in the previous scenarios from the richest to the poor. Here, the deviations of the LES, IAS, and CDE are larger and amount to MADs of between 0.3pp and 0.4pp and to maximum ADs of between 0.3pp and 0.4pp. In this scenario, the difference between the NL and HL results naturally turns out to be the most extreme as national aggregate income does not change and thus there is no effect in the NL simulations.

5.3 Equivalent Income-based Measures

The impact of a change in some item's price on equivalent income depends on the substitutability of the item and is limited naturally by the two extremes of perfect substitutability and per-

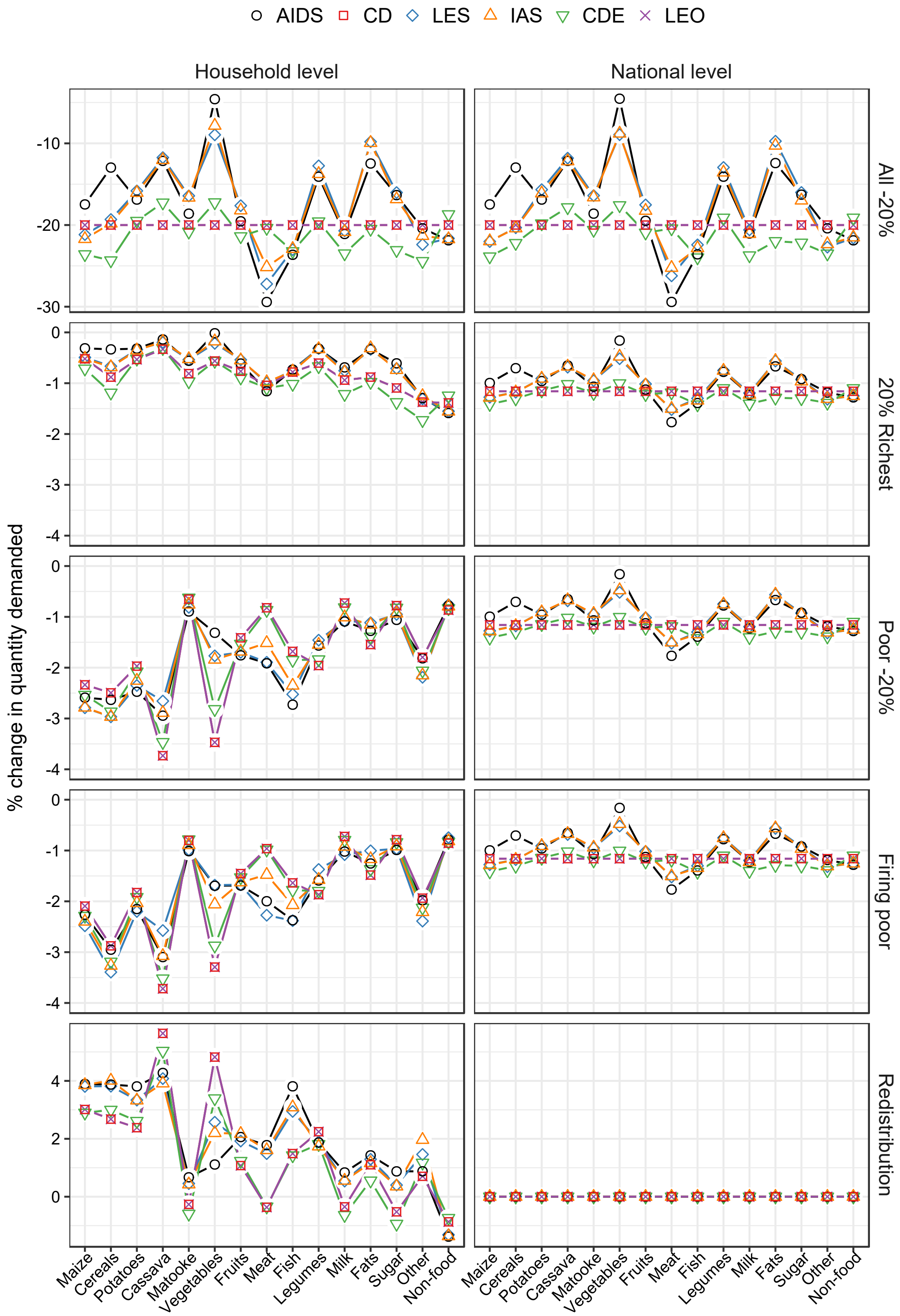

Figure 3: Percentage changes in national quantity demanded of all items resulting from the various income scenarios.

fect complementarity. Perfect complementarity is embodied in the LEO which assumes constant consumption quantity ratios and cross-price elasticities of $-w_j$. Since it represents the worst-case impact on equivalent income, it is a useful benchmark for the analysis of equivalent income-based outcome measures. The CD assumes all uncompensated cross-price elasticities to be zero. Also, the LES, IAS, and CDE are inflexible and do not allow arbitrary cross-price effects. Instead, the cross-price elasticities are a direct implication of their respective functional form (see section 3). Only the AIDS has enough independent parameters to depict arbitrary cross-price effects. Hence, the specification choice will influence monetary welfare measures like the equivalent or compensated income. As has been shown in section 5.2, cross-price effects are only sizable in the case of the AIDS whereas those of the other specifications are near zero.

Teasing out substitution effects requires examining the impacts of changes in relative prices between items. However, choosing a single set of price changes across the items would not yield a representative result because substitutability varies between items. Thus, we adopt the following strategy to get a comprehensive impression of the specification choice's influence on equivalent income-based outcome measures. We construct 100 scenarios in which prices are increased randomly in such a way that the loss in purchasing power of national income, as measured by a Laspeyres price index, amounts to 20%.[20] In each simulation, all households are subject to the same percentage price shock but just as consumption bundles and elasticities differ between households, so, too, do the impacts on their real incomes differ.

The discussion starts with the results on aggregate national equivalent income before turning to poverty, inequality, and cost-of-living indices.

National equivalent income. Figure 4 summarizes the national equivalent income results in case of HL specifications in terms of pp deviations from the HL AIDS results and in case of NL specifications in terms of pp deviations from their HL counterparts' results. The summary for each specification consists of the mean (represented by geometric shapes), the range between the 2.5 ($Q_{2.5\%}$) and the 97.5 ($Q_{97.5\%}$) percentile (the extent of the vertical line), henceforth referred to as the inter-percentile range (IPR), and the minimum and maximum deviation (the short horizontal lines) over all 100 simulations. The table above the graph shows the total national equivalent income effects of the HL AIDS simulations as percentage changes to provide an anchor for assessing the relative magnitude of the deviations.

Starting with the HL results, the means highlight that the impacts on the national equivalent income of the various specifications are all in a narrow range with mean deviations of between 0pp for the CD and -0.4pp for the worst-case Leontief specification. Leaving the LEO aside, the absolute deviations implied by the minimum and maximum deviations are smaller than 1pp in all cases. Considering the deviation extremes, and particularly the IPR, the LES and IAS by far outperform the other specifications. The CDE performs a good deal better than

[20] For each scenario and item i, a random number z_i is drawn from a uniform distribution. Then, the new price p_i' is calculated as $p_i' = p_i \cdot \left(\frac{z_i \cdot \left(\frac{1}{0.8} - 1\right) \cdot \sum_{j \in I} p_j q_j}{\sum_{j \in I} z_i p_i q_i} + 1 \right)$.

HL AIDS simulations: percentage changes in equivalent income

Min.	$Q_{2.5\%}$	Mean	$Q_{97.5\%}$	Max.
-19.9	-19.9	-19.4	-17.9	-17.5

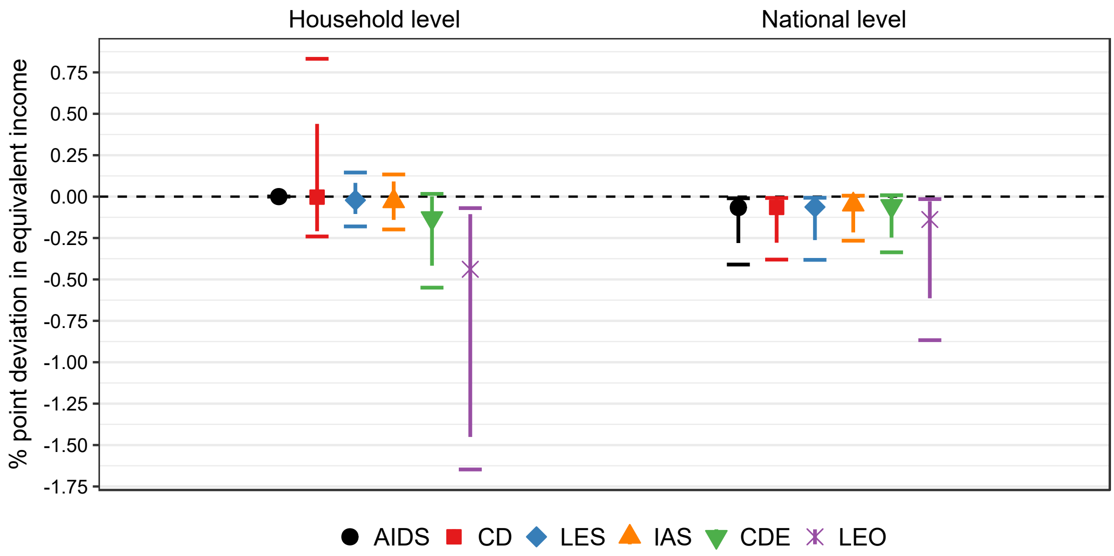

Figure 4: Impacts of a 20% purchasing power loss on equivalent income. The left-hand graph displays the deviations of the HL specifications from the HL AIDS results summarized in the table above, the right-hand the deviations of the NL specifications from the corresponding HL AIDS results. The statistics displayed are the mean, IPR, min. and max. deviation. Source: Own computation from simulation results.

the CD. It is interesting to observe that for this kind of outcome measure and shock, the naïve CD assumption improves markedly over the LEO assumption, since the LEO is identical to the common practice of a simple price mapping, as sometimes implied by the assumptions behind "first-order welfare analysis" or "microaccounting" approaches and the CD and LEO have the same data requirements. However, one advantage of the LEO is that it reliably overestimates the negative impacts and thus yields lower bounds for the equivalent incomes. But interestingly, the maximum deviation of the CDE is only tiny so that it rather reliably yields tighter lower bounds than the LEO.

If the same simulations are performed with a single national demand system, the means of the NL results show that the negative impacts are always overestimated compared to the corresponding HL results. However, the mean AD is very small for all specifications except the LEO and is, in fact, far less than 0.1pp. Nevertheless, in the extreme, the LES, AIDS, and CD can deviate by up to -0.4pp and the IAS and CDE by up to -0.3pp. By definition, the NL LEO always results in a decrease of the equivalent income of the full 20%.

Comparing the magnitude of these deviations relative to the total size of the change in equivalent income from the HL AIDS simulations in the table at the top of Figure 4, the specification choice (excluding LEO) can cause deviations of up to 5% in the case of the CD but in the vast majority of cases deviations are likely less than 1%. Choosing an NL aggregation mostly causes even less deviation in comparison to the corresponding HL specification.

Whether the conclusions change in a more detailed view across income levels is investigated by means of Figure 5, which groups the households by income quintiles, where 1 is the poorest and 5 the richest quintile, and summarizes their equivalent income results. The LEO shows deviations from the AIDS results continuously decreasing with increasing income quintile, which likely reflects that the demands of richer households tend to be less elastic and hence makes the AIDS behave more similarly to the LEO. The AD of the CD can reach 1.2pp for the poorer households. In tendency, the likely deviations from the AIDS results and thereby the errors created in comparison to the AIDS are the larger the poorer the households are. Only for the CDE does the pattern not strictly hold. The performance ordering among the specifications relative to the AIDS remains roughly unchanged from the aggregate national equivalent income case discussed above. The LES and IAS consistently perform best across all quintiles with a slight advantage for the LES. In summary, these results indicate that the specification choice tends to gain importance for income distribution analysis with the decreasing income of the households in the focus of the analysis.

HL AIDS simulations: percentage changes in equivalent income

Quintile	Min.	$Q_{2.5\%}$	Mean	$Q_{97.5\%}$	Max.
1	-24.5	-23.3	-19.6	-17.5	-17.3
2	-23.9	-23.3	-19.7	-17.8	-17.2
3	-22.7	-22.5	-19.7	-17.9	-17.3
4	-21.0	-20.5	-19.5	-18.5	-18.3
5	-20.9	-20.8	-19.2	-14.7	-13.2

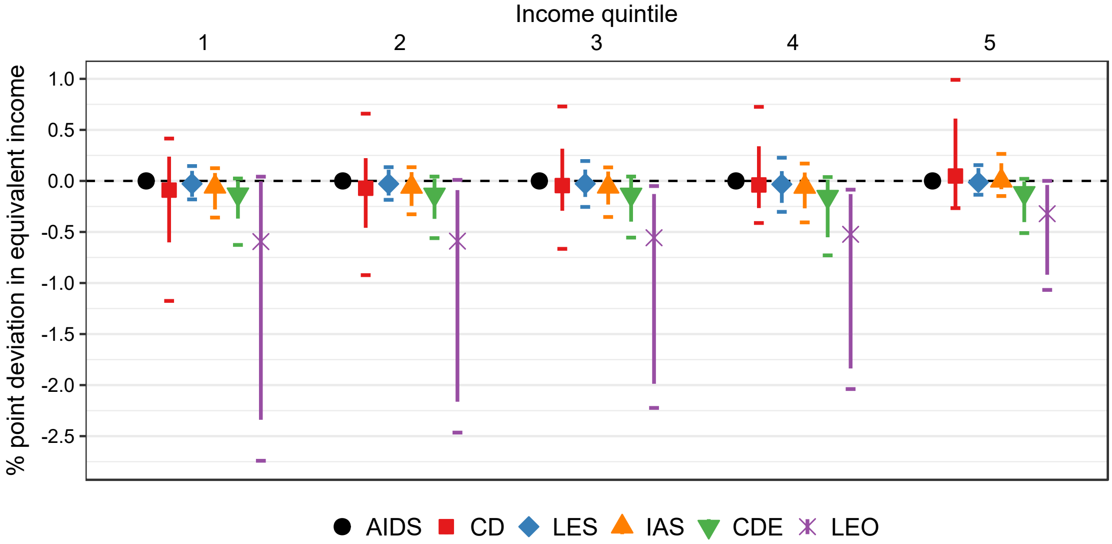

Figure 5: Impacts of a 20% purchasing power loss on equivalent income.The graph displays the deviations of the HL specifications by income quintile from the corresponding HL AIDS results summarized in the table above. The statistics displayed are the mean, IPR, min. and max. deviation. Source: Own computation from simulation results.

Poverty, inequality, and cost-of-living. Finally, we seek to illuminate the relevance of the two choices for indices calculated on the basis of the equivalent income. To this end, poverty, inequality, and cost-of-living effects are calculated based on the simulated equivalent incomes and are illustrated in Figure 6. As a result of the tendency of all specifications to bias the equivalent income negatively in comparison to the AIDS, they also tend to overestimate the poverty headcount and gap. The LEO results serve as upper bounds on the poverty headcounts of which the mean is 0.5pp above the AIDS headcount. The LES registers the smallest mean deviation from the AIDS and the IAS comes second but the mean deviations from the AIDS are actually smaller than 0.11pp for all specifications apart from the LEO. Nevertheless, the maximum deviation of the CD from the AIDS is larger than 1pp and is thus substantial in comparison to the mean of the total effect measured for the AIDS of 11.6% (table of Figure 6). But also the maximum difference of the CDE is 0.5pp higher than that of the best-performing LES, showing that the choice of specification is relevant here. Considering the IPR, the LES is the most reliable for yielding headcount results close to those of the AIDS, closely followed by the IAS, the CDE, and finally the CD. The CD is again substantially better than the LEO. The poverty gap results reveal a qualitatively equivalent pattern.

Turning to inequality as measured by the Gini index, the main difference is that the CDE outperforms the IAS and is thus second best. It is also worth noting that the LEO does not provide worst-case bounds for inequality measures, such as the Gini, and that again the CD fares better. The final measure, the true cost-of-living index is, in fact, very similar to the HL national equivalent income results reported in Figure 4 with the subtle difference being that here households are weighted democratically. However, the outcome is largely the same, showing that the measured purchasing power loss can differ notably between the specifications but will predominantly turn out very similarly. The LES and IAS results most closely resemble those of the AIDS and the IPR of the CD compares favorably to the CDE.

Results from HL AIDS simulations

		Point difference from Base				
Index	Base	Min.	$Q_{2.5\%}$	Mean	$Q_{97.5\%}$	Max.
Poverty headcount	19.7	9.8	10.2	11.6	14.5	14.9
Poverty gap	5.5	3.4	3.4	4.0	5.1	5.5
Gini	42.7	-1.0	-0.8	0.2	2.4	3.1
True cost-of-living	100.0	-21.3	-21.1	-19.6	-18.6	-18.2

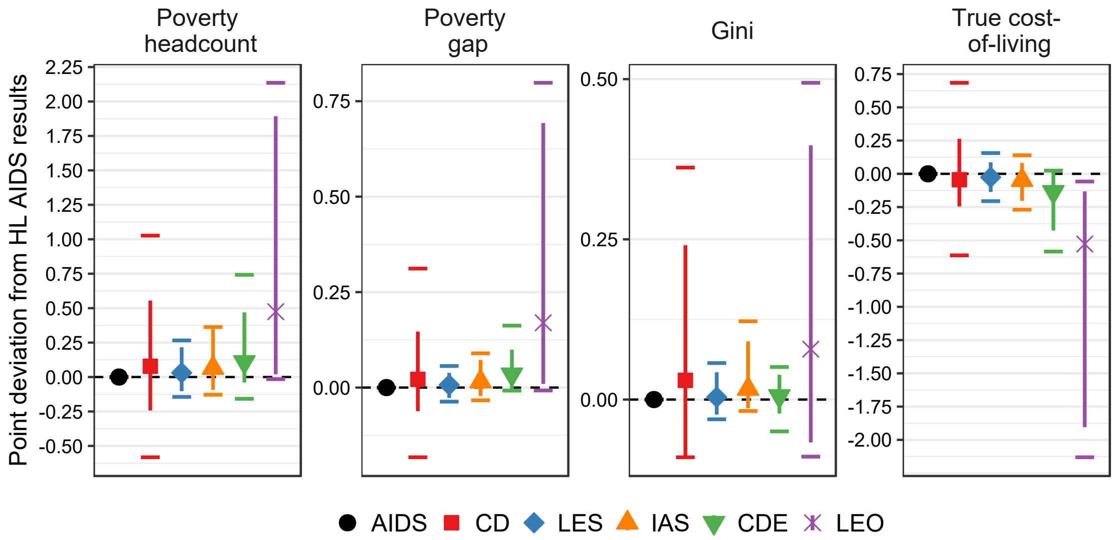

Figure 6: Impacts of a 20% purchasing power loss on poverty and inequality indices. The graph displays the deviations of the HL specifications from the HL AIDS results summarized in the table above. The statistics displayed are the mean, IPR, min. and max. deviation. Source: Own computation from simulation results.

6 Conclusions

Impact evaluation analysis using simulation models has grown into a major domain of economics both in academic research and in policy decision support. But although the corresponding models frequently rely on some kind of representation of consumption demand behavior, little is known about the sensitivity of these models' simulation results – and thus also of the conclusions derived therefrom including policy recommendations – with respect to the choices of the demand system specification and the level of aggregation across consumers.

We investigate these issues with an example of a unique real-world household demand dataset from Uganda. It includes demand data together with calibrated demand system parameters for the CD, LEO, LES, IAS, CDE, and AIDS specifications for all 6,887 individual households from a representative household survey and one aggregate national household. Employing this dataset, various stereotypical scenarios of price and income changes are simulated using household- and national-level demand system models of the different specifications and the results are evaluated with respect to a multitude of outcome measures.

Of all specifications examined, only the AIDS and LES models are not guaranteed to comply with consumer demand theory across the entire price-income space. Thus, it is necessary to check for the distortion they introduce by violating theoretical conditions. Violations indeed occur for both specifications in several household-level simulations. However, although the distortion for an individual household can be large, a measure of the distortions suggests that results are distorted only marginally in aggregate. Moreover, symptoms of the violations are negative budget shares and positive compensated own-price elasticities which can cause impact analysis models to become infeasible, for example, if the demand system is a component of a CGE model. As the AIDS is the most flexible specification considered here and distortions are very small, the AIDS results are taken as approximations of the "true" demand behavior and other specifications are judged against those.

The answer to the research questions raised in the introduction is that both specification and aggregation choices strongly influence the outcome measures, but the magnitude differs by type of shock and outcome measure. The specification matters, whether there is a direct interest in the quantity impacts or the quantity changes induce second-round effects in larger model systems, such as partial or general equilibrium models. Indeed, the specification choice is shown to be more important for the evaluation of quantity impacts from price shocks than a high level of household disaggregation, but aggregation might also cause substantial bias for some products. The LES performed best for this kind of setting while the IAS came second and the CD, CDE, and particularly the LEO lagged far behind. Among the specifications examined, only the AIDS depicts meaningful cross-price effects. The cross-price effects of the remaining specifications turned out to be negligibly small. The aggregation choice matters very little for the cross-price effect in comparison to the specification.

If the main economic shock is on household incomes, the CD and LEO behave identically and are particularly bad choices for the analysis of quantity impacts but there is no clear favorite between the LES, IAS, and CDE. However, while the specification is more important than the aggregation choice if household incomes are uniformly affected, it is the other way around if household incomes change non-uniformly. The poorer the affected households are, the more aggregation matters.

Monetary welfare changes due to random price increases seem surprisingly little affected by the specification choice in the mean. Nevertheless, in extreme cases the deviation can be marked. The LEO, embodying perfect complementarity between all products, provides a lower bound for the welfare effect and all specifications greatly improve on that. The performance of the LES and IAS is clearly superior to the CDE and particularly the CD. The choice of the specification is again more important the poorer the households analyzed are. The ranking between specifications also transfers to the poverty, inequality, and cost-of-living indices calculated from the monetary welfare measure. The change in poverty headcount simulated, for instance, can vary quite strongly between specifications. The importance of the aggregation choice is lower but is still considerable here. Of course, the calculation of poverty and other indices mentioned above is impossible without household disaggregation.

Although the results presented are for a particular case, some conclusions can be generalized and applied to the wider range of cases where income inequality is high and demand elasticities are large. This is the case, for instance, for many developing and emerging economies. As a rule of thumb, demand elasticities are the larger the narrower the consumption items are defined, the more disaggregated the representation of the consumers, and the poorer the consumers. The differences between the various specification or aggregation choices are expected to be much smaller if demand is predominantly inelastic with respect to prices, as is the case for developed countries or for strongly aggregated consumption items or consumers, and if the elasticities have been estimated from aggregate consumption data. In such cases, the implied bounds on price elasticities are likely irrelevant. An example of such a setting is the global CGE model GTAP where national consumption demand is represented by a single national household and demand elasticities have been estimated from cross-country data. There, all items' demands are price-inelastic (see Hertel and van der Mensbrugghe 2016) and thus the restrictiveness of the CDE employed with respect to the price elasticities plays no role so that the CDE might possibly outperform the LES and IAS due to its greater flexibility.

A first conclusion is that even though the AIDS is not globally regular, the resulting distortions on aggregate outcomes have been assessed to be small for the rather large, non-marginal shocks applied here using these particular sets of calibrated AIDS model parameters. As the calibrations of AIDS models to prior demand elasticities resulted in a far better fit than the other specifications, the AIDS seems to be a good choice for micro-level demand simulations if the model system in which it is applied can cope with possible irregularities and small violations of regularity, and if corresponding distortions of the results are tolerable. But since the magnitude of the distortions is not known in advance, and it depends on the size of the shock as well as on the particular set of AIDS model parameter values, assessing the validity of results from an AIDS model-based simulation always requires verification with respect to regularity and potential outcome distortions. However, if regularity over a wide region of the price-income space is imperative, as, for example, in CGE models, the AIDS is not an option. Moreover, often the researcher has at hand only income and possibly own-price elasticities which calls for parsimonious specifications. Among those, the LES and IAS surprisingly turned out to be the best, with the exception of income shocks where they were not the best but were only similar to the others. The IAS's performance is similar to the LES, has the same parsimonious data needs but, in contrast, is also globally regular. Thus, the IAS is a sensible, pragmatic specification choice for use within model systems.

The aggregation choice is particularly important if non-uniform income changes are simulated and poorer households are affected, but has a notable influence in most simulations. Results also show that the CD is a substantially better choice for simulating the quantity, welfare, poverty, or inequality impacts of price changes than the widely applied practice of simple price mapping as embodied in the LEO. Thus, adopting the CD for such studies would improve the derived policy recommendations without requiring additional data, and the LEO could in addition be used to provide worst-case bounds.

It should be noted that the dataset used has been calibrated giving equal weight to income and price elasticities. A different weighting, for instance, only on income, could trade off performance in price for performance in income shock simulations.

To date, the LEO, CD, LES, and CDE are, with few exceptions, the consumption demand specifications predominantly found in economic impact analysis simulation models. And most models use little or no household disaggregation. This study has shown in which settings these choices have notable consequences for the results and the policy recommendations derived from those. Nevertheless, one also needs to bear in mind that in an applied context the choices hinge not only on what the "best" demand representation is, but also on practical considerations of data, numerical solvability, and time constraints for the implementation. Our work has pointed to some pragmatic specification choices which promise some improvement over the current practice.

In the search for a practicable improvement over the current specifications commonly applied in impact analysis models, future research will extend the analysis to a number of recent demand systems which combine flexibility with effectively global regularity, such as those introduced by Cooper et al. (2015) and McLaren and Yang (2016). Another extension will aim at rank three demand systems which more realistically model how income elasticities change with the level of income. These allow a richer demand representation at the cost of additional complexity and data needs.

References

Anderson, K., Martin, W. and van der Mensbrugghe, D. (2006). Would multilateral trade reform benefit Sub-Saharan Africans?, *J. African Economies* **15**(4): 626–670.

Banks, J., Blundell, R. and Lewbel, A. (1996). Tax reform and welfare measurement: Do we need demand system estimation?, *Econ. J.* **106**(438): 1227–41.

Banks, J., Blundell, R. and Lewbel, A. (1997). Quadratic Engel curves and consumer demand, *Rev. Econ. Statist.* **79**(4): 527–539.

Barnett, W. A. and Serletis, A. (2008). Consumer preferences and demand systems, *J. Econometrics* **147**(2): 210–224.

Blundell, R. and Stoker, T. M. (2007). Models of aggregate economic relationships that account for heterogeneity, *in* J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Vol. 6, Elsevier, Amsterdam, pp. 4609–4666.

Bourguignon, F., da Silva, L. A. P. and Bussolo, M. (eds) (2008). *The impact of macroeconomic policies on poverty and income distribution : macro-micro evaluation techniques and tools*, The World Bank and Palgrave MacMillan, Washington, DC.

Boysen, O. (2016). Food demand characteristics in Uganda: Estimation and policy relevance, *South African J. Econ.* **84**(2): 260–293.

Boysen, O., Miller, A. C. and Matthews, A. (2016). Economic and household impacts of projected policy changes for the Irish agri-food sector, *J. Agr. Econ.* **67**(1): 105–129.

Caves, D. W. and Christensen, L. R. (1980). Global Properties of Flexible Functional Forms, *Amer. Econ. Rev.* **70**(3): 422–32.

Cooper, R. J. and McLaren, K. R. (1996). A system of demand equations satisfying effectively global regularity conditions, *Rev. Econ. Statist.* **78**(2): 359–64.

Cooper, R. J., McLaren, K. R., Rehman, F. and Szewczyk, W. A. (2015). Economic welfare evaluation in an era of rapid technological change, *Econ. Letters* **131**: 38–40.

Cranfield, J. A. L., Eales, J. S., Hertel, T. W. and Preckel, P. V. (2003). Model selection when estimating and predicting consumer demands using international, cross section data, *Empirical Econ.* **28**(2): 353–364.

Dawkins, C., Srinivasan, T. and Whalley, J. (2001). Calibration, *in* J. J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Vol. 5, Elsevier, Amsterdam, pp. 3653 – 3703.

de Boer, P. (2010). Modeling household behavior in a CGE model: Linear expenditure system or indirect addilog?, *DEGIT Conference Paper c015_059*, Dynamics, Economic Growth, and International Trade, DEGIT - XV Conference, 3 to 4 September 2010, Frankfurt, Germany.

de Boer, P. and Paap, R. (2009). Testing non-nested demand relations: linear expenditure system versus indirect addilog, *Statistica Neerlandica* **63**(3): 368–384.

de Janvry, A. and Sadoulet, E. (2010). The global food crisis and Guatemala: What crisis and for whom?, *World Devel.* **38**(9): 1328 – 1339.

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

Deaton, A. S. and Muellbauer, J. (1980b). An almost ideal demand system, *Amer. Econ. Rev.* **70**(3): 312–26.

Decoster, A. and Schokkaert, E. (1990). Tax reform results with different demand systems, *J. Public Econ.* **41**(3): 277–296.

Fisher, D., Fleissig, A. R. and Serletis, A. (2001). An empirical comparison of flexible demand system functional forms, *J. Appl. Econometrics* **16**(1): 59–80.

Foster, J., Greer, J. and Thorbecke, E. (1984). A class of decomposable poverty measures, *Econometrica* **52**(3): 761–66.

Gohin, A. (2005). The specification of price and income elasticities in computable general equilibrium models: An application of latent separability, *Econ. Modelling* **22**(5): 905–925.

Golan, A., Judge, G. and Miller, D. (1996). *Maximum Entropy Econometrics*, Wiley, New York.

Gorman, W. M. (1981). Some Engel curves, *in* A. Deaton (ed.), *Essays in the Theory and Measurement of Consumer Behaviour in Honor of Sir Richard Stone*, Cambridge University Press, Cambridge, UK, pp. 7–29.

Hanoch, G. (1975). Production and demand models with direct or indirect implicit additivity, *Econometrica* **43**(3): 395–419.

Hertel, T. (2013). Global applied general equilibrium analysis using the Global Trade Analysis Project framework, *in* P. B. Dixon and D. W. Jorgenson (eds), *Handbook of Computable General Equilibrium Modeling*, Vol. 1, Elsevier, Amsterdam, chapter 12, pp. 815 – 876.

Hertel, T. and van der Mensbrugghe, D. (2016). Behavioral parameters, *in* B. Narayanan, A. Aguiar and R. McDougall (eds), *GTAP 9 Data Base Documentation*, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University, West Lafayette, IN, chapter 14.

Hertel, T. W. (ed.) (1997). *Global Trade Analysis: Modeling and Applications*, Cambridge University Press, Cambridge, UK.

Houthakker, H. S. (1960). Additive preferences, *Econometrica* **28**(2): 244–257.

Jensen, B., de Boer, P., van Daal, J. and Jensen, P. (2011). Global restrictions on the parameters of the CDES indirect utility function, *J. Econ.* **102**(3): 217–235.

King, M. A. (1983). Welfare analysis of tax reforms using household data, *J. Public Econ.* **21**(2): 183 – 214.

Klevmarken, N. (1979). A comparative study of complete systems of demand functions, *J. Econometrics* **10**(2): 165–191.

Lewbel, A. (1991). The rank of demand systems: Theory and nonparametric estimation, *Econometrica* **59**(3): 711–730.

Madden, D. (1996). Marginal tax reform and the specification of consumer demand systems, *Oxford Econ. Pap.* **48**(4): 556–567.

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

McLaren, K. R. and Yang, O. (2016). A class of demand systems satisfying global regularity and having complete rank flexibility, *Empirical Econ.* **51**(1): 315–337.

Parks, R. W. (1969). Systems of demand equations: An empirical comparison of alternative functional forms, *Econometrica* **37**(4): 629–650.

Perroni, C. and Rutherford, T. F. (1998). A comparison of the performance of flexible functional forms for use in applied general equilibrium modelling, *Computational Econ.* **11**(3): 245–63.

Pollak, R. A. (1971). Additive utility functions and linear Engel curves, *Rev. Econ. Stud.* **38**(4): 401–414.

Pollak, R. A. and Wales, T. J. (1995). *Demand System Specification and Estimation*, Oxford University Press, New York.

Ray, R. (1986). Sensitivity of 'optimal' commodity tax rates to alternative demand functional forms, *J. Public Econ.* **31**(2): 253–268.

Ray, R. (1999). Marginal and non-marginal commodity tax reforms with rank two and rank three demographic demand systems, *Oxford Econ. Pap.* **51**(4): 689–712.

Ryan, D. L. and Wales, T. J. (1998). A simple method for imposing local curvature in some flexible consumer-demand systems, *J. Bus. Econ. Statist.* **16**(3): 331–38.

Simler, K. R. (2010). The short-term impact of higher food prices on poverty in Uganda, *Policy Research Working Paper 5210*, The World Bank, Washington, DC.

Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand, *Econ. J.* **64**(255): 511–527.

Sydsæter, K., Hammond, P. and Strøm, A. (2012). *Essential mathematics for economic analysis*, 4th edn, Pearson Education Limited, Harlow, UK.

UBOS (2014). *Uganda National Household Survey 2012/2013*, Uganda Bureau of Statistics, Kampala, Uganda.

Wodon, Q. and Zaman, H. (2010). Higher food prices in Sub-Saharan Africa: Poverty impact and policy responses, *World Bank Res. Observer* **25**(1): 157–176.

Wolff, H., Heckelei, T. and Mittelhammer, R. (2010). Imposing curvature and monotonicity on flexible functional forms: An efficient regional approach, *Computational Econ.* **36**(4): 309–339.

Yu, W., Hertel, T. W., Preckel, P. V. and Eales, J. S. (2004). Projecting world food demand using alternative demand systems, *Econ. Modelling* **21**(1): 99–129.

Appendices

A The Quadratic Almost Ideal Demand System (QUAIDS)

The QUAIDS was introduced by Banks et al. (1997) and is an extension of the AIDS due to Deaton and Muellbauer (1980b, see section 3.5). It is defined by the budget share eq. (37), the price indices $a(p)$ (38) and $b(p)$ (39), and restrictions to impose adding up (40), homogeneity (41), and symmetry (42). The income and price elasticities are defined by eqs. (43) and (44).

$$\begin{aligned} w_i =& \alpha_i + \sum_{j \in I} \gamma_{ij} \ln p_j + \beta_i \ln \left(\frac{M}{a(p)} \right) \\ &+ \frac{\lambda_i}{b(p)} \left(\ln \left(\frac{M}{a(p)} \right) \right)^2 \end{aligned} \qquad \forall i \in I \qquad (37)$$

$$\ln a(p) = \alpha_0 + \sum_{i \in I} \alpha_i \ln p_i + \frac{1}{2} \sum_{i \in I} \sum_{j \in I} \gamma_{ij} \ln p_i \ln p_j \qquad (38)$$

$$b(p) = \prod_{i \in I} p_i^{\beta_i} \qquad (39)$$

$$\sum_{i \in I} \alpha_i = 1, \quad \sum_{i \in I} \beta_i = 0, \quad \sum_{i \in I} \lambda_i = 0, \quad \sum_{i \in I} \gamma_{ij} = 0 \qquad \forall j \in I \qquad (40)$$

$$\sum_{j \in I} \gamma_{ij} = 0 \qquad \forall i \in I \qquad (41)$$

$$\gamma_{ij} = \gamma_{ji} \qquad \forall i, j \in \{I | i < j\} \qquad (42)$$

$$\eta_i = 1 + \frac{1}{w_i} \left[\beta_i + \frac{2\lambda_i}{b(p)} \ln \left(\frac{M}{a(p)} \right) \right] \qquad \forall i \in I \qquad (43)$$

$$\begin{aligned} \varepsilon_{ij} =& - \delta_{ij} + \frac{1}{w_i} \left[\gamma_{ij} - \left(\beta_i + \frac{2\lambda_i}{b(p)} \ln \left(\frac{M}{a(p)} \right) \right) \right. \\ & \left. \left(\alpha_j + \sum_{k \in I} \gamma_{jk} \ln p_k \right) - \frac{\lambda_i \beta_j}{b(p)} \left(\ln \left(\frac{M}{a(p)} \right) \right)^2 \right] \end{aligned} \qquad \forall i, j \in I \qquad (44)$$

with Kronecker δ_{ij} (see section 3.3). The QUAIDS has $(n^2+3n)/2-1$ independent parameters and is thus flexible. In comparison to the rank two AIDS that restricts budget shares to be linear in the log of income, the QUAIDS is of rank three and adds flexibility to the Engel curves representable by adding a term quadratic in the log of income. As a consequence, the QUAIDS allows income elasticities to vary with the level of income such that an item might be a luxury for low and a necessity for high income levels. The QUAIDS specification does not guarantee compliance with the monotonicity or curvature properties and thus might not be regular even for the prices-income points of the sample observations. McLaren and Yang (2016) state that its regularity properties are not better than those of the AIDS. The QUAIDS does not impose any bounds on the value ranges of the income and price elasticities.

B Generalized Cross-Entropy Estimation Model

The goal of this estimation is to generate a theoretically consistent set of parameters for a specific demand system which precisely replicates the given budget shares $w_i = \frac{p_i q_i}{M}$ (q_i denotes the quantity demanded) at a given income (M) and a vector of prices p_i while fitting a given set of prior income $\overline{\eta}_i$ and price elasticities $\overline{\varepsilon}_{ij}$ as closely as possible, where $i, j \in I$ and I denotes the set of consumption items.

Eqs. (45) to (50) form the core model of the GCE estimation which then needs to be amended with the specific constraints of the particular demand system model:

$$\begin{aligned} \min H(\pi) = & \sum_{i \in I} w_i \sum_{l \in L} \left(\pi_{il}^{\eta} \ln \frac{\pi_{il}^{\eta}}{\overline{\pi}_{il}^{\eta}} + \pi_{iil}^{\varepsilon} \ln \frac{\pi_{iil}^{\varepsilon}}{\overline{\pi}_{iil}^{\varepsilon}} \right) \\ & + \left(\sum_{i \in I} \sum_{j \in \{I | j \neq i\}} w_i w_j \sum_{l \in L} \pi_{ijl}^{\varepsilon} \ln \frac{\pi_{ijl}^{\varepsilon}}{\overline{\pi}_{ijl}^{\varepsilon}} \right) \end{aligned} \tag{45}$$

subject to

$$\eta_i = \sum_{l \in L} \pi_{il}^{\eta} z_{il}^{\eta} \qquad \forall i \in I \tag{46}$$

$$\varepsilon_{ij} = \sum_{l \in L} \pi_{ijl}^{\varepsilon} z_{ijl}^{\varepsilon} \qquad \forall i, j \in I \tag{47}$$

$$\sum_{l \in L} \pi_{il}^{\eta} = 1 \qquad \forall i \in I \tag{48}$$

$$\sum_{l \in L} \pi_{ijl}^{\varepsilon} = 1 \qquad \forall i, j \in I \tag{49}$$

$$\pi_{il}^{\eta}, \pi_{ijl}^{\varepsilon} \geq 0 \qquad \forall i, j \in I, \, l \in L \tag{50}$$

The cross-entropy objective function (45) minimizes the distance – as measured by the Kullback-Leibler measure (see, e.g., Golan et al. 1996) of information divergence – of the estimated probability distribution $\pi_{i(j)l}^{s}$ from their prior probabilities $\overline{\pi}_{i(j)l}^{s}$ where $s \in S = \{\eta, \varepsilon\}$ indicates parameters associated with the income and price elasticities, respectively. The prior probabilities are chosen such that their implied expected value is equal to the prior elasticity. The importance of each item's elasticity with respect to the entire demand system is assumed to be proportional to the corresponding budget share w_i which enters the objective function as a weight. Since the number of cross-price elasticities is quadratically related to the number of items, these are weighted by the product of the budget shares of the two items involved $w_i w_j$.[21] Eqs. (46) and (47) link the income and price elasticity variables η_i and ε_{ij} to their expected values, defined as the sum over the support points $z_{i(j)l}^{s}$ (indexed by $l \in L$) multiplied by their respective probabilities $\pi_{i(j)l}^{s}$. Equations (48) and (49) ensure that probabilities sum to one. In

[21] Note that the cross-price elasticity variables could be reduced to the upper or lower triangle of the matrix without loss of information if the estimates are known to satisfy the symmetry property.

addition, restrictions (50) confine probability variables to the non-negative domain.

The vector of support points associated with a specific elasticity defines the support range in which the estimated elasticity is allowed to fall. Since the estimated elasticity, for example, η_i, is calculated as the sum of all support points multiplied by their associated probabilities (eq. 46), the prior elasticity is chosen as the midpoint of the support point vector z_{il} and the support points extend equidistantly from each other and symmetrically around the midpoint. For example, if the prior income elasticity is $\eta_i = 1.2$, a support space of $|L| = 5$ points is chosen and it is assumed that the error around the elasticity is uniformly distributed (the *prior distribution*) $\pi_i^\eta = \overline{\pi}_i^\eta = (0.2, 0.2, 0.2, 0.2, 0.2)$, then $z_i^\eta = (0.2, 0.7, 1.2, 1.7, 2.2)$ so that eq. (46) initially evaluates to the prior income elasticity $\overline{\eta}_i$. Alternatively, the error might be assumed to be normally distributed around the prior elasticity, which can be accomplished by assigning different probabilities to $\overline{\pi}_i^\eta$ which approximate a normal distribution function. This attributes a higher probability to a support point the closer it is to the midpoint of the support range and thus assumes there is information about the distribution of the error around the prior elasticity.

To complete the GCE demand system estimation model, the core GCE model needs to be complemented by equations which define the demand system (budget share equations and parameter domain definitions) and its income and price elasticities.