Essays in Economic Measurement and Consumer Behavior

Serhat Ugurlu

To my p.w. Melisa . . .

Ingvild Almås
The IIES, Stockholm University, Stockholm, Sweden.
FAIR, Norwegian School of Economics (NHH), Bergen, Norway.
Almås acknowledges valuable support from Vetenskapsrådet (the Swedish Research Council), ESOP, University of Oslo, The Choice Lab, Norwegian School of Economics, and the Young Research Talents program of the Norwegian Research Council.

Thomas F. Crossley
University of Essex and Institute for Fiscal Studies, London, UK.
Crossley gratefully acknowledges support for this research from the Economic and Social Research Council (ESRC) through the Research Centre on Micro-Social Change (MiSoC) at the University of Essex (grant number ES/L009153/1) and Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (grant number ES/M010147/1).

Mandeep Grewal
Norwegian School of Economics (NHH), Bergen, Norway.

Marielle Hvide
Norwegian School of Economics (NHH), Bergen, Norway.

Serhat Ugurlu
Norwegian School of Economics (NHH), Bergen, Norway.
Ugurlu acknowledges support from Professor Wilhelm Keilhaus Minnefond Fund for Economic Research.

Acknowledgements

I would like to express my great appreciation to Prof. Ingvild Almås and Prof. Thomas F. Crossley, my research supervisors, for their valuable and constructive suggestions, encouragement, contributions to my work and willingness to discuss my ideas. This thesis would not have been possible without their expert guidance.

I would like to offer my gratitude to all members of FAIR at the Norwegian School of Economics. I am particularly grateful to Prof. Alexander Wright Cappelen, Prof. Erik Øiolf Sørensen, and Prof. Bertil Tungodden for their counseling, and Janina Ramona Juranek for her support.

I would like to extend my gratitude to all members of the Department of Economics. Especially, I thank to Dr. Lassi Ahlvik, Dr. Thomas de Haan, Christel Florvåg, Prof. Gernot Peter Doppelhofer, Irene Grønningsæter, Trude Gudmundset, Dr. Torfinn Harding, Dagny Hanne D. Kristiansen, Prof. Eirik Gaard Kristiansen, Dr. Yuanhao Li, Dr. Krisztina Molnar, Prof. Kjell Gunnar Salvanes, Prof. Fred Schroyen, Dr. Vincent Somville, Prof. Siri Pettersen Strandenes, Prof. Øystein Thøgersen, and Dr. Håkon Tretvoll for discussions, suggestions, and lunch and coffee-talks, which made my studies at NHH truly memorable.

I am particularly grateful for the assistance given by Mandeep Grewal and Marielle Hvide. Their excellent work formed a basis for the second chapter of this thesis.

I would like to offer my special thanks to Thor Andreas Aursland, Antonio Dalla-Zuanna, Luca Picariello, and Ingrid Hoem Sjursen for their support and friendship. With their excellent focus on high-quality research, they have been an outstanding cohort that continually aided me to strive to achieve the best. I also thank Fehime Ceren Ay, Ranveig Falch, Ingar Kyrkjebø Haaland, Stefan Meissner, Felix Morency-Lavoie, Charlotte Ringdal, Erling Risa, Ingvild Lindgren Skarpeid, Timothy G. A. Wyndham, and Xiaogeng Xu for being supportive colleagues.

I am thankful to A. Muge Polat Sayli, Melisa Sayli, Mustafa Sayli, Begum Ugurlu, Hicran Ugurlu, and Suleyman Ugurlu, my family, for their endless support during my studies.

I gratefully acknowledge valuable support from Professor Wilhelm Keilhaus Minnefond Fund for Economic Research.

Preface

It is a time of great opportunities for empirical work across many disciplines. Thanks to developments in data collection and analysis, researchers today have access to a substantial number of data sets in unprecedented detail. It feels natural to speculate that, in the near future, there will be better, richer, and larger data sets that allow researchers to formulate new empirical problems or to approach some existing problems with different points of views. Jean et al. (2016) and Athey and Imbens (2017) are some already available examples.

In terms of economic research, benefits of these developments are especially observed in collection of price information (Cavallo and Rigobon, 2016; Cavallo et al., 2018). Previously difficult and costly to obtain prices can be collected with high frequency and in enormous batches using online tools or scanners. Such data sets provide detailed and comparable information about economic environments which individuals and nations face (Cavallo, 2017). Such information, by itself, has a considerable importance for statistical offices. In addition, for academics and policymakers, who are interested in counterfactual policy analyses, prices are an important element of an individual's decision mechanism.

Even though data is becoming increasingly available, in order to explain an economic phenomena, not only availability of information but also guidance of economic theories is required (Frisch, 1933). With this vision, this thesis is a collection of works that attempts to lay grounds for connecting behavior of different individuals (consumers) with measurement of aggregate price differences across societies. The underlying motivation for this combination follows from the observation that, when consumers, who form a society, have different tastes towards the same goods, there is no single measure of aggregate price differences that can truly explain impacts of these differences on each individual (Prais, 1959; Nicholson, 1975). Unquestionably, a change in prices is a movement from one set of numbers to another, which is a single statistical fact. However, when individuals with heterogeneous tastes respond differently to changes in prices, individuals experience disproportionate impacts on their welfare levels. Thus, different consumers require different adjustments to their budgets as a result of the same price change. In such environments, where no single index is perfect, an economic theory is required to develop an aggregate price index that respects these differences and that still has a clear interpretation.

Understanding impacts of different prices or policies when individuals differ requires well-defined theories of individual behavior (King, 1983). A focus on individuals, which are the smallest decision units of an economy, relates to the aforementioned developments in data collection and analysis. This is because such a focus requires knowledge, or elicitation, of an individual's preferences over different feasible choices. Of course, a focus on individuals instead of more aggregate decision units has been commonplace in economic research. However, increasing size and detail of micro-level information through nation-wide surveys, or scanner data sets, combined with prices and new empirical approaches to work with such data sets, provide us with new insights in terms of quantifying potentially different preferences of consumers.

Therefore, the first step to form a connection between individual behavior and price measurement is to have an empirical model to elicit preferences so that a counterfactual analysis of individual behavior is possible. Such models allow us to estimate impacts of price differences across individuals with different tastes. Thus, in the first chapter of this thesis, I revisit the literature on micro-econometric analysis of consumer behavior. I emphasize the focus of this literature on functional form flexibility while maintaining a link between theoretically expected behavior and empirical methodology. Agreeing with this focus, I develop an estimation strategy that provides certain improvements over existing approaches in terms of functional approximation, scalability, inclusion of observable heterogeneity, and theoretical regularity. I achieve this by combining individual behavior as characterized by the neoclassical theory of consumer demand and a semi-nonparametric estimation strategy from the toolbox of machine learning and statistics. The underlying idea in this combination is to benefit from developments in non-parametric applications without losing the viewpoint of an economic theory (see, e.g., Mullainathan and Spiess, 2017). Thus, I wish to provide some structure to an otherwise entirely data-driven methodology, so that an individual's behavioral predictions adhere to our expectations about how a rational individual behaves. In other words, the purpose of this work is to obtain an empirical model of consumer behavior without any assumptions other than those the theory provides. In the chapter, my primary aim is to develop the approach and to illustrate its feasibility through comparisons with some prominent approaches. Further work would signify its benefits in high dimensional non-separable choice settings and in maintaining theoretical regularity.

Once we have a theoretically and statistically consistent way to elicit preferences of an individual, the second step is to illustrate a use of price measures across societies, or purchasing power parity indices, so that the importance of accurate cross-country comparisons of prices is clear to the reader. For this purpose, as an example, the second chapter of this thesis revisits the PPP approach to analysis of currency misalignments. In this context, a purchasing power parity is employed to obtain price levels, or real exchange rates, for each country so that misalignment of the value of a currency with respect to its long-run value is measured. We specifically evaluate alleged undervaluation of the Chinese renminbi against the US dollar with arguably the most suitable data set of internationally comparable prices: ICP 2011 (World Bank, 2014). This study shows that there is no empirical evidence to suggest that the renminbi is undervalued. Furthermore, it provides an example of why price comparisons across societies are important economic measures; they help us to deflate incomparable nominal values to comparable real values.

Price indices that we obtain from ICP 2011 are constructed using a leading price index methodology GEKS. GEKS is an approach that takes a widely used price index (Fisher) to measure price differences across two societies as input and returns a suitable index for international price comparisons. This means satisfying some index properties. An excellent property of the Fisher index is its consistency with individual behavior for certain types of preferences (Diewert, 1976). However, when individuals differ, this consistency is violated. Nonetheless, one can obtain a single index number with an explicit interpretation by focusing on impacts of price differences on each individual in a society. Chapter three surveys different approaches that can be used to obtain constant-social-welfare price indices. We develop these indices by aggregating individual costs of maintaining consumption bundles, living standards, or relative income levels at home and foreign prices. A starting point of this chapter is the literature on social inflation measures, notably works of Pollak (1980) and Crossley and Pendakur (2010). This literature develops measures of price differences with heterogeneous individuals; however, its focus on incorporating individual preferences into temporal price indices causes a limited consistency with index properties that are imperative for purchasing power parities. Thus, while developing spatial price measures in the same spirit with the temporal literature, we also consider how some index properties can be maintained. Our results provide price indices with explicit interpretations and consistency with thought-experiments involving optimizing individuals

Thus, chapters of this thesis discuss importance and illustrate feasibility of combining structural economic models and a variety of empirical tools to develop modular, flexible, and still interpretable approaches to measure important economic values, such as elasticities or real consumption, that shape policy mechanisms. Developments in empirical literatures across many different disciplines of science address important issues in terms of reliability of numerical outcomes in empirical research. However, it is the consistency with economic theories that provides a clear interpretation to these outcomes, for example in terms of welfare. Thus, after many decades of Frisch (1933), it is as important to keep a focus on theories while embedding the state-of-the-art in policy-informative empirical economic research.

Contents

Chapter 1
An Analysis of Consumer Behavior with Feedforward Neural Networks

Serhat Ugurlu

Abstract To obtain theoretically and statistically consistent estimates of consumers' demand functions, I suggest a functional approximation approach by combining a feedforward neural network (FNN) estimation with the neoclassical theory of consumer demand. I illustrate the viability of this approach by providing comparisons with the parametric demand models and the semi-nonparametric FNN models for a system of non-separable demand equations for ten aggregate food categories. I present estimates of demand equations using two parametric models: AIDS (Deaton and Muellbauer, 1980a) and QUAIDS (Banks et al., 1997), and two FNN models: an atheoretical model that does not consider theoretical shape restrictions (ADN), and a theoretical model that satisfies these restrictions (TDN). I present estimated elasticities and welfare impacts of price changes with confidence intervals. My empirical application with a UK household-level dataset shows that all models yield similar estimates of demand functions and elasticities for aggregate food categories.

Key words: consumer demand, nonparametric, neural network, elasticity

1.1 Introduction

Answers to many policy related questions require an understanding of how consumers respond to changes in prices and incomes, such as impacts of tax policies on state revenues and average welfare (see, e.g., Abramovsky et al., 2015), calculations of cost of living indices (Neary, 2004; Crossley and Pendakur, 2010), and measurement of inequality (Almås, 2012; Aguiar and Bils, 2015). The neoclassical theory of consumer demand provides a theoretical understanding of how consumers behave (Deaton and Muellbauer, 1980b). In order to have an empirical understanding, different approaches are developed to evaluate consumers' responses to changes in prices and incomes as elasticities (Barnett and Serletis, 2008). Specifically, parametric empirical approaches feature a close relationship with the theory of consumer demand (Stone, 1954; Theil, 1965; Christensen et al., 1975; Deaton and Muellbauer, 1980a; Banks et al., 1997). They provide suitable estimation techniques that are designed to test, or impose, individual rationality in a tractable way (Blundell, 1988).[1] However, parametric approaches impose not only rationality but also certain functional forms that are either ad hoc but fit well to data in application, or derived from utility (or expenditure) functions that individuals are assumed to have. Theoretically, there is no reason for observed behavior to be consistent with imposed functional forms. When functional forms are incorrect, estimates of elasticities are subject to a specification bias and they are statistically (asymptotically) inconsistent.

One alternative to solve functional form uncertainty is to adopt a non-parametric approach (see, e.g., Lewbel, 1991; Hausman and Newey, 1995). However, although non-parametric approaches yield statistically consistent estimates of elasticities, they introduce additional challenges. First, the connection between the theory and the estimation strategy is not as clear as it is for parametric approaches. To address this issue, methodologies to combine some of the rationality conditions with certain types of non-parametric estimators have been proposed (Haag et al., 2009; Blundell et al., 2012, 2016). However, no non-parametric estimation that satisfies all rationality conditions simultaneously within the support of a dataset has been established. Second, without restrictive separability assumptions, demand functions are estimated simultaneously to gain statistical efficiency and because they integrate to the same utility function (Lewbel, 1997). As a result, demand functions share coefficients to satisfy cross-equation rationality restrictions. For systems of demand equations with many commodities, which could also be enriched with other observable variables such as demographics, it is even harder to achieve a simultaneous non-parametric estimation that respects theoretical restrictions.

In this paper, in order to obtain asymptotically consistent estimates of theoretically consistent systems of non-separable demand equations, I take a data driven approach to estimate price and income elasticities with a semi-nonparametric functional approximation. To this end, I combine the theory of consumer demand from the economic toolbox and feedforward neural networks (FNN) from the machine learning toolbox. I demonstrate that this combination offers an intermediate alternative to parametric and non-parametric approaches to microeconometric modeling of consumer behavior by delivering a system of non-separable demand equations without a priori functional form assumptions while maintaining the theoretical benefits of a parametric approach.[2] I illustrate the feasibility of the proposed methodology with state-of-the-art optimization algorithms.

A feedforward neural network (FNN), among some alternative machine learning approaches that can also be used to have an understanding of consumer behavior

[1] Rationality conditions provide necessary properties that a demand function must exhibit if it is derived from the utility maximization problem of a rational individual. These conditions are homogeneity, adding up, symmetry, and negativity (Hurwicz and Uzawa, 1971).

[2] This is related in spirit to Elbadawi et al. (1983), who show that a consistent estimation of price and income elasticities is possible without functional form restrictions using a Fourier flexible form (see also Gallant, 1981). In fact, the Fourier flexible form is a special case of a feedforward neural network design (Gallant and White, 1988).

(see, e.g., Bajari et al., 2015), provides the following advantages. First, FNNs yield a continuous functional approximation with a data dependent explicit functional representation. This structure addresses the functional form uncertainty and provides a differentiable functional approximation with favorable approximation properties (see, e.g., Cybenko, 1989; Hornik et al., 1990; Hornik, 1991).[3] As a result, price and income elasticities are easy to obtain as functions of partial derivatives of the estimated empirical demand functions. Second, estimation of an FNN is scalable to large numbers of quantitative inputs and outputs. Indeed, for high dimensional input vectors, FNNs have more favorable approximation properties than traditional series and non-parametric curve based methodologies (Barron, 1994). Therefore, an FNN is an ideal candidate to obtain estimates of demand functions with many commodities and observable variables. Third, FNNs yield asymptotically consistent estimates if hyper-parameters that control complexity are allowed to change with sample size, which can be achieved with cross-validation (White, 1990; Geman et al., 1992).

Despite these advantages, artificial neural networks have not been used as an off-the-shelf estimation strategy for economic research. One reason for this is the "black-box" nature of these methodologies; in empirical economics, researchers are usually interested in finding interpretable estimates for parameters of interest. However, artificial neural networks do not provide such interpretable coefficients. In this paper, I demonstrate that if the aim is to obtain consistent estimates of price and income elasticities, the limitation on interpretation is not restrictive because the elasticities are obtained as functions of derivatives of the estimated demand functions. More generally, the limitation on interpretability is not restrictive when the research question is not related to the coefficients per se but to the output of the estimation (see, e.g., Altman et al., 1994; Kuan and Liu, 1995; Swanson and White, 1997; Jean et al., 2016; Mullainathan and Spiess, 2017). Another reason is the difficulty of obtaining FNN estimations using derivative-free or derivative-based optimization routines. However, with recent developments in optimization algorithms and new gradient-based FNN solvers, FNNs are widely applied to solve a variety of regression and classification problem with remarkable success (Rios and Sahinidis, 2013; Choromanska et al., 2015; LeCun et al., 2015). Therefore, a good match between the research question and the estimation methodology, and developments in solving these estimators make FNNs a viable alternative for an empirical modeling of consumer behavior.[4]

In fact, there are previous examples of FNN applications to model consumer behavior. Joerding and Li (1994) develop a modified simulated annealing algorithm (a derivative-free optimization algorithm) and illustrate an application of their algorithm by estimating demand equations for three aggregate categories using country-

[3] For every continuous function defined on a compact set, and $\varepsilon > 0$, there exists an FNN $\hat{f}$ such that $\| f - \hat{f} \| < \varepsilon$. Conditions that generate this approximation property are outlined by the universal approximation theorem of FNNs.

[4] The literature of feedforward neural network estimation dates back a few decades. However, with the advancements in computational capacity, distributed computations, data size and variety, research on estimations of neural networks, and related developments in optimization algorithms make feedforward neural networks a computationally viable option today.

level time-series data. They derive rationality restrictions as constraints that are suitable for their optimization algorithm. McAleer et al. (2008) use a neural networks based approach to augment some existing parametric approaches for more flexibility in the effects of income and demographic variables on consumer behavior.

I contribute to this literature by approaching the estimation of a system of non-separable demand equations semi-nonparametrically with a fully specified FNN and with modern derivative-based optimization routines, by integrating neoclassical restrictions to such routines to obtain theoretically and statistically consistent estimates, and by demonstrating feasibility of these methodologies to microeconometric modeling of consumer demands with household-level data. To this aim, I present two easy to apply FNN estimators: first, an atheoretical demand network (ADN) that does not impose functional shape restrictions of rational behavior and presents a semi-nonparametric fit to a system of demand equations; second, a theoretical demand network (TDN) that imposes shape restrictions of rational behavior through a network's architecture and constraints on optimization procedures.

Imposing rationality restrictions and obtaining theoretically consistent estimates of demand equations are important for a variety of reasons. First, imposing these restrictions generates an integrable demand system, which makes a welfare analysis possible. Thus, using such a demand system, one can make an evaluation of impacts of changes in prices and budgets on an individual's welfare. Second, integrating a priori knowledge of how demand functions should look like into network architectures and estimation procedures, or giving a structure to these semi-nonparametric estimators, yields well-behaved estimates, for example non-negative expenditure levels, that may improve estimation accuracy (see also Blundell et al., 2012).

To illustrate feasibility of estimation methodologies using FNNs, I estimate a demand system for ten aggregate food categories using a UK household-level dataset. ADN and TDN estimates of demand systems are obtained and compared with estimates from prominent parametric alternatives AIDS (Deaton and Muellbauer, 1980a) and QUAIDS (Banks et al., 1997). Evaluated at the budget axis, my empirical application suggests a mostly linear system of demand equations, which does not exhibit non-linearities that parametric methodologies are unable to capture. Even though the relationships are mostly linear, results suggest that ADN and TDN approaches perform as well as the parametric alternatives. These illustrations show that using ADN or TDN as an alternative to estimate consumers' responses to changes in prices and incomes is indeed feasible; however, it may not always be necessary. Future research will show whether this holds up for other applications. One would expect that if some inputs affect consumption decision in a highly non-linear fashion, the differences between the methods may be larger.

The rest of the article proceeds as follows. In section 1.2, I formalize the functional estimation problem, model an FNN architecture that is suitable for an analysis of consumer behavior, discuss the integrability conditions, illustrate their implementation, and characterize the estimation procedure. In section 1.3, I describe the dataset, present results from four estimations, and compare them using mean squared errors on an independent test set. I also highlight differences, if any, in predicted welfare effects of hypothetical price changes. Section 1.4 presents a discussion of the results. Section 1.5 concludes.

1.2 Methodology

In order to demonstrate feasibility of a theoretically and statistically consistent FNN estimation that is suitable for a microeconometric analysis of consumer demand, I first outline properties that a demand function must satisfy for consistency with behavior of a rational individual. Second, I briefly describe a general architecture of FNNs. Then, I present two FNN estimators: atheoretical and theoretical demand networks.

1.2.1 Theoretical Framework

Let $\mathbf{q}$ be a vector of quantities of G commodities and $u(\mathbf{q})$ be a utility function representing locally non-satiated and continuous preferences of a consumer. If the consumer has budget m and faces a price vector $\mathbf{p}$ with elements p_g, the consumer solves the maximization problem

$$\max_{\mathbf{q}} u(\mathbf{q})$$
$$s.t.\, \mathbf{p}^T\mathbf{q} \leq m$$

in order to maximize utility.

Let $v(\mathbf{p}, m)$ be the indirect utility function that yields the maximum utility that the consumer can attain given prices $\mathbf{p}$ and budget m. Roy's identity provides the relationship between the indirect utility function $v(\mathbf{p}, m)$ and demand function of good g in budget shares:

$$f_g(\mathbf{p}, m) = -\frac{\partial \log v(\mathbf{p}, m)/\partial \log p_g}{\partial \log v(\mathbf{p}, m)/\partial \log m}.$$

If budget share demand functions $f_g(\mathbf{p}, m)$ are derived from the utility maximization problem of a rational consumer, they must satisfy four rationality conditions, also known as the integrability conditions (Hurwicz and Uzawa, 1971):

Homogeneity: Demand functions $f_g(\mathbf{p}, m)$ are homogeneous of degree zero in prices and budgets. If a demand function satisfies homogeneity, $f_g(c\mathbf{p}, cm) = f_g(\mathbf{p}, m)$ holds for any $\mathbf{p}$, m, and scalar c. Homogeneity follows from the linearity of the budget share equation in the utility maximization problem of the individual: when prices and budget are scaled proportionately with a positive scalar

c, the set of feasible consumption bundles is the same. As a result, the quantity demanded must be the same.

Adding up: Value of the optimum quantity vector $\mathbf{q}$ evaluated at prices $\mathbf{p}$ must be equal to the individual's budget m. Adding up follows from non-satiation: there is always a better consumption bundle that the individual can attain with a small increase in cost. Therefore, the optimal bundle must lie on the budget constraint. With a linear budget constraint, by dividing both sides of equality $\mathbf{pq} = m$ with m, adding up is satisfied if share demand functions $f_g(\mathbf{p}, m)$ add up to one.

Symmetry: Let $\mathsf{S}(\mathbf{p}, m)$ be the Slutsky substitution matrix of an individual, which is a matrix of compensated demand derivatives with elements s_{ij} in its i^{th} row and j^{th} column. The cross-diagonal elements of $\mathsf{S}(\mathbf{p}, m)$ must be equal. In other words, the Slutsky substitution matrix of a rational individual must be symmetric.

Negativity: The Slutsky substitution matrix of a rational individual $\mathsf{S}(\mathbf{p}, m)$ must be negative-semi-definite. Symmetry and negativity are mathematical results that follow from an individual's cost minimization problem

$$\min \mathbf{p}^T \mathbf{q}$$
$$s.t. \, u \geq \bar{u}$$

for a reference level of utility $\bar{u}$. Let $e(\mathbf{p}, u)$ be an expenditure function that yields the solution to the individual's cost minimization problem. The Slutsky substitution matrix is also the Hessian of the expenditure function $e(\mathbf{p}, u)$. Therefore, if demand functions are theoretically consistent, then an expenditure minimization problem exists as the dual of the utility maximization problem and the Slutsky matrix must be symmetric. If the Slutsky matrix is also negative semi-definite, then the expenditure function is concave in prices. Concavity of the expenditure function in prices is an important property that intuitively follows from the consumer's expected behavior. As relative prices change, the consumer can attain a given utility level by spending, at most, the value of the original bundle evaluated at new prices. If the individual also substitutes relatively expensive goods for cheaper alternatives, cost of attaining the same utility level is lower.

1.2.2 Atheoretical and Theoretical Demand Networks

In terms of statistical consistency, a cross-validated FNN is a natural candidate to estimate a system of demand equations semi-nonparametrically by the universal approximation theorem of FNNs, which demonstrates their favorable approximation properties (White, 1990; Hornik, 1991; Geman et al., 1992; Sonoda and Murata, 2017). The fundamental idea of an FNN estimation is to generate derived variables as linear combinations of input variables and then to model an output variable as non-linear functions of these derived variables (see, e.g., Hastie et al., 2009; Goodfellow et al., 2016). An FNN architecture with a single non-linear transformation of derived variables, except the output transformation, has the following functional representation:

$$\hat{f}(\mathbf{x}) = \beta\left(\theta_1^2 + \sum_{k=1}^{K}\theta_{1,k}^2\,\alpha\left(\theta_k^1 + \sum_{i=1}^{I} x_i\theta_{k,i}^1\right)\right), \tag{1.1}$$

where x_i are input variables, θs are FNN coefficients to generate linear combinations of input variables, $\alpha(\cdot)$ is a prespecified non-linear transformation function for nested linear combinations, and $\beta(\cdot)$ is an output transformation function to obtain prediction $\hat{f}(\mathbf{x})$.[5] This structure provides a semi-nonparametric approach with two adjustable hyper-parameters: number of nested non-linear transformations with function $\alpha(\cdot)$, or layers (l), and number of derived variables K_l in each layer l, i.e., nodes. Despite an explicit functional form, FNNs are best represented by network diagrams. A diagram that illustrates the neural network terminology is available in appendix A.1.

Flexibility of an FNN architecture allows to estimate multiple non-separable outputs simultaneously (Kuan and White, 1994). Demand functions are estimated as systems of equations, instead of separate functions, to gain statistical accuracy, to overcome restrictive separability assumptions, and to impose cross-equation rationality restrictions (Lewbel, 1997). Therefore, being able to extend to multiple quantitative outputs seamlessly is a valuable property of FNN estimators for a microeconometric analysis of consumer demand. A system of equations formed by a single layer FNN would have the following functional structure:

$$\begin{gathered}
\hat{f}_1(\mathbf{x}) = \beta\left(\theta_1^2 + \sum_{k=1}^{K}\theta_{1,k}^2\,\alpha\left(\theta_k^1 + \sum_{i=1}^{I} x_i\theta_{k,i}^1\right)\right), \\
\vdots \\
\hat{f}_G(\mathbf{x}) = \beta\left(\theta_G^2 + \sum_{k=1}^{K}\theta_{G,k}^2\,\alpha\left(\theta_k^1 + \sum_{i=1}^{I} x_i\theta_{k,i}^1\right)\right).
\end{gathered}$$

Figure 1.1 is a network diagram that illustrates a more general case of above system of equations with multiple nested non-linear transformations (layers).

Atheoretical Demand Network (ADN)

A natural first step to estimate demand functions with an FNN is without considering rationality conditions. An FNN without rationality restrictions yields a semi-nonparametric estimate to a system of equations; an atheoretical demand network (ADN).

[5] Independence of the approximation properties of FNNs to a selection of non-linear transformation functions is documented (Hornik, 1991; Sonoda and Murata, 2017).

Even though ADN is estimated without considering rationality conditions, it is meaningful to expect two conditions to hold by definition; predicted budget shares must add up to one and they must be non-negative. Both conditions can be globally imposed by a multinominal logit transformation function

$$\beta(a_g) = \frac{\exp(a_g)}{\sum_{g=1}^{G} \exp(a_g)}$$

at the output layer, where a_g is the neural network output in the g^{th} node of the output layer before a multinominal logit transformation.[6] Therefore ADN globally only satisfies adding up by construction. With this modification, I define an ADN as a cross-validated FNN with a multinominal logit output activation function.

Since theoretical restrictions are not imposed, an ADN solution can be characterized as the solution of an unconstrained optimization problem

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g(\Theta, \mathbf{x}_n) \right]^2 \tag{1.2}$$

where $w_{g,n}$ is an observed budget share, $\hat{f}_g(\Theta, \mathbf{x_n})$ is the ADN prediction for budget share of good g for observation $\mathbf{x}_n$ from a set of observations $\mathscr{X}$ with coefficient vector Θ.

Theoretical Demand Network

Adding up, homogeneity, symmetry and negative semi-definiteness of the Slutsky matrix are consequences of the neoclassical demand theory that provides a priori information on shapes of estimated demand equations (Deaton and Muellbauer, 1980a). Because ADN estimates are statistically consistent, if data are consistent with integrability conditions, one would expect ADN to yield demand equations that also satisfy these conditions. However, in finite datasets, this expectation does not necessarily hold. Hence, estimated demand functions may be theoretically inconsistent. Therefore, imposing theoretical a priori information on shapes of possible demand equations provides a theoretically consistent estimation structure; a theoretical demand network (TDN). Moreover, imposing a priori information on estimation improves accuracy and generalizability of non-parametric estimations (Gallant and Golub, 1984; Joerding and Meador, 1991; Blundell et al., 2012). For these purposes, TDN builds upon ADN and overcomes lack of theoretical consistency by globally satisfying homogeneity and adding up by construction, and locally satisfying symmetry and negativity by imposing these restrictions on a set of, or on all, observations during optimization.

---

[6] Multinominal logit activation function is commonly used in the output layer for classification tasks where the output is interpreted as probabilities. In this sense, using multinominal logit function in the output layer forms a direct analogy between probabilistic outcomes and budget shares. For a similar analogy within the literature of demand analysis, see Theil (1965).

To satisfy homogeneity, I adjust price and budget inputs a priori and estimate budget share equations on adjusted price vectors and budgets. Because multiplication of a vector with a scalar c is an adjustment to the length of the vector while keeping its direction constant, I adjust each price vector and budget tuple by dividing them to the length of that price vector measured by a norm function $\left(\bar{\mathbf{p}}_n = \mathbf{p}_n \big/ \|\mathbf{p}_n\|, \bar{m}_n = m_n \big/ \|\mathbf{p}_n\|\right)$. Hence, input price vectors are scaled to unit length while relative prices are preserved and budgets are adjusted accordingly.[7] In this way, homogeneity is implemented by adding a pre-input layer that takes price and budget observations and feeds adjusted prices and budgets to the input layer.

Adding up is satisfied by ADN with a multinominal logit transformation function at the output layer. Another way to impose adding up would be dropping one commodity from the estimation procedure and obtaining the budget share of this commodity by subtracting the sum of estimated budget shares from one. However, imposing adding up via a multinominal logit function also guarantees positive budget shares.

Symmetry is satisfied if an empirical Slutsky matrix $\hat{\mathsf{S}}(\Theta, \mathbf{x}_n)$ is a symmetric matrix, i.e., if cross diagonal estimates of compensated demand derivatives $\hat{s}_{ij,n}$ and $\hat{s}_{ji,n}$ are equal for all goods i and j at all possible observations n. Let

$$\sum_{i=1}^{G-1} \sum_{j=i+1}^{G} \left(\hat{s}_{ij,n} - \hat{s}_{ji,n}\right)^2 = \begin{cases} 0 & \text{if symmetry holds} \\ > 0 & \text{if symmery is violated} \end{cases}$$

be a measure of deviation from symmetry for an estimated Slutsky matrix at an observation n. Then

$$R^{\mathscr{S}}(\Theta, \mathscr{X}^{\mathscr{S}}) = \sum_{n=1}^{N^{\mathscr{S}}} \sum_{i=1}^{G-1} \sum_{j=i+1}^{G} \left(\hat{s}_{ij,n} - \hat{s}_{ji,n}\right)^2$$

is a measure of total deviation from symmetry at the set of symmetry imposed observations $\mathscr{X}^{\mathscr{S}}$ with size $N^{\mathscr{S}}$ (see also Joerding and Li, 1994; Cardell et al., 1995). If estimated Slutsky matrices for all observations in $\mathscr{X}^{\mathscr{S}}$ satisfies symmetry, then $R^{\mathscr{S}}(\Theta, \mathscr{X}^{\mathscr{S}})$ is equal to zero. Therefore, a symmetry restricted functional approximation can be characterized as the solution to the following constrained optimization problem:

[7] Similar approaches are available. For example, selecting a numeraire good g and writing prices of all the other goods in terms of relative prices to the numeraire good $\left(\bar{\mathbf{p}} = \mathbf{p} \big/ p_g\right)$, and adjusting budget by the price of the numeraire good $\left(\bar{m} = m \big/ p_g\right)$ is another way to achieve homogeneity. While this approach is a scaling in the price space along the direction of the price of the numeraire good while keeping the direction of the price vector constant, the approach I use is a scaling along the direction of the price vector itself. I follow a scaling along the direction of the price vector because it proves to be computationally more efficient in calculation of the Slutsky matrix, which helps in imposing symmetry and negativity.

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g(\Theta, \mathbf{x}_n) \right]^2$$
$$s.t. R^{\mathscr{S}}(\Theta, \mathscr{X}^{\mathscr{S}}) = 0 .$$

Negativity holds if $\hat{\mathsf{S}}(\Theta, \mathbf{x}_n)$ is a negative semi-definite matrix. To impose negativity, I use a constraint indicator

$$\lambda(\Theta, \mathbf{x}_n) \equiv \max_{\mathbf{v}_n} \left\{ -\mathbf{v}_n^T \hat{\mathsf{S}}(\Theta, \mathbf{x}_n) \mathbf{v}_n : \hat{\mathbf{f}}(\Theta, \mathbf{x}_n)^T \mathbf{v}_n = 0, \mathbf{v}_n^T \mathbf{v}_n = 1 \right\},$$

where $\mathbf{v}_n$ is a vector of scalars and $\hat{\mathbf{f}}(\Theta, \mathbf{x}_n)^T$ is a vector of predicted budget shares. The differentiable constraint indicator $\lambda(\Theta, \mathbf{x}_n)$ is non-negative only if the matrix $\hat{\mathsf{S}}(\Theta, \mathbf{x}_n)$ is negative semi-definite (Gallant and Golub, 1984).[8] Let $\min(0, \lambda(\Theta, \mathbf{x}_n))^2$ be a measure of deviation of an estimated Slutsky matrix from negativity at an observation n. Then

$$R^{\mathscr{N}}(\Theta, \mathscr{X}^{\mathscr{N}}) \equiv \sum_{n=1}^{N^{\mathscr{N}}} \min(0, \lambda(\Theta, \mathbf{x}_n))^2$$

is the sum of squared total deviations from negativity across a set of negativity imposed observations $\mathscr{X}^{\mathscr{N}}$. If estimated Slutsky matrices at all observations in $\mathscr{X}^{\mathscr{N}}$ satisfy negativity, then $R^{\mathscr{N}}(\Theta, \mathscr{X}^{\mathscr{N}})$ is equal to zero. Hence, a negativity restricted functional approximation is the solution to the following constrained optimization problem:

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g(\Theta, \mathbf{x}_n) \right]^2$$
$$s.t. R^{\mathscr{N}}(\Theta, \mathscr{X}^{\mathscr{N}}) = 0 .$$

Therefore, combining these restrictions, TDN is characterized as a cross-validated multilayer FNN with a pre-input layer for homogeneity adjustment, a multinominal logit output layer, and the solution to the following constrained optimization problem:

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g(\Theta, \mathbf{x}_n) \right]^2 \tag{1.3}$$
$$s.t. R^{\mathscr{S}}(\Theta, \mathscr{X}^{\mathscr{S}}) = 0$$
$$s.t. R^{\mathscr{N}}(\Theta, \mathscr{X}^{\mathscr{N}}) = 0$$

[8] Formulation of the constraint indicator $\lambda(\Theta, \mathbf{x}_n)$ and its derivative are described in detail by Gallant and Golub (1984). See also Lau (1978) and Diewert et al. (1981) for the relationship between negative semi-definiteness and the constraint indicator.

Derivations of these constraints for a single hidden layer FNN are provided in appendix A.5.

Estimation Procedures

Efficient approaches to obtain FNN estimators have been controversial. Today, the state of the art FNN models are estimated using stochastic optimization approaches, gradient updates through backpropagation (Rumelhart et al., 1986), and algorithms that efficiently apply these ideas to large samples.[9]

FNN solvers are designed to solve unconstrained optimization problems, for example to find the function that minimizes the sum of squared errors among some permissible class of functions with measures against overfitting. Hence, any FNN solver would be suitable to obtain an ADN estimate.

However, the constrained minimization problem that characterizes a TDN estimate incorporates highly non-linear equality constraints. Hence, standard optimization algorithms for FNNs are not suitable to obtain a TDN estimate. Furthermore, in fact, symmetry and negativity restrictions contain a continuum of restrictions because these conditions have to hold at all possible observations. This would pose a problem for standard non-linear optimization algorithms. For example, in case of a constrained optimization problem with d coefficients and $\mathscr{N}^{\mathscr{S}} + \mathscr{N}^{\mathscr{N}}$ restrictions, a standard non-linear optimization algorithm would search for a solution at $d - \mathscr{N}^{\mathscr{S}} - \mathscr{N}^{\mathscr{N}}$ dimensional coefficient subspace, which diminishes in size as the number of restrictions increases.

Penalty methods to optimization provide a way to address both issues by embedding restrictions into an unconstrained optimization problem to approximate a constrained optimization problem. Transforming a constrained optimization problem to an unconstrained problem with addition of a penalty term allows the use of modern optimization routines of neural networks to obtain a TDN estimate. Moreover, because penalty approaches search for a solution in d dimensional coefficient space instead of a subspace, they allow imposing all restrictions across the support of a dataset.

In order to transform the constraint optimization problem that characterizes a TDN, I define a differentiable penalty function

$$P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}}) \equiv \frac{1}{N^{\mathscr{S}}} R^{\mathscr{S}}(\Theta, \mathscr{X}^{\mathscr{S}}) + \frac{1}{N^{\mathscr{N}}} R^{\mathscr{N}}(\Theta, \mathscr{X}^{\mathscr{N}}),$$

which is zero if both terms in the summation are zero. This holds only if symmetry is satisfied at all observations in $\mathscr{X}^{S}$ and negativity is satisfied at all observations in $\mathscr{X}^{N}$. Otherwise, $P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{N})$ is positive. By using $P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}})$, I formulate the following unconstrained minimization problem:

[9] Backpropagation is a backward-derivation approach to calculate the derivatives of an objective function with respect to each parameter. For an in-depth discussion of neural network estimation and optimization algorithms, see, e.g., Goodfellow et al. (2016).

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g(\Theta, \mathbf{x}_n) \right]^2 + \gamma P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}}), \tag{1.4}$$

where γ is an adjustable weight parameter that assigns a relative importance to the penalty of violating constraints with respect to functional fit. For $\gamma = 0$, violating constraints impose no penalty and solution of the approximated constrained problem 1.4 is equal to the solution of the unconstrained problem 1.2. For a small γ, cost of violating the constraints is small. Hence, optimization is primarily driven by the goodness of fit of estimated demand functions $\hat{f}_g(\Theta, \mathbf{x}_n)$. As γ increases, penalty incurred by constraint violations increases.

Two issues related to the use of penalty functions are important to consider. The first issue is about how well a penalty approach approximates a constrained optimization problem. The second issue is about empirical feasibility of using the approach. Luenberger and Ye (2016) provide an in-depth discussion of these issues. Specifically, for a monotonically increasing series of penalty weights $\{\gamma_s\}$, any limit point of the coefficient vector sequence $\{\Theta_s\}$ provides a solution to the constrained optimization problem. Furthermore, this problem can be solved using first order gradient based methods.

This formulation of a penalty approach is very similar to the formulation of a cost function in many machine learning applications where γ is a regularization coefficient and $P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}})$ is a regularization cost (see appendix A.4). However, application of a penalty method is different to a standard regularization procedure because a penalty method gradually increases the cost of deviations from the constraints until the coefficient vector sequence converges. Contrary to regularization, which increases γ until prediction accuracy no longer increases, optimization with a penalty approach aims to obtain solution of a constrained problem which may or may not increase prediction accuracy. Therefore, I obtain a TDN estimate by defining a monotonically increasing series $\{\gamma_s\}$, solving the unconstrained optimization problem 1.4 with weight γ_1 and an appropriate FNN optimization algorithm, and repeating the process with the other ordered γ values until the coefficient vector sequence $\{\Theta_s\}$ converges.

1.3 Empirical Application

The purpose of this empirical application is to illustrate the feasibility of using an FNN estimator and the proposed imposition of the restrictions. To this end, I obtain ADN and TDN estimates of a system of ten share demand equations, and compare these estimates with those of AIDS and QUAIDS models, which are widely used parametric approaches in microeconometric analyses of consumer demand. I evaluate performances of each estimator by measuring their generalizability to test sets in terms of a mean squared error loss function and by plotting each estimated system of demand equations.

With these tests, two measures of comparison are aimed for. First, sum of squared errors on a test set is a measure of generalization error of each estimated system of demand equations. For a given level of estimation variance, if the parametric approaches induce biases due to their functional forms, one would expect the unbiased FNN estimators to yield systematically lower mean squared errors on a test set. However, if imposed functional forms do not yield biased estimates, then, the mean squared errors of the parametric approaches on test sets would be similar to those of the FNN estimators. Second, in fact, as common to non-parametric approaches, one would expect the FNN estimators to have larger variances compared to the parametric approaches. In this case, even if there is bias induced by functional forms of the parametric approaches, a mean squared error indicator may fail to capture that simply because of the differences in estimation variances. If that is the case, estimated systems of demand equations should still have significantly different shapes. Such differences are illustrated by plotting all equations along the budget dimension.

1.3.1 Data

I create a main sample by combining two datasets. First, I obtain household-level information from the UK Living Costs and Food Survey (LCF) 2008-2012, which is a repeated cross-section. LCF provides information on households' consumption patterns and observable characteristics. In this empirical illustration, I use households' expenditures in ten aggregate food categories, locations, time of survey, child/adult compositions, and number of economically active individuals in each household.[10]

Second, I use price information from the UK Consumer Price Inflation Item Indices and Price Quotes, which collects location specific monthly average price information on a large variety of disaggregate product categories. I obtain prices of aggregate food categories as unweighted geometric means of disaggregate product categories. Aggregate and disaggregate product categories are matched using COICOP Division 01.[11] I combine household-level observations with price vectors using time of survey and location. Therefore, in my sample, each household that is interviewed at the same time (month-year) and location (geographic region) is assigned the same price vector. After dropping observations with zero expenditures in any aggregate food category, the main sample has 6673 household-level observa-

[10] Aggregate food categories are breads-cereals, meat, fish, milk-cheese-egg, oils-fats, fruit, vegetables, sugar-sweets, other foods, and non-alcoholic beverages.

[11] Classification of Individual Consumption According to Purpose (COICOP) is a UN guideline to categorize consumer expenditures. Item indices in LCF and price surveys are compatible with COICOP categories.

tions with 576 distinct price vectors.[12] Table 1.1 presents some descriptive statistics of my main sample.

From the main sample, I draw 200 bootstrapped samples with replacement. Then, each bootstrapped sample is divided into three sub-samples: estimation, cross-validation, and test samples. In order to generate these sub-samples, a random selection of 70% of observations from each bootstrapped sample is assigned to an estimation sub-sample, 15% to a cross-validation sub-sample, and 15% to a test sub-sample. Estimation sub-samples of each bootstrapped sample are used to estimate demand equations. Cross-validation sub-samples are used to adjust hyper-parameters of ADN and TDN estimates. Test sub-samples are used to evaluate generalizability of estimated systems of demand equations. I refer to estimation (e), cross-validation (c), and test (t) samples with subscripts in the remainder of the section.

1.3.2 Estimation

For each bootstrapped sample, I use the estimation sub-sample to obtain AIDS, QUAIDS, ADN, and TDN estimates of a system of demand equations. I calculate mean squared error of estimated systems of equations on a test set t

$$C(\Theta, \mathscr{X}_t | \mathscr{X}_e, \mathscr{X}_c) \equiv \frac{1}{N_t} \sum_{n=1}^{N_t} \sum_{g=1}^{G} \left[w_{n,t}^g - \hat{f}_g(\Theta, \mathbf{x}_{n,t} | \mathscr{X}_e, \mathscr{X}_c) \right]^2$$

as a measure of generalizability to an independent test sample ($\mathscr{X}_t$) given estimation ($\mathscr{X}_e$) and cross-validation samples ($\mathscr{X}_c$), where N_t is the size of the test set, $w_{n,t}^g$ is observed budget share of good g for observation n of $\mathscr{X}_t$, $\hat{f}_g(\Theta, \mathbf{x}_{n,t} | \mathscr{X}_e, \mathscr{X}_c)$ is the predicted budget share of g^{th} good for the n^{th} observation of $\mathscr{X}_t$ given $\mathscr{X}_e$ and $\mathscr{X}_c$.[13]

I estimate AIDS and QUAIDS models using an iterative algorithm (Blundell and Robin, 1999). In this way, AIDS and QUAIDS models satisfy adding up, homogeneity, and symmetry restrictions. Details on the functional forms of these models are in appendix A.2.

To estimate ADN and TDN models, I first select the rectifier function as a non-linear transformation for hidden layers.[14] Second, as a gradient-based optimization

---

[12] With this adjustment, my main sample is not representative of the entire UK population. I argue that this not a critical issue in this empirical application where the focus is on applying an estimation methodology and comparing it with some parametric alternatives. The results I present can be interpreted as conditional on this sample selection but not the entire population.

[13] Note that I do not make use of cross-validation sets when estimating and calculating the costs of AIDS and QUAIDS models. Cross-validation samples are only used to tune the hyper-parameters of ADN and TDN estimations. Therefore, for AIDS and QUAIDS estimations, $\hat{f}_g(\Theta, \mathbf{x}_{n,t} | \mathscr{X}_e, \mathscr{X}_c) = \hat{f}_g(\Theta, \mathbf{x}_{n,t} | \mathscr{X}_e)$.

[14] $\alpha(x) = \max(0, x)$. Rectifier function is used in many machine learning applications and is suggested as a default alternative (Goodfellow et al., 2016).

algorithm to solve the minimization problems of ADN and TDN, I use the stochastic adaptive moment (Adam) algorithm (Kingma and Ba, 2014). Alternative non-linear transformation functions and optimization algorithms can also be used to obtain ADN and TDN estimates. As a robustness check of my results to these selections, results with an alternative combination of transformation function and optimization algorithm are illustrated in appendix A.3.[15]

For each ADN and TDN estimation, I select optimum hyper-parameters that control model complexity, e.g., number of hidden layers and nodes, using cross-validation. Cross-validation procedure starts by setting a low value to hyper-parameters and estimating a model using an estimation set. Then, cost on a cross-validation set, $C(\Theta, \mathscr{X}_c|\mathscr{X}_e)$, is calculated and stored, hyper-parameters are increased, and the procedure is repeated. To stop the cross-validation, I use a stopping algorithm that calculates the tendency of the last few cross-validation costs and stops the algorithm if cross-validation cost tends to increase as allowed complexity increases. The optimum hyper-parameters are selected as the integers that yield the lowest cross-validation cost.

Intuition of this procedure is as the following: Increasing hyper-parameters makes it possible for an FNN estimator to yield more complex functions, which may also lead to overfitting (Barron, 1994). As model complexity increases, an estimated system of demand equations yields a better fit to an estimation sample. However, this is not necessarily true for a cross-validation sample, which is not used in the estimation process. Therefore, during the hyper-parameter grid search, function space is enlarged up to the point where the estimated functions start providing worse generalizations to a cross-validation set.

In order to estimate TDN and ADN, I use a mixture of stochastic and mini-batch gradient optimization techniques for speed and as a safeguard against local minima (Goodfellow et al., 2016). Algorithms are initiated by a stochastic optimization until a relaxed convergence criterion is met. Then, a mini-batch optimization is applied with a stricter convergence criterion. For TDN estimates, once convergence is achieved with the initial penalty coefficient γ_1, the procedure continues using the other ordered values of $\{\gamma_s\}$ with mini-batch optimization until the coefficient vector sequence $\{\Theta_s\}$ converges.

Pseudo-codes that illustrate these steps in detail are provided in appendix A.6.

1.3.3 Results

Results that are reported in this section are obtained using a relaxed convergence criterion 10^{-6}, a strict convergence criterion 10^{-8}, a mini-batch size of 256, and $\gamma_{s+1} = 1.5\gamma_s$ with $\gamma_1 = 1$ as the monotonically increasing series $\{\gamma_s\}$.

The first set of results that I present are estimated using log-prices and log-expenditures as inputs, and budget shares of ten food categories as output variables.

---

[15] All analyses reported in this paper are done with the Tensorflow library in Python (Abadi et al., 2015).

TDN is estimated imposing homogeneity, adding up, and symmetry to be comparable to AIDS and QUAIDS models. Results of an application with both symmetry and negativity restrictions are available in appendix A.3.

Figure 1.2 depicts distributions of mean squared errors of estimated systems of demand equations on test sets. Violin graphs that illustrate these distributions are remarkably similar, which indicates that all estimated demand systems provide similar fits to test sets.

Figure 1.3 presents estimated demand functions for ten food categories. Predicted budget shares are calculated at median prices. ADN and TDN estimates are close to linear in log budgets for nine goods. For vegetables, ADN and TDN captures a non-linear pattern between budget shares and log budgets. Predicted budget shares for vegetables increase until middle income households, and then flattens for high income households. AIDS is unable to capture this non-linear pattern because of its linear relationship between log budget levels and budget shares. QUAIDS is able to capture this relationship; however, due to its quadratic form, we see that the response is captured as a decrease. Significant differences in predicted budget shares are observed at the tails of the log budget distribution.

Figure 1.4 illustrates estimated income elasticities for a hypothetical household facing median prices with the median budget level. Even though the 95% confidence intervals mostly coincide, ADN and TDN estimates of elasticities have a higher variance, which is perhaps not surprising for results of a semi-nonparametric approach with a moderate sample size.

Similar figures for price elasticities are not produced due to the high number of outputs (G^2). However, tables 1.2 to 1.4 report estimated price elasticities and Slutsky matrices for four categories. Cross-price elasticities of AIDS, QUAIDS, ADN and TDN are mostly similar. ADN and TDN estimates indicate slightly stronger responses to changes in own-prices.

The second set of results are obtained using log-prices, log-expenditures, and household specific demographics as input variables, and budget shares as output variables.[16] Demographic variables are included in ADN and TDN estimations as any other variable. Therefore, impact of observable heterogeneity on consumption choice is left to the estimation procedure to discover. Because demographic variables include relatively large values compared to log inputs, they are standardized using their main sample means and standard deviations before the estimation procedure. The reason for this standardization is to obtain a proper initialization of coefficients and gradient updates during the optimization process (Hastie et al., 2009). I include demographic variables in AIDS and QUAIDS estimations as additive taste shifters by also modifying the price deflators to maintain theoretical consistency (see appendix A.2).

Figure 1.5 illustrates distributions of mean squared errors of estimated demand functions with demographic variables on test sets. Similar to the previous case, violin graphs indicate similar model performances. Figure 1.6 shows lowess-smoothed estimated demand functions with demographics.

---

[16] Demographic variables are number of children, number of adults, and number of economically active individuals in a household.

To elicit effects of demographic variables on estimated demand functions, figures 1.7 and 1.8 depict QUAIDS and TDN estimates of four demand functions for households with different numbers of children.[17] These figures show that the semi-nonparametric TDN approach also captures a similar level effect to the way that demographics are incorporated into QUAIDS estimation as taste shifters.

1.3.4 Calculating Welfare Impacts of Price Changes

In order to estimate welfare impacts of price changes, I use a differential equation approach developed by Hausman and Newey (1995), and calculate equivalent variation of a price change with an ordinary differential equation.

Let $\mathbf{p}(t)$ be a price path from $\mathbf{p}(0)$ to $\mathbf{p}(1)$. Let $S(t,m) = m - e\left(\mathbf{p}(t), u^1\right)$ denote the equivalent variation of a price change from $\mathbf{p}(t)$ to $\mathbf{p}(1)$ at utility level u^1. Then, $S(0,m)$ is obtained as the solution to the following ordinary differential equation:

$$\partial S(t,m)/\partial t = -q\left(\mathbf{p}(t), m - S(t,m)\right)^T \frac{\partial \mathbf{p}(t)}{\partial t}$$

with the initial condition $S(1,m) = 0$, and where $m - S(t,m)$ is compensated income. This problem can be solved with conventional numerical solution methods for ordinary differential equations (see also Blundell et al., 2012; Hausman and Newey, 2016).[18]

Figure 1.9 illustrates the equivalent variation of a 10% increase in the price of each good separately. The welfare impacts predicted by the parametric AIDS and QUAIDS models, and the semi-nonparametric TDN model are relatively similar across the observed budget distribution. However, the TDN model does not exhibit non-monotonic relationships between impacts of price changes and log of total food expenditures that the QUAIDS model suggests for some goods.

1.4 Discussion

Results of the empirical application illustrate two main findings. First, it is indeed feasible to obtain theoretically and statistically consistent estimates of demand functions using a mixture of economic theory and machine learning toolbox. Second, it may not always be necessary; there may be empirical settings where existing parametric models of consumer demand work well.

[17] The number of goods and the demographic variable are selected due to illustration purposes. The plots for the other goods, and other demographic categories, are available upon request.

[18] The results that I present are obtained using an ordinary differential equation solver from the Scipy package of Python (Jones et al., 2001).

If parametric approaches to demand estimations do not induce significant biases due to specification errors, one would expect parametric models to perform well. If not, ADN and TDN estimates provide statistically consistent estimates to unobserved demand functions by being free of functional form specifications. Hence, even though network estimates require large sample sizes to have small variances, they would be most beneficial when parametric models are inadequate in capturing non-linearities or complex relationships that data may exhibit (Geman et al., 1992).

Similar model performances on test sets in my empirical example suggest that biases induced by functional forms of AIDS and QUAIDS models are not significant. One reason for this functional form match may be the aggregate nature of the estimation setting and low substitution between the consumption categories.[19] On the other hand, another implication of similar performances, given the simplicity of estimated demand functions, is in favor of ADN and TDN estimations: even though estimates do not suggest significant non-linearities, ADN and TDN still performs as well as the parametric alternatives. Therefore, results of my empirical application provide evidence in favor of both the existing parametric approaches and the network models to elicit preferences for aggregate levels of consumption goods.

In addition to providing theoretical consistency and allowing for welfare analysis, use of theoretical restrictions in TDN also has practical benefits. A combination of the restrictions with the estimation procedure through a penalty approach resembles regularization approaches in machine learning. The primary purpose of regularization is to safeguard against overfitting, which is a concern for non-parametric estimators. The penalty term in the optimization procedure of TDN can also be interpreted as a protection against overfitting: it eliminates theoretically inconsistent functions from allowed function space so there are less functions to choose from (see also Gallant and Golub, 1984).[20] Moreover, if data supports theory, restrictions in the penalty term are not binding as sample size increases, so penalty converges to zero. In this case, a penalty term that is constructed from theoretical restrictions works by reducing variance without harming consistency of an estimator. Hence, introducing theoretical restrictions in FNN estimators with a penalty approach provides a regularization with an economic judgment. In principle, this idea is also applicable in other economic research with theoretical conditions that can be written as a penalty function.

Apart from overfitting, which is addressed with cross-validation and inclusion of theoretical restrictions, two issues associated with use of ADN and TDN approaches are worth discussing. First, even though ADN and TDN estimates present feasible alternatives, they are time-wise costlier to obtain compared to AIDS and QUAIDS

[19] With a simulation analysis, Barnett and Seck (2008) show that the AIDS model performs well when estimating preferences for aggregate commodities that are usually characterized with low substitution due to loss of information in the aggregation process.

[20] The difference in application of a regularization versus a penalty approach comes from their purposes. While standard regularization approaches aim to introduce bias to reduce variance up to the point where prediction accuracy no longer increases, a penalty approach aims to estimate coefficients that strictly impose the restrictions in a numerical sense. Hence, the penalty coefficient γ is increased until convergence instead of being selected with cross-validation, which is a standard practice to set a regularization coefficient.

estimates, especially a TDN estimate. Stochastic optimization approaches and distributed computing tremendously help in terms of speed but relying on bootstrapping to calculate confidence intervals of estimated demand functions and elasticities make these approaches even costlier. Therefore, if time or computational capacity is an issue, a logical approach would be to start with a parametric estimator, to test robustness of functional specification using an ADN estimate, and to consider the necessity of a TDN estimate depending on the results.

Second, the TDN approach imposes restrictions in terms of numerical approximation. Theoretically, a penalty approach yields results that are equal to the results of a constrained optimization problem only as a penalty coefficient approaches to infinity. Therefore, in practice, estimates may slightly differ from what restrictions may require. For example, imposing Slutsky symmetry requires cross-diagonal elements of a Slutsky matrix to be equal. In my empirical illustration, cross-diagonal Slutsky terms for breads and fats, $s_{bread,fat}$ and $s_{fat,bread}$, are estimated as 0.05719 and 0.04208 with ADN. With TDN, these estimates are 0.04678 and 0.04712. As this illustration shows, inclusion of restrictions work in an intended way. However, numerical equality of the results will depend on the tolerance level used in the optimization procedure. Therefore, a stricter and costlier, or a more relaxed and less costly, imposition of constraints on estimation is possible. Desired accuracy of the estimates would depend on the difference in the economic interpretation of the results.

1.5 Conclusion

Existing parametric approaches to model consumer demand impose certain functional form hypotheses on the shape of demand functions. However, the only information on shapes of demand functions that are provided by the preference-based neoclassical demand theory are the integrability conditions. Non-parametric approaches address the functional form uncertainty; however, imposing shape restrictions is necessary to obtain theoretically consistent functional estimates, for example, to be able to understand welfare impacts of price changes.

In this paper, I suggest two empirical models that are intermediate alternatives to parametric and non-parametric approaches. Both the theoretically consistent TDN model and the atheoretical ADN model address the functional form uncertainty, and naturally extend to high dimensional environments with many goods or other explanatory variables. Moreover, the TDN model preserves the close relationship of the parametric approaches to the demand theory.

The ADN model is very straightforward to estimate using any machine learning software library that supports a feedforward neural network estimation.[21] The TDN model builds upon the ADN model by adding calculation of Slutsky matrices at every optimization step for all observations on which the constraints are imposed.

[21] As of today, alternative libraries are available in Python, R, Matlab, and SPSS, which are some of the most frequently used software for statistical analyses.

Results of my empirical application show that there are environments in which the parametric models work well. However, even in such environments, ADN and TDN still perform equally well. Therefore, true benefits of these flexible models in microeconometric modeling of consumer demand would be more apparent when there are non-linearities that parametric models fail to capture. An example of this is illustrated with vegetable demand, which exhibits a non-linear pattern in a system of equations that otherwise contain linear relationships between budget shares and log budget levels.

In principle, structuring non-parametric estimators using theoretically guided restrictions is also an idea that can be useful in other economic research where researchers are unsure about the exact shape of a functional relationship but have prior information about some properties of these functions. This paper illustrates such estimations are easy to obtain using state-of-the-art statistical software and optimization algorithms.

Table 1.1 Descriptive Statistics

Category	Variable	Min	Max	Mean	St. Dev.
Prices	Bread	0.66	0.99	0.83	0.06
	Meat	4.02	6.41	5.20	0.44
	Fish	4.22	6.14	5.15	0.37
	Milk	1.53	2.12	1.83	0.11
	Fat	1.49	2.13	1.74	0.12
	Fruit	0.52	1.13	0.83	0.13
	Vegetable	0.72	0.95	0.84	0.04
	Sugar	0.81	1.13	0.98	0.06
	Other	0.85	1.71	1.30	0.21
	Non-alcoholic	1.06	1.63	1.36	0.12
Budget Shares	Bread	0.013	0.448	0.175	0.06
	Meat	0.003	0.638	0.231	0.09
	Fish	0.001	0.680	0.054	0.04
	Milk	0.002	0.473	0.138	0.05
	Fat	0.001	0.170	0.026	0.01
	Fruit	0.001	0.385	0.075	0.05
	Vegetable	0.003	0.492	0.124	0.05
	Sugar	0.001	0.413	0.065	0.04
	Other	0.001	0.363	0.031	0.02
	Non-alcoholic	0.001	0.377	0.081	0.04
Other	Total Expenditure	20.11	208.76	71.11	30.79
	# of Econ. Active	0	6	1.4	1.06
	# of Children	0	7	0.7	1.04
	# of Adults	1	7	2.1	0.72

Note: Table provides descriptive statistics of input and output variables that are employed in the empirical illustration.

Table 1.2 Uncompensated Price Elasticities

	Bread	Fat	Fish	Fruit
AIDS				
Bread	**-0.74**	0.03	0.05	**-0.19**
	(-0.99, -0.54)	(-0.15, 0.20)	(-0.05, 0.15)	(-0.33, -0.04)
Fat	-0.01	**-0.34**	-0.07	-0.05
	(-0.14, 0.12)	(-0.59, -0.10)	(-0.16, 0.02)	(-0.18, 0.08)
Fish	0.18	-0.27	**-0.80**	0.09
	(-0.17, 0.52)	(-0.64, 0.17)	(-1.14, -0.51)	(-0.23, 0.39)
Fruit	**-0.23**	-0.03	0.03	**-0.53**
	(-0.40, -0.04)	(-0.25, 0.18)	(-0.09, 0.14)	(-0.79, -0.28)
QUAIDS				
	Bread	Fat	Fish	Fruit
Bread	**-0.74**	0.03	0.05	**-0.19**
	(-0.99, -0.54)	(-0.15, 0.19)	(-0.05, 0.15)	(-0.33, -0.04)
Fat	-0.01	**-0.34**	-0.07	-0.05
	(-0.14, 0.12)	(-0.60,-0.11)	(-0.15, 0.03)	(-0.19, 0.07)
Fish	0.19	-0.26	**-0.81**	0.10
	(-0.16, 0.55)	(-0.64, 0.20)	(-1.16, -0.51)	(-0.22, 0.41)
Fruit	**-0.23**	-0.04	0.03	**-0.53**
	(-0.40, -0.05)	(-0.26, 0.17)	(-0.08, 0.15)	(-0.79, -0.29)
ADN				
	Bread	Fat	Fish	Fruit
Bread	**-0.93**	0.06	0.07	-0.05
	(-1.04,-0.79)	(-0.04, 0.20)	(-0.02, 0.19)	(-0.23, 0.03)
Fat	-0.04	**-0.92**	0.04	-0.08
	(-0.16, 0.05)	(-1.03, -0.78)	(-0.06, 0.16)	(-0.25, 0.03)
Fish	0.05	-0.05	**-1.04**	0.05
	(-0.09 ,0.21)	(-0.28, 0.11)	(-1.24, -0.88)	(-0.12, 0.24)
Fruit	0.01	0.01	0.03	**-0.93**
	(-0.12, 0.10)	(-0.11, 0.15)	(-0.09, 0.16)	(-1.05, -0.75)
TDN				
	Bread	Fat	Fish	Fruit
Bread	**-0.98**	0.01	0.0	0.01
	(-1.01, -0.94)	(-0.06, 0.04)	(-0.01, 0.01)	(-0.02, 0.02)
Fat	-0.02	**-0.96**	-0.01	-0.02
	(-0.06, 0.01)	(-1.03, -0.82)	(-0.04, 0.00)	(-0.07, 0.01)
Fish	0.02	-0.02	**-0.98**	0.03
	(-0.01, 0.06)	(-0.12, 0.04)	(-1.00, -0.93)	(-0.01, 0.07)
Fruit	0.01	-0.01	0.01	**-0.97**
	(-0.02, 0.03)	(-0.08, 0.05)	(-0.01, 0.02)	(-1.00, -0.91)

Note: Statistically non-zero estimates are illustrated in **bold**.

Table 1.3 Compensated Price Elasticities

	AIDS			
	Bread	Fat	Fish	Fruit
Bread	**-0.57**	**0.25**	0.10	-0.05
	(-0.82, -0.37)	(0.07, 0.42)	(-0.01, 0.20)	(-0.19, 0.08)
Fat	**0.19**	-0.08	-0.02	0.10
	(0.05, 0.31)	(-0.33, 0.14)	(-0.10, 0.08)	(-0.02, 0.23)
Fish	0.33	-0.07	**-0.75**	0.21
	(-0.02, 0.68)	(-0.44, 0.37)	(-1.1, -0.46)	(-0.10, 0.51)
Fruit	-0.07	0.17	0.08	**-0.40**
	(-0.24, 0.11)	(-0.04, 0.39)	(-0.04, 0.2)	(-0.66, -0.15)
	QUAIDS			
	Bread	Fat	Fish	Fruit
Bread	**-0.57**	**0.25**	0.10	-0.06
	(-0.82, -0.37)	(0.06, 0.42)	(-0.02, 0.20)	(-0.19, 0.08)
Fat	**0.19**	-0.09	-0.01	0.10
	(0.05, 0.31)	(-0.34,0.13)	(-0.10, 0.09)	(-0.03, 0.23)
Fish	0.35	-0.06	**-0.76**	0.22
	(-0.01, 0.71)	(-0.44, 0.4)	(-1.11, -0.46)	(-0.09, 0.53)
Fruit	-0.07	0.17	0.08	**-0.41**
	(-0.24, 0.11)	(-0.05, 0.38)	(-0.03, 0.19)	(-0.66, -0.16)
	ADN			
	Bread	Fat	Fish	Fruit
Bread	**-0.75**	**0.30**	**0.12**	0.08
	(-0.86,-0.62)	(0.19, 0.44)	(0.02, 0.24)	(-0.08, 0.19)
Fat	**0.17**	**-0.65**	0.10	0.08
	(0.04, 0.28)	(-0.76, -0.50)	(-0.01, 0.22)	(-0.08, 0.19)
Fish	**0.21**	0.16	**-0.99**	**0.18**
	(0.05, 0.39)	(-0.08, 0.32)	(-1.2, -0.83)	(0.01, 0.38)
Fruit	**0.18**	**0.23**	0.08	**-0.79**
	(0.04, 0.27)	(0.11, 0.38)	(-0.03, 0.22)	(-0.91, -0.62)
	TDN			
	Bread	Fat	Fish	Fruit
Bread	**-0.80**	**0.24**	0.06	**0.14**
	(-0.82, -0.76)	(0.18, 0.29)	(-0.04, 0.07)	(0.12, 0.16)
Fat	**0.18**	**-0.68**	**0.04**	**0.14**
	(0.14, 0.22)	(-0.77, -0.53)	(0.02, 0.06)	(0.09, 0.17)
Fish	**0.19**	**0.20**	**-0.92**	**0.16**
	(0.16, 0.24)	(0.09, 0.27)	(-0.95, -0.88)	(0.12, 0.21)
Fruit	**0.18**	**0.23**	**0.06**	**-0.83**
	(0.15, 0.21)	(0.14, 0.29)	(0.04, 0.08)	(-0.86, -0.77)

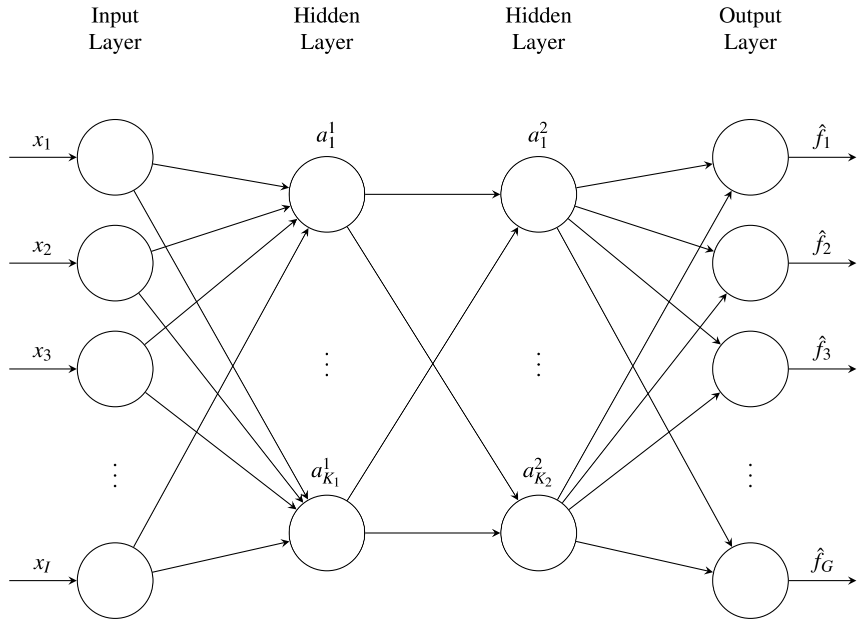
Note: Statistically non-zero estimates are illustrated in **bold**.

Table 1.4 Slutsky Matrices

	AIDS			
	Bread	Fat	Fish	Fruit
Bread	**-0.10**	**0.04**	0.01	-0.01
	(-0.14, -0.06)	(0.01, 0.07)	(-0.01, 0.03)	(-0.03, 0.01)
Fat		-0.01	0.00	0.02
		(-0.07, 0.03)	(-0.02, 0.02)	(-0.01, 0.05)
Fish			**-0.04**	0.01
			(-0.05, -0.02)	(-0.01, 0.02)
Fruit				**-0.05**
				(-0.09, -0.02)
	QUAIDS			
	Bread	Fat	Fish	Fruit
Bread	**-0.10**	**0.04**	0.01	-0.01
	(-0.14, -0.06)	(0.01, 0.07)	(-0.00, 0.03)	(-0.03, 0.01)
Fat		-0.02	-0.00	0.02
		(-0.07, 0.03)	(-0.02, 0.02)	(-0.00, 0.05)
Fish			**-0.04**	0.01
			(-0.05, -0.02)	(-0.00, 0.02)
Fruit				**-0.05**
				(-0.09, -0.02)
	ADN			
	Bread	Fat	Fish	Fruit
Bread	**-0.14**	**0.05**	**0.02**	0.01
	(-0.16, -0.11)	(0.03, 0.08)	(0.00, 0.04)	(-0.01, 0.03)
Fat	**0.04**	**-0.16**	0.02	0.02
	(0.01, 0.06)	(-0.19, -0.12)	(-0.00, 0.05)	(-0.02, 0.04)
Fish	**0.01**	0.00	**-0.05**	**0.01**
	(0.00, 0.02)	(-0.00, 0.01)	(-0.06, -0.04)	(0.00, 0.02)
Fruit	**0.02**	**0.03**	0.01	**-0.11**
	(0.00, 0.04)	(0.01, 0.05)	(-0.00, 0.03)	(-0.13, -0.09)
	TDN			
	Bread	Fat	Fish	Fruit
Bread	**-0.15**	**0.04**	**0.01**	**0.02**
	(-0.16, -0.14)	(0.03, 0.06)	(0.00, 0.01)	(0.02, 0.03)
Fat	**0.04**	**-0.17**	**0.01**	**0.03**
	(0.03, 0.06)	(-0.19, -0.14)	(0.00, 0.01)	(0.02, 0.04)
Fish	**0.01**	**0.01**	**-0.05**	**0.01**
	(0.00, 0.01)	(0.00, 0.01)	(-0.06, -0.04)	(0.00, 0.01)
Fruit	**0.02**	**0.03**	**0.01**	**-0.12**
	(0.02, 0.03)	(0.02, 0.04)	(0.00, 0.01)	(-0.13, -0.11)

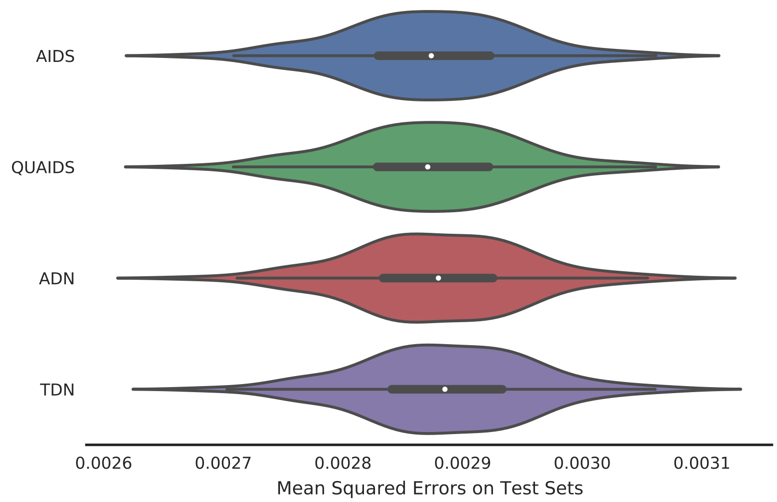
Note: Statistically non-zero estimates are illustrated in **bold**. Symmetry depends on the tolerance, e.g. $s_{bread,fat} = 0.0398$, $s_{fat,bread} = 0.0402$ when $\varepsilon_s = 1e-3$. Lower-diagonal elements of the Slutsky matrices for AIDS and QUAIDS estimations are not reported due to imposed symmetry. Lower-diagonal elements of the Slutsky matrix for TDN estimation are reported to illustrate the impact of numerical imposition of symmetry through a penalty function and for comparison with the ADN estimates of Slutsky terms.

Fig. 1.1 Multiple Output FNN Diagram



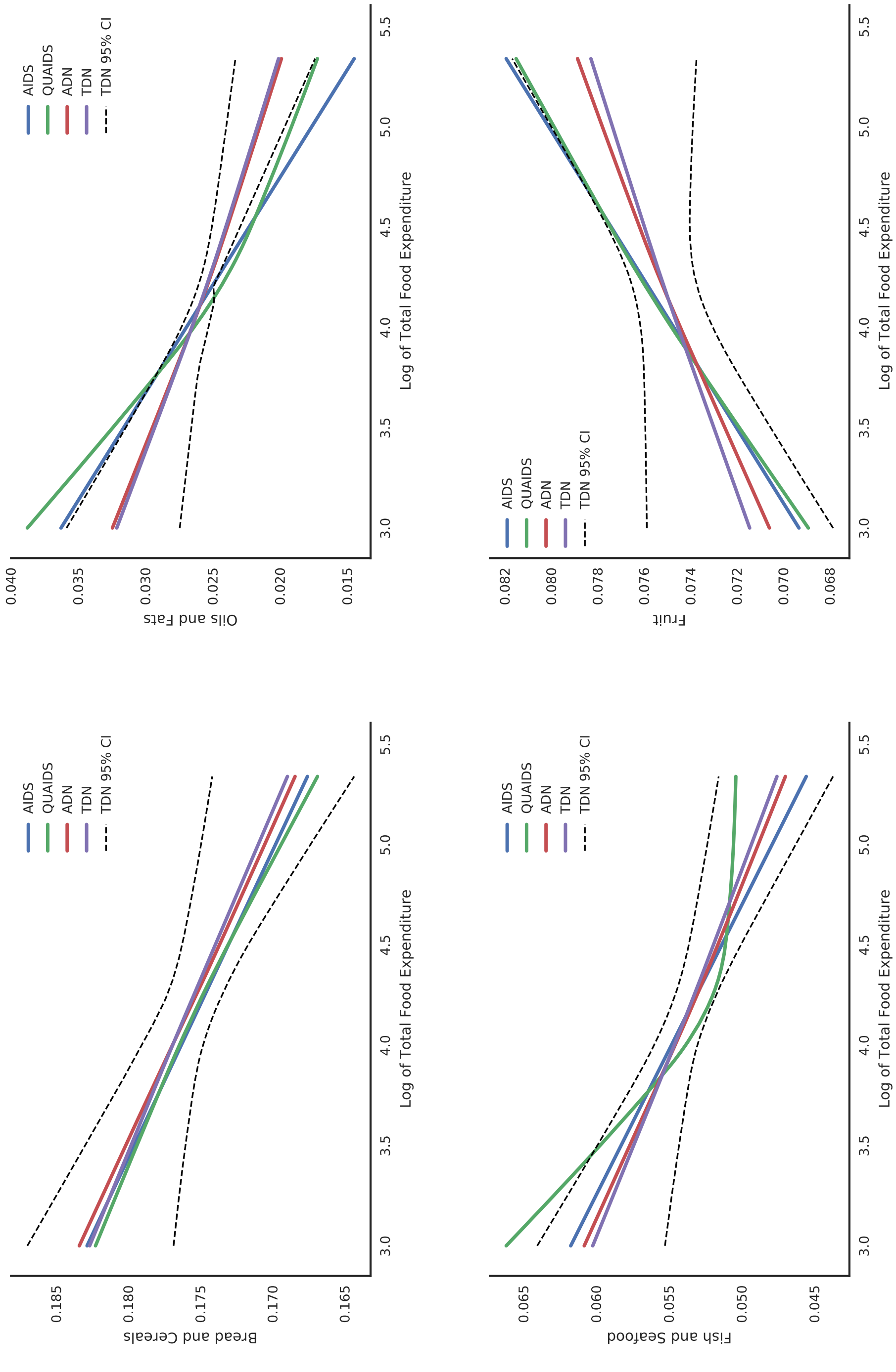
Note: The figure illustrates a network diagram for a multiple hidden layer - multiple output FNN.

Fig. 1.2 Distributions of Test Costs



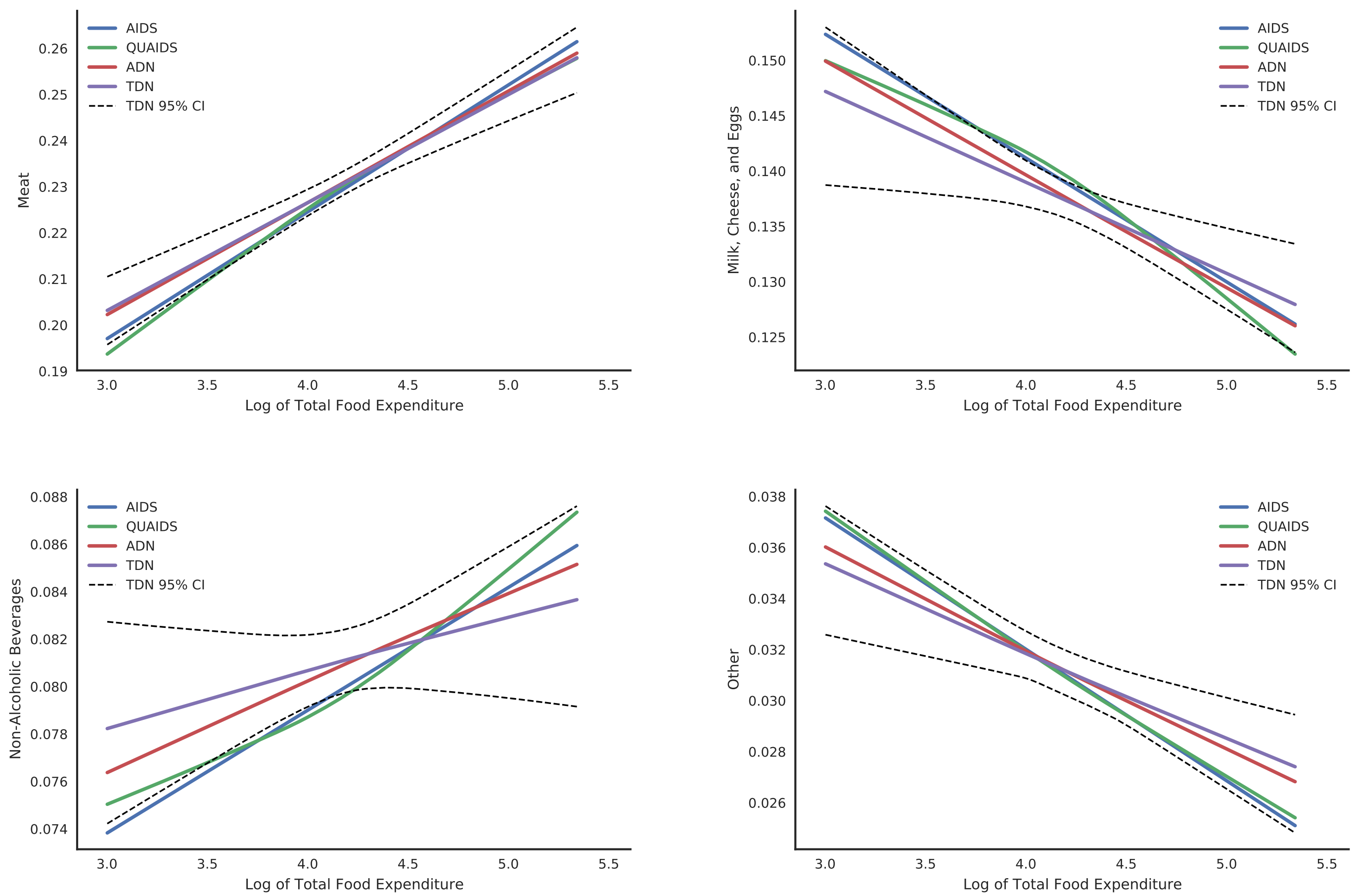
Note: The figure illustrates distributions of mean squared errors of four demand estimations on test sets, i.e., $C(\Theta, \mathscr{X}_t|\mathscr{X}_e, \mathscr{X}_c)$ for AIDS, QUAIDS, ADN, and TDN estimates. Violin plots indicate similar generalization performances for all estimations. White mid-points are means of test cost distributions. Bold black lines are interquartile ranges. Thin black lines within violins are 95% confidence intervals. Black curves are kernel density estimates of test cost distributions.

Fig. 1.3 Estimated Demand Functions



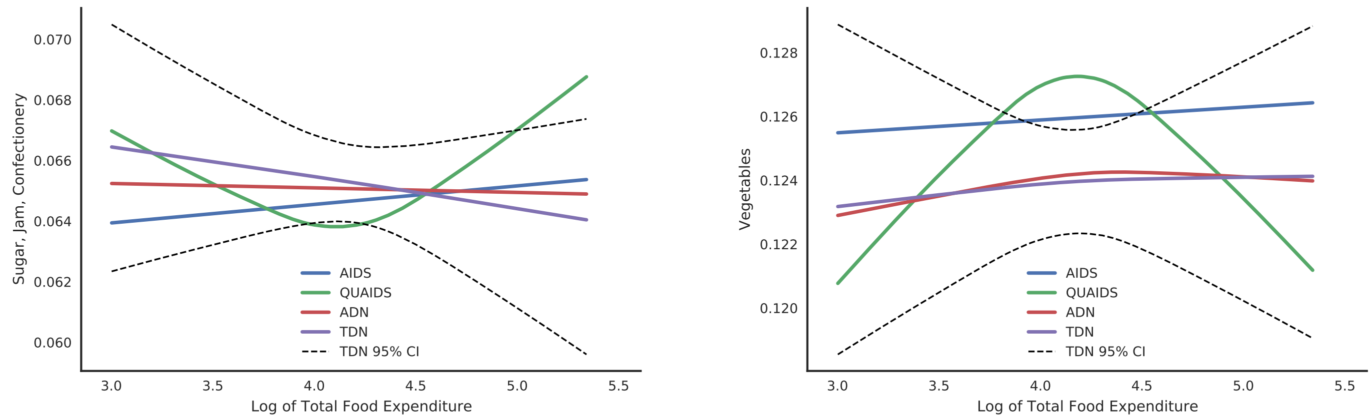
Continued on the next page.

Fig. 1.3 Estimated Demand Functions



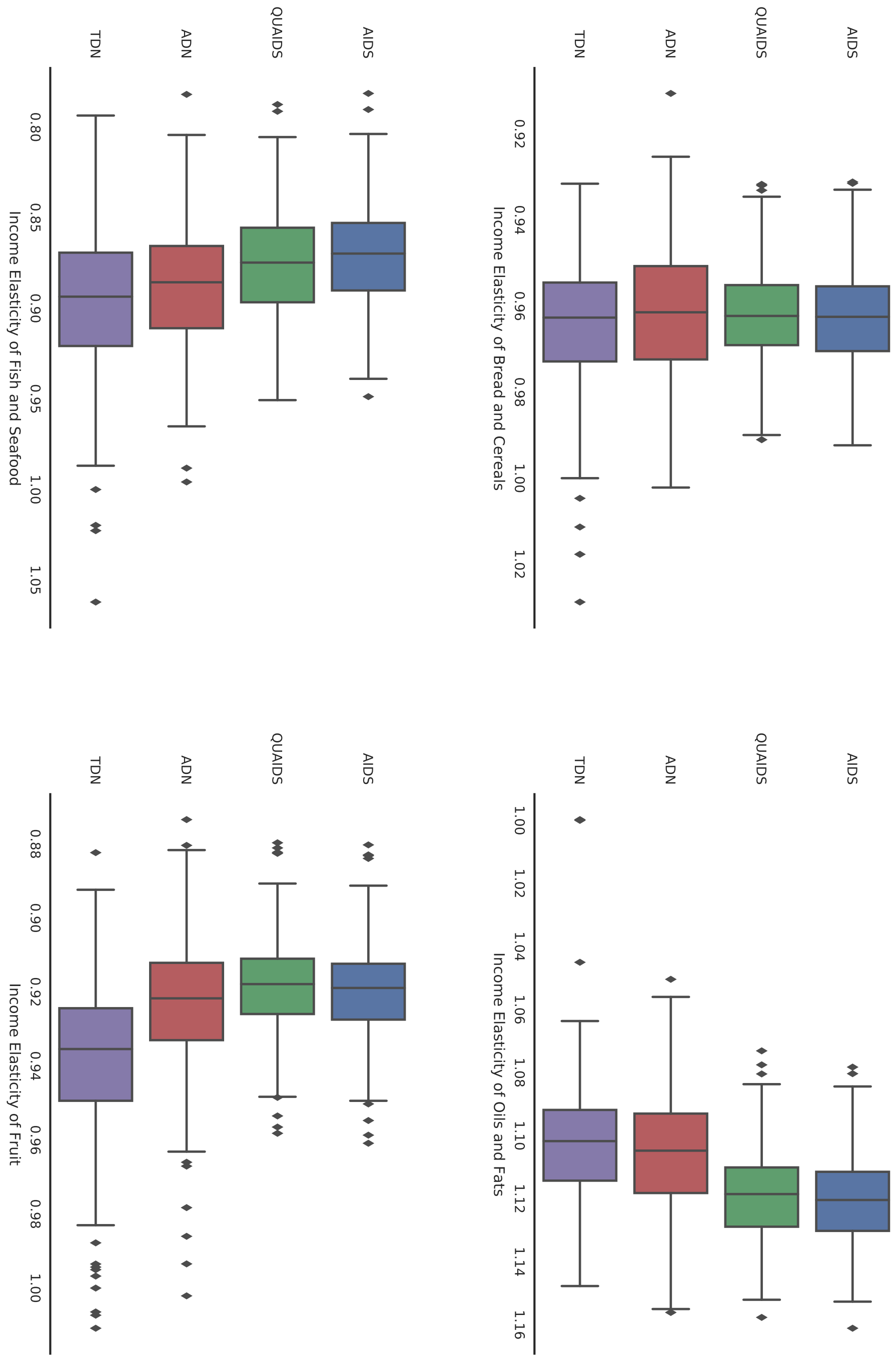
Continued on the next page.

Fig. 1.3 Estimated Demand Functions



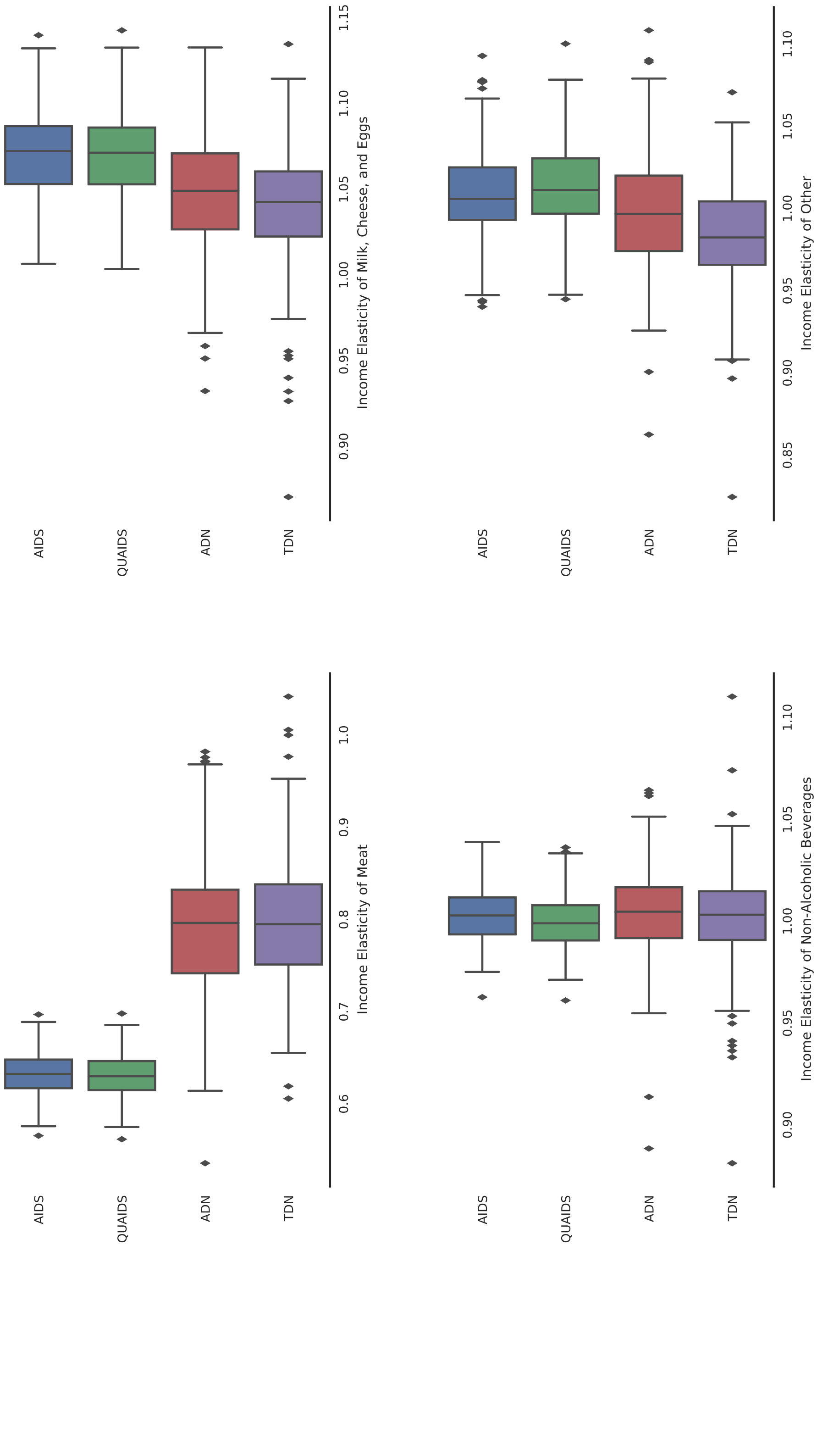
Note: Figures plot predicted budget shares of AIDS, QUAIDS, ADN, and TDN estimations on y-axes and log of total food expenditure on x-axes. Budget shares are estimated using all inputs from the main sample.

Fig. 1.4 Income Elasticities



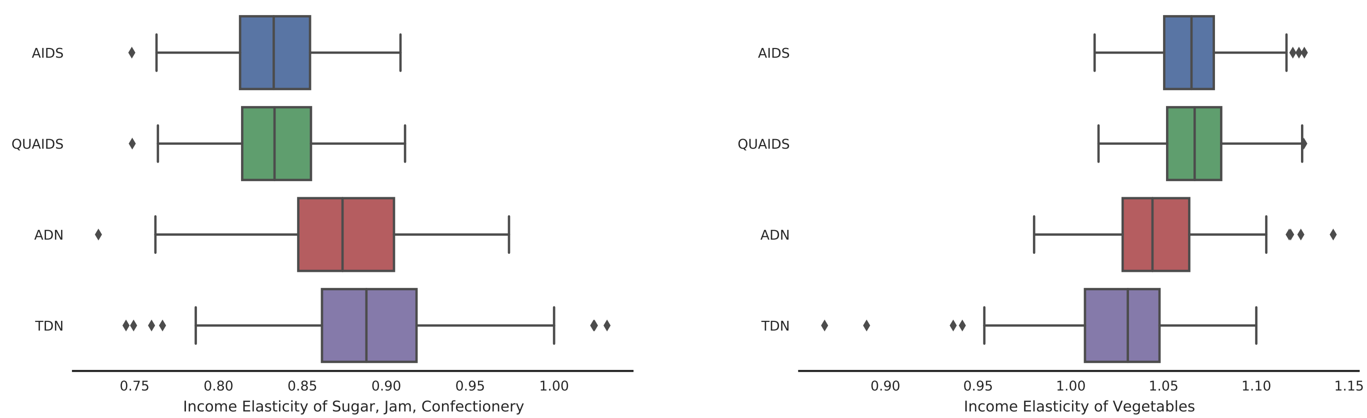
Continued on the next page.

Fig. 1.4 Income Elasticities



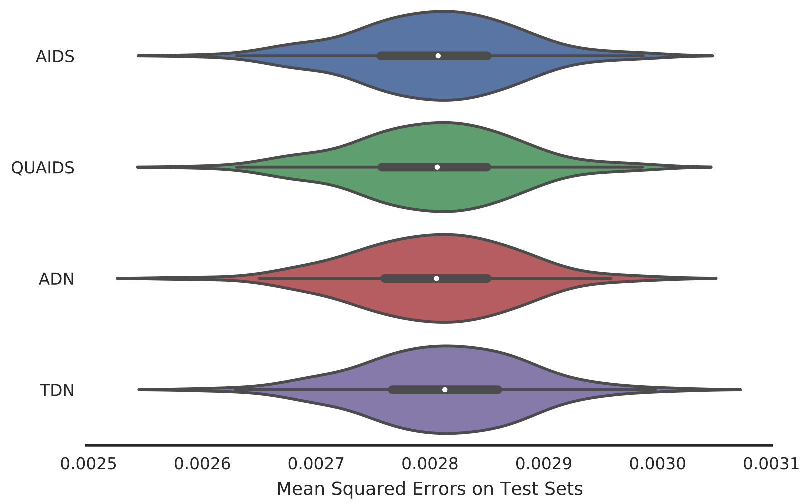
Continued on the next page.

Fig. 1.4 Income Elasticities



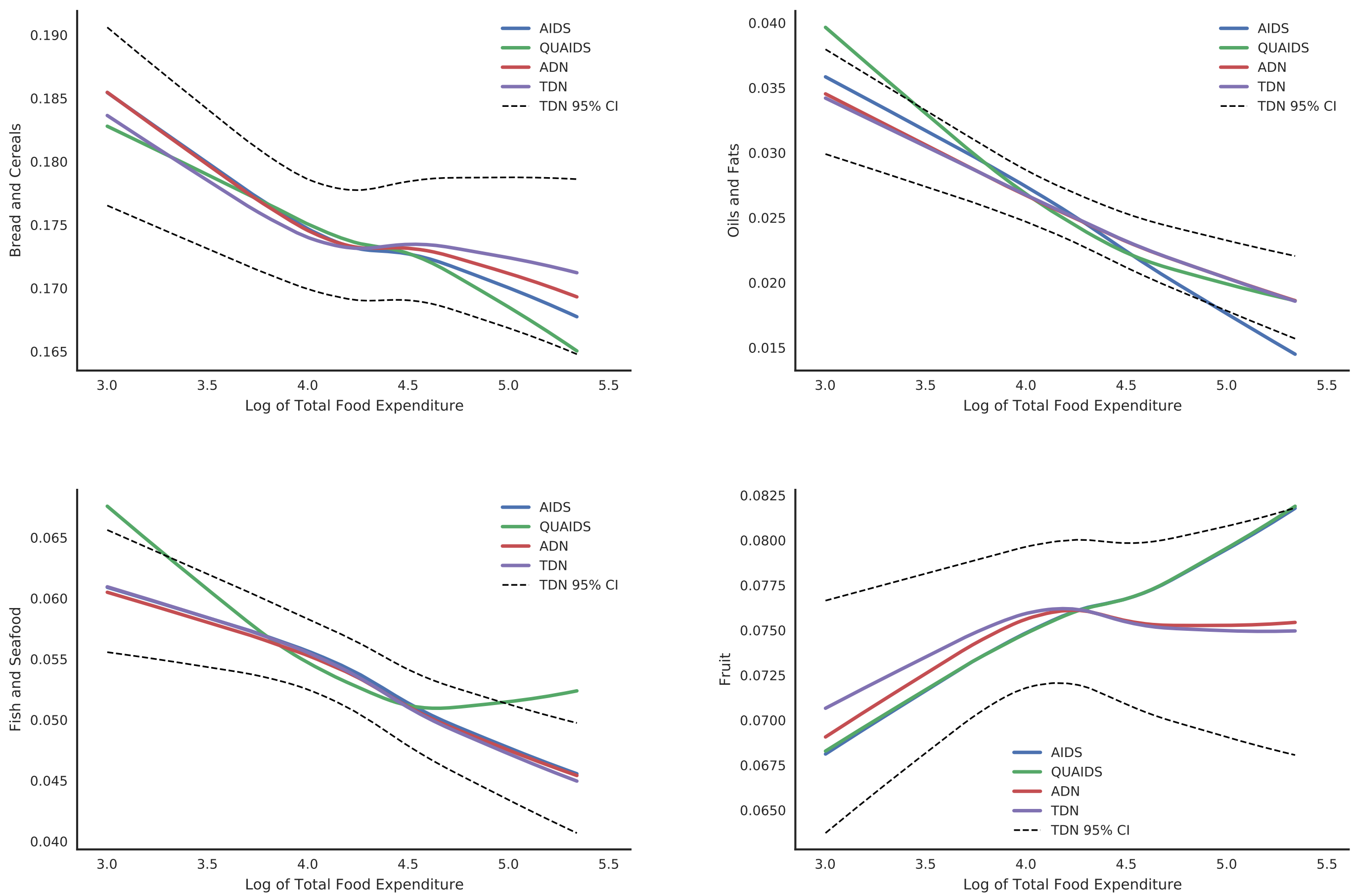
Note: Box-plots illustrate distributions of income elasticities.

Fig. 1.5 Distributions of Test Costs with Demographics



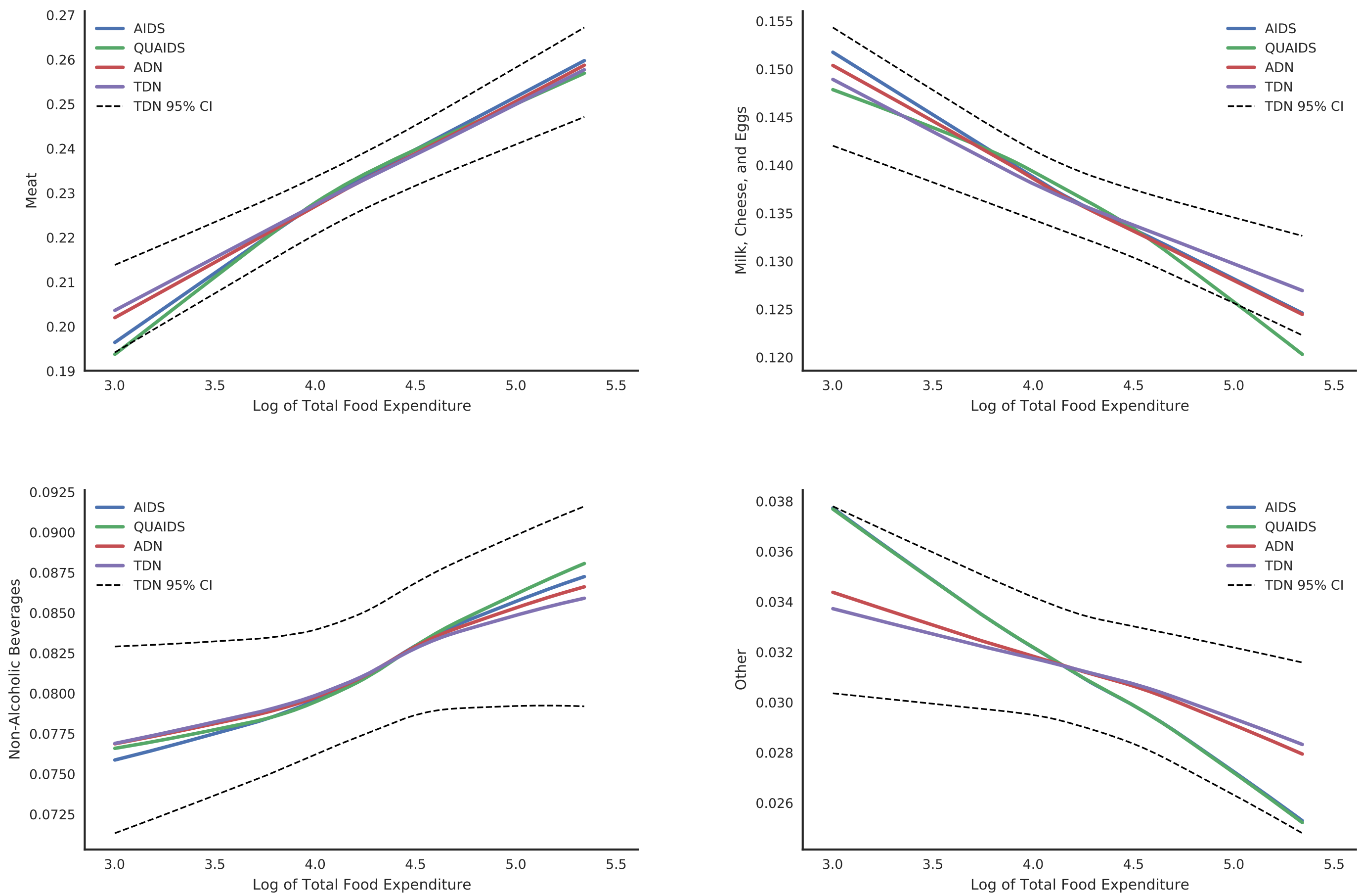
Note: The figure illustrates distributions of mean squared errors of four demand estimations on test sets, i.e., $C(\Theta, \mathscr{X}_t | \mathscr{X}_e, \mathscr{X}_c)$ for AIDS, QUAIDS, ADN, and TDN estimates. Violin plots indicate similar generalization performances for all estimations. White mid-points are means of test cost distributions. Bold black lines are interquartile ranges. Thin black lines within violins are 95% confidence intervals. Black curves are kernel density estimates of test cost distributions.

Fig. 1.6 Estimated Demand Functions with Demographics



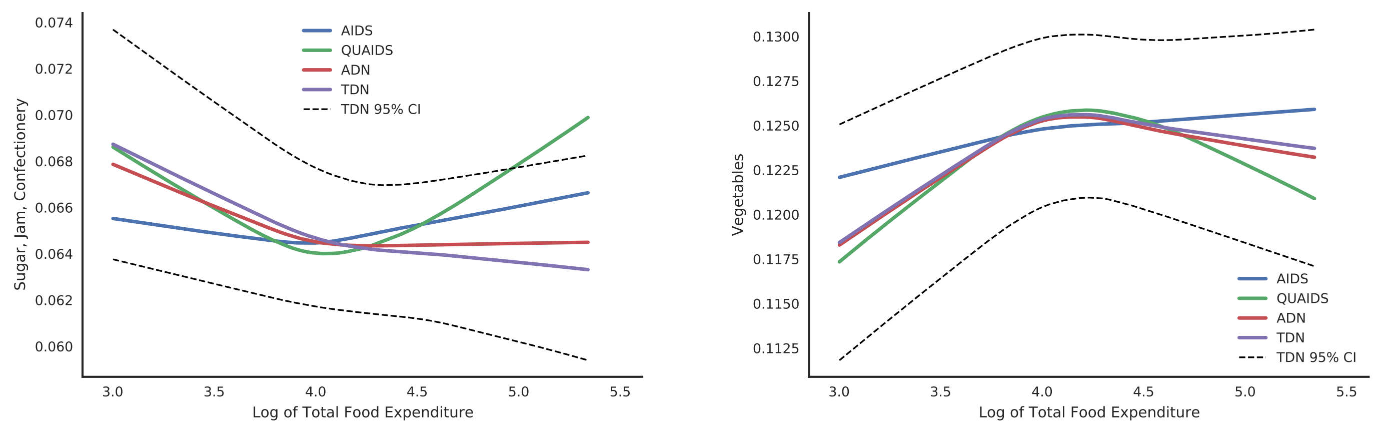
Continued on the next page.

Fig. 1.6 Estimated Demand Functions with Demographics



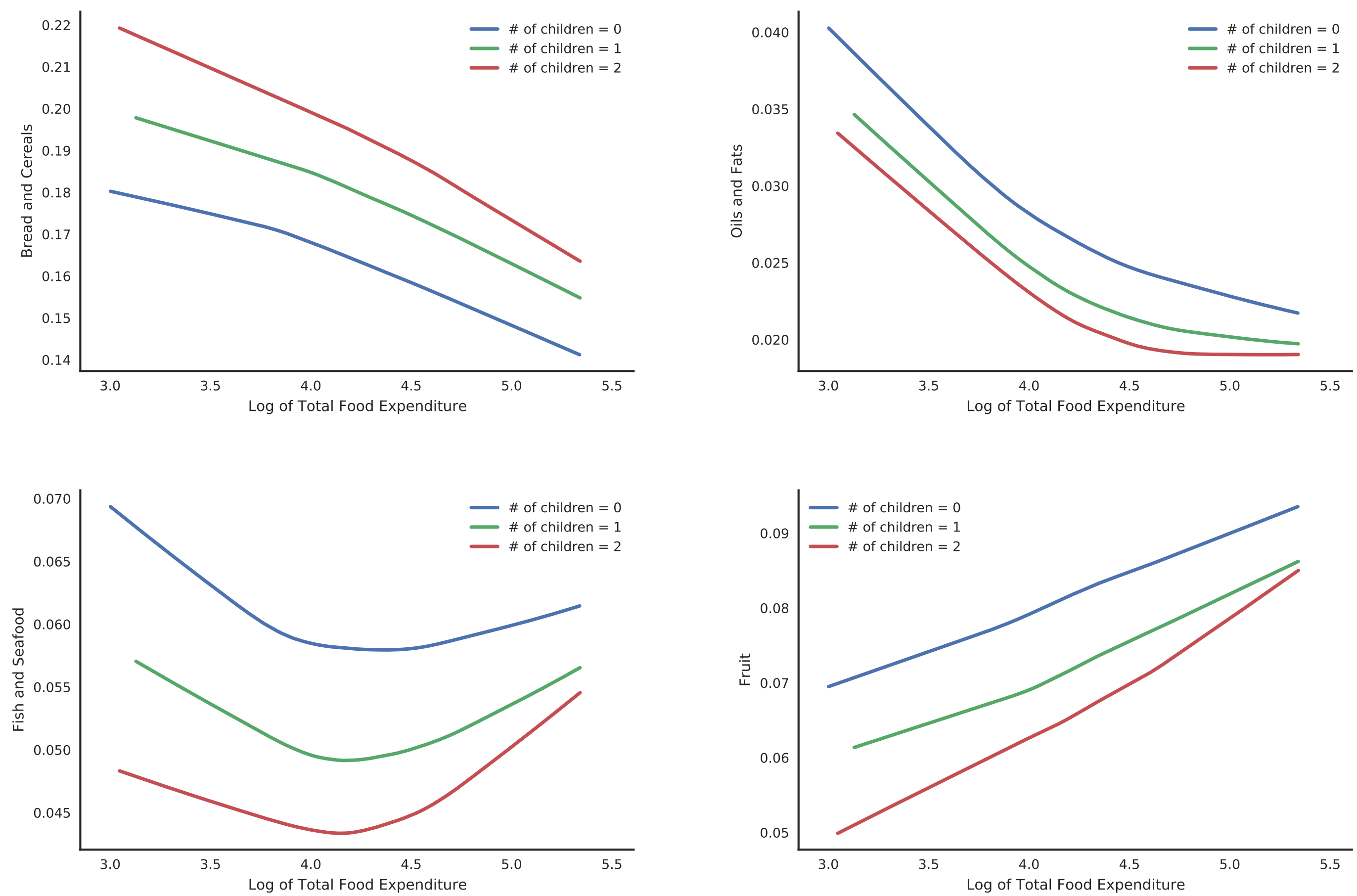
Continued on the next page.

Fig. 1.6 Estimated Demand Functions with Demographics



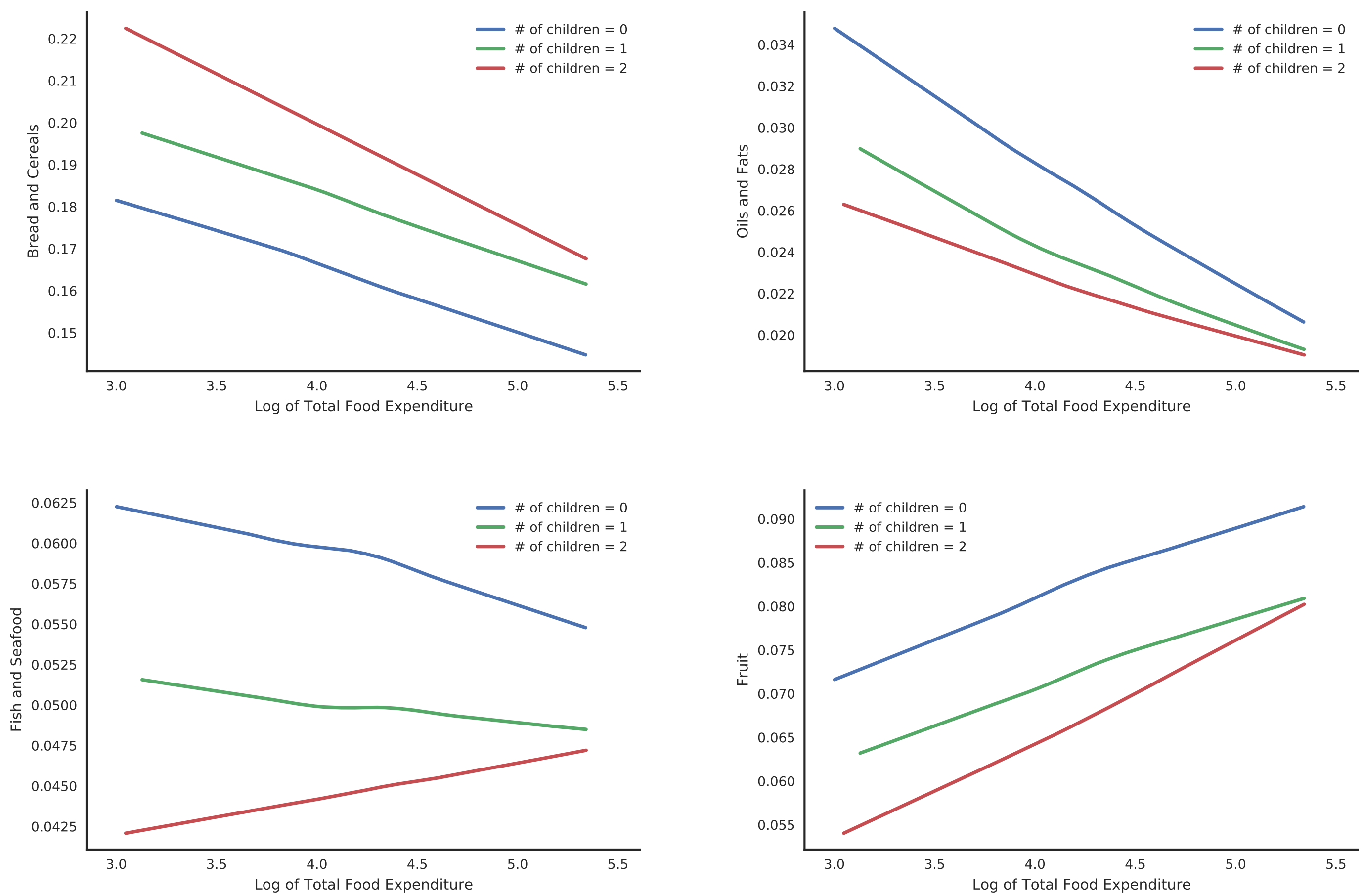
Note: Figures plot lowess-smoothed predicted budget shares of AIDS, QUAIDS, ADN, and TDN estimations on the y-axis, and log of total food expenditure on the x-axis. Plots are drawn using all observations from the main sample.

Fig. 1.7 Estimated QUAIDS Functions with Number of Children



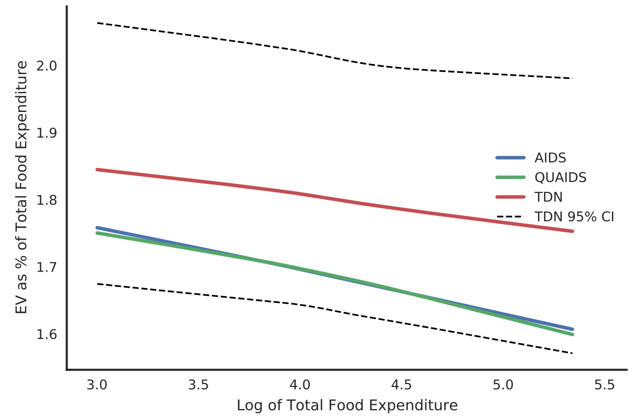
Note: Figures plot the lowess-smoothed predicted budget shares of QUAIDS estimations on the y-axis for households with different numbers of children and log of total food expenditure on the x-axis.

Fig. 1.8 Estimated TDN Functions with Number of Children

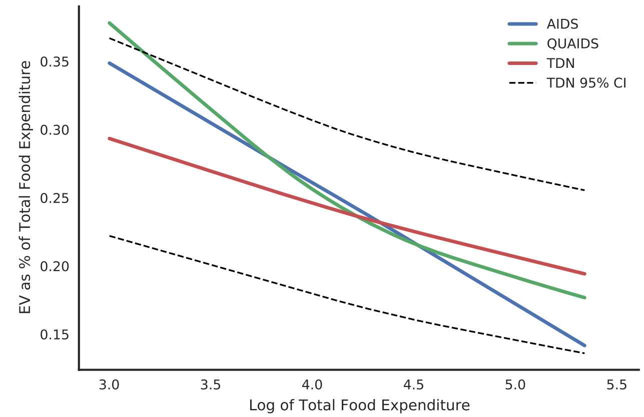


Note: Figures plot the lowess-smoothed predicted budget shares of TDN estimations on the y-axis for households with different numbers of children and log of total food expenditure on the x-axis.

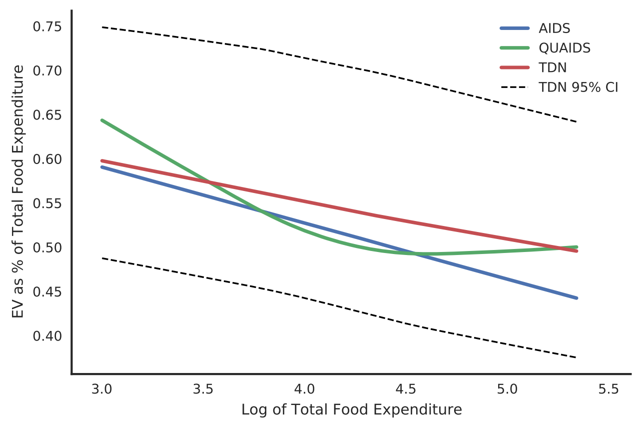
Fig. 1.9 Welfare Analysis



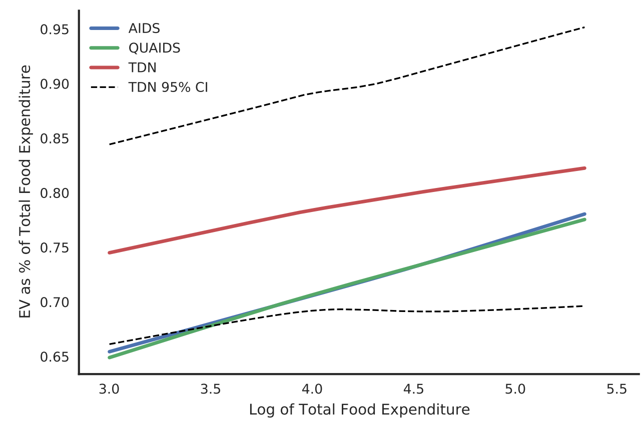
Increase in Bread and Cereals Prices



Increase in Oil and Fat Prices



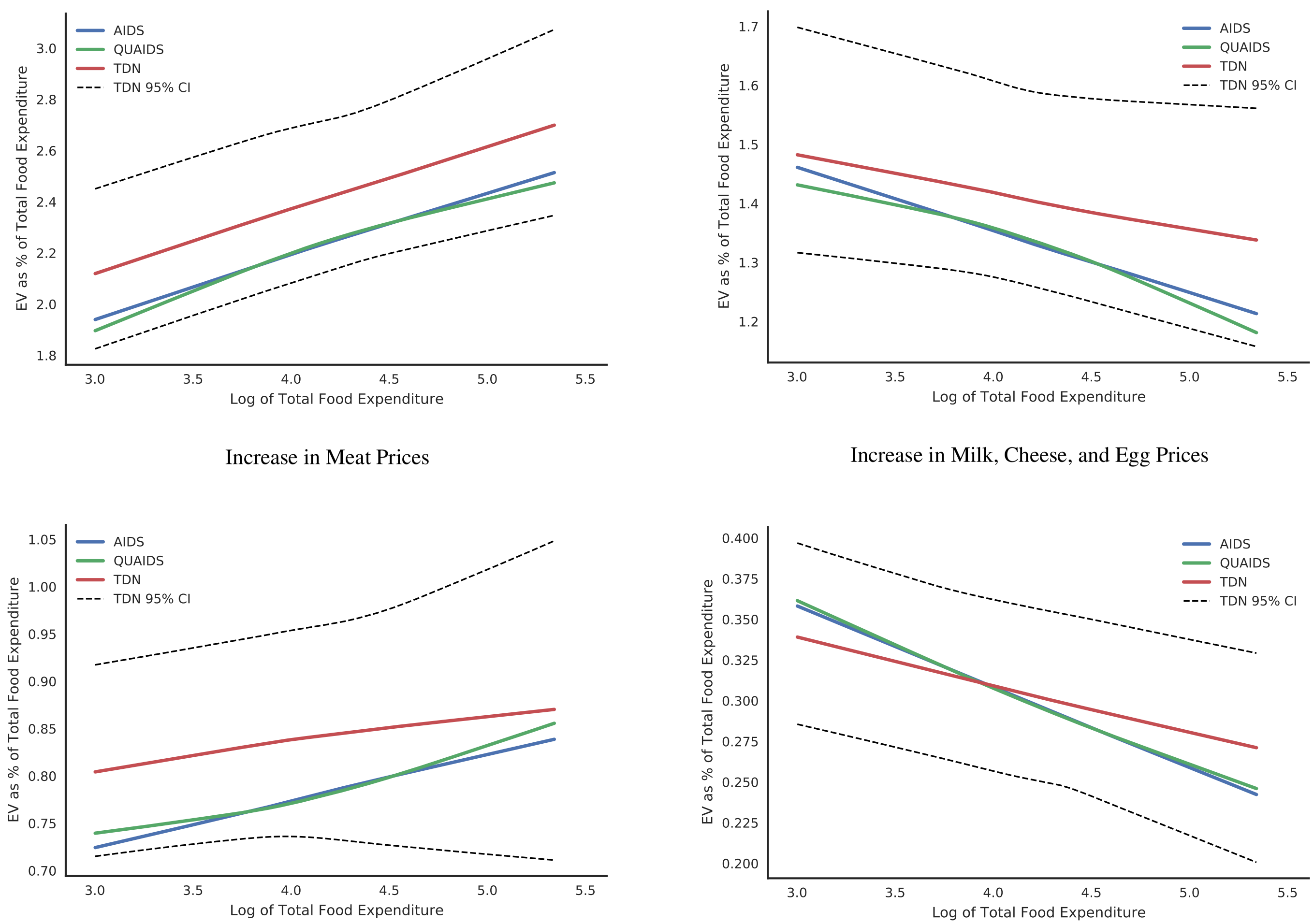
Increase in Fish and Seafood Prices



Increase in Fruit Prices

Continued on the next page.

Fig. 1.9 Welfare Analysis



Increase in Meat Prices

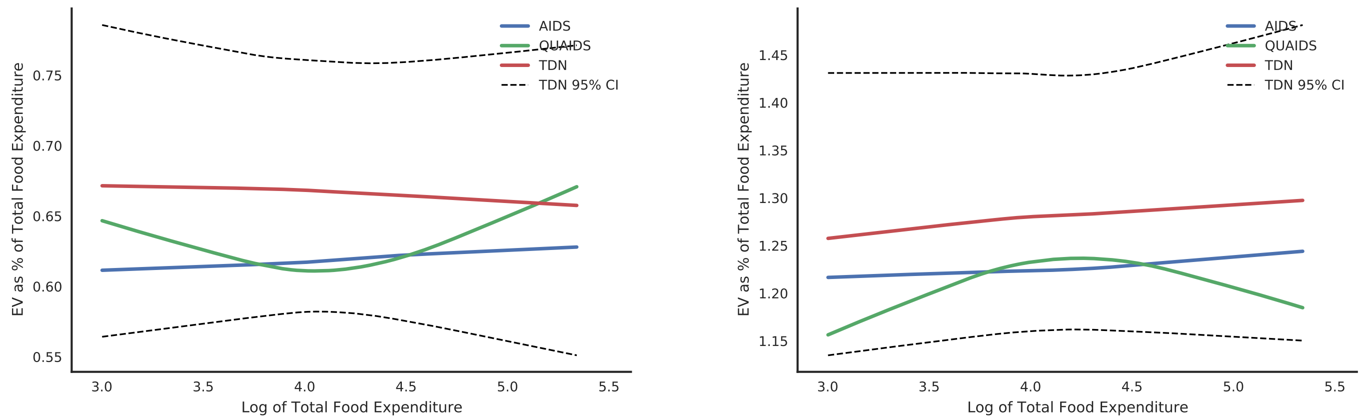
Increase in Milk, Cheese, and Egg Prices

Increase in Non-Alcoholic Beverage Prices

Increase in Other Goods Prices

Continued on the next page.

Fig. 1.9 Welfare Analysis



Increase in Sugar Prices

Increase in Vegetable Prices

Note: Figures illustrate equivalent variations of a 10% change in the price of each good as % of total food expenditure on the y-axis and log of total food expenditure on the x-axis.

Chapter 2
The PPP Approach Revisited: A study of RMB valuation against the USD

Ingvild Almås, Mandeep Grewal, Marielle Hvide, Serhat Ugurlu

Abstract We analyze the alleged undervaluation of the Chinese renminbi against the US dollar through an application of the relative PPP hypothesis; the PPP approach. The PPP approach measures the relative misalignment of a currency by estimating the relationship between log price levels and log per capita real incomes from a cross section of countries. We estimate this relationship by using ICP 2011 and incorporating model selection tests. Our results confirm that price level-real income relationship is best approximated by a quadratic functional form. We show that, using this functional form, the PPP approach does not reveal any evidence of renminbi undervaluation as of 2011, and this result is robust to various sensitivity tests.

Key words: renminbi, currency misalignment, Penn effect

Note: This chapter is published in *Journal of International Money and Finance, 2017*. When referring to this work, please cite the published article at `https://doi.org/10.1016/j.jimonfin.2017.06.006`.

2.1 Introduction

China has a history of exchange rate regimes with strict capital controls and foreign exchange interventions. Although the Chinese government allowed for a gradual appreciation of the renminbi (RMB) between 2005 and 2015, it has been argued that China still limited the movement of its exchange rate to maintain a trade advantage over its trading partners.[1] The United States, China's leading trade partner, has made the most prominent allegations. They have long argued that the renminbi is significantly undervalued against the US dollar, and this undervaluation played a

---

[1] In fact, in August 2015, China devalued the renminbi by around 3.5%. However this devaluation does not affect the discussion of this paper as we analyze the potential undervaluation in 2011.

part in creating their large annual trade deficits with China (Morrison and Labonte, 2013). This argument was re-emphasized by the United States Treasury Department in reports released in 2014 and 2015 (U.S. Department of the Treasury Office of International Affairs, 2014, 2015).

A much used approach to investigate whether a currency is undervalued against another is based on the relative purchasing power parity (PPP) hypothesis: in order to obtain a long run measure of the real exchange rates, the relationship between price levels[2] and per capita real incomes is estimated. The relationship has been estimated to be upward sloping, and this is referred to as the Penn effect (Samuelson, 1994; Rogoff, 1996). The theoretical arguments are given by the Balassa-Samuelson hypothesis predicting an upward sloping relationship. In this approach, currency misalignments are taken to be the difference between a country's observed price level and its predicted price level (see, e.g., Rodrik, 2008). If the predicted price level is higher than the observed price level, the currency is considered overvalued, and vice versa. If there is no statistical difference between the estimated and the observed price levels, the currency is considered not to be misaligned.

In this paper, we revisit the PPP approach to currency misalignment, and discuss whether this approach provides any evidence for the alleged undervaluation of the renminbi against the US dollar in 2011 using ICP 2011. The PPP approach has previously been applied to analyze currency misalignments. Much of the literature has been devoted to the question of whether the renminbi is undervalued or not. A large strand of the literature argues that the renminbi has historically been undervalued against the US dollar. Most of this research is based on cross sectional data. Some examples are Frankel (2006), who uses cross sectional data from the Penn World Table 6.1 and suggests that the renminbi was undervalued by 36% in 2000, and Subramanian (2010), who uses ICP 2005, and suggests that the renminbi was undervalued by 30% in 2005. On the other hand, other researchers who use panel data sets suggest different results. For example, Zhang and Chen (2014) use a panel data set to test renminbi undervaluation and suggest that the renminbi was undervalued by 38% in 2011; however, if a control variable for net financial assets is included, then they conclude that the renminbi was overvalued by 8%. By extending the cross sectional estimation that is used by Frankel (2006), Cheung et al. (2007) apply the PPP approach on panel data and suggest that the undervaluation of the renminbi was not statistically significant in 2004. They also show that, if serial correlation is controlled for, renminbi undervaluation was never statistically significant between 1976 and 2004.

The standard way to estimate the price-income relationship has been to use a log-linear functional form, and to regress countries' price levels on their per capita real income levels. However, Hassan (2016) shows that this relationship is better approximated by a log-quadratic functional form. Independent research by Cheung et al. (2017) also analyses the implications of this non-linearity in the Penn effect

[2] Let E, P and $P^{\star}$ denote the nominal exchange rate, local, and foreign aggregate prices, respectively. Then, the price level (PL) of a country, defined as $PL = PPP/E$, is equal to its real exchange rate (RER), $RER = P/EP^{*}$, by definition via the following relation: $P/EP^{*} = (P/P^{*})/E = PPP/E \Longleftrightarrow RER = PL$.

by estimating renminbi misalignment with a quadratic form, and show that the misalignment result may vary substantially depending on the data set regardless of the functional form.[3] Our analysis contributes to the ongoing debate. First, we make use of the latest ICP round directly in a cross sectional analysis (see also Kessler and Subramanian, 2014), which arguably provides the most reliable PPP measure for China, account for a possible non-linearity with the suggested log-quadratic specification, and measure misalignment using statistical bounds on the price level estimates. Second, we present results with alternative estimations, such as kernel weighted local polynomials and a principal component regression. This is in addition to more standard robustness checks using different control variables, such as net foreign asset positions and sample stratifications.[4] Thus, our research provides a comprehensive analysis that includes different samples, functional forms, and estimation techniques. We select our preferred models using standard model selection tests and residual plots. Our results suggest that the preferred specifications provide no evidence for the alleged undervaluation of the Chinese renminbi as of 2011.[5]

Other approaches have been suggested and used to assess the alleged undervaluation of the renminbi.[6] For example, Bénassy-Quéré et al. (2004), and Cline and Williamson (2010) use a fundamental equilibrium exchange rate (FEER) approach to assess the misalignment of the Chinese renminbi. Their findings suggest that the renminbi was undervalued by 44% in 2003 and by 41% in 2009, respectively. Cui (2013) uses a behavioral equilibrium exchange rate (BEER) approach with monthly data and suggests that renminbi undervaluation varied between 25% to 35% during the course of 2011. By using different approaches, the IMF Consultative Group on Exchange Rates suggests that renminbi undervaluation was between 3% and 23% in 2011 (IMF, 2011).

As different methodologies and data sets can be used to study potential misalignment, and as different approaches sometimes yield different results, the search for a consensus on whether the renminbi is undervalued continues. We choose to use the PPP approach to study the alleged misalignment and consider this approach to be an important complement to alternative existing methodologies. The PPP approach, contrary to the other approaches, conducts a cross country analysis to estimate currency misalignments, and it has the benefit of being a general equilibrium approach that does not allow over or undervaluation of all exchange rates simultaneously (see, e.g., Lee et al., 2008; Subramanian, 2010).

The rest of the article proceeds as follows. Section 2 describes ICP 2011, and discusses why it is more appropriate than previous releases of the International

[3] According to their results, as of 2011, PWT 8.1 suggests that the renminbi was fairly valued, World Development Indicators suggest that the renminbi was overvalued around 30%, and this difference holds for both log-linear and log-quadratic specifications.

[4] We also estimate misalignment of the renminbi with PWT 9.0, using expenditure data from ICP 2011, excluding oil exporters from our estimations, and grouping countries by income levels, temperatures, income inequality levels, and institutional quality levels.

[5] On the other hand, our estimation using PWT 9.0 confirms that the renminbi was undervalued in 2005, which is the previous benchmark year for the ICP.

[6] See Coudert and Couharde (2009) for a review.

Comparison Program to assess exchange rate misalignments, especially for the renminbi. Section 3 summarizes the empirical methodology, and discusses the Balassa-Samuelson effect, which constitutes the theoretical foundation for the PPP approach. In section 4, we report the results of our functional form and misalignment analysis. Section 5 presents sensitivity tests, which show that our main result is robust to a variety of controls including different data sets, additional variables, and sample stratifications of ICP 2011. Section 6 concludes.

2.2 Data

PPPs, per capita real incomes, and yearly average nominal exchange rates are obtained from ICP 2011 that is the latest release of the International Comparison Program, which is organized by the World Bank irregularly every few years in collaboration with various academic, statistical, and economic institutions. The program specifically aims to provide worldwide estimates of purchasing power parities to obtain an internationally comparable measure of price levels and real incomes. Figure 2.1 plots the relationship between price levels and per capita real incomes in ICP 2011.

ICP 2011, with its increased coverage, provides PPP exchange rates in various aggregate levels for 177 participating countries.[7] The data set also reports population sizes, nominal exchange rates, nominal GDPs, and their breakdown values for the participating countries. The PPPs that are provided in the International Comparison Program are based on localized price surveys. The remaining information is collected from the collaborating institutions and countries' own national accounts.

ICP 2011 introduces several methodological improvements over the previous versions. As a result, the data provided by this latest release is argued to have a better quality (World Bank, 2014). Moreover, improvements in the representativeness of the underlying data and in the estimation techniques are documented by several authors such as Deaton and Aten (2017) and Inklaar and Prasada Rao (2017).

ICP 2011 addresses a major concern about previously available internationally comparable data on Chinese prices. This concern is a result of the lack of representativeness of Chinese data in ICP 2005, which caused a downward revision of China's real income level by 40% compared to its projections. These problems are caused by China's limited participation in the previous round of the International Comparison Program. In 2005, China participated in the price surveys with only 11 of its 34 provinces, and provided price data from only urban areas of these provinces. This limited participation resulted in a considerable urban bias in the Chinese price level (Hill and Syed, 2010). Supporting this argument, Deaton and Heston (2010) and Feenstra et al. (2013) argued that the Chinese price level was overestimated by about 20% to 30% in ICP 2005. However in 2011, China participated with 30 of its 34 provinces, and provided price data from both urban and rural areas (Asian

[7] Our estimations that utilize the full sample have a sample size of 176 because the Cuban peso is not traded internationally.

Development Bank, 2014). As a result, ICP 2011 provides the most reliable internationally comparable Chinese price data as of date. To the best of our knowledge, we are the first to utilize this data to analyze whether the renminbi is undervalued against the US dollar.

In the sensitivity analyses, we make use of various other data sets. One of these data sets is a data set on country expenditure levels, which is provided by ICP 2011. The ICP data provides the aggregate expenditure level of each participating country, as well as the PPP exchange rates for expenditures. These expenditure levels are obtained from the participating countries' own national accounts. The PPP exchange rates for expenditures are constructed by using the same methodologies that are used to construct the standard PPP exchange rates. Explanations of all the other data sets we use are available in appendix B.1. We report the results on the largest samples for which data are available.

2.3 Methodology

The standard way to apply the PPP approach is to estimate the relationship between price levels and per capita real incomes by using a log linear regression (Rogoff, 1996). This functional form is specified in equation 2.1 where PL_i is the ratio of the PPP exchange rate to the nominal exchange rate for country i and $GDPP_i$ is the real gross domestic product (GDP) per capita.

$$\ln PL_i = \beta_0 + \beta_1 \ln GDPP_i + \varepsilon_i \tag{2.1}$$

Misalignment with the long run real exchange rate is calculated as the deviation from the estimated price level. In other words, percentage misalignment of the price level is the difference between log of its observed level and log of its predicted level if the difference is statistically significant.

$$\%\ misalignment = \ln PL_i - \ln \hat{PL}_i \tag{2.2}$$

However, as Hassan (2016) describes, a log linear relationship between real incomes and price levels is not empirically apparent. Therefore the standard log linear functional form may be misspecified. Hence we analyze the suitability of a log quadratic functional form as a more flexible functional form. To this end, we estimate a log linear functional form, as specified in equation 2.1, and a log quadratic functional form, as specified in equation 2.3, by applying an OLS regression with robust standard errors, and compare the results. Furthermore, an application with a nonparametric estimation can be found in appendix B.2.

$$\ln PL_i = \beta_0 + \beta_1 \ln GDPP_i + \beta_2 \ln GDPP_i^2 + \varepsilon_i \tag{2.3}$$

We test whether the log linear functional form is misspecified by applying model selection tests. We analyze the results of standard model comparison tests: a log-likelihood ratio test, Akaike, and Bayesian information criteria (AIC and BIC). A log likelihood ratio test analyses if the parameter vector of an unconstrained model satisfies some smooth constraint by comparing its likelihood with that of a constrained model. If we reject the null hypothesis of the test, that is if the constraint is binding, then the preferred model is the unconstrained model. In our context, a rejection means that the log quadratic functional form is the preferred functional form. In case of AIC and BIC comparisons, we choose the model with the lowest criterion as the preferred model (Greene, 2012). We also report residual-regressor plots of each estimation to see whether there is a significant relationship between the residuals and the regressors.

As sensitivity tests, we conduct various analyses by considering the theoretical explanation for the expected positive relationship between price levels and per capita real incomes. The most prominent theoretical explanation is the Balassa-Samuelson hypothesis (Balassa, 1964; Samuelson, 1964), which explains the relationship through international differences in relative productivity between the tradable and the non-tradable sectors. The hypothesis states that price levels are lower in low-income countries because of their relatively lower productivity in the tradable goods sector. Balassa and Samuelson argue that the tradable sector is dominated by production of goods while the non-tradable sector is mainly comprised of services which cannot be exported. Prices of the tradable goods are therefore determined in the world markets, whereas prices of the non-tradable goods are determined domestically. Moreover, with perfect labor mobility between the tradable and the non-tradable sectors, wages are equalized, and they are paid according to the marginal productivity.[8] The hypothesis suggests that as a country develops, the marginal productivity of labor in the tradable sector increases relative to the marginal productivity of labor in the non-tradable sector as a result of technological spillovers and "know how" through international trade. Hence, the aggregate price of the non-tradable goods increases relative to the aggregate price of the tradable goods. Since the price level of a country is given by a combination of the prices of both its tradable and non-tradable goods, the country experiences an increase in its aggregate price level and hence an appreciation of its real exchange rate.

First, note that the Balassa-Samuelson hypothesis is a productivity based explanation for higher price levels in more developed countries. Given that prices in the tradable sector are exogenous, increasing productivity results in an increase in wages of the tradable sector, and subsequently an increase in the price level of the non-tradable sector. Although, in the literature, income is used to test this relationship, it is documented that expenditure data may be more reliable because of various measurement reasons (see, e.g., Deaton, 1997; Meyer and Sullivan, 2011). Therefore, we include a robustness check where we analyze the relationship between log price level of expenditure and per capita real expenditures.[9]

[8] Wage in each sector is defined as the product of the marginal productivity of labor and the price level in that sector.

[9] Another way of conducting this analysis would be to instrument per capita real income using per capita real expenditure, and then regressing log price level of income on the predicted per capita income levels. Such an analysis yields very similar results in terms of misalignment.

Second, it is important to note that if the productivity increase in the tradable sector does not fully translate into higher prices in the non-tradable sector, the increase in the aggregate price level will be weaker. The Balassa-Samuelson hypothesis assumes that the translation occurs through wages in the tradable and the non-tradable sectors being equalized as a result of full labor mobility. However existence of wage differentials and imperfect sectoral labor mobility between the tradable and the non-tradable sectors have been documented in the previous literature (see, among others, Strauss and Ferris, 1996; Lee, 2006; Schmillen, 2013). As a result, countries may have very different price levels even if they have similar per capita income levels. This observation indicates that additional controls may be required to obtain a better price level prediction. We present additional sensitivity tests to account for this issue.

2.4 Results

A robust OLS estimation using the linear functional form suggests that the coefficient of log per capita real income is positive and significant ($p-value < 0.001$, see table 2.1). Log of the price level of China is estimated as -0.532 with a 95% confidence interval (-0.576, -0.489). The difference with the observed value suggests an 8% undervaluation of the Chinese renminbi compared to the US dollar as of 2011. Therefore, if we were to analyze only the outcome of the log linear functional form, we would conclude that the Chinese renminbi is undervalued against the US dollar in 2011. Figures 2.2 and 2.3 graphically represent these results.

On the other hand, estimating the relationship between price levels and per capita real incomes with a robust OLS estimation using the log quadratic functional form suggests that the coefficients of log per capita real income and its squared are significant (see table 2.1). Log of the price level of China is estimated as -0.614 with a 95% confidence interval (-0.671, -0.556). Thus the observed price level of China falls within the confidence interval, and hence the renminbi is not undervalued as of 2011. Figures 2.5 and 2.6 graphically represent these results.

Therefore, two functional forms yield different conclusions, and we need to determine which form better represents the relationship between price levels and per capita real incomes in ICP 2011. As the log linear functional form can be defined as a constrained form of the log quadratic functional form, we can use a log-likelihood ratio test to compare the two specifications. The result of the log-likelihood ratio test reports the test statistic as 16.40 which has a $\chi^2(1)$ distribution with a p-value indicating significant rejection of the null hypothesis that the coefficient of the quadratic term is zero ($p-value < 0.001$). Thus the log quadratic model is preferred. Other common tests for model selection includes checking Akaike information criterion (AIC) and Bayesian information criterion (BIC) for both models. Table 2.2 summarizes these tests. Both criteria suggest that the log quadratic functional form is preferable. Therefore all tests reveal that the log quadratic functional form is preferred.

Yet another way to investigate how well a model represents the underlying relationship in the data is to analyze the residual plots. Figure 2.4 plots residuals from the regression using the log linear functional form. According to the figure, residuals follow a nonlinear relationship with the regressors. On the other hand, figure 2.7 plots residuals from the regression using the log quadratic functional form. The residuals, except for those at the highest income quartile, show no significant relationship with per capita real incomes.

As we have shown, the results of the analysis depends on the functional form chosen for the estimation. An alternative route to the parametric one taken in the main analysis of this paper would be to use a nonparametric approach. Such an estimation also reveals that the Chinese renminbi is not undervalued as of 2011. Details of this nonparametric estimation are available in appendix B.2.

2.5 Sensitivity Analyses

In this section, we will first provide two sensitivity analyses to test the robustness of our results to sample selection. Then, we will discuss the results with various control variables and stratifications of ICP 2011.

2.5.1 Sensitivity to Sample Selection

The Relationship Between Price Levels and Per Capita Real Expenditures

Figure 2.8 plots log price levels of expenditures[10] and log per capita real expenditures.

Table 2.3 summarizes the results of robust OLS estimations using both functional forms. All coefficients are significant ($p-value < 0.001$). Estimating the relationship between price levels and real expenditures by using the log linear functional form suggests no misalignment whereas using the log quadratic functional form suggests that the Chinese renminbi is 17% overvalued. Estimation results using both functional forms are presented in figures 2.9 and 2.12.

Therefore, similar to the results for the estimations using the relationship between price levels and per capita real incomes, the two functional forms suggest different conclusions. A log-likelihood ratio test, where the log linear model is nested in the log quadratic model, yields a test statistic of 47.37 which has a $\chi^2(1)$ distribution indicating significant rejection of the null hypothesis that the coefficient of the quadratic term is zero ($p-value < 0.001$). AIC and BIC of the log quadratic model are also lower. Therefore standard model selection tests suggest that the log

[10] Price levels of expenditures are calculated similarly by using *PPP* for individual expenditures instead of *PPP* for GDP.

quadratic functional form is the preferred functional form when we analyze the empirical relationship using expenditure levels as well.

Moreover, figure 2.11 and figure 2.14 present the residual plots of the regressions for the two different functional forms. The log linear functional form reveals a nonlinear relationship whereas the log quadratic functional form reveals no significant relationship between the regressor and the residuals.

These analyses suggest that the log quadratic functional form is preferable when we estimate the relationship between price levels of expenditures and per capita real expenditures. The results confirm our main result that the Chinese renminbi is not undervalued against the US dollar in 2011.

The Relationship Between Price Levels and Per Capita Real Incomes - Using The Penn World Table 9.0 (PWT 9.0)

The PWT provides a panel of PPP exchange rates for GDP and its components incorporating the current and the past ICP rounds (see Feenstra et al., 2015a). In the construction of the PWT, different techniques are used to obtain the PPP estimates[11], and these estimates are extrapolated outside the benchmark years, which are the years without an ICP study, using country specific price indices from national accounts. In this latest release, that is the PWT 9.0, data from ICP 2011 is incorporated, the data from ICP 2005 is methodologically improved (Feenstra et al., 2015b; Inklaar and Prasada Rao, 2017), and the urban bias for the Chinese PPP estimates in ICP 2005 is accounted for (Feenstra et al., 2013).

In our analysis, we use the yearly cross section from PWT 9.0 to assess the misalignment of renminbi in 2011. Considering that the PWT 9.0 uses the benchmark data from ICP 2011, this analysis indicates whether the techniques used in obtaining PPP estimates affect our results. For the interested readers, PWT 9.0 also allows us to analyze the misalignment of renminbi in 2005, another benchmark year. Figure 2.15 shows the cross sections for the years 2005 and 2011 from PWT 9.0. We estimate the price-income relationship separately for each year using the corresponding cross sections, and thus refrain our analysis from the problems associated with PPP extrapolations based on national accounts (see, e.g., McCarthy, 2013).[12]

For both years, the preferred functional form is the quadratic functional form. The results of the log-likelihood ratio tests suggest significant rejections of the null hypotheses that the coefficients of the quadratic terms are zero for both years (see

[11] To obtain PPP estimates, ICP uses Gini-Éltetö-Köves-Szulc (GEKS) whereas PWT uses a combination of GEKS and Geary-Khamis (GK) aggregation techniques.

[12] Results for 1970 - 2014 can be provided upon request.

also table 2.5 for AIC and BIC tests).[13] The results suggest that the renminbi is not undervalued against the US dollar in 2011, confirming our main analysis with ICP 2011. However, the renminbi is estimated to be 40% undervalued in 2005.[14] Figures 2.16 and 2.19 graphically represent these results.

2.5.2 Additional Sensitivity Analyses

In this section, we further analyze the robustness of our result to the addition of various control variables, and stratifications of ICP 2011.[15]

Our first analysis is related to the assumption that the Balassa-Samuelson hypothesis rests upon, which is the full labor mobility. One difficulty in incorporating the degree of sectoral labor mobility into our estimation is that it is unobservable. Artuc et al. (2015) show that the cost of labor mobility is negatively correlated with various educational indicators.[16] Therefore, mean years of education may be a proxy to account for the degree of sectoral labor mobility. This analysis reports that the renminbi is not undervalued against the US dollar as of 2011.

Another control may be applied by incorporating net foreign asset positions. Zhang and Chen (2014) show that the undervaluation of the Chinese renminbi, which they obtain by using the PPP approach and a panel of 19 large economies, is not robust to the addition of a control variable for net foreign asset positions. Moreover, Cline and Williamson (2008) suggest that controlling for financial variables, such as net foreign asset positions, is an improvement over the standard PPP approach. Our analysis[17] indicates the renminbi is not undervalued, showing that our initial result is robust to the inclusion of net foreign assets as a control variable.

Considering that the misalignment estimates are sensitive to the data at hand, we also test our result with different subsamples. One such subsample could be obtained with the exclusion of oil exporting countries because oil exporting economies may exhibit different dynamics between price levels and real incomes (see, e.g., Korhonen and Juurikkala, 2009). Our analysis with the subsample that excludes the top oil

---

[13] The BIC test reports very similar values for both functional forms in 2005, and the log-likelihood ratio test report a p-value of 0.026, which suggests a significance rejection of the null hypothesis that the coefficient of the quadratic term is zero at 5% level of significance. The significance of the quadratic term is stronger if the analyses is repeated with a sub-sample that does not include some high income outlier countries that do not exhibit the Penn effect. If so, the renminbi is estimated 33% undervalued as of 2005. The quadratic functional form and the result for 2011 is robust to this change.

[14] The linear specification suggests that the renminbi was undervalued by 56% in 2005, and by 10.2% in 2011.

[15] We report the results of all estimations. Details of some of these analyses are in B.2. Details of the other analyses can be made available upon request.

[16] See figure 2.22 for the relationship between labor mobility costs, as estimated by Artuc et al. (2015), and mean years of education for 2011 from our education data set, which confirms their analyses.

[17] Net foreign asset positions of 154 countries as of 2011 are obtained from the World Bank.

exporters[18] shows that the renminbi is not undervalued against the US dollar, and confirms our initial analysis.

Other stratifications can be obtained by analyzing the renminbi in its own quartile and half of the world per capita income, average temperature, income inequality, and institutional quality distributions.[19] All of these results, as well as our previous results are summarized in table 2.6.[20] The linear functional form is the preferred functional form for some stratifications with small sample sizes. However, the quadratic functional form becomes the preferred choice when we repeat these analyses with a grouping that yields larger data sets[21], which is the sample half. In the preferred specifications, none of the analyses indicate an undervaluation of the Chinese renminbi.

2.6 Conclusion

In this paper, we analyze the alleged undervaluation of the Chinese renminbi against the US dollar by revisiting the PPP approach. By utilizing a new data set, ICP 2011, and following the discussion on the functional form of the Balassa-Samuelson effect, we find no evidence for the alleged undervaluation of the Chinese renminbi against the US dollar as of 2011. The existence of an undervaluation of the Chinese renminbi against the US dollar, which has been the focus of the literature so far, is only revealed if we use the log linear functional form, whose significant rejection as a preferred functional form is suggested by model selection tests. Our main results, as well as some of our robustness tests, suggest that the Chinese renminbi is neither undervalued nor overvalued whereas the other robustness tests suggest that the Chinese renminbi may be overvalued. Hence, we conclude that the PPP approach robustly shows no evidence for the alleged undervaluation of the renminbi as of 2011.

[18] See appendix B.3 for a list of these countries.

[19] Income distribution is obtained over the per capita incomes as reported by ICP 2011. Note that China is very close to being in the third quartile. Our nonparametric analysis already provides an analysis for China's exact place in the distribution. Data sources for other groupings are provided in B.1. To obtain institutional quality, the first principle component of the five institutional indicators (see appendix B.1) is used.

[20] Details and figures of all estimations can be provided upon request.

[21] The sample sizes for the quartile estimations are around 40.

Table 2.1 OLS Estimation Results with Income Data

	Log Linear	Log Quadratic
$\ln GDPP_i$	0.216^{***}	-0.732^{**}
$\ln GDPP_i^2$		0.053^{***}
Constant	-2.522^{***}	1.661^{*}
$\ln \hat{PL}_{China}$	-0.532	-0.614
95% CI	$(-0.576, -0.488)$	$(-0.671, -0.556)$
% misalignment	-8.0%	No Misalignment
R^2	0.448	0.497
Root MSE	0.299	0.286
Observations	176	176

$^{***}, ^{**}, ^{*}$ significant at 1%, 5% and 10%, respectively

Note: $GDPP_i$ stands for the country i's per capita real income. $\hat{PL}_{China}$ is the estimated price level of China. % misalignment is calculated as the difference between the log of the observed price level of China, which is -0.612, and the log of $\hat{PL}_{China}$ if the difference is statistically significant.

Table 2.2 Model Selection Criteria with Income Data

Functional Form	AIC	BIC
Log Linear	76.549	82.890
Log Quadratic	62.146	71.657

Note: AIC stands for Akaike information criterion, and BIC stands for Bayesian information criterion. Both criteria suggest the functional form with the lowest score as the preferred functional form.

Table 2.3 OLS Estimation Results with Expenditure Data

	Log Linear	Log Quadratic
$\ln exp_i$	0.277^{***}	-1.76^{***}
$\ln exp_i^2$		0.121^{***}
Constant	-2.837^{***}	5.553^{***}
$\ln \hat{PL}_{China}^{exp}$	-0.593	-0.728
95% CI	$(-0.641, -0.545)$	$(-0.785, -0.670)$
% misalignment	No Misalignment	17%
R^2	0.506	0.623
Root MSE	0.293	0.257
Observations	176	176

***significant at 1%

Note: exp_i stands for the country i's per capita real expenditure. $\hat{PL}_{China}^{exp}$ is the estimated price level of expenditure for China. % misalignment is calculated as the difference between the log of the observed price level of expenditure for China, which is -0.558, and the log of $\hat{PL}_{China}^{exp}$ if the difference is statistically significant.

Table 2.4 Model Selection Criteria with Expenditure Data

Functional Form	AIC	BIC
Log Linear	69.050	75.391
Log Quadratic	23.676	33.187

Note: AIC stands for Akaike information criterion, and BIC stands for Bayesian information criterion. Both criteria suggest the functional form with the lowest score as the preferred functional form.

Table 2.5 Model Selection Criteria with PWT 9.0

Model	2005 AIC	2005 BIC	2011 AIC	2011 BIC
Log Linear	149.295	155.681	123.817	130.203
Log Quadratic	146.368	155.947	114.559	124.137

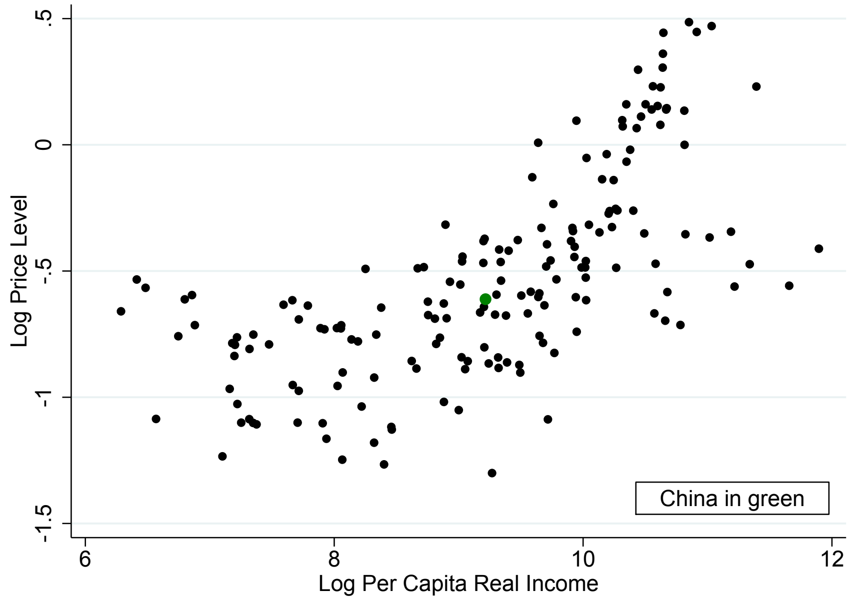
Note: AIC stands for Akaike information criterion, and BIC stands for Bayesian information criterion. Both criteria suggest the functional form with the lowest score as the preferred functional form.

Table 2.6 Results Summary

Price Level Data	Controls	Functional Form	Misalignment
ICP 2011		Log Quadratic	No Misalignment
ICP 2011		Non-parametric	No Misalignment
ICP 2011 (expenditure)		Log Quadratic	Overvalued (17%)
PWT 9.0		Log Quadratic	No Misalignment
ICP 2011	Labor Mobility	Log Quadratic	Overvalued (11%)
ICP 2011	Labor Mobility	Non-parametric	No Misalignment
ICP 2011	Net Foreign Assets	Log Quadratic	No Misalignment
ICP 2011 (w/o Top Oil Exporters)		Log Quadratic	No Misalignment
ICP 2011 (2^{nd} Income Quartile)		Log Linear	No Misalignment
ICP 2011 (1^{st} Income Half)		Log Quadratic	No Misalignment
ICP 2011 (2^{nd} Gini Quartile)		Log Quadratic	Overvalued (11%)
ICP 2011 (1^{st} Gini Half)		Log Quadratic	Overvalued (13%)
ICP 2011 (1^{st} Temperature Quartile)		Log Quadratic	Overvalued (11%)
ICP 2011 (1^{st} Temperature Half)		Log Quadratic	Overvalued (10%)
ICP 2011 (2^{nd} Institutional Quartile)		Log Linear	Overvalued (10%)
ICP 2011 (1^{st} Institutional Half)		Quadratic	Overvalued (15%)

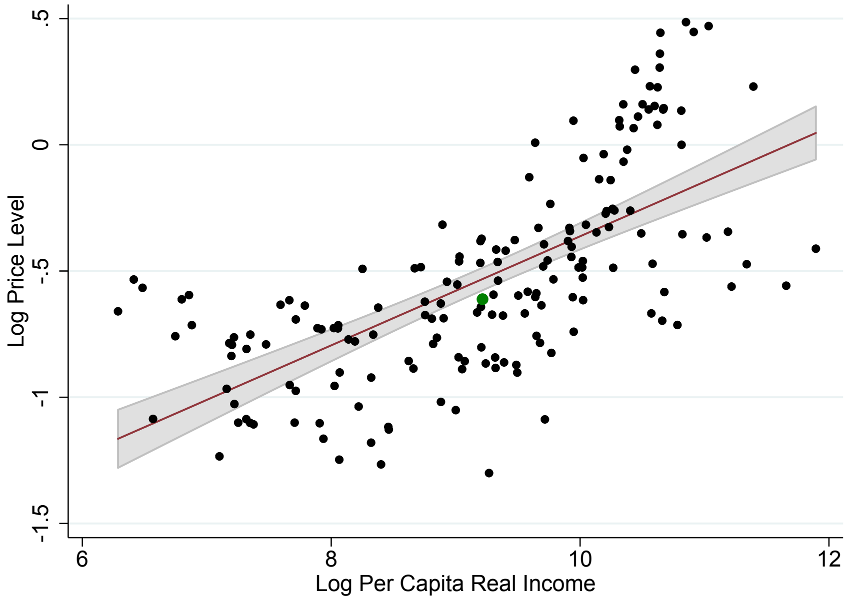
Note: The table presents a summary of misalignment analyses. The first column documents the data set used. (expenditure) indicates use of expenditure data instead of income data. “w/o” stands for without. Quartiles and halves show which subsamples of the indicated data set is used for each analysis, e.g. in which subsample China is. The second column shows additional control variables. The third column reports the misalignment results, and the distance of the log predicted price level to the log observed price level if the price level is statistically misaligned. The fourth column reports the functional form used in the analysis, which are the preferred functional forms for the parametric analyses.

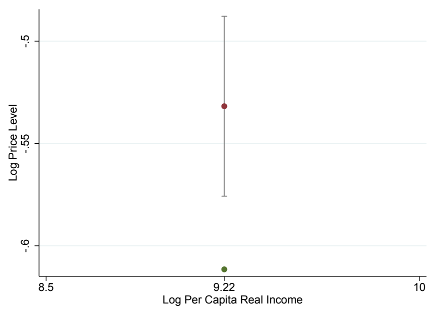
Fig. 2.1 Price Levels and Per Capita Real Incomes in ICP 2011



Note: The figure displays the log price levels, calculated as the ratio of the PPP exchange rates to the nominal exchange rates, and the per capita real incomes of each country in ICP 2011. China is in the middle of the distribution, and is represented by a green data point.

Fig. 2.2 OLS Estimation with Income Data - Log Linear Functional Form





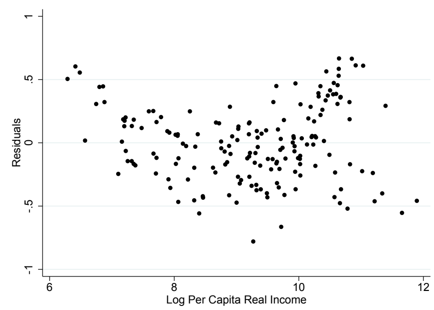
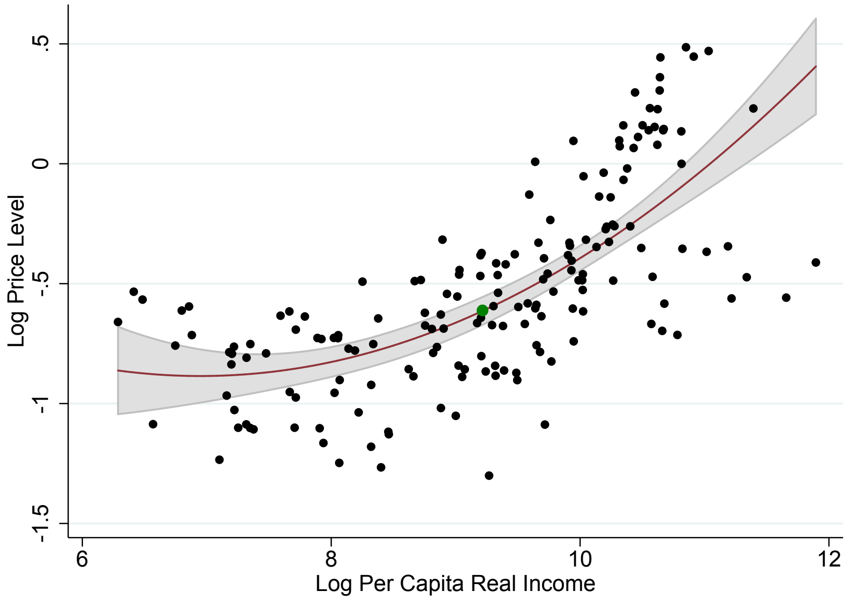


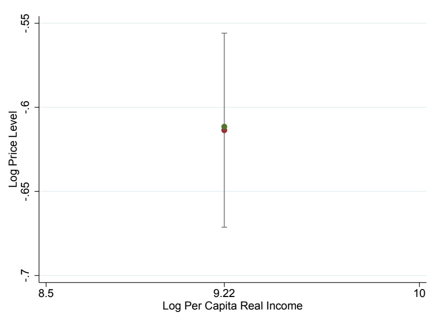
Fig. 2.3 Log Linear - China

Fig. 2.4 Log Linear - Residuals

Note: Figure 2.2 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure 2.3 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure 2.4 plots the residuals from the OLS estimation using the log linear functional form against the regressor.

Fig. 2.5 OLS Estimation with Income Data - Log Quadratic Functional Form





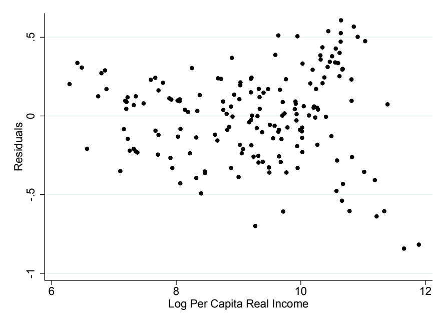
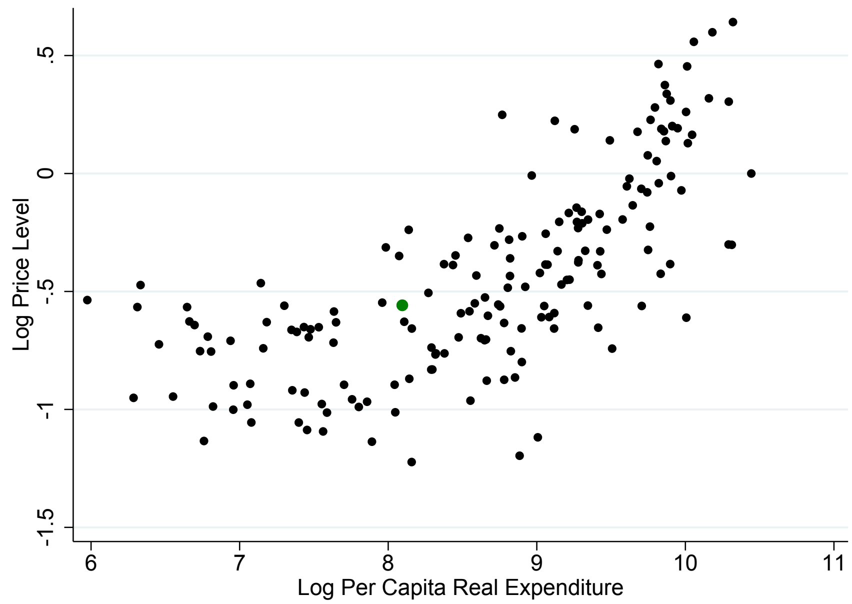


Fig. 2.6 Log Quadratic - China

Fig. 2.7 Log Quadratic - Residuals

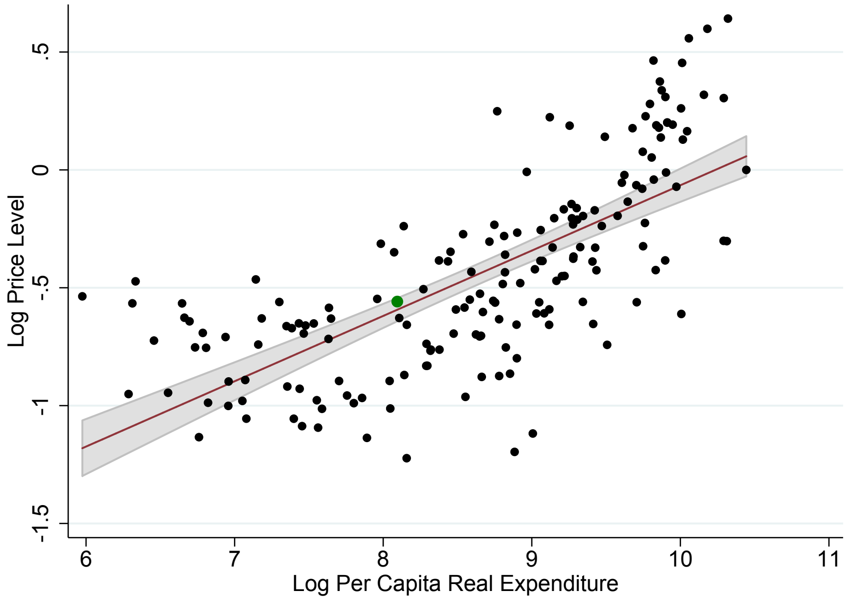
Note: Figure 2.5 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure 2.6 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure 2.7 plots the residuals from the OLS estimation using the log quadratic functional form against the regressor.

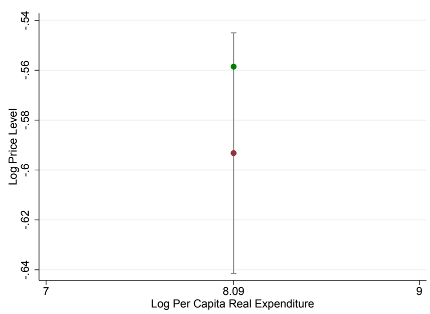
Fig. 2.8 Price Levels and Per Capita Real Expenditures in ICP 2011



Note: The figure displays the log price levels of expenditures, calculated as the ratio of the PPP exchange rates for expenditures to the nominal exchange rates, and the per capita real expenditures of each country in ICP 2011. China is close to the middle of the distribution, and is represented by a green data point.

Fig. 2.9 OLS Estimation Results with Expenditure Data - Log Linear Functional Form





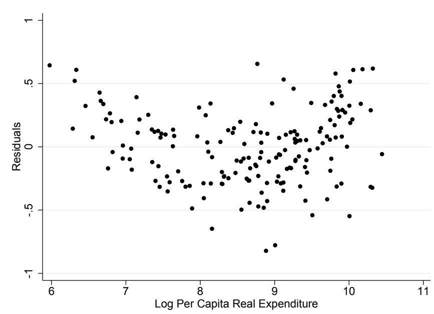
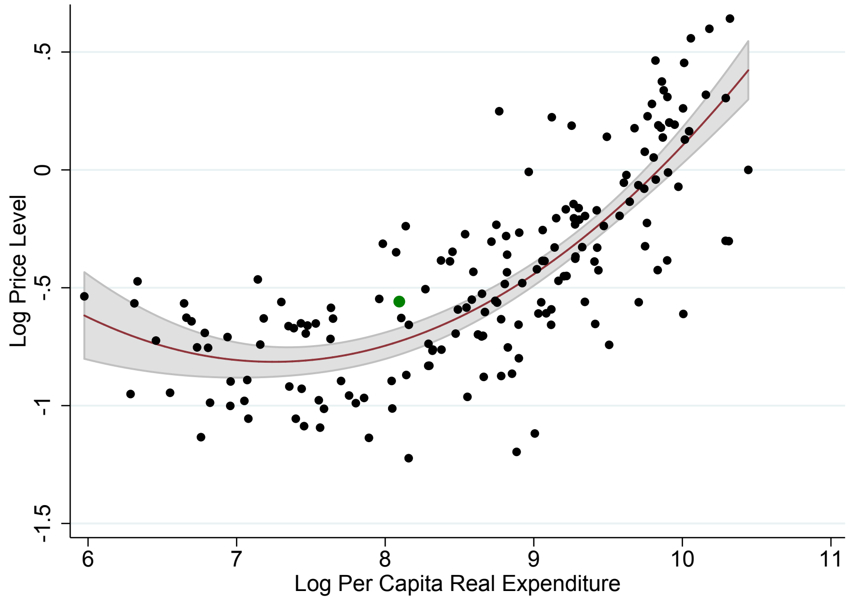


Fig. 2.10 Log Linear - China

Fig. 2.11 Log Linear - Residuals

Note: Figure 2.9 shows the estimated price levels of expenditures in red, and the 95% confidence interval in gray. The observed price level of expenditure for China is represented by the green data point. Figure 2.10 emphasizes the observed price level of expenditure for China, the estimated price level of expenditure for China, and the confidence interval. Figure 2.11 plots the residuals from the OLS estimation using the log linear functional form against the regressor.

Fig. 2.12 OLS Estimation Results with Expenditure Data - Log Quadratic Functional Form



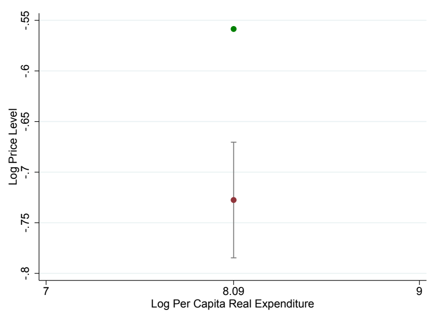


Fig. 2.13 Log Quadratic - China

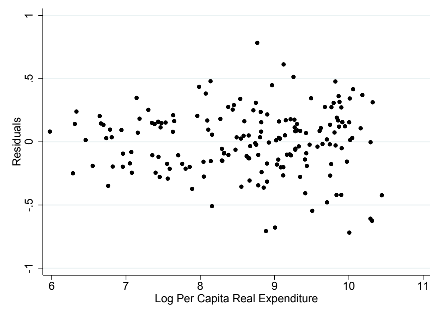
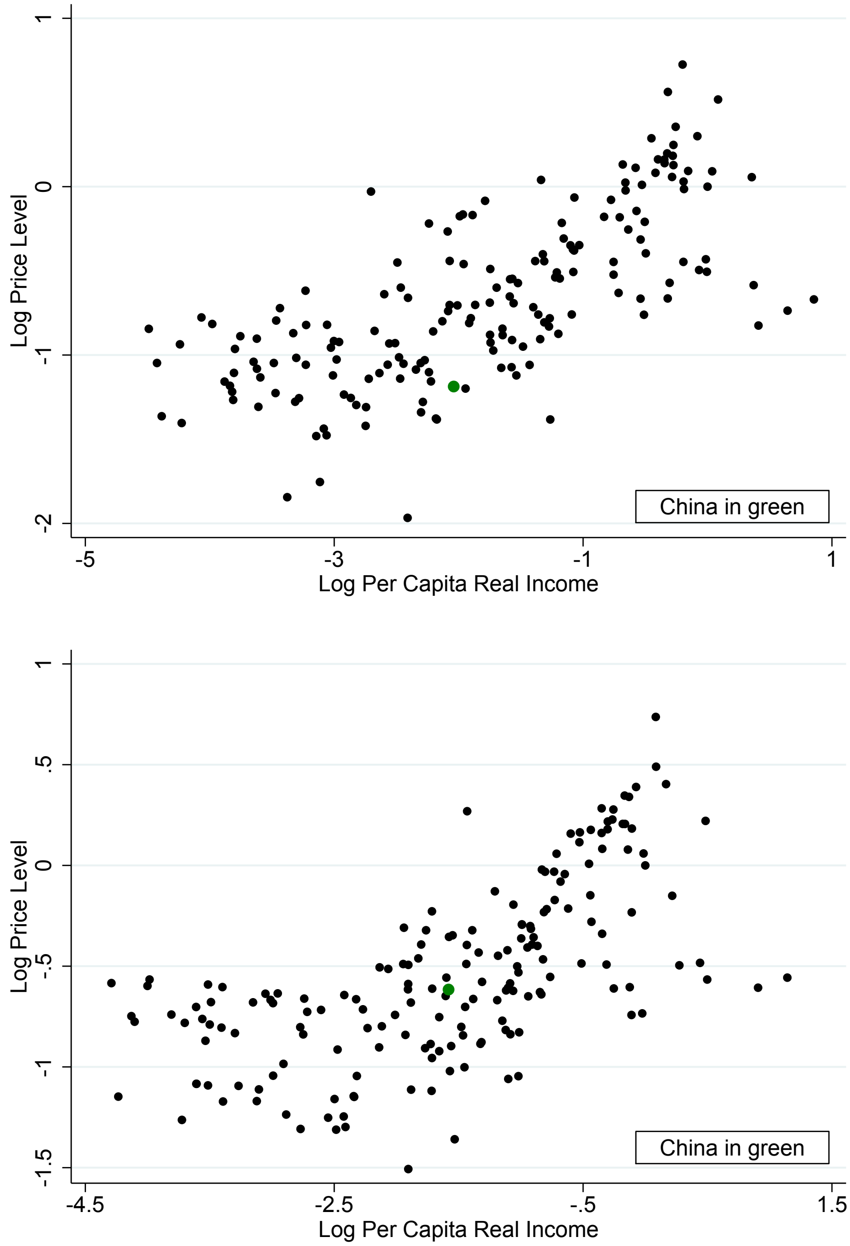


Fig. 2.14 Log Quadratic - Residuals

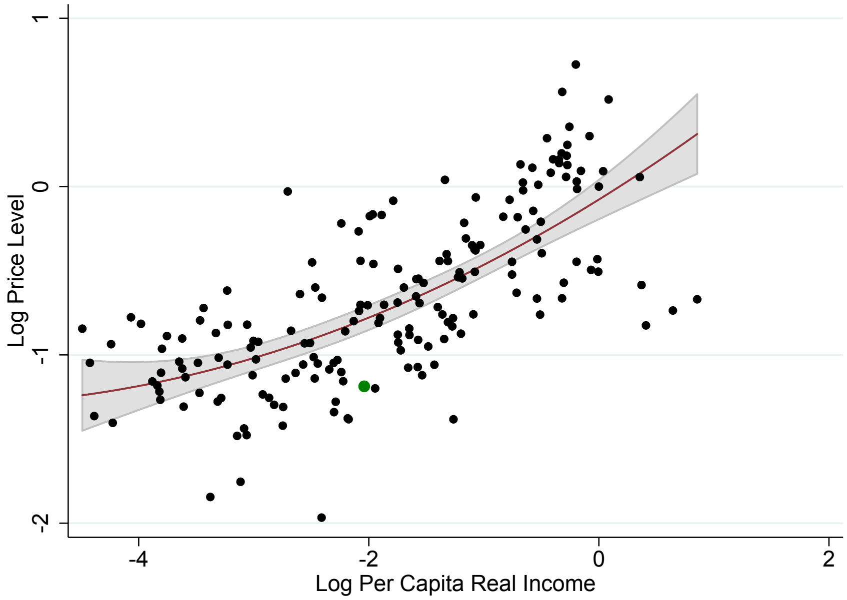
Note: Figure 2.12 shows the estimated price levels of expenditures in red, and the 95% confidence interval in gray. The observed price level of expenditure for China is represented by the green data point. Figure 2.13 emphasizes the observed price level of expenditure for China, the estimated price level of expenditure for China, and the confidence interval. Figure 2.14 plots the residuals from the OLS estimation using the log quadratic functional form against the regressor.

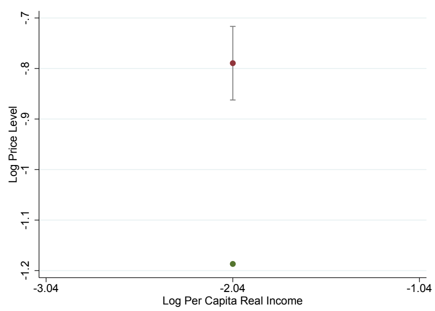
Fig. 2.15 PWT 9.0 2005 and 2011 Cross Sections



Note: The top figure shows the price-income relationship for the 2005 cross section of PWT 9.0. The bottom figure shows the same relationship for the 2011 cross section of PWT 9.0. In both panels, the price level of China is depicted with a green data point.

Fig. 2.16 OLS Estimation Results with PWT 9.0 2005





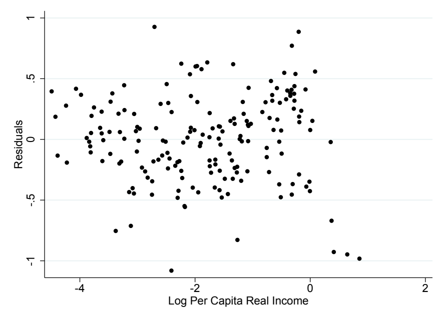
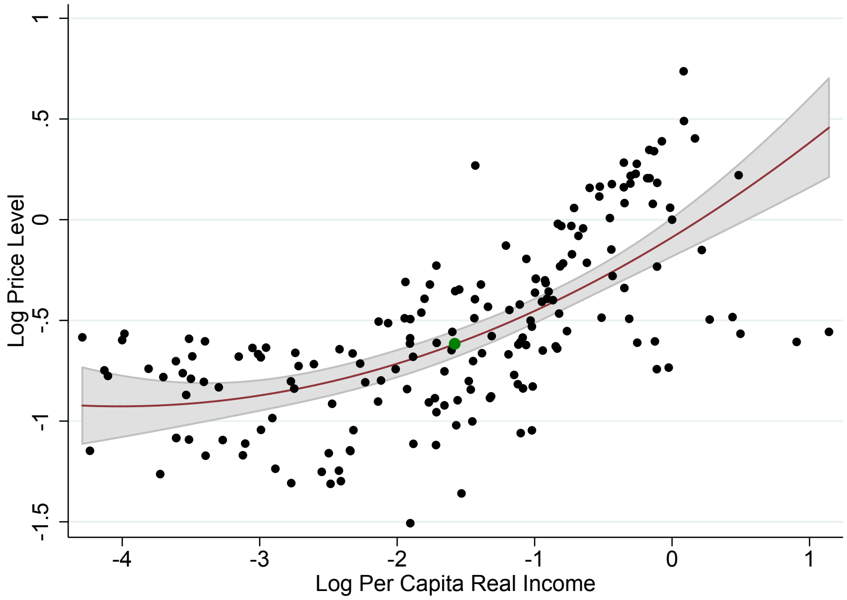


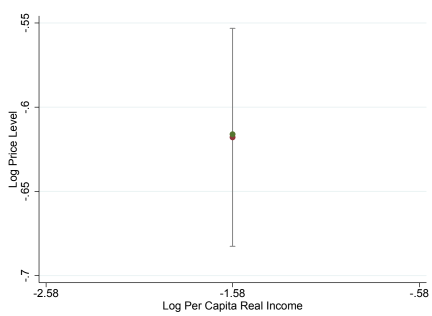
Fig. 2.17 China 2005

Fig. 2.18 Residuals 2005

Note: Figure 2.16 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure 2.17 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure 2.18 plots the residuals from the OLS estimation using the quadratic functional form against the regressor.

Fig. 2.19 OLS Estimation Results with PWT 9.0 2011





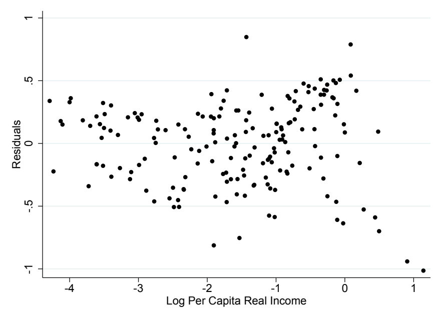
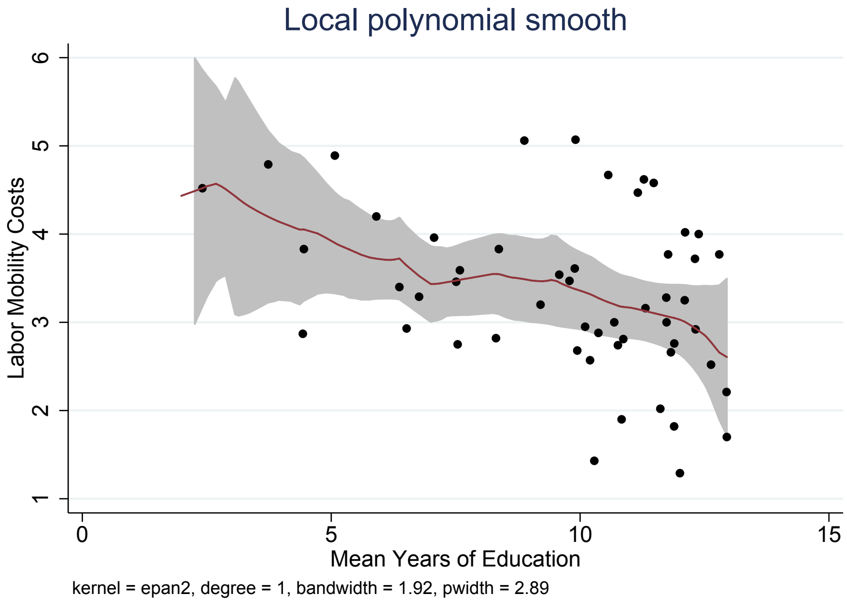


Fig. 2.20 China 2011

Fig. 2.21 Residuals 2011

Note: Figure 2.19 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure 2.20 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure 2.21 plots the residuals from the OLS estimation using the quadratic functional form against the regressor.

Fig. 2.22 Labor Mobility Cost and Mean Years of Education



Note: The figure displays the relationship between the labor mobility costs and the mean years of education for 2011. The data is plotted for 53 countries that are common to Artuc et al. (2015) and ICP 2011.

Chapter 3
The Cost of Nations

Ingvild Almås, Thomas F. Crossley, Serhat Ugurlu

Abstract Traditional approaches to international price comparisons yield a single set of symmetric and transitive (purchasing power parity) indices. Although these indices are useful measures of aggregate differences in goods prices, in general they lack welfare foundations. In this paper, we discuss how we can create indices that measure a nation's cost of living by using a social welfare function. We develop indices that capture the cost of keeping a country at the same social welfare level at foreign prices. These "Cost of Nations" indices aggregate explicitly from optimizing individuals and allow for heterogeneity and non-homothetic preferences. As with the analogous individual-level Konüs cost of living index, Cost of Nations indices are not, in general, symmetric and transitive. However, we propose two methods of imposing symmetry and transitivity. In contrast to existing approaches, our approach allows us to characterize the size of the required adjustments to a social-welfare-constant index. Conversely, we can document the deviations from symmetry and transitivity required to have a social-welfare-constant Cost of Nations index. We apply our ideas to micro data from the U.S., the U.K and Spain.

Key words: price index, international comparisons, purchasing power parity, individual optimization, heterogeneity, non-homotheticity

3.1 Introduction

People are different. Individuals make different consumption choices not only because they have different levels of income, but also because they have different tastes. As a result, price differences impact people differently. Yet, when we make cross country comparisons, and aim to correct for price differences across countries, we usually employ a single scalar price index. Such indices are widely used in economic research and for policy analysis, they are, e.g., used to deflate nominal values for a country, such as GDP or total national consumption, and to deflate individual or household nominal incomes.

Many such indices can be shown to measure *goods* prices in a desirable way (i.e., they may have good axiomatic price index properties, such as transitivity and symmetry). However, if we want price indices to represent the welfare effects of price differences, standard indices have limitations. On the other hand, the economic approach to price indices aim to measure cost of living for individuals that minimize costs of attaining a certain level of utility (see, e.g., Rao, 2016). But, when people are different, no single price index can represent cost of living for all individuals separately, and we face a challenge as to how to aggregate individual cost of living to represent a nation's cost (Prais, 1959; Samuelson and Swamy, 1974; Nicholson, 1975).

This paper develops the concept of welfare-based "Cost of Nations" indices that are constructed to measure the cost of keeping the nation at the same social welfare level at domestic and foreign prices, respectively. However, holding social welfare constant does not, alone, pin down the cost of a nation, as there are many different ways one could reach this social welfare level. The direct analogue to the individual case (the Konüs cost-of-living index) would be to have nations minimize cost given social welfare and prices. But as we consider the nation's objectives and instruments to be different than those of individuals, this is not our preferred index. Instead we consider three alternative approaches to index construction (all of which keep social welfare constant): i. one that preserves observed distributions of individual quantities, ii. one that preserves relative individual incomes, and iii. one that preserves individual welfare.

We first construct bilateral, constant-social-welfare Cost of Nations indices that aggregate explicitly from individual cost-minimizing agents and allow for heterogeneity and non-homothetic preferences. We then study their index number properties. As with the analogous individual-level Konüs cost of living indices, these Cost of Nations indices are not, in general, symmetric and transitive. However, we propose two methods of imposing symmetry and transitivity on a set of Cost of Nations indices. One of these holds global welfare constant.

In contrast to the existing literature on international comparisons, our approach allows us to characterize, for particular sets of price vectors, the size of the adjustments to a social-welfare-constant index that are required to achieve symmetry and transitivity. Conversely, we can document the deviations from symmetry and transitivity required to have a social-welfare-constant Cost of Nations index. We apply our ideas to prices and micro consumption data from the U.S., the U.K. and Spain, and show that our social-welfare-constant Cost of Nations indices require only modest adjustments to satisfy symmetry and transitivity. Moreover, the Cost-of-Nations indices can be much more representative of the center of the underlying distribution of individual-constant-welfare (Konüs) indices than are standard multilateral indices.

We compare our welfare consistent indices to a standard multilateral index: the Gini-Éltetö-Köves-Szulc method (GEKS), used by the World Bank and Eurostat and also in recent versions of the Penn World Table (World Bank, 2014; Feenstra et al., 2015a). Another standard multilateral index is the Geary-Khamis method

(GK) (Geary, 1958). Both of these indices can be shown to have attractive axiomatic index properties, i.e., they satisfy axioms such as transitivity and symmetry. The GK also satisfies adding-up which is useful for analyses of national economic production and, e.g., decomposition of gross national product or income (GDP and GNI). However, these indices only capture welfare effects of price changes under very restrictive assumptions. GK represents welfare costs only if preferences are Leontief and homogeneous GEKS represents welfare costs only if preferences can be represented by a homogeneous quadratic utility function. In this case, it is equal to a Fisher index (Neary, 2004). In addition to the standard multilateral indices, we also make reference to another multilateral economic index, namely the Geary-Allen International Accounts (GAIA) (Neary, 2004). GAIA is constructed to be a welfare based measure for international comparisons, but even so, it only represents welfare costs of price differences as long as preferences are identical within and across countries and there are no income differences within countries (or alternatively, preferences are homothetic).

The idea of holding social welfare constant in constructing a price index has been suggested in the literature on temporal price indices (Pollak, 1980; Crossley and Pendakur, 2010). However, the problem of spatial comparisons brings particular challenges. First, international comparisons naturally suggest an important additional dimension of preference heterogeneity, across countries. Second, symmetry and transitivity are essential in the spatial context and consequently these properties have played a central role in the literature on international price comparisons (see, e.g., Diewert and Nakamura, 1993). Thus, one novel contribution of this paper is to combine the emphasis on symmetry and transitivity from the literature on multilateral indices with individual optimization, heterogeneity, and rich preferences.

The rest of the paper proceeds as follows. Section 2 describes our theoretical framework, discusses imposing symmetry and transitivity, and outlines our theoretical results. Section 3 provides an empirical illustration of our bilateral indices and evaluates deviations caused by imposing symmetry and transitivity. Section 4 discusses the implications of our results. Section 5 concludes.

3.2 Theoretical Framework

Let $\mathbf{p}_h = \left(p_h^1, \ldots, p_h^G\right)^T$ be a vector of prices in a home society h with elements p_h^g, where g is an index of consumption commodities. Define an individual i_h from society h as an optimizing agent with a locally non-satiated, convex, and continuous preference relation $\gtrsim_{i_h}$. Given a home society budget level $m_h^{i_h}$, the optimizing individual solves a utility maximization problem

$$\max_{\mathbf{q}^{i_h} > 0} u^{i_h}(\mathbf{q}^{i_h})$$
$$s.t.\ (\mathbf{p}_h)^T \mathbf{q}^{i_h} \leq m_h^{i_h}$$

to choose a quantity vector $\mathbf{q}^{i_h} = \left(q^{1,i^h}, \ldots, q^{G,i^h}\right)^T$ and obtain utility level u^{i_h}. Let $\mathbf{q}_h^{i_h}$ be the solution to the individual's utility optimization problem at home society prices and budget level, and $u_h^{i_h}$ be the corresponding utility level.

Let $e^{i_h}(\mathbf{p}, u^{i_h})$ be an expenditure function, which yields the cost of the individual's optimal bundle $\mathbf{q}^{i_h}$ in any price vector $\mathbf{p}$ and utility level u^{i_h}. If we desired an index that measures the cost of price differences from $\mathbf{p}_h$ to $\mathbf{p}_f$ for an individual, a natural index would be the Konüs cost of living index (Konüs, 1939), which captures the ratio of the individual's minimum cost of obtaining a reference level of utility $(u_r^{i_h})$ at $\mathbf{p}_h$ and $\mathbf{p}_f$.[1]

$$\pi_{h,f}^{i_h,K} = \frac{e^{i_h}(\mathbf{p}_f, u_r^{i_h})}{e^{i_h}(\mathbf{p}_h, u_r^{i_h})}$$

Define a society as a collection of individuals, for example, as a nation for international comparisons. Let $\mathbf{m}_h = \left(m_h^{1_h}, \ldots, m_h^{I_h}\right)^T$ be an actual (observed) budget distribution for society h and $M_h^h = \sum_{i_h=1}^{I^h} m_h^{i_h}$ be the aggregate, or social, cost that the society incurs at home prices. We are interested in obtaining a scalar price index $\Pi_{h,f}$ such that $M_f^h = \Pi_{h,f} M_h^h$ is the cost of a society h at the price vector of society f to attain the social welfare level it attained at observed prices and budgets.

If preferences are homogeneous and homothetic, then, expenditure function of an individual can be decomposed as $e^{i_h}(\mathbf{p}, u^{i_h}) = \alpha(\mathbf{p})u^{i_h}$ for a homogeneous function $\alpha(\cdot)$ of prices. In this case, Konüs price indices for all individuals

$$\pi_{h,f}^{i_h,K} = \frac{\alpha\left(\mathbf{p}^f\right)}{\alpha\left(\mathbf{p}^h\right)}$$

become equivalent to each other. Thus, $\pi_{h,f}^{i_h,K} = \pi_{h,f}^K$ becomes a suitable measure to calculate the cost of not just an individual but a society at the price vector $\mathbf{p}_f$, i.e., $M_f^h = \pi_{h,f}^K M_h^{h,}$. If preferences are heterogeneous or non-homothetic, $\pi_{h,f}^{i_h,K}$ differ across individuals or income levels. Therefore, it is no longer a suitable measure for aggregate level comparisons.

Let $W^h\left(u^{1^h}, \ldots, u^{I^h}\right)$ be a continuous, monotonic, and anonymous Bergson-Samuelson social welfare function that is used to evaluate the social welfare level of country h at observed prices and budget distribution. By analogy to Konüs price index, an aggregate price index $\Pi_{h,f}$ can be defined as a constant-social-welfare price index. For an optimizing society, this analogy leads to the Social Cost-of-Living Index (SCOLI) (Pollak, 1980)

$$\Pi_{h,f}^{\text{SCOLI}} = \frac{E^h\left(\mathbf{p}^f, U^h\right)}{E^h\left(\mathbf{p}^h, U^h\right)},$$

[1] This index will be symmetric and transitive if preferences are homothetic, or if we use an arbitrary reference utility level rather than the "home" $(u_h^{i_h})$ or "foreign" $(u_f^{i_h})$ utility level.

where $E(\cdot)$ is a social expenditure function that calculates the minimum total expenditure required to attain the evaluated social welfare level $U^h = W^h\left(u^{1^h}, \ldots, u^{I^h}\right)$ at home and foreign prices.

For a non-optimizing society, alternative ways to pin down expenditure levels that attain the same social indifference curve is required. Below, we present three thought-experiments that present such alternatives. Within an international comparisons context, our thought-experiments yield answers to the following questions: What is the cost of a nation if all individuals are allocated costs of their actual bundles (aggregate Laspeyres), if all individuals attain the same individual utility levels (constant utility), and if relative budget distributions are the same (CS-SCOLI) in a foreign nation's prices?

3.2.1 Asymmetric Bilateral Comparisons

Even though we characterize our price indices using thought-experiments with optimizing individuals, for meaningful purchasing power parity interpretations, it is important for the indices to satisfy some structural properties. We evaluate our indices with respect to the following axioms discussed by Eichhorn and Voeller (1976), Eichhorn (1978), and Balk (2008):

Axiom 1 *$\Pi_{h,f}$ is increasing in $\mathbf{p}_f$ and decreasing in $\mathbf{p}_h$.*

Axiom 2 *$\Pi_{h,f}$ is homogeneous of degree one in foreign prices.*

Axiom 3 *$\Pi_{h,f} = 1$ if $\mathbf{p}_h = \mathbf{p}_f$.*

Axiom 4 *$\Pi_{h,f}$ is homogeneous of degree zero in both prices.*

Axiom 5 *$\Pi_{h,f}$ is invariant to changes in the units of measurement of commodities.*

All indices that we discuss satisfy these axioms. Proofs are available in appendix C.2.

Aggregate Laspeyres Index (Π^L)

In this thought-experiment, we are interested in a social cost level that a society attains when all individuals have cost levels equal to the cost of their observed quantities evaluated at new prices. We are interested in this case for a few reasons. First, as we elaborate below, it is directly observed from data without any knowledge of preferences and social welfare functions, which has practical importance. Second, the observed quantity vectors are also optimal at new prices for preferences that do not imply substitution towards cheaper products.

Since prices $\mathbf{p}_f$ and home society quantities $\mathbf{q}_h^{i_h}$ are observed, cost that an individual incurs at new prices $\mathbf{p}_f$ to be able to afford $\mathbf{q}_h^{i_h}$ is calculated as $m_f^{i_h} = \mathbf{p}_f^T \mathbf{q}_h^{i_h}$.

Therefore the social cost is $M_f^h = \mathbf{p}_f^T \sum_{i_h=1}^{I_h} \mathbf{q}_h^{i_h}$ and the aggregate price index is

$$\Pi_{h,f}^L = \frac{\mathbf{p}_f^T \sum_{i_h=1}^{I_h} \mathbf{q}_h^{i_h}}{M_h^h}, \tag{3.1}$$

where $M_h^h = \mathbf{p}_h^T \sum_{i_h=1}^{I_h} \mathbf{q}_h^{i_h}$ and $\mathbf{p}_h^T$ is the transpose of the price vector $\mathbf{p}_h$. Therefore, $\Pi_{h,f}^L$ is a Laspeyres index with aggregate quantities.[2]

When optimizing individuals are allocated a budget level that is equivalent to the value of their home society quantities evaluated at new prices, they can always afford their home society quantities and individual utility levels. Therefore, individuals are never worse off at new prices. However, they may also do better in terms of individual welfare by substituting more expensive products for cheaper products at new prices. Since this is true for all individuals, this thought-experiment corresponds to a possibly higher social welfare level at new prices except when there is no substitution. Therefore, $\Pi_{h,f}^L$ is a constant-social-welfare index only with a strict restriction on individuals' preferences. We return to this case in the next thought-experiment.

Constant Utility Index (Π^{CU})

In this thought-experiment, we are interested in a social cost level when all individuals attain their actual utility levels $u_h^{i_h}$ with minimum individual costs. This thought-experiment yields a constant-social-welfare index because inputs to the Bergson-Samuelson social welfare function are kept constant at home and foreign prices.

Let $v^{i_h}(\mathbf{p}, m)$ be an indirect utility function that represents an individual's preferences $\gtrsim_{i_h}$. For each individual i_h facing prices $\mathbf{p}_h$ with a home country budget $m_h^{i_h}$, we look for a budget level in foreign prices, $m_f^{i_h}$, such that $v^{i_h}(\mathbf{p}_f, m_f^{i_h}) = v^{i_h}(\mathbf{p}_h, m_h^{i_h})$. This budget level

$$m_f^{i_h} = e^{i_h}\left(\mathbf{p}_f, v^{i_h}(\mathbf{p}_h, m_h^{i_h})\right)$$

is also referred as an equivalent income, or money-metric utility. Use of equivalent income measures was previously suggested in this context by, e.g., Samuelson and Swamy (1974); Samuelson (1974); Fleurbaey and Gaulier (2009).

If all individuals are allocated their equivalent incomes in foreign prices, then,

$$M_f^h = \sum_{i_h=1}^{I_h} e^{i_h}\left(\mathbf{p}_f, v^{i_h}(\mathbf{p}_h, m_h^{i_h})\right)$$

yields the social cost of society h at prices of society f; and a constant utility index is

[2] Social cost at foreign prices, M_f^h, is equivalent to sum of individual costs that are calculated with individual Laspeyres indices $\pi_{h,f}^{i_h,L} = (\mathbf{p}_f)^T \mathbf{q}_h^{i_h} \Big/ (\mathbf{p}_h)^T \mathbf{q}_h^{i_h}$.

$$\Pi^{CU}_{h,f} = \frac{\sum_{i_h=1}^{I_h} e^{i_h}\left(\mathbf{p}_f, v^{i_h}(\mathbf{p}_h, m_h^{i_h})\right)}{M_h^h}.$$

$\Pi^{CU}_{h,f}$ can also be defined in terms of compensating variations. By definition, compensating variation of individual i_h is

$$CV^{i_h} = e^{i_h}\left(\mathbf{p}_f, v^{i_h}(\mathbf{p}_h, m_h^{i_h})\right) - e^{i_h}\left(\mathbf{p}_h, v^{i_h}(\mathbf{p}_h, m_h^{i_h})\right),$$

which implies

$$\Pi^{CU}_{h,f} = 1 + \frac{\sum_{i_h=1}^{I_h} CV^{i_h}}{M_h^h} \tag{3.2}$$

is a constant-social-welfare price index that is consistent with a thought experiment of constant utilities.

In order to obtain $\Pi^{CU}_{h,f}$, information on preferences is required so that individuals' compensating variations can be calculated. However, information on the shape of the social welfare function is not required. When all individuals attain the same indifference curves, inputs to the Bergson-Samuelson social welfare function are constant. Hence, the social welfare in home prices and foreign prices are equal for any evaluation function.

Similar to the relationship between individual Laspeyres indices and $\Pi^{L}_{h,f}$, $\Pi^{CU}_{h,f}$ is an index that analyses the social cost level when all individuals have budgets at a new price vector that is consistent with their individual Konüs price indices $\pi^{i_h,K}_{h,f}$. An equivalent definition of $\Pi^{CU}_{h,f}$ is

$$\Pi^{CU}_{h,f} = \frac{\sum_{i_h=1}^{I_h} \pi^{i_h,K}_{h,f} m_h^{i_h}}{M_h^h}.$$

There exists a well-known relationship between individual Konüs price indices and individual Laspeyres indices when individuals substitute more expensive products with cheaper products.

Lemma 1. *If at new prices, an individual incurs a cost equal to the cost of the home quantities evaluated at new prices, then the individual consumes different quantities, and necessarily enjoys a higher utility level.*

Proof. See Konüs (1939).

This relationship between individual Laspeyres and Konüs price indices may also be extended for Leontief preferences, in which case the utility level at foreign prices would be equal to the utility level at home prices. Thus, cost level at new prices an individual attains with $\pi^{i_h,L}_{h,f}$ never yields a lower standard of living for the individual than cost level at new prices with $\pi^{i_h,K}_{h,f}$, which leads to the theorem below.

Theorem 1. *$\Pi^{L}_{h,f}$ is an upper bound for $\Pi^{CU}_{h,f}$, i.e. $\Pi^{L}_{h,f} \geq \Pi^{CU}_{h,f}$.*

Proof. See appendix.

If one assumes a cardinal comparability of interpersonal utility levels,[3] then, $\Pi^{CU}_{h,f}$ is a comparison of two social cost levels that respect the existing distribution of individual welfare levels in a society.

Common Scaling Index (Π^{CS})

In this thought-experiment, we are interested in a social cost level when relative budgets across individuals is constant at home and foreign prices. Keeping relative budget levels across individuals constant in two price vectors preserves the place of each individual along the actual individual budget distribution. In a temporal context, this case is described in detail by Crossley and Pendakur (2010) as the common-scaling social cost-of-living) index (CS-SCOLI) (see also Donaldson and Pendakur, 2012). We analyze common-scaling in budget levels within our framework of spatial price comparisons across different societies. The primary difference between temporal and spatial cases comes from the additional layer of heterogeneity inherent in spatial comparisons; not only preferences within societies but also preferences across societies are different. Hence, Π^{CS} is not symmetric (or transitive) because of changing reference preferences of each bilateral comparison.

Let $\mathbf{m}^h_h = \left(m^{1_h}_h, \ldots, m^{I_h}_h\right)$ be the actual budget distribution for society h. A budget distribution in foreign prices that respects the actual relative budgets across individuals is a positive common scaling μ to all individual budget levels: $\mathbf{m}^h_f = \mu \mathbf{m}^h_h$.

Let $B^h(\mathbf{p}_h, \mathbf{m}^h_h)$ be an indirect social welfare function for society h defined as

$$B^h(\mathbf{p}_h, \mathbf{m}^h_h) = W^h\left(u^{1_h}(\mathbf{q}^{1_h}_h), \ldots, u^{I_h}(\mathbf{q}^{I_h}_h)\right)$$

and yield the social welfare level at a price vector and a vector of individual budget levels. Then a constant-social-welfare constant-relative-budget index $\Pi^{CS}_{h,f}$ is defined as the solution to the following equality:

$$B^h(\mathbf{p}_h, \mathbf{m}^h_h) = B^h(\mathbf{p}_f, \Pi^{CS}_{h,f}\mathbf{m}^h_h), \tag{3.3}$$

and the social cost that respects the actual distribution of relative budget levels is $M^h_f = \Pi^{CS}_{h,f} M^h_h$.

Constant utility and common scaling price indices for a society have a constant-social-welfare interpretation by their definitions. However, they refer to different budget distributions to attain the actual social welfare level. When we consider constant utility levels at new prices, individuals who are the most adversely (favorably) affected by price changes are allocated more (less) budgets at new prices compared to the other individuals. Therefore, some individuals would become relatively richer (poorer) at foreign prices compared to their relative budgets at home prices,

[3] In other words, if statements such as individual i_h has a higher welfare level than individual j_h are meaningful.

and would experience an upward (downward) shift along the income distribution. On the contrary, constant scaling of budget levels imply a proportionate adjustment to everybody's budget levels, and has no effect on placement of individuals along the income distribution. However, proportionate budget adjustments do not imply proportionate welfare adjustment. Hence, at scaled budgets, some individuals may have lower (higher) welfare levels at foreign prices than they have at home prices. Therefore, even though both thought experiments specify social costs that maintain a standard of living for a society, they have different interpretations in terms of how this standard of living is achieved. The only case when constant utility and common scaling indices coincide, or when they refer to the same budget distribution at foreign prices, is characterized in the following theorem:

Theorem 2. $\Pi^{CS}_{h,f} = \Pi^{CU}_{h,f}$ *if preferences are homogeneous and homothetic.*

Proof. See appendix.

Therefore, when preferences are not identical, among all possible social cost levels that are required to attain a social welfare level, none are consistent with thought experiments that respect the existing welfare and income inequalities simultaneously.

Evaluating Symmetry of Π^L, Π^{CU}, Π^{CS}

In a bilateral comparison between a society h and a society f, there are two purchasing power parities that one can consider: a purchasing power parity to compare prices at country h to prices at country f, $\Pi_{h,f}$, and another purchasing power parity to compare prices at country f to prices at country h, $\Pi_{f,h}$. These indices are symmetric if $\Pi_{h,f} = 1\big/\Pi_{f,h}$ holds.

Conceptually, price indices that are consistent with thought experiments with heterogeneous individuals need not be symmetric. $\Pi_{h,f}$ is a purchasing power parity that is consistent with different tastes of individuals forming society h whereas $\Pi_{f,h}$ is a purchasing power parity that is consistent with different tastes of individuals forming society f. Therefore, even if level differences between $\mathbf{p}_h$ and $\mathbf{p}_f$ can be summarized with a symmetric measure, impacts of these differences vary across societies.

Consequently, social purchasing power parities that we characterize are from the viewpoint of a single home society h and takes only the preferences of a home society into account. Thus, Π^L, Π^{CU}, and Π^{CS} are base dependent and asymmetric.[4]

A symmetric treatment of indices may be desired for practical purposes. Symmetric indices are base society independent, and in this sense, they provide a unique comparison of prices of different societies: the price index does not depend on which

[4] $\Pi^L_{f,h}$ is symmetric if $\mathbf{Q}^f = c\mathbf{Q}^h$ where $\mathbf{Q}^f$ and $\mathbf{Q}^h$ are vectors of aggregate quantities in foreign and home countries and $c > 0$ is a scalar. However, equality of relative aggregate quantities of two societies is a very unrealistic case.

society is chosen as the base society. However, since none of our indices are symmetric when preferences are non-identical within and across borders, a symmetric index is necessarily inconsistent with constant-social-welfare thought experiments.

3.2.2 Symmetric Bilateral Comparisons

There are alternative ways to obtain symmetric indices taking asymmetric indices as base indices. We focus on geometric averaging (see, e.g., Balk, 2008). Laspeyres price index, which evaluates cost of home society quantities at home prices and foreign prices, and Paasche price index, which evaluates cost of foreign society quantities at home prices and foreign prices, are important examples of asymmetric indices. Their geometric average, the Fisher ideal index, is a very common symmetric price index. Using the same approach, we obtain

$$s_{h,f} = \sqrt{\Pi_{h,f}\left(1/\Pi_{f,h}\right)}$$

as a symmetric price index where $\Pi_{h,f}$ is one of our asymmetric indices.

Asymmetric social purchasing power parities have a constant social welfare interpretation[5], and are consistent with thought-experiments with individual-level optimization at a foreign price vector. On the other hand, the symmetric index is an average of asymmetric indices. Thus, the symmetric index does not have a constant social welfare interpretation at the thought-experiment with which a base asymmetric index is consistent. In other words, social cost level of a society h at a foreign price vector $\mathbf{p}_f$, for example $M_f^h = s_{h,f} M_h^h$, does not necessarily attain the actual social welfare level. To see this, without loss of generality, assume $\Pi_{h,f} > 1/\Pi_{f,h}$. Then $s_{h,f}$ is less than $\Pi_{h,f}$, which is a constant-social-welfare price index. In this case, social cost that a symmetric index suggests will be less than a social cost that an asymmetric index suggests,which attains the actual social welfare level. The opposite is true if $\Pi_{h,f} < 1/\Pi_{f,h}$.

For any other methodology that could be used to impose symmetry on bilateral price comparisons in terms of averaging, it is also easy to see that symmetric price indices are different than asymmetric indices by definition. Therefore, imposing symmetry in bilateral price comparisons introduces ambiguity in the interpretation of a price index even if asymmetric price indices that form a basis for symmetry have an explicit cost of living interpretation. Cost of a society is over, or under, allocated depending on the relationship between $\Pi_{h,f}$ and $\Pi_{f,h}$. Asymmetric purchasing power parities provide benchmarks to evaluate this ambiguity.

Nevertheless, it is possible to use asymmetric indices and build a symmetric index with a different welfare interpretation than a constant-social-welfare interpretation on which we base our asymmetric indices. Since this case also yields a multilateral index, we describe it in the following section.

[5] Π_L has this interpretation only if preferences are Leontief.

3.2.3 Multilateral Comparisons

Multilateral international comparisons are comparisons of more than two countries at a time ($C \geq 3$) with a transitive system of purchasing power parities. A purchasing power parity index S is transitive if $S_{h,f}S_{f,k} = S_{h,k}$ for all societies h, f, and k.

None of the bilateral indices that we characterize satisfies transitivity due to changing reference preferences of each comparison. Hence, they do not form a system of transitive purchasing power parities that is suitable for multilateral comparisons. In such cases, multilateral indices are obtained with aggregations of the bilateral indices, for example with a Gini-Éltetö-Köves-Szulc (GEKS) aggregation, which is a geometric average of all possible symmetric bilateral comparisons (World Bank, 2013).[6] In our case,

$$S_{h,f}^{GEKS} = \left(\prod_{c=1}^{C} \Pi_{h,c} \Pi_{c,f} \right)^{1/C}$$

is obtained as a multilateral index, which is based on the symmetric bilateral index $s_{h,f}$.

Similar to the Fisher aggregation, GEKS aggregation of our asymmetric indices are not consistent with our thought-experiments. Hence, a multilateral index S^{GEKS} does not have a constant social welfare interpretation; social cost level associated with the price index $S_{h,f}^{GEKS}$ may not attain the actual social welfare level. Therefore, imposing symmetry or transitivity in terms of averaging introduces practical convenience for uses of price indices at a cost: the explicit interpretation of asymmetric price indices is lost.

We turn our focus to the following question: is it possible to form another thought-experiment that is consistent with our asymmetric comparisons, and yields a multilateral index with an explicit cost of living interpretation? Imagine that countries h, f, and k have observed social welfare levels U^h, U^f, and U^k at their own price vectors $\mathbf{p}^h, \mathbf{p}^f$, and $\mathbf{p}^k$ with actual social cost levels M_h^h, M_f^f, and M_k^k. If we know asymmetric price indices, we can calculate three global cost levels at each price vector: $\sum_{i\in\{h,f,k\}} \Pi_{i,h} M_i^i$, $\sum_{i\in\{h,f,k\}} \Pi_{i,f} M_i^i$, $\sum_{i\in\{h,f,k\}} \Pi_{i,k} M_i^i$.[7] At each of these global cost levels, because all societies attain their actual social welfare levels, we can consider a constant global welfare interpretation. Hence, a comparison of these global cost levels is a comparison of global cost levels that attain the same global welfare level.

---

[6] A set of multilateral price indices is also possible to obtain as a solution to systems of equations that equate observed global expenditures to a world expenditure at hypothetical world prices. Geary-Khamis (GK) (Geary, 1958), GAIA is also a GK type approach. These methods do not rely on pre-specified bilateral price indices, and directly yield a set of multilateral price indices. For other multilateral approaches, see also Balk (2008).

[7] Each converted social cost level $M_j^i = \Pi_{i,j} M_i^i$ is a constant-social-welfare social cost level in prices of society j.

More generally, let $\{1,2,\dots,C\}$ be an index set of societies in a multilateral comparison. Let $W^0(U^1,\dots,U^C)$ be a global welfare function that is used to evaluate a global welfare level, which is defined over social welfare levels of C countries. Let $\gamma(\mathbf{p}_h)=\sum_{c=1}^{C}\Pi_{c,h}M_c^c$, $\Pi_{c,h}\in\{\Pi_{c,h}^{L},\Pi_{c,h}^{CU},\Pi_{c,h}^{CS}\}$ be the global cost level at prices of country h. At each converted sum of social cost levels $\gamma(\mathbf{p}_c)$, $c\in[1,C]$, the same global welfare level is obtained because all countries attain their observed social welfare levels. Then, we define a multilateral price index S^{CGW} as the ratio of global costs at prices of country f to global cost in prices of country h:

$$S_{h,f}^{CGW}=\frac{\gamma(\mathbf{p}_f)}{\gamma(\mathbf{p}_h)}.$$

Proposition 1. *S^{CGW} satisfies symmetry and transitivity.*

Proof. See appendix.

The first step in calculating this index is to obtain the asymmetric indices that are consistent with our bilateral thought-experiments. Then, global cost levels in the price vector of each society are obtained using social cost levels of each society in that price vector. Since social costs are calculated using consistent asymmetric indices, interpretation of the multilateral index depends on which asymmetric index is utilized.

Similar to a constant utility index, the constant global social welfare index does not depend on the shape of the social welfare function since all inputs to the function are constant. However, a constant global welfare interpretation implicitly requires assuming existence of a global indifference curve; different relative social costs of societies may yield the same global welfare. Our index compares ratios of global costs with social cost levels that respect an observed distribution with respect to quantities, individual welfare levels, or actual relative costs within each society.

3.3 Empirical Illustrations

In this section, we calculate asymmetric, symmetric, and multilateral indices. To this aim, we use comparable consumer expenditure surveys from Spain, the UK, and the US.[8] Table 3.1 provides a list of the surveys that we use.

From these surveys, we obtain information on household-level expenditures in seven broad consumption categories: food and non-alcoholic beverages, alcoholic beverages and tobacco, clothing and footwear, health, transportation, communication, and miscellaneous expenditures. In addition, we obtain information on household demographics, sample weights, and time of the reported transactions as month-year.

[8] Surveys that we use in this research collect household-level expenditure information by adhering to COICOP guidelines of the United Nations Statistics Division.

We obtain price data from ICP 2011, which reports 2011 average prices (normalized to the US prices) for each consumption category that we use. Prices are extrapolated monthly using category-time-location specific consumer price indices (CPI) from statistical institutes of each country. We match extrapolated prices with households using information on location and time of the reported transactions.

To calculate our price indices, we only use the 2011 subsamples of our datasets with ICP 2011 average prices (without extrapolation).

Asymmetric Bilateral Indices

The aggregate Laspeyres index Π^L is calculated from observed quantity and price information.

In order to calculate the constant utility index Π^{CU} and the common scaling index Π^{CS}, we estimate demand functions for households in each country by using an almost ideal demand system (AIDS), an iterated least squares estimator, and imposing homogeneity, adding up, and symmetry (Deaton and Muellbauer, 1980a; Blundell and Robin, 1999, see also appendix A.2). We use number of children, number of adults, and number of female in a household as observable heterogeneity in demand patterns. We allow these demographic variables as taste shifters (Abramovsky et al., 2015). Then, Π^{CU} is calculated by evaluating compensating variations of all individuals with estimated preference parameters.

To calculate Π^{CS}, in addition to estimated preferences, we also use a utilitarian social welfare function to evaluate the actual social welfare.[9] With a utilitarian social welfare function, the actual welfare is evaluated as $U^h(u^{1_h},\ldots) = \sum_{i_h=1}^{I_h} V(\mathbf{p}_h, m_h^{i_h}, \mathbf{z}^{i_h})$, where $V(\mathbf{p}_h, m_h^{i_h}, \mathbf{z}^{i_h})$ is an indirect utility function for household i_h facing prices $\mathbf{p}_h$ with budget $m_h^{i_h}$ and observables $\mathbf{z}^{i_h}$. Then, using AIDS indirect utility function and a utilitarian social welfare function, we obtain the following equality

$$\sum_{i_h=1}^{I_h} \frac{\log m_h^{i_h} - \log a^{i_h}(\mathbf{p}_h)}{\log b^{i_h}(\mathbf{p}_h) - \log a^{i_h}(\mathbf{p}_h)} = \sum_{i_h=1}^{I_h} \frac{\log \Pi_{h,f}^{CS} m_h^{i^h} - \log a^{i_h}(\mathbf{p}_f)}{\log b^{i_h}(\mathbf{p}_f) - \log a^{i_h}(\mathbf{p}_f)}$$

where a^{i_h} and b^{i_h} are price aggregators with preference parameters of household i_h. From this equivalence, an explicit formula for the common scaling index

[9] Π^{CS} can be calculated with other types of social welfare functions as well. However, assuming cardinal unit comparability, the utilitarian social welfare function is the only social welfare function that yields a social ranking that is invariant to scaling of utility levels (d'Aspermont and Gevers, 1977).

$$\Pi_{h,f}^{CS} = \exp\left\{ \frac{\sum_{i_h=1}^{I^h} \left[\frac{\log m_h^{i_h} - \log a^{i_h}(\mathbf{p}_h)}{\log b^{i_h}(\mathbf{p}_h) - \log a^{i_h}(\mathbf{p}_h)} - \frac{\log m_h^{i_h} - \log a^{i_h}(\mathbf{p}_f)}{\log b^{i_h}(\mathbf{p}_f) - \log a^{i_h}(\mathbf{p}_f)} \right]}{\sum_{i=1}^{I^h} \left[\log b^{i_h}(\mathbf{p}_f) - \log a^{i_h}(\mathbf{p}_f) \right]^{-1}} \right\}$$

is obtained.

Table 3.2 reports all asymmetric price indices.

Symmetric Bilateral Indices

For any bilateral comparison with an asymmetric index $\Pi_{h,f} \in \left\{ \Pi_{h,f}^{L}, \Pi_{h,f}^{CU}, \Pi_{h,f}^{CS} \right\}$, we calculate a symmetric index $s_{h,f}$ as $s_{h,f} = \sqrt{\Pi_{h,f} / \Pi_{f,h}}$. Table 3.3 reports these geometric averages as symmetric indices.

Multilateral Indices

Using the asymmetric bilateral indices, we calculate a multilateral price index in two ways. First, we use symmetric indices $s_{h,f}$ that are based on asymmetric indices and apply GEKS formula to obtain a transitive set of index numbers for multilateral comparisons S^{GEKS}.

Second, we calculate a constant-global-welfare (CGW) index. To this aim, we calculate costs of societies at all possible actual price vectors using asymmetric indices, and obtain aggregate global cost levels that attains the actual global welfare level. Because all societies attain their actual social welfare levels at cost levels that are evaluated at asymmetric indices, each global cost level is a global cost at a price vector that attains the actual global welfare level. Using these cost levels, we obtain multilateral indices S^{CGW}.

Table 3.4 reports transitive index numbers taking the UK as the base country.

3.4 Discussion

The results illustrate how we obtain price indices that are consistent with our thought-experiments by using micro-level data from comparable consumer surveys. Table 3.5 summarizes our results taking the UK as the home (or the base for multilateral indices) country.

An implication of our approach, which aggregates individual cost levels to obtain social cost levels with proportionate (CS) or non-proportionate (L or CU) budget adjustments to all individuals in a society, is being able to compare not just average cost levels across nations but whole budget distributions at each country's price vector. The first plot in figure 3.1 illustrates budget distributions of each society in their own prices. Of course, since these budgets are nominal values in prices of different countries, in our case in GBP, EUR and USD, they are incomparable. The other plots illustrate budget distributions that are denominated in GBP using purchasing power parities from our thought-experiments.

A comparison of distributions instead of mean values provides a deeper understanding of income comparisons across societies. For example, the aggregate Laspeyres index indicates that, respecting observed quantities at British prices, up to upper middle income earners, British households incur a higher cost to be able afford their consumption bundles than Spanish households. Moreover, only the top income earner US households have higher cost levels than British households. Therefore, conditional on the commodities in our empirical illustration, all asymmetric indices suggest that mean British households are richer than mean American and Spanish households whereas this result switches in favor of Spanish households at the top 25%, and in favor of American households at the top few percentiles of budget distributions.[10]

We discussed that the cost of imposing symmetry and transitivity on asymmetric comparisons due to changing reference preferences is the loss of explicit interpretations and consistencies of price indices. Because common ways to impose these index number properties are averaging, such symmetric and transitive indices can still be interpreted as averages of price differences across societies. Therefore, it is important to analyze deviations between our purchasing power parities and their symmetric/multilateral counterparts.

Figure 3.2 plots symmetric bilateral price indices (with a geometric averaging) along its y-axis, asymmetric bilateral price indices along its x-axis, and a 45-degree line. The plot indicates that imposing symmetry clearly moves some indices away from the line of equality. If we measure the cost of imposing symmetry with a % difference (in absolute value) between a symmetric index and its asymmetric base, we observe that deviations from an asymmetric aggregate Laspeyres index range from 0.9% to 3.7%. For asymmetric constant utility and common scaling indices, deviations range from 0.3% to 3.1% and from 0.1% to 2%, respectively. Economic significance of these deviations may depend on the application. For comparison, deviations between our asymmetric indices and nominal exchange rates range from 3.1% to 20.7%. A comparison with a standard Fisher index, a geometric average of Laspeyres and Paasche indices, shows deviations ranging from 0.2% to 10.8%. These significant differences in ranges indicate that cost of not accounting for individual preferences may be much larger than cost of imposing symmetry on consistent price indices.

Figure 3.3 repeats the same comparison for multilateral indices with the UK as the base country and all asymmetric indices. GEKS aggregated asymmetric indices and constant global welfare (CGW) aggregated asymmetric indices yield similar

[10] Our current empirical application does not consider spending on many important elements such as housing. Therefore, these orderings of budget levels should not be interpreted as a measure of well-being in terms of full consumption.

outputs. Deviations of multilateral indices from their asymmetric bases range from 0.2% to 2.4% for GEKS and 0.06% to 3.8% for CGW. For comparison, deviations between our asymmetric indices and a standard GEKS index that is not consistent with preferences range from 0.2% to 7.1%.

Analyzing costs of price differences on households allow us to identify where consistent asymmetric indices lie on the distribution of household-specific Konüs indices, which are proper measures of impacts of price changes on households. Figures 3.4 to 3.7 illustrate these distributions for different asymmetric and symmetric bilateral comparisons. Figures also include a Fisher index, which does not respect preferences, for comparison. Indices to compare prices at Spain and the UK suggest that deviations between symmetric indices and a Fisher index are small. However, indices to compare prices at the UK and the US suggest significant deviations.

A similar pattern repeats itself when we analyze location of multilateral indices on the distribution of household-specific Konüs indices. Figures 3.8 to 3.11 illustrate these comparisons. These figures also include GEKS price indices[11] and nominal exchange rates for comparison. It is well known that nominal exchange rates are not good measures of price differences and there is a need for purchasing power parities (see, e.g., Rogoff, 1996). Clearly, our results illustrate how far nominal exchange rates could lie from consistent indices and a standard PPP approach yields more relevant results with respect to an underlying distribution of individual Konüs indices even without respecting preferences. Yet, the relationship between our multilateral indices and a GEKS index is similar to the relationship between our asymmetric indices and a Fisher index. Deviations may be small in some comparisons and large in the others.

3.5 Conclusion

There is a large literature on index number approaches to purchasing power parities, which discusses statistical measures of price differentials. In contrast, economic approaches define price indices as ratios of costs that an individual incurs to maintain a living standard at different prices. Yet, these approaches to social comparisons are consistent with economic behavior only if tastes are identical across individuals or if they differ with income levels. However, individuals differ not only in their incomes but also in their other observables and in their tastes. Due to these differences, a single statistical change in prices has non-identical impacts on individuals' living situations.

In this paper, we analyze how one can obtain social measures of price differences by respecting heterogeneity in tastes, thus, by considering non-identical impacts on individuals. We argue, if societies optimize, Pollak's index yields an elegant approach to obtain such a measure. However, without optimizing societies, we argue indices of price differences can be obtained by respecting observable or measurable

[11] GEKS price indices are obtained imposing transitivity on symmetric Fisher indices with GEKS formula.

certain elements of a society that are in relation with an individual's utility maximization: quantities, utilities, and relative budgets. We survey three such methodologies with an explicit focus on individuals: we question what impact does a price difference have on each individual in a society, i.e., costs of individuals. We build up from these costs and obtain bilateral measures that indicate impacts of price differences on nations, i.e., costs of nations.

In discussing these indices, we stress the importance of symmetry and transitivity for spatial comparisons. However, when tastes differ, no such single index could truly reflect impacts of price differences on all individuals. Therefore, we consider costs of imposing symmetry and transitivity on our bilateral comparisons. We measure these costs as deviations from asymmetric bases. Our empirical results suggest costs of imposing symmetry or transitivity may be small for some comparisons but large for the others. Nonetheless, costs of not accounting for preferences to retain index properties, in terms of relevance with respect to an underlying distribution of individual Konüs indices, may be more significant. A conclusive analysis of this issue requires an application that compares a large number of countries. Our illustrations provide evidence that such comparisons are indeed possible.

We also stress the importance of differentiating interpretations of bilateral and multilateral indices and argue that each of our indices are answers to different questions. Our asymmetric indices correspond to social cost levels that attain a measured social welfare level by respecting an existing element of a society. However, imposing symmetry or transitivity with common index number methodologies yields average measures of price differences that do not have such an explicit interpretation. Thus, we develop a multilateral thought-experiment with a constant global welfare interpretation using asymmetric indices as base indices. With this index, we illustrate that the desired properties of symmetry and transitivity are possible to obtain by being consistent with a thought-experiment involving optimizing individuals.

Table 3.1 Data Sets

Country	Survey	Year	Sample Size
Spain	Household Budget Survey	2006-2015	216426
US	Consumer Expenditure Survey	2009-2013	134874
UK	Living Costs and Food Survey	2008-2013	33266

Note: Table reports names and years of data sets used in the empirical illustration.

Table 3.2 Asymmetric Bilateral Indices

	Foreign (f) UK	US	Spain
Home (h)			
	Aggregate Laspeyres $\Pi^{L}_{h,f}$		
UK	1	1.344 (0.744)	1.116 (0.896)
US	0.789 (1.268)	1	0.869 (1.151)
Spain	0.913 (1.095)	1.240 (0.806)	1
	Constant Utility $\Pi^{CU}_{h,f}$		
UK	1	1.326 (0.754)	1.100 (0.909)
US	0.759 (1.317)	1	0.853 (1.172)
Spain	0.902 (1.110)	1.226 (0.816)	1
	Common Scaling $\Pi^{CS}_{h,f}$		
UK	1	1.353 (0.739)	1.109 (0.901)
US	0.742 (1.348)	1	0.843 (1.187)
Spain	0.896 (1.116)	1.238 (0.808)	1

Note: Tables report asymmetric aggregate Laspeyres $\Pi^{L}_{h,f}$, constant utility $\Pi^{CU}_{h,f}$, and common scaling $\Pi^{CS}_{h,f}$ price indices. Each row represents a home society. Each column represents a foreign society. $\Pi_{h,f}$ is interpreted as a constant-social-welfare price index that is consistent with one of the thought-experiments. Values in parentheses are reciprocals of the index values to illustrate deviations from symmetry. As an example $\Pi^{L}_{UK,US} = 1.344$, i.e., $1/\Pi^{L}_{UK,US} = 0.744$. If Π^{L} were symmetric, then $\Pi^{L}_{US,UK} = 1/\Pi^{L}_{UK,US}$ must have held. However, $\Pi^{L}_{US,UK} = 0.789$.

Table 3.3 Symmetric Bilateral Indices

	Foreign (f) UK	US	Spain
Home (h)			
	With Aggregate Laspeyres $s^L_{h,f}$		
UK	1	1.305 (0.766)	1.105 (0.905)
US	0.766 (1.305)	1	0.837 (1.195)
Spain	0.905 (1.105)	1.195 (0.837)	1
	With Constant Utility $s^{CU}_{h,f}$		
UK	1	1.322 (0.757)	1.105 (0.905)
US	0.757 (1.322)	1	0.834 (1.199)
Spain	0.905 (1.105)	1.199 (0.834)	1
	With Common Scaling $s^{CS}_{h,f}$		
UK	1	1.35 (0.741)	1.112 (0.899)
US	0.741 (1.35)	1	0.825 (1.212)
Spain	0.899 (1.112)	1.212 (0.825)	1

Note: Tables report symmetric price indices based on asymmetric aggregate Laspeyres $\Pi^L_{h,f}$, constant utility $\Pi^{CU}_{h,f}$, and common scaling $\Pi^{CS}_{h,f}$ price indices. Each row represents a home society. Each column represents a foreign society. $s_{h,f}$ is does not have an explicit interpretation but it provides an average measure of price differences across societies. Values in parentheses are reciprocals of the index values to illustrate symmetry. As an example $s^L_{UK,US} = 1.305$, $1/s^L_{UK,US} = 0.766$. Since $s^L_{UK,US}$ is symmetric, then $s^L_{UK,US} = 1/s^L_{US,UK}$ must hold. Indeed, $s^L_{US,UK} = 0.766$.

Table 3.4 Multilateral Indices

Country	$S_{h,UK}^{GEKS,L}$	$S_{h,UK}^{GEKS,CU}$	$S_{h,UK}^{GEKS,CS}$	$S_{h,UK}^{CGW,L}$	$S_{h,UK}^{CGW,CU}$	$S_{h,UK}^{CGW,CS}$
UK	1	1	1	1	1	1
US	1.310	1.323	1.349	1.292	1.325	1.354
Spain	1.101	1.104	1.113	1.103	1.118	1.129

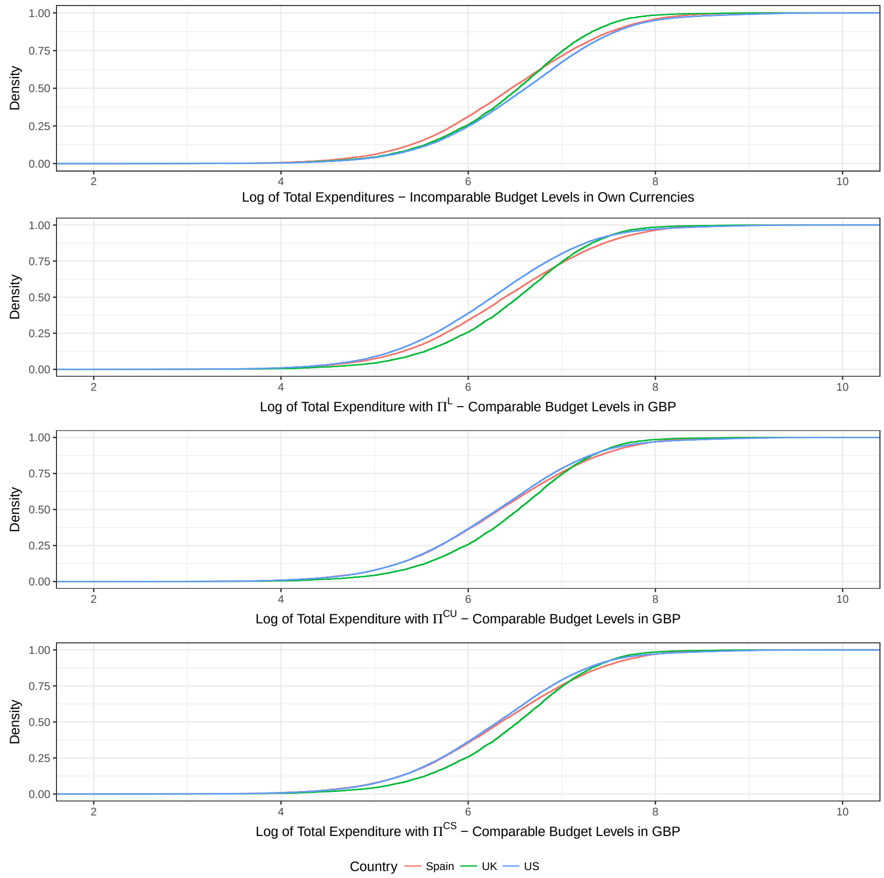
Note: Tables report transitive price indices based on asymmetric price indices $\Pi_{h,f}^{L}$, $\Pi_{h,f}^{CU}$, and $\Pi_{h,f}^{CS}$ and taking the UK as the base country. Each row represents a society. S^{GEKS} is a multilateral aggregate price index that is inconsistent with our thought-experiments but indicates an average of price differences. S^{CGW} is interpreted as a constant global welfare price index that is consistent with a thought-experiment.

Table 3.5 Summary - UK Indices

UK (h)	Asymmetric Bilateral			Symmetric Bilateral			GEKS Multilateral			CGW Multilateral			Other		
Foreign (f)	Π_{L}^{fh}	Π_{CU}^{fh}	Π_{CS}^{fh}	$\Pi_{S,L}^{fh}$	$\Pi_{S,CU}^{fh}$	$\Pi_{S,CS}^{fh}$	$s_{GEKS,L}^{fh}$	$s_{GEKS,CU}^{fh}$	$s_{GEKS,CS}^{fh}$	$s_{CGW,L}^{f,h}$	$s_{CGW,CU}^{fh}$	$s_{CGW,CS}^{fh}$	XR	Fisher	GEKS
US	1.344	1.326	1.353	1.305	1.322	1.350	1.310	1.323	1.349	1.292	1.325	1.354	1.613	1.422	1.421
Spain	1.116	1.100	1.109	1.105	1.105	1.112	1.101	1.104	1.113	1.103	1.118	1.129	1.161	1.103	1.103

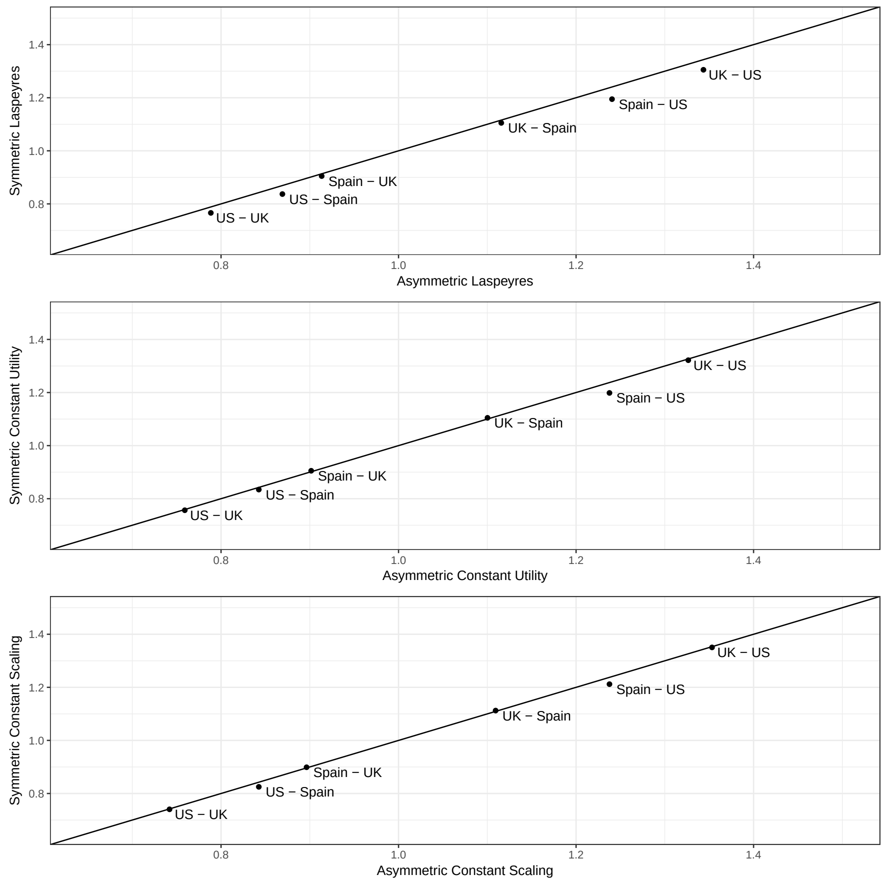
Note: Table summarizes all price indices by taking the UK as the home (or the base for multilateral comparisons) country. Each row represents a foreign society f. XR are 2011 average nominal exchange rates. Fisher and GEKS are the Fisher ideal indices (square root of Laspeyres and Paasche indices) and their multilateral conversions for the expenditure categories that we consider in our estimations (without adjustment for differences in preferences within and across borders).

Fig. 3.1 Budget Comparisons



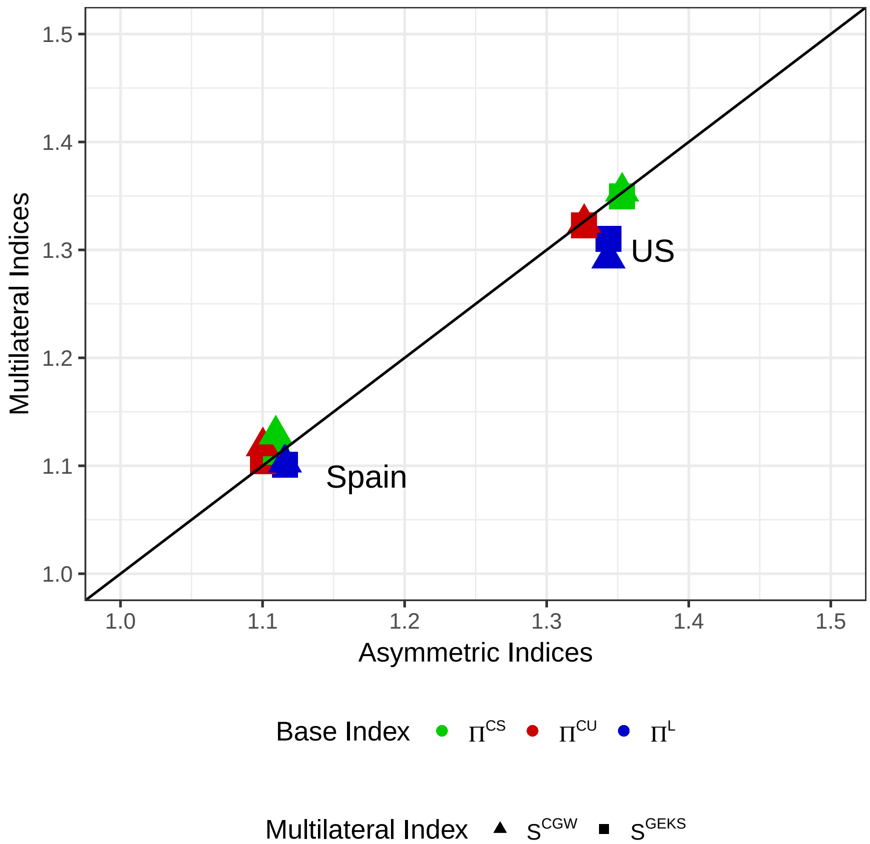
Note: Figures indicate distributions of log budget levels in each country. The top figure illustrates incomparable distributions in each country's own price vectors. The other figures illustrate comparable distributions using each of our thought-experiments. Second and third figures, which use Π^L and Π^{CU}, have different shapes than the distributions in the top figure because households have disproportionate budget adjustments in these thought-experiments. On the other hand, the last figure preserves the place of each household along budget distributions of all countries because of a proportionate adjustment to all income levels. Figures indicate, even though mean income earner UK households have higher budget levels than American and Spanish households, top income earner UK households have lower budget levels. This result indicates that comparing distributions of cost levels across nations provides a deeper understanding of differences in international income levels.

Fig. 3.2 Deviations from Symmetry - Symmetric Indices



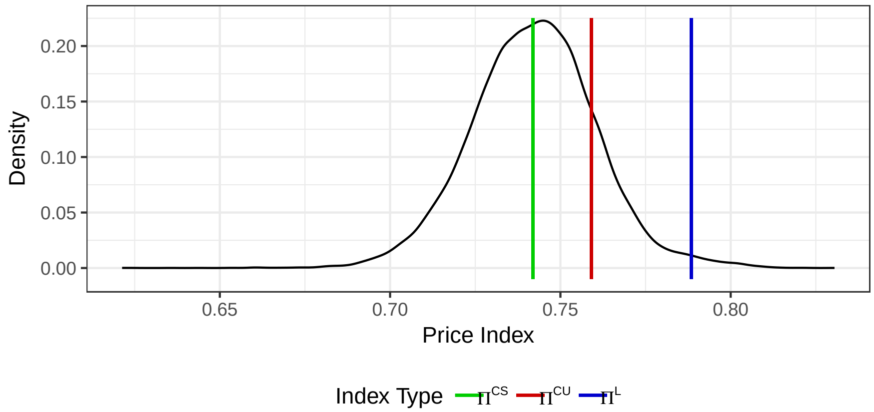
Note: Figures illustrate deviations between each base index and their symmetric conversion with a geometric averaging. As illustrated in comparisons between Spain and US, or UK and US, some symmetric conversions may lie significantly away from the line of equality.

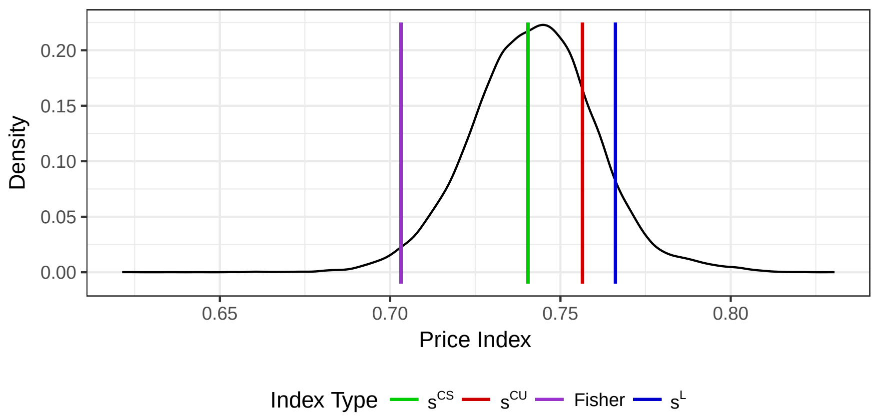
Fig. 3.3 Deviations from Symmetry - Multilateral Indices



Note: Figure illustrates deviations between each base index and their multilateral conversion with a GEKS formula taking the UK as the base country. As illustrated in the figures, a comparison between US and UK lies significantly away from the line of equality for Π^{L} as the base asymmetric index. Note that US-UK symmetric comparisons are also comparisons that lie away from the line of equality in figure 3.2.

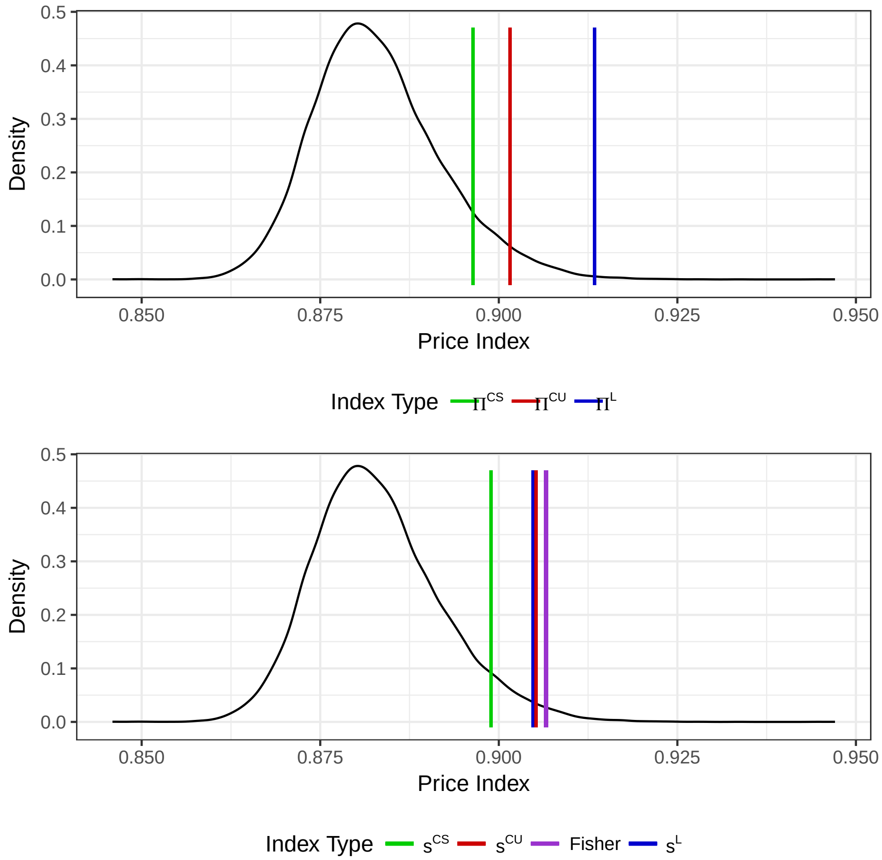
Fig. 3.4 Individual Konüs Price Indices - Home: US, Foreign: UK, $\pi_{US,UK}^{i_{US},K}$





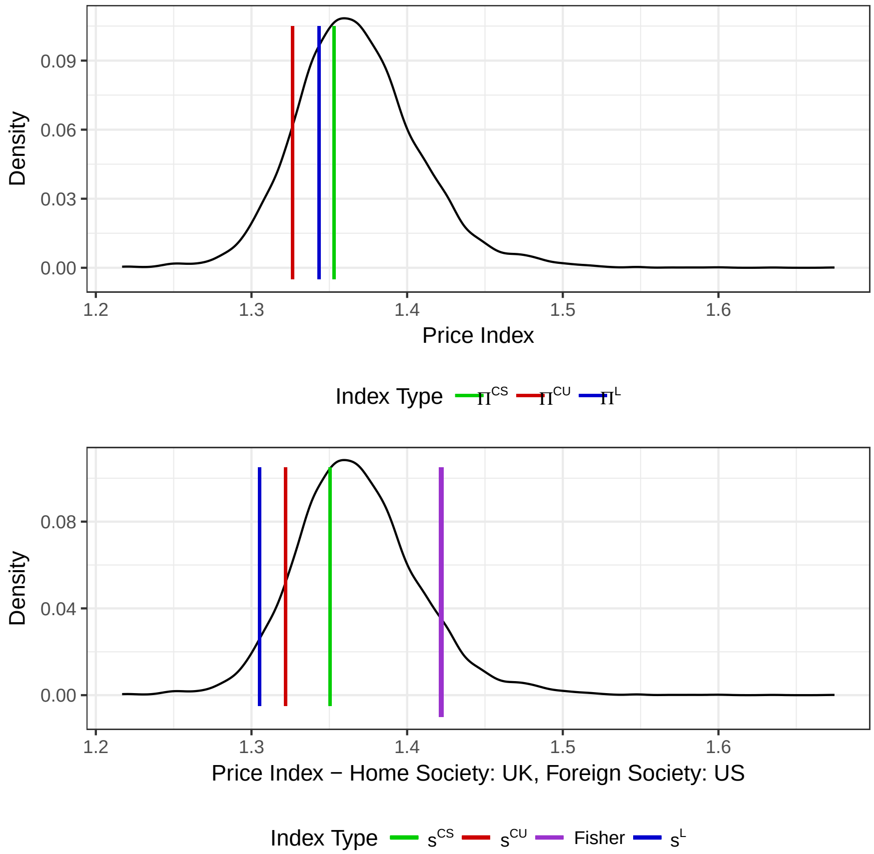
Note: Figures illustrate distributions of individual Konüs indices for American households at UK prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of symmetric indices (with a geometric averaging) along the same distribution. The Fisher index is calculated as the geometric average of Laspeyres and Paasche indices that are inconsistent with heterogeneous preferences.

Fig. 3.5 Individual Konüs Price Indices - Home: Spain, Foreign: UK, $\pi^{i_{Spain},K}_{Spain,UK}$



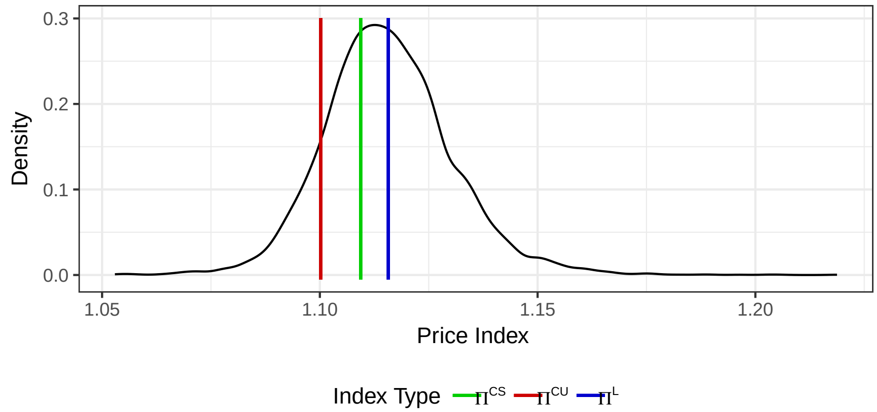
Note: Figures illustrate distributions of individual Konüs indices for Spanish households at UK prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of symmetric indices (with a geometric averaging) along the same distribution. The Fisher index is calculated as the geometric average of Laspeyres and Paasche indices that are inconsistent with heterogeneous preferences.

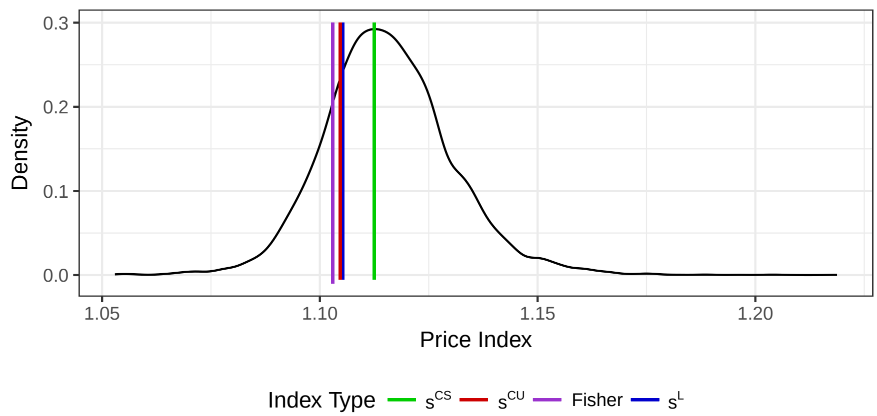
Fig. 3.6 Individual Konüs Price Indices - Home: UK, Foreign: US, $\pi_{UK,US}^{i_{UK},K}$



Note: Figures illustrate distributions of individual Konüs indices for British households at US prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of symmetric indices (with a geometric averaging) along the same distribution. The Fisher index is calculated as the geometric average of Laspeyres and Paasche indices that are inconsistent with heterogeneous preferences.

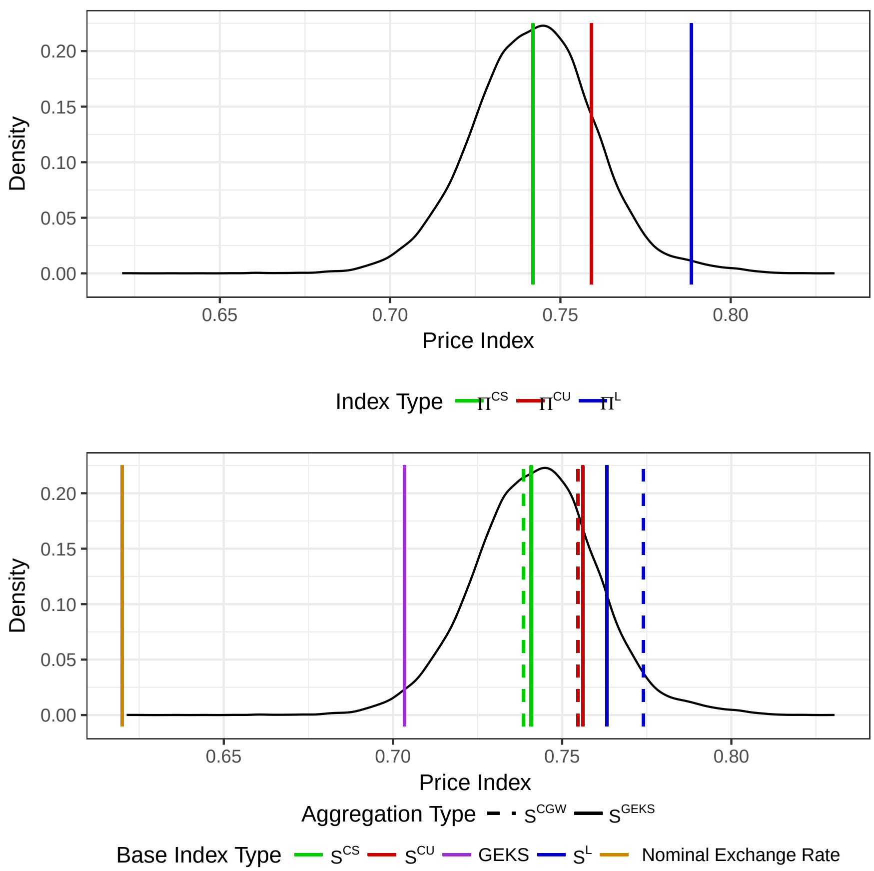
Fig. 3.7 Individual Konüs Price Indices - Home: UK, Foreign: Spain, $\pi^{i_{UK},K}_{UK,Spain}$





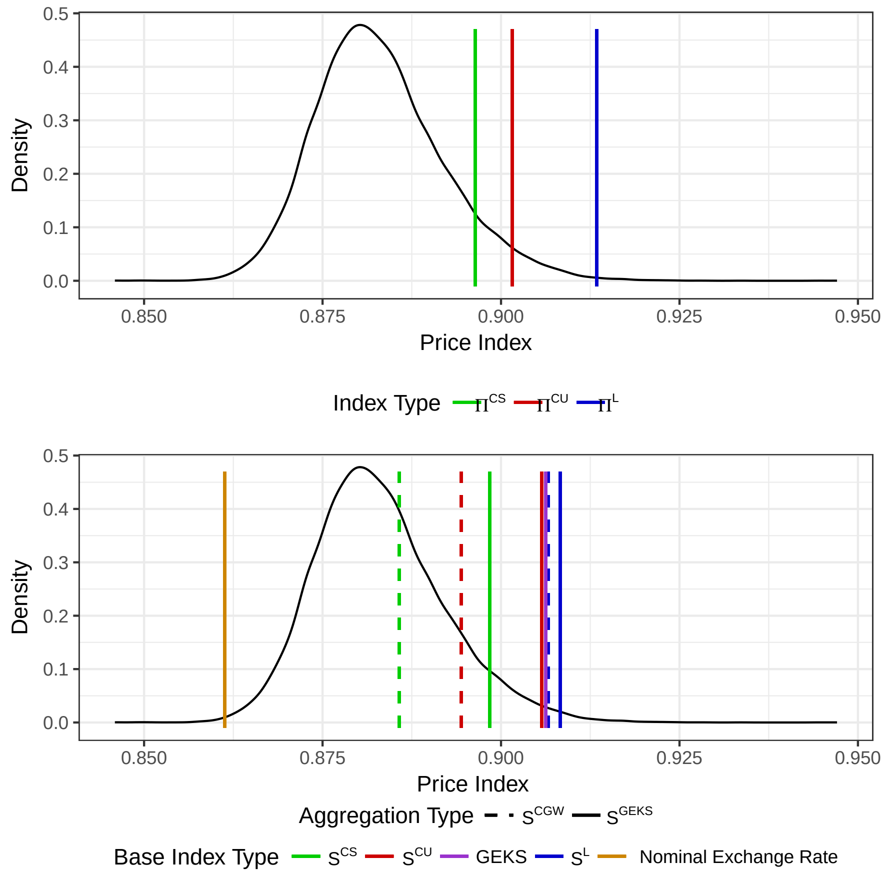
Note: Figures illustrate distributions of individual Konüs indices for UK households at Spanish prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of symmetric indices (with a geometric averaging) along the same distribution. The Fisher index is calculated as the geometric average of Laspeyres and Paasche indices that are inconsistent with heterogeneous preferences.

Fig. 3.8 Individual Konüs Price Indices - Home: US, Foreign: UK, $\pi_{US,UK}^{i_{US},K}$



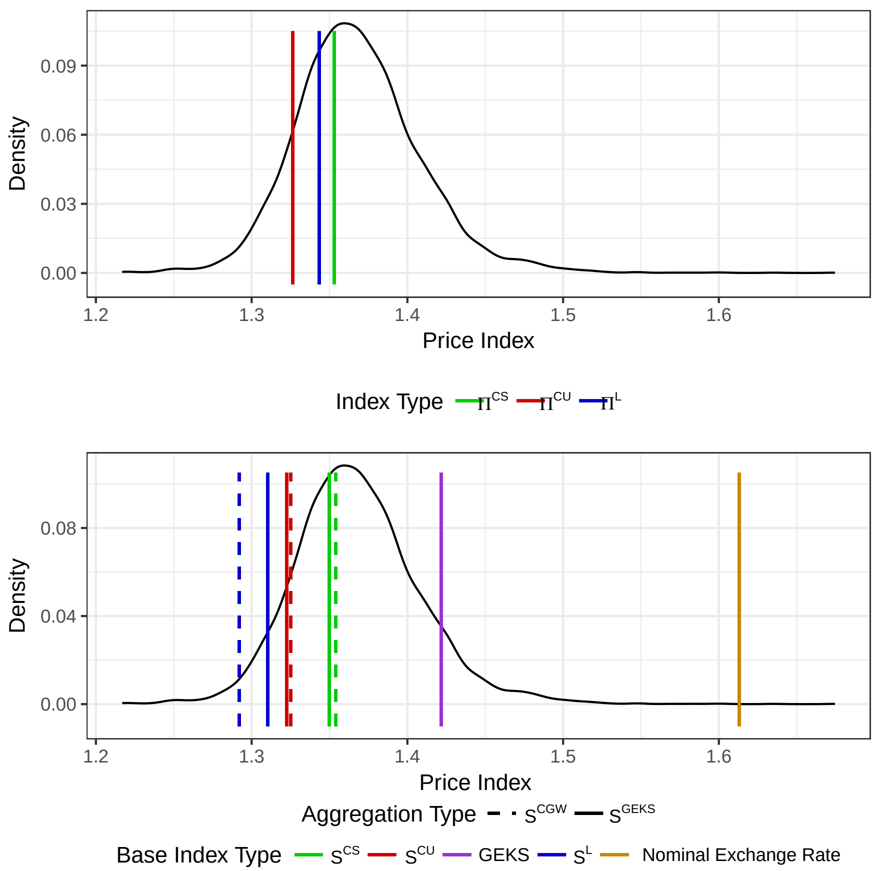
Note: Figures illustrate distributions of individual Konüs indices for American households at UK prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of multilateral indices along the same distribution. The GEKS index is calculated using Fisher indices as the base symmetric indices.

Fig. 3.9 Individual Konüs Price Indices - Home: Spain, Foreign: UK, $\pi_{Spain,UK}^{i_{Spain},K}$



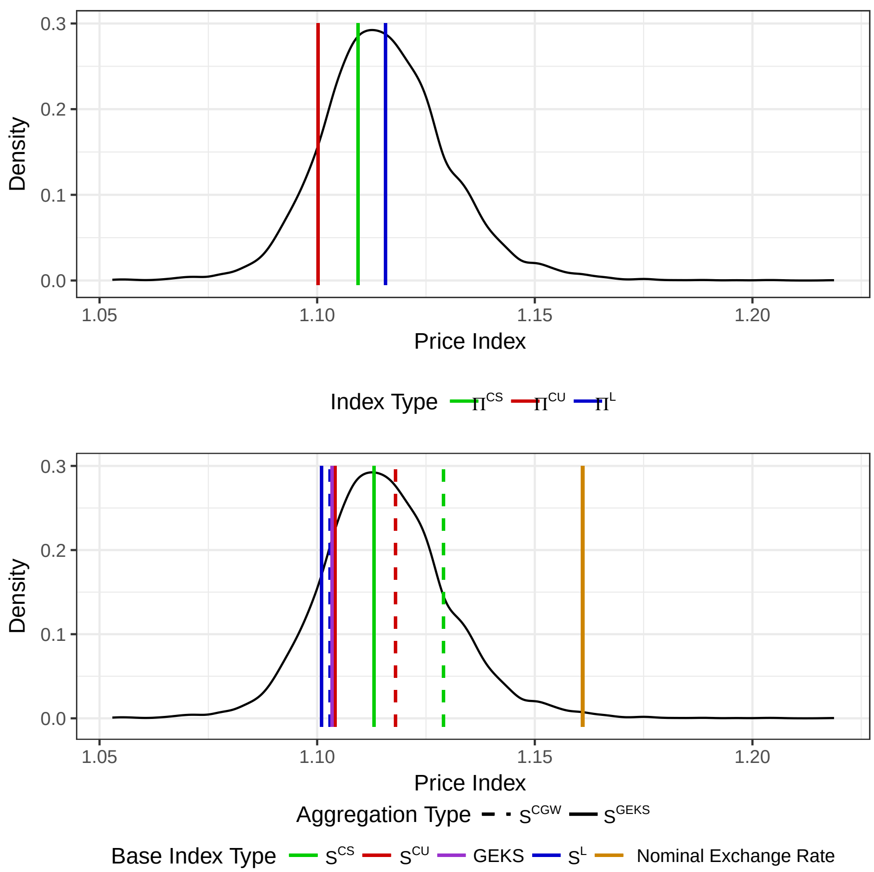
Note: Figures illustrate distributions of individual Konüs indices for Spanish households at UK prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of multilateral indices along the same distribution. The GEKS index is calculated using Fisher indices as the base symmetric indices.

Fig. 3.10 Individual Konüs Price Indices - Home: UK, Foreign: US, $\pi_{UK,US}^{i_{UK},K}$



Note: Figures illustrate distributions of individual Konüs indices for British households at UK prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of multilateral indices along the same distribution. The GEKS index is calculated using Fisher indices as the base symmetric indices.

Fig. 3.11 Individual Konüs Price Indices - Home: UK, Foreign: Spain, $\pi^{i_{UK},K}_{UK,Spain}$



Note: Figures illustrate distributions of individual Konüs indices for British households at Spanish prices. The vertical lines in the top figure illustrate the place of asymmetric bilateral social purchasing power parities along the distribution of individual purchasing power parities. The vertical lines in the bottom figure illustrate the place of multilateral indices along the same distribution. The GEKS index is calculated using Fisher indices as the base symmetric indices.

Appendix A
Appendix to Chapter 1

A.1 A Multilayer FNN Diagram

A network diagram is an illustrative tool to plot functional relationships. For example, a familiar linear regression with two inputs

$$y = \theta_1 x_1 + \theta_2 x_2$$

would have the network diagram that is illustrated in figure A.1.

In this diagram, each node in the input layer represents an input x, and the node at the output layer represents the predicted output y. Arrows that feed the input information to output represents coefficients θ_1 and θ_2.

In order to represent a network diagram for a multilayer FNN, imagine having intermediate layers between the input and output layers. These intermediate layers generate derived variables from input variables with weighted addition and non-linear transformation. For example, the network diagram in figure A.2 represents an FNN with four layers. Layers are categorized into three types: an input layer, hidden layers (intermediate layers), and an output layer. Each layer consists of a number of nodes that are connected to the nodes in the previous and the next layer. Nodes are responsible for processing information provided from the previous layers, and passing the processed information to the next layers. Each node in the input layer, i.e., $x_1, \ldots, x_I$, represents an independent variable. Consequently, the number of nodes in the input layer is equal to the number of independent variables plus one for the intercept term. Similarly, nodes in the output layer represent a dependent variable. Therefore the number of nodes in input and output layers are generally predetermined by the purpose of the analysis. Number of hidden layers and number of nodes in each hidden layer are data dependent adjustable hyper-parameters of an FNN. Arrows represent coefficients θ.

The following weighted addition and non-linear transformation occur in the k^{th} node of the first hidden layer of a feedforward neural network (see figure A.2) for all $k_1 \in [1, K_1]$:

$$a^1_{k_1} = \alpha\left(\sum_{i=1}^{I+1} \theta^1_{k_1,i} x_i\right)$$

where $\alpha(\cdot)$ is a non-linear activation function, $a^1_{k_1}$ is the output of the k^{th} node of the first hidden layer, $\theta^1_{k_1,i}$ is the coefficient of x_i in the summation in the k^{th} node of the first hidden layer.[1] A same type of weighted addition and transformation occur in the second hidden layer for all $k_2 \in [1, K_2]$:

$$a^2_{k_2} = \alpha\left(\sum_{k_1=1}^{K_1+1} \theta^2_{k_2,k_1} a^1_{k_1}\right)$$

The output is obtained by feeding $a^2_{k_2}$, $\forall k_2 \in [1, K_2]$, to the output layer:

$$y = \beta\left(\sum_{k_2=1}^{K_2+1} \theta^3_{1,k_2} a^2_{k_2}\right)$$

where $h(\cdot)$ is a suitable transformation.

A.2 AIDS and QUAIDS Functional Approximations

AIDS demand functions are derived from PIGLOG class preferences with the following indirect utility function:

$$\log v(p,m) = \frac{\log m - \log a(p)}{b(p)}$$

which yields the following functional approximation for $f_g(\cdot)$

$$f_g^{AIDS}(\cdot) = \alpha_g + \sum_{g'=1}^{G} \gamma_{gg'} \log p_j + \beta_g \log\left[\frac{m}{a(p)}\right]$$

where $a(p)$ is a translog price index:

$$\log a(p) = \alpha_0 + \sum_{g=1}^{G} \alpha_g \log p_g + \frac{1}{2}\sum_{g=1}^{G}\sum_{g'=1}^{G} \gamma_{gg'} \log p_g \log p_{g'}.$$

QUAIDS model provides a more flexible approximation for $f_g(\cdot)$ by assuming the following indirect utility function

[1] Input size $I+1$ is because of addition of an intercept term.

$$\log v(p,m) = \left\{ \left[\frac{\log m - \log a(p)}{b(p)} \right]^{-1} + \lambda(p) \right\}^{-1}$$

and by introducing a quadratic term for log of price adjusted income:

$$f_g^{QUAIDS}(\cdot) = \alpha_g + \sum_{g'=1}^{G} \gamma_{gg'} \log p_j + \beta_g \log\left[\frac{m}{a(p)}\right] + \frac{\lambda_g}{b(p)} \left\{ \log\left[\frac{m}{a(p)}\right] \right\}^2$$

where $b(p)$ is a Cobb-Douglas price aggregator

$$b(p) = \prod_{g=1}^{G} p_g^{\beta_g}$$

and $a(p)$ is the same translog price index.

A.3 Robustness Tests

Selection of Transformation Function and Optimization Algorithm

In this section, I present robustness of the main functional approximations to the choice of activation function and optimization algorithm. To this aim, I build ADN and TDN models with a logit non-linear transformation function[2] and obtain estimates using a gradient descent with Nesterov momentum optimization algorithm (Nesterov, 2004; Sutskever et al., 2011).

Figure A.3 illustrates distributions of test costs for ADN and TDN estimates with robustness test specification (ADN-R and TDN-R). Distributions of test costs from the specification with a rectifier transformation function and Adam optimizer, which is the specification that yields the main results of this paper, are also illustrated for comparison. The figure shows that generalization performances of all estimations are similar.

Imposing Negativity

Figure A.4 illustrates distributions of test costs of AIDS, QUAIDS, ADN, TDN, and TDN-N estimates. TDN is estimated imposing homogeneity, adding up, and symmetry. TDN-N is estimated imposing homogeneity, adding up, symmetry, and negativity. Imposing negativity in addition to other restrictions does not affect the generalizability compared to TDN estimate for this sample.

[2] $\alpha(x) = 1/(1+\exp(-x))$.

Figure A.5 plots estimated demand functions of all estimations. Fits provided by TDN and TDN-N estimates along the budget axis are nearly identical.

A.4 About Statistical Consistency of ADN and TDN

Statistical consistency of an estimated function $\hat{f}(\cdot)$ is defined in the usual way, i.e., the asymptotic convergence of an estimator to a target object. Consistency of semi-nonparametric feedforward neural network estimators to a target function in a mean squared sense when hyper-parameters are allowed to vary with sample size is documented by White (1990) and Geman et al. (1992).

Therefore, ADN, as a feedforward neural network with a certain output structure, satisfies the consistency requirements and asymptotically converges to the unobserved theoretical demand function:

$$\lim_{N\to\infty} E\left(\left[E\left(\mathbf{w}|\mathbf{x}\right) - \hat{f}^{ADN}(\mathbf{x}|\mathscr{X}_N)\right]^2\right) = 0,$$

where $\mathscr{X}_N$ is a sample with size N, when functional complexity of $\hat{f}^{ADN}(\mathbf{x}|\mathscr{X}_N)$ is adjusted with cross-validation.

When estimating a theoretically constrained network TDN, it is important to preserve this consistency. Recall that a TDN is defined as the solution to the following minimization problem:

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g^{TDN}(\Theta, \mathbf{x}_n)\right]^2 + \gamma P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}}),$$

which has an additive term that penalizes the optimization algorithm for deviations from the theoretical restrictions. Therefore, if $P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}}) \to 0$ as $N \to \infty$, then TDN is also asymptotically consistent. It is easy to see that this condition holds only if the theoretical restrictions that are represented in the penalty function $P(\Theta, \mathscr{X}^{\mathscr{S}}, \mathscr{X}^{\mathscr{N}})$ are true. Otherwise, TDN is destined to asymptotically converge to an incorrect function. Note that this is independent of the penalty coefficient $\gamma \gg 0$.

The intuition of this procedure is straightforward. Consistency of FNN estimators relies on their ability to find a suitable function from a function space with adjustable functional complexity. The role of the penalty term is to guide this search procedure towards certain functional shapes that are allowed by the neoclassical theory of consumer demand. If data supports rationality, as $N \to \infty$, an ADN would capture integrability conditions and yield a function that lies within the class of theoretically allowed functions. However, if data does not support rationality, then unobserved demand functions are not within the class of functions that the penalty term enforces. Hence, ADN and TDN are different; consistency is violated. Therefore, TDN is asymptotically consistent if theoretical restrictions are supported by empirical observations.

In application, functional search imposed by the penalty term is, in a way, similar to a common practice in the estimations of the neural network algorithms. This practice, norm regularization, uses L_1 or L_2 norm of the coefficient vector to penalize an optimization algorithm, and is commonly applied to remedy overfitting problems associated with neural networks. If I were to use a similar approach, for example, with L_2 norm, an FNN network with a norm regularizer would be the solution to the following minimization problem:

$$\min_{\Theta} \sum_{n=1}^{N} \sum_{g=1}^{G} \left[w_{g,n} - \hat{f}_g(\Theta, \mathbf{x}_n) \right]^2 + \lambda L_2(\Theta)$$

where a regularization coefficient λ is chosen with cross-validation.[3] Clearly, the regularization term $L_2(\Theta)$ does not asymptotically converge to zero. Hence, the solution is not a consistent estimation of the target function. The idea behind regularization and selection of λ with cross-validation is to introduce bias to the estimation to reduce variance up to a point that the bias introduced does not exceed the variance reduced. As a result, predictive accuracy is increased.

A penalty term $P(\cdot)$ works differently than a norm regularizer. TDN is estimated over a sequence of penalty coefficients to obtain convergence to a constrained minimization problem. Alternatively, similar to the regularization coefficient λ, the penalty coefficient γ can be selected with cross-validation, which could be interpreted as a partial imposition of the theoretical restrictions. Such an approach may yield demand functions with higher generalizability even when the theoretical constraints are rejected. However, theoretical properties of the estimates become ambiguous. Therefore, such an approach is not preferred in this paper.

A.5 Derivations

In this section, I present derivations of elasticities and the neoclassical restrictions for a single hidden layer multiple output neural network. The single hidden layer structure is illustrated for notational simplicity.

Elasticities

It can be shown that uncompensated price elasticities e^u_{ij}, compensated price elasticities e^c_{ij}, income elasticities e_i, and slutsky terms s_{ij}, are defined as

[3] $L_2(\Theta) = \sqrt{\sum_{k=1}^{K} |\theta_k|^2}$.

$$
\begin{aligned}
e_i &= \frac{\mu_i}{w_i} + 1 \\
e^u_{ij} &= \frac{\mu_{ij}}{w_i} - \delta_{ij} \\
e^c_{ij} &= e^u_{ij} + e_i w_j \\
s_{ij} &= w_i e^c_{ij}
\end{aligned}
$$

where w_i and w_j are the predicted budget shares of the good i and j, δ_{ij} is the kronecker delta, $\mu_{ij} = \partial w_i \big/ \partial \ln p_j$, and $\mu_i = \partial w_i \big/ \partial \ln m$ (see also Banks et al., 1997).

Let $\mathbf{x}$ be a vector of I input variables with elements x_l, and $\mathbf{w}$ be a vector of G predicted budget shares with elements w_g. Let K be the number of nodes in the hidden layer. Let the following notation represent the flow in an FNN model:

$$
\begin{aligned}
z^1_k &= \sum_{i=1}^{i} \theta^1_{k,i} x_l + \theta^1_k \\
a^1_k &= \alpha(z^1_k) \\
z^2_g &= \sum_{k=1}^{K} \theta^2_{g,k} a^1_k + \theta^2_g \\
w_g &= \beta(z^2_g)
\end{aligned}
$$

where $\alpha(\cdot)$ is the hidden layer transformation function and $\beta(\cdot)$ is a multinominal logit transformation function

$$
\beta(z^2_g) = \frac{\exp(z^2_g)}{\sum_{g=1}^{G} \exp(z^2_g)}.
$$

Define the norm of the price vector with the Euclidean norm as $||\mathbf{p}|| = \sqrt{\sum_{g=1}^{G} p^2_g}$.

Below, I derive μ_i and μ_{ij}. Then, equations of elasticities are obtained straightforwardly.

Derivation of $\mu_i = \partial w_i \big/ \partial \ln m$

Let w_i be the i^{th} element of the multinominal logit output.

$$
\frac{\partial w_i}{\partial \ln m} = \sum_{g=1}^{G} \frac{\partial w_i}{\partial z^2_g} \frac{\partial z^2_g}{\partial \ln m} \tag{A.1}
$$

$$
\frac{\partial z^2_g}{\partial \ln m} = \sum_{k=1}^{K} \theta^2_{g,k} \frac{\partial a^1_k}{\partial \ln m}
$$

$$\frac{\partial a_k^1}{\partial \ln m} = \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \frac{\partial z_k^1}{\partial \ln m} \tag{A.2}$$

Combining A.1 to A.2, by chain rule :

$$\mu_i = \sum_{g=1}^{G} \frac{\partial w_i}{\partial z_g^2} \left[\sum_{k=1}^{K} \theta_{g,k}^2 \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \frac{\partial z_k^1}{\partial \ln m} \right]$$

With a multinominal logit output layer

$$\frac{\partial w_i}{\partial z_g^2} = w_g \delta_{ig} - w_i w_g$$

where w_g is the predicted budget share of good g, and δ_{ig} is the kronecker delta, which is zero if $i = g$. Then,

$$\mu_i = w_i \sum_{g=1}^{G} (\delta_{ig} - w_g) \left[\sum_{k=1}^{K} \theta_{g,k}^2 \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \frac{\partial z_k^1}{\partial \ln m} \right]$$

Furthermore,

$$\frac{\partial z_k^1}{\partial \ln m} = \theta_{k,m}^1 \frac{\partial x_m}{\partial \ln m}$$

where $l = m$ is the index of the log expenditure in the input vector $\mathbf{x}$. For ADN, $X_m = \ln m$, thus $\partial X_m / \partial \ln m = 1$, and μ_i^{ADN} is obtained as:

$$\mu_i^{TDN} = w_i \sum_{g=1}^{G} (\delta_{ig} - w_g) \left[\sum_{k=1}^{K} \theta_{g,k}^2 \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \theta_{k,m}^1 \right]$$

For TDN, $X_m = \ln \left(m / ||p|| \right)$, thus $\partial X_m / \partial \ln m = 1$, and $\mu_i^{TDN} = \mu_i^{ADN}$.

Derivation of $\mu_{ij} = \partial w_i / \partial \ln p_j$

The derivation of μ_{ij} follows the derivation of μ_i and it can be shown that

$$\mu_{ij} = w_i \sum_{g=1}^{G} (\delta_{ig} - w_g) \left[\sum_{k=1}^{K} \theta_{g,k}^2 \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \frac{\partial z_k^1}{\partial \ln p_j} \right]$$

where

$$\frac{\partial z_k^1}{\partial \ln p_j} = \sum_{l=1}^{I} \theta_{k,l}^1 \frac{\partial X_l}{\partial \ln p_j}$$

For ADN, $X_l = \ln p_l$. Therefore, $\partial X_l / \partial \ln p_j = \delta_{lj}$. Hence μ_{ij}^{ADN} is obtained as:

$$\mu_{ij}^{ADN} = w_i \sum_{g=1}^{G} (\delta_{ig} - w_g) \left[\sum_{k=1}^{K} \theta_{g,k}^2 \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \theta_{k,j}^1 \right]$$

For TDN, $X_l = \ln(p_l/||p||)$. Hence,

$$\frac{\partial X_l}{\partial p_j} = \frac{\partial \ln p_l}{\partial \ln p_j} - \frac{\partial \ln ||p||}{\partial \ln p_j}$$
$$\frac{\partial X_l}{\partial p_j} = \delta_{lj} - (p_j/||p||)^2 .$$

Hence, μ_{ij}^{TDN} is obtained as:

$$\mu_{ij}^{TDN} = w_i \sum_{g=1}^{G} (\delta_{ig} - w_g) \left[\sum_{k=1}^{K} \theta_{g,k}^2 \frac{\partial \alpha(z_k^1)}{\partial z_k^1} \sum_{l=1}^{G+1} \theta_{k,l}^1 \left(\delta_{lj} - (p_j/||p||)^2 \right) \right]$$

where the first G elements of input vector $\mathbf{x}$ are homogeneity adjusted log prices and the input with index $G+1$ is the log of adjusted expenditure.

Income and price elasticities are straightforward to compute by plugging μ_i and μ_{ij} in definitions provided at the beginning of this section.

TDN Constraints

Symmetry

The symmetry constraint is characterized as

$$R^s(\Theta, \mathscr{X}^{\mathscr{S}}) = \sum_{n=1}^{N^s} \sum_{i=1}^{G-1} \sum_{j=i+1}^{G} [\hat{s}_{ij,n} - \hat{s}_{ji,n}]^2$$

where a Slutsky term $s_{ij,n}$ is equal to $w_{i,n} e_{ij,n}^c$. Once a Slutsky matrix is calculated, then it is straightforward to obtain the upper diagonal elements of the symmetric matrix $\left(\hat{\mathsf{S}}_n - \hat{\mathsf{S}}_n^T\right)^2$ for all $\mathbf{x}_n \in \mathscr{X}^{\mathscr{S}}$. Their sum across all observations that the symmetry is imposed yields the cost of deviation from symmetry. Note that the summation includes only the upper diagonal elements of the symmetric matrix $\left(\hat{\mathsf{S}}_n - \hat{\mathsf{S}}_n^T\right)^2$ in order to reduce the computational cost.

Negativity

A Slutsky matrix, $\hat{\mathsf{S}}_n$ is negative semi-definite if the constraint indicator

$$\lambda(\Theta, \mathbf{x}_n) \equiv \max_{\mathbf{v}_n} \left\{ -\mathbf{v}_n^T \hat{\mathsf{S}}(\Theta, \mathbf{x}_n) \mathbf{v}_n : \hat{\mathbf{f}}(\Theta, \mathbf{x}_n)^T \mathbf{v}_n = 0, \, \mathbf{v}_n^T \mathbf{v}_n = 1 \right\}$$

is non-negative. Gallant and Golub (1984) provide the necessary steps to calculate $\lambda(\Theta, \mathbf{x}_n)$ and its gradient with respect to the elements of the coefficient vector Θ.

A.6 Pseudo-codes of the Optimization Process

Pseudo-codes in this section outline the optimization processes that I use in this paper. Table A.1 summarizes the definitions of the relevant variables in these algorithms. The function $Update(\cdot)$ refers to the optimization algorithm that is chosen to solve an ADN or a TDN model. These algorithms call the relevant optimization problems and minimize $C(\Theta, \mathscr{X}_e)$.

Briefly, algorithms 1 and 2 summarize the ADN and TDN processes in general. They call stochastic, i.e., coefficient updates with individual observations, and mini-batch, i.e., coefficient updates with random groups of observations, optimization procedures to perform coefficient updates. The random assignment of observations into the mini-batches are ensured by the random assignment of observations into the estimation samples. These optimization procedures are in algorithms 3 and 4. They stop cross-validation by calling an early stopping procedure that is outlined in algorithm 5. The TDN optimization algorithm also calls for a penalty method in order to impose the neoclassical restrictions. This procedure is outlined in algorithm 6.

Table A.1 provides definitions of all terms that are used in the pseudo-codes.

Algorithm 1: ADN Optimization Algorithm

Data: An $n \times I$ matrix of input observations $\mathscr{X}_e$ with rows $\mathbf{x}_{n,e}^T$, and an $n \times G$ matrix of output observations $\mathscr{Y}_e$ with rows $\mathbf{w}_{n,e}^T$, $n = 1, 2, \dots, N$.

Result: Estimated ADN model: hyper-parameters and estimated coefficients.

begin Estimation
- Define objective function as equation 1.2
- Define $hh\_end = hh\_start + window$
- **begin** Cross-Validation
  - **while** $hh\_size \leq hh\_end$ **do**
    - $k = hh\_size$
    - Initialize Θ_0^k
    - **begin** Stochastic optimization
      - **do** ***Algorithm 3*** *Stochastic optimization*
    - **end**
    - $\Theta_0^k = \Theta^k$
    - **begin** Mini Batch optimization
      - **do** ***Algorithm 4*** *Mini Batch optimization*
    - **end**
    - **if** $hh\_size == hh\_end$ **then**
      - **do** ***Algorithm 5*** *Cross Validation Check*
    - **end**
    - $hh\_size = hh\_size + 1$
  - **end**
- **end**
- **return** k *s.t.* $k = \text{argmin}_k C(\Theta^k, \mathscr{X}_c | \mathscr{X}_e)$, $\forall k \in [hh\_start, hh\_end]$ *and* Θ^k

end

Algorithm 2: TDN Optimization Algorithm

```
Data: An n × I matrix of input observations X_e with rows x^T_{n,e}, and an n × G
      matrix of output observations Y_e with rows w^T_{n,e}, n = 1,2,...,N.
Result: Estimated TDN model: hyper-parameters and estimated coefficients.
begin Estimation
    Define objective function as equation 1.4
    Define a sequence of γ_s s.t. γ_{s+1} = γ_s, s = 1,2,...
    Define hh_end = hh_start + window
    begin Cross-Validation
        while hh_size ≤ hh_end do
            k = hh_size
            s = 1
            Initialize Θ^k_0
            begin Stochastic Optimization
                do Algorithm 3 Stochastic Optimization
            end
            Θ^k_0 = Θ^k
            begin Mini Batch Optimization
                do Algorithm 4 Mini Batch Optimization
            end
            Θ^{k,s}_0 = Θ^k
            begin Penalty Method
                do Algorithm 6 Penalty Method
            end
            if hh_size == hh_end then
                do Algorithm 5 Cross Validation Check
            end
            hh_size = hh_size + 1
        end
    end
    return k s.t. k = argmin_k C(Θ^k, X_c | X_e), ∀k ∈ [hh_start, hh_end] and Θ^k
end
```

Algorithm 3: Stochastic Optimization

```
Data: An n × I matrix of input observations 𝒳_e with rows x^T_{n,e}, and an n × G
      matrix of output observations 𝒴_e with rows w^T_{n,e}, n = 1, 2, ..., N, a
      number of nodes k, and a vector of coefficients Θ^k_0.
Result: A vector of updated coefficients Θ^k
begin
    epoch = 1
    while epoch ≤ stochastic_grad_epoch_limit do
        Θ^k_{epoch,0} ⟵ Θ^k_{epoch−1}
        foreach n in 1, 2, ..., N do
            Θ^k_{epoch,n} ⟵ Update(Θ^k_{epoch,n−1}, x_{n,e}, w_{n,e})
        end
        Θ^k_{epoch} ⟵ Θ^k_{epoch,N}
        if |C(Θ^k_{epoch}, 𝒳_c|𝒳_e) - C(Θ^k_{epoch−1}, 𝒳_c|𝒳_e)| ≤ tol_stoc then
            break
        else
            epoch ⟵ epoch + 1
        end
    end
    return Θ^k = Θ^k_{epoch}
end
```

Algorithm 4: Mini Batch Optimization

Data: An $n \times I$ matrix of input observations $\mathscr{X}_e$ with rows $\mathbf{x}_{n,e}^T$, and an $n \times G$ matrix of output observations $\mathscr{Y}_e$ with rows $\mathbf{w}_{n,e}^T$, $n = 1, 2, \ldots, N$, a number of nodes k, a vector of coefficients Θ_0^k, and a *batch_size*.

Result: A vector of coefficients Θ^k

begin
- epoch = 1
- **while** $epoch \leq mini\_batch\_grad\_epoch\_limit$ **do**
  - $\Theta^k_{epoch,0} \longleftarrow \Theta^k_{epoch-1}$
  - **foreach** n *in* $1, 2, \ldots, int(N / batch\_size)$ **do**
    - $batch\_x = \mathscr{X}_e[n \times batch\_size : (n+1) \times batch\_size]$
    - $batch\_y = \mathscr{Y}_e[n \times batch\_size : (n+1) \times batch\_size]$
    - $\Theta^k_{epoch,n} \longleftarrow Adam(\Theta^k_{epoch,n-1}, batch\_x, batch\_y)$
  - **end**
  - $\Theta^k_{epoch} \longleftarrow \Theta^k_{epoch,int(N / batch\_size)}$
  - **if** $N \ \% \ batch\_size \neq 0$ **then**
    - $batch\_x = \mathscr{X}_e[N - (N \ \% \ batch\_size) : N]$
    - $batch\_y = \mathscr{Y}_e[N - (N \ \% \ batch\_size) : N]$
    - $\Theta^k_{epoch} \longleftarrow Adam(\Theta^k_{epoch}, batch\_x, batch\_y)$
  - **end**
  - **if** $|C(\Theta^k_{epoch}, \mathscr{X}_c|\mathscr{X}_e)$ - $C(\Theta^k_{epoch-1}, \mathscr{X}_c|\mathscr{X}_e)| \leq tol\_batch$ **then**
    - **break**
  - **else**
    - $epoch \longleftarrow epoch + 1$
  - **end**
- **end**
- **return** $\Theta^k = \Theta^k_{epoch}$

end

Algorithm 5: Cross Validation Check

Data: Cross-validation costs of estimated coefficients Θ^k

Result: *hh_end*

begin
- $y = \left(C(\Theta^{hh\_end-window}, \mathscr{X}_c|\mathscr{X}_e), \ldots, C(\Theta^{hh\_end}, \mathscr{X}_c|\mathscr{X}_e)\right)$
- $x = (hh\_end - window, \ldots, hh\_end)$
- **estimate** β *s.t.* $y = \alpha + \beta x$
- **if** $\beta \leq -tol\_batch$ *and* $hh\_end \leq hh\_max$ **then**
  - $hh\_end = hh\_end + 1$
- **end**
- **return** *hh_end*

end

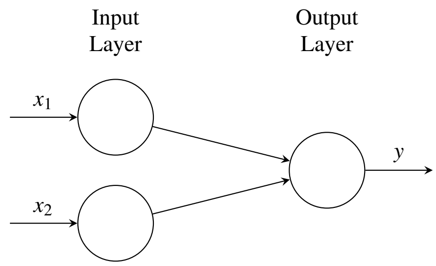
Algorithm 6: Penalty Method

Data: A vector of coefficients $\Theta_0^{k,s}$
Result: An updated vector of coefficients Θ^k
begin
  while $s \leq penalty\_limit$ **do**
    s = s + 1
    $\Theta_0^k = \Theta^{k,s-1}$
    begin Mini Batch optimization
      do ***Algorithm 4*** *Mini Batch optimization*
    end
    $\Theta^{k,s} = \Theta^k$
    if $\max|\Theta^{k,s} - \Theta^{k,s-1}| \leq tol\_coef$ **then**
      break
    end
  end
  return $\Theta^k = \Theta^{k,s}$
end

Table A.1 Codebook

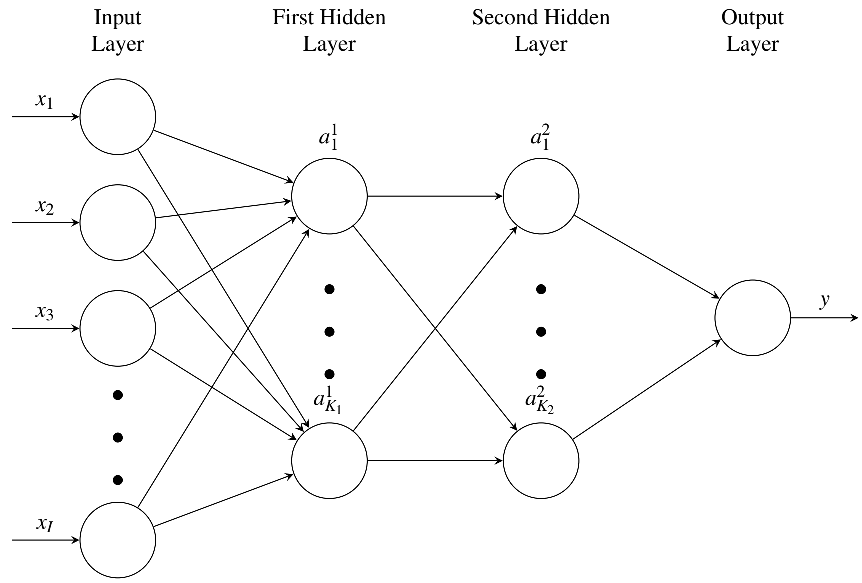
Variable	Definition
k	Number of nodes
tol_batch	Tolerance for mini-batch convergence
tol_stoc	Tolerance for stochastic convergence
stochastic_grad_epoch_limit	Max number of epochs for stochastic updates
mini_batch_grad_epoch_limit	Max number of epochs for mini-batch updates
batch_size	Sample size of each mini-batch
hh_start	Starting number of nodes in the hidden layer
window	Window size to calculate cross-validation tendency
hh_max	Max number of nodes for cross-validation
γ	Penalty weight
c	A coefficient to generate an increasing γ sequence
%	Remainder after division

Fig. A.1 OLS Network Diagram



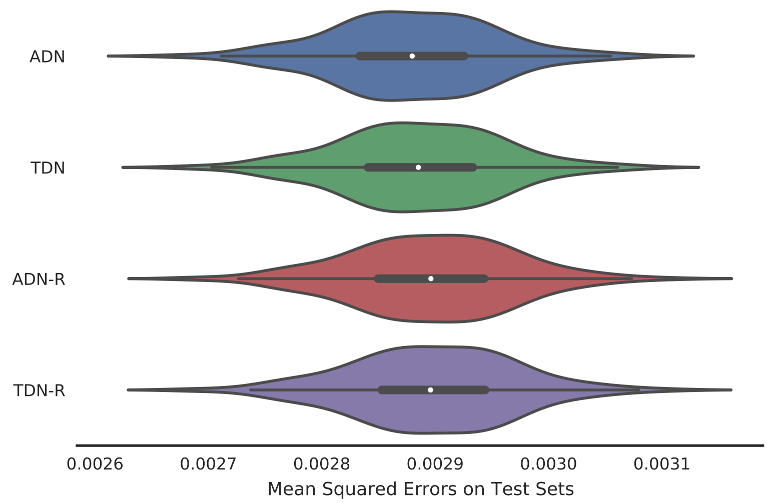
Note: The figure illustrates a network diagram that represents an OLS regression with two inputs. Each circle in the input layer is an input variable. Arrows represent coefficients that multiply the input variables: θ_1 and θ_2. The circle represents the predicted output $y = \theta_1 x_1 + \theta_2 x_2$.

Fig. A.2 Multilayer FNN Network Diagram



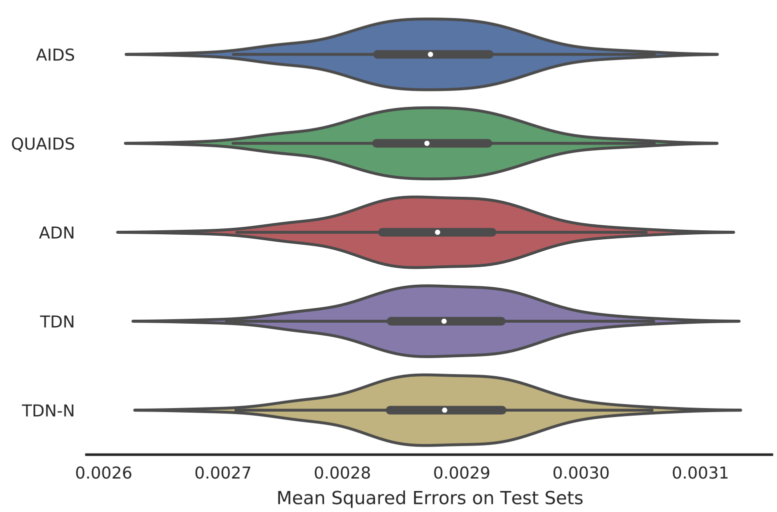
Note: The figure illustrates a network diagram that represents a four layer single output functional structure. Each circle in the input layer is an input variable. Arrows represent coefficients. The intermediate layers are called ”hidden layers”. The circle at the output layer represents the predicted output.

Fig. A.3 Distributions of Test Costs with Robustness Specification



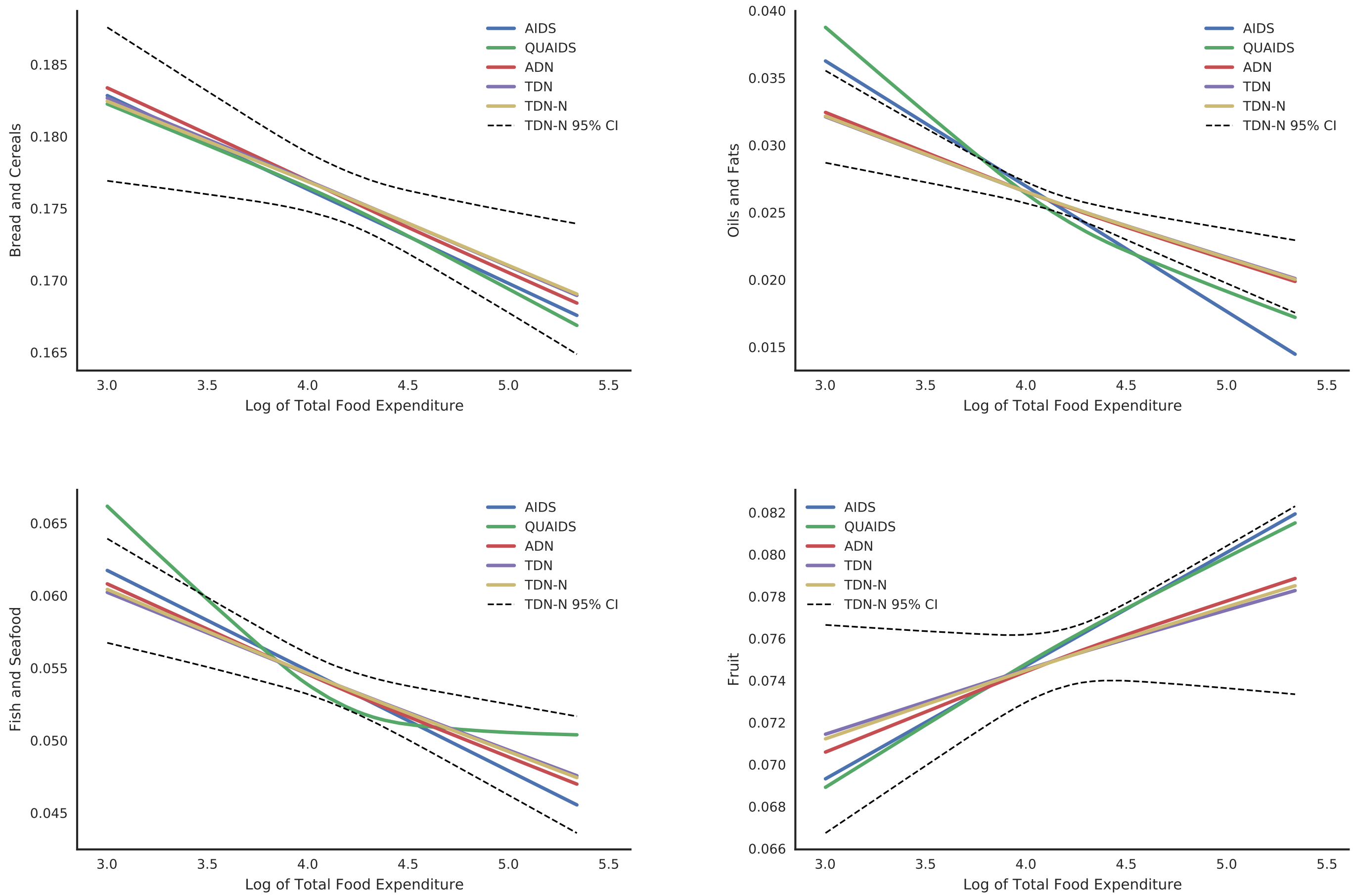
Note: The figure illustrates distributions of mean squared errors of four demand estimations on test sets, i.e., $C(\Theta, \mathscr{X}_t | \mathscr{X}_e, \mathscr{X}_c)$ for ADN, TDN, ADN-R (ADN with robustness test specification), and TDN-R (TDN with robustness test specification) estimates. Violin plots indicate similar generalization performances for all estimations. White mid-points are means of test cost distributions. Bold black lines are interquartile ranges. Thin black lines within violins are 95% confidence intervals. Black curves are kernel density estimates of test cost distributions.

Fig. A.4 Distributions of Test Costs with Negativity



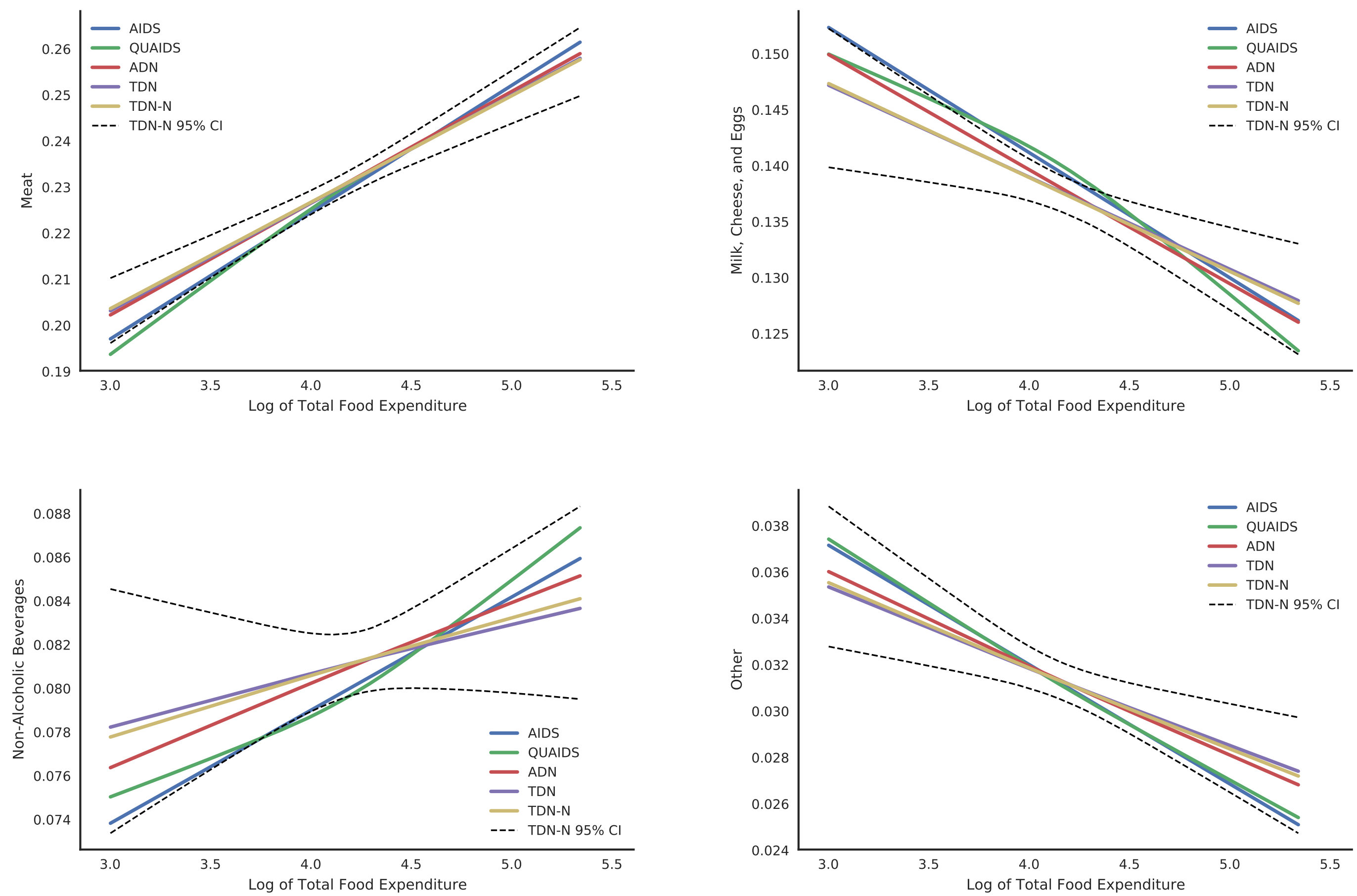
Note: The figure illustrates distributions of mean squared errors of five demand estimations on test sets, i.e., $C(\Theta, \mathscr{X}_t | \mathscr{X}_e, \mathscr{X}_c)$ for AIDS, QUAIDS, ADN, TDN (homogeneity, adding up, symmetry), and TDN-N (homogeneity, adding up, symmetry, negativity) estimates. Violin plots indicate similar generalization performances for all estimations. White mid-points are means of test cost distributions. Bold black lines are interquartile ranges. Thin black lines within violins are 95% confidence intervals. Black curves are kernel density estimates of test cost distributions.

Fig. A.5 Estimated Demand Functions with Negativity



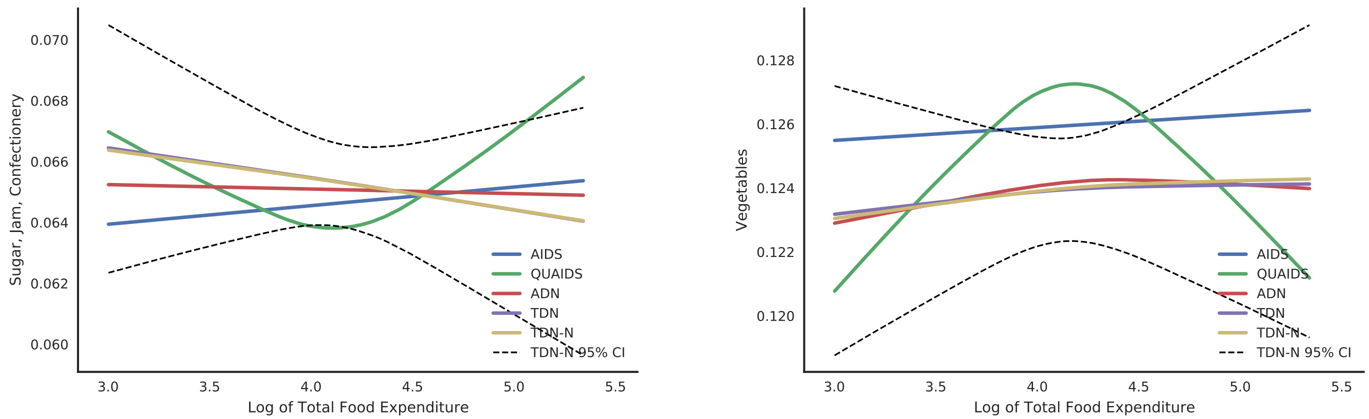
Continued on the next page.

Fig. A.5 Estimated Demand Functions with Negativity



Continued on the next page.

Fig. A.5 Estimated Demand Functions with Negativity



Note: Figures plot predicted budget shares of AIDS, QUAIDS, ADN, TDN, TDN-N estimations on y-axes, and log of total food expenditure on x-axes. Budget shares are estimated using all inputs from the main sample.

Appendix B
Appendix to Chapter 2

B.1 Additional Data Sets

To obtain mean years of education, we use the education data provided by the Human Development Report 2014 (UNDP, 2014). The data set provides the average number of years of education that is received by people ages 25 and older for 156 countries. The data is collected by UNESCO Institute for Statistics and prepared by using the methodology of Barro and Lee (2013).

As indicators of income inequality, we used the 2011 GINI indices as reported by the World Bank.[4] For the countries that do not have a 2011 GINI estimate, we use the closest GINI estimate within 5 years of 2011. In total, we have GINI indices for 126 countries.

We obtain the data for average country-wise temperatures from the World Bank's Climate Change Knowledge Portal,[5] which reports mean temperatures for 151 countries in our sample from 1961 to 1999.

Our institutional indicators are obtained from the Worldwide Governance Indicators database (see Kaufmann et al., 2010). We use 2011 estimates of five governance indicators as provided by the data set for 170 countries. These indicators are voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, control of corruption.

PWT 9.0 is explained in section 2.5.1.

[4] `http://databank.worldbank.org/data/reports.aspx?source=2&series=SI.POV.GINI&country=`

[5] `http://data.worldbank.org/data-catalog/cckp_historical_data`

B.2 Additional Sensitivity Analyzes

Nonparametric Estimation: Using Per Capita Real Income Levels

We estimate the relationship between price levels and per capita real incomes with a nonparametric estimation by using the following specification

$$\ln PL_i = f(\ln GDPP_i) + \varepsilon_i, \tag{B.1}$$

where $f(\cdot)$ is a smooth function, and the ε_i's are normally distributed disturbances with mean zero and standard deviation σ. The value of $f(\cdot)$ is estimated by using a kernel weighted local polynomial estimator of degree one (Yatchew, 2003). Epanechnikov kernel is used as the weighting function because it minimizes the mean squared error of the local polynomial estimator (Gijbels and Fan, 1996). The reported results are calculated by using the rule of thumb bandwidth.[6] However, our results are robust to the use of the cross validated bandwidth, and any other reasonable bandwidth in this regard.[7]

Log of the price level of China is estimated as -0.665 with a 95% confidence interval (-0.740, -0.590). Recalling that the observed log price level of China is -0.612, a nonparametric estimation suggests that the Chinese renminbi is not undervalued against the US dollar as of 2011. Hence the result that is suggested by the nonparametric fit supports the result provided by the quadratic functional form. Figure B.1 represents these results.

Controlling for Sectoral Labor Mobility

In order to avoid multicollinearity, which is caused by using correlated explanatory variables per capita real income and mean years of education (see figure B.4), we estimate a robust OLS regression of log price level on the first principle component of per capita real income and mean years of education[8] (see, e.g., Joliffe, 2002). Figure B.5 plots log price levels and the first principle component.

Table B.1, figures B.6 and B.9 summarize the results of OLS estimations. The coefficients of the principle component variables are positive and significant $(p-value < 0.001)$. A linear functional form suggests that the Chinese renminbi is undervalued by 5.4% whereas a quadratic functional form suggests that it is overvalued by 11%. A log-likelihood ratio test yields a test statistic of 50.78 which has a $\chi^2(1)$ distribution indicating significant rejection of the null hypothesis that the

[6] Rule of thumb bandwidth is calculated by obtaining a pilot estimate, and using statistics from this estimate to calculate the asymptotically optimal bandwidth.

[7] Cross validated bandwidth is the bandwidth that minimizes the mean squared error of the nonparametric estimation. For more information, see, e.g., Gijbels and Fan (1996).

[8] The first principle component explains 89.94% of the variation in both variables.

coefficient of the quadratic term is zero ($p - value < 0.001$), and the AIC and BIC of the log quadratic model are again lower (see table B.2).

Figure B.8 and figure B.11 present residual plots of regressions using both functional forms. The log linear functional form suggests a nonlinear relationship between the regressor and the residuals whereas the log quadratic functional form suggests no significant relationship.

A nonparametric estimation using the first principle component of per capita real income and mean years of education as the explanatory variable estimates log of the price level for China as -0.678 with a 95% confidence interval (-0.753, -0.603). Therefore, a nonparametric estimation controlling for sectoral labor mobility suggests that the Chinese renminbi is not undervalued against the US dollar as of 2011. Figure B.12 represents these results.

Controlling for Net Foreign Asset Positions

Table B.3 summarizes the results of the robust OLS estimations using both functional forms. nfa_i is significant in both estimations. However, the log linear functional form still reveals undervaluation whereas the log quadratic functional form does not reveal any evidence of undervaluation. A log-likelihood ratio test reports the test statistic as 13.75 which has a $\chi^2_{(1)}$ distribution with a p-value indicating significant rejection of the null hypothesis that the coefficient of the quadratic term is zero ($p - value < 0.001$). Table B.4 shows that the log quadratic functional form has the lower AIC and BIC values. Therefore, again, the log quadratic functional form is the preferred functional form, and we find no evidence of renminbi undervaluation.

Exclusion of the Top Oil Exporting Economies

The top oil exporting countries are selected as the top 15 oil exporters according to The United States Energy Information Administration (The United States Energy Information Administration, 2011). The list of these countries is in appendix B.3.

Table B.5 summarizes the results of the robust OLS estimations using both functional forms. All coefficients are significant. The log linear functional form suggests that the renminbi is undervalued whereas the log quadratic functional form suggests that the renminbi is not undervalued. A log-likelihood ratio test reports the test statistic as 27.75 which has a $\chi^2_{(1)}$ distribution indicating significant rejection of the null hypothesis that the coefficient of the quadratic term is zero ($p - value < 0.001$). Table B.6 shows that the log quadratic functional form has the lower AIC and BIC values. Therefore our choice of the functional form and the misalignment result are robust to the exclusion of the top oil exporters.

B.3 Lists of Countries

Countries in the main sample: Albania, Algeria, Angola, Anguilla, Antigua and Barbuda, Armenia, Aruba, Australia, Austria, Azerbaijan, the Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bermuda, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Cape Verde, Cayman Islands, Central African Republic, Chad, Chile, China, Colombia, Comoros, Democratic Republic of the Congo, Republic of the Congo, Costa Rica, Croatia, Curacao, Cyprus, Czech Republic, Cote d'Ivoire, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Fiji, Finland, France, Gabon, the Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Republic of Korea, Kuwait, Kyrgyzstan, Lao PDR, Latvia, Lesotho, Liberia, Lithuania, Luxembourg, Macao, Macedonia, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Montenegro, Montserrat, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Palestinian Territory, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Singapore, Sint Maarten, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and Grenadines, Sudan, Suriname, Swaziland, Sweden, Switzerland, Sao Tome and Principe, Taiwan, Tajikistan, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Turks and Caicos Islands, Uganda, Ukraine, the United Arab Emirates, the United Kingdom, the United States, Uruguay, Venezuela, Vietnam, Virgin Islands, Yemen, Zambia

Countries without education data: Anguilla, Aruba, the Bahamas, Bermuda, Cape Verde, Cayman Islands, Curacao, Iran, Lao PDR, Macao, Macedonia, Montserrat, Sint Maarten, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sao Tome and Principe, Taiwan, Turks and Caicos, Virgin Islands

Countries without gini data: Algeria, Anguilla, Antigua and Barbuda, Aruba, the Bahamas, Bahrain, Barbados, Belize, Bermuda, Brunei Darussalam, Cape Verde, Cayman Islands, Comoros, Curacao, Dominica, Egypt, Equatorial Guinea, Gabon, the Gambia, Ghana, Grenada, Hong Kong, Iraq, Jamaica, Jordan, Kenya, Republic of Korea, Kuwait, Macao, Malta, Montserrat, Myanmar, New Zealand, Oman, Palestinian Territory, Qatar, Saudi Arabia, Singapore, Sint Maarten, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sao Tome and Principe, Taiwan, Trinidad and Tobago, Turks and Caicos Islands, the United Arab Emirates, Virgin Islands, Yemen

Countries without temperature data: Anguilla, Antigua and Barbuda, Aruba, Bahrain, Barbados, Bermuda, Cape Verde, Cayman Islands, Curacao, Dominica, Grenada, Hong Kong, Macao, Maldives, Malta, Montserrat, Palestinian Territory, Seychelles, Singapore, Sint Maarten, St. Kitts and Nevis, St. Lucia, Sao Tome and Principe, Taiwan, Turks and Caicos Islands

Countries without institutional data: Cape Verde, Curacao, Montserrat, Palestinian Territory, Sint Maarten, Turks and Caicos Islands

Top oil exporting countries: Algeria, Angola, Canada, Iran, Iraq, Kazakhstan, Kuwait, Mexico, Nigeria, Norway, Qatar, Russia, Saudi Arabia, the United Arab Emirates, Venezuela

Table B.1 OLS Estimation Results Controlling Sectoral Labor Mobility

	Log Linear	Log Quadratic
PC_i	0.207^{***}	0.258^{***}
PC_i^2		0.090^{***}
Constant	-0.546^{***}	-0.708^{***}
$\ln \hat{PL}_{China}$	-0.558	-0.722
95% CI	$(-0.603, -0.513)$	$(-0.777, -0.667)$
% misalignment	-5.4%	11%
R^2	0.480	0.625
Root MSE	0.289	0.247
Observations	156	156

***significant at 1%.

Note: PC_i stands for the first principle component value for country i. $\hat{PL}_{China}$ is the estimated price level of China. % misalignment is calculated as the difference between the log of the observed price level of China, which is -0.612, and the log of $\hat{PL}_{China}$ if the difference is statistically significant.

Table B.2 Model Selection Criteria Controlling Sectoral Labor Mobility

Model	AIC	BIC
Log Linear	57.722	63.821
Log Quadratic	8.937	18.087

Note: AIC stands for Akaike information criterion, and BIC stands for Bayesian information criterion. Both criteria suggest the functional form with the lowest score as the preferred functional form.

Table B.3 OLS Estimation Results Controlling NFA Positions

	Log Linear	Log Quadratic
$\ln GDPP_i$	0.195^{***}	-0.670^{**}
$\ln GDPP_i^2$		0.048^{**}
nfa_i	$-4.04e-16^{***}$	$-2.99e-16^{**}$
Constant	-2.340^{***}	1.462
$\ln \hat{PL}_{China}$	-0.558	-0.630
95% CI	$(-0.604, -0.512)$	$(-0.685, -0.575)$
% misalignment	-10.7%	No Misalignment
R^2	0.422	0.471
Root MSE	0.288	0.276
Observations	154	154

***,** significant at 1%, 5%, respectively.

Note: $GDPP_i$ stands for the country i's per capita real income, and nfa_i stands for the country i's net foreign asset position. $\hat{PL}_{China}$ is the estimated price level of China. % misalignment is calculated as the difference between the log of the observed price level of China, which is -0.612, and the log of $\hat{PL}_{China}$ if the difference is statistically significant..

Table B.4 Model Selection Criteria Controlling NFA Positions

Model	AIC	BIC
Log Linear	54.332	60.406
Log Quadratic	42.580	51.691

Note: AIC stands for Akaike information criterion, and BIC stands for Bayesian information criterion. Both criteria suggest the functional form with the lowest score as the preferred functional form.

Table B.5 OLS Estimation Results Excluding Top Oil Exporters

	Log Linear	Log Quadratic
$\ln GDPP_i$	0.226^{***}	-1.080^{**}
$\ln GDPP_i^2$		0.073^{***}
Constant	-2.604^{***}	3.104^{**}
$\ln \hat{PL}_{China}$	-0.520	-0.629
95% CI	$(-0.565, -0.474)$	$(-0.687, -0.571)$
% misalignment	-9.2%	No Misalignment
R^2	0.479	0.561
Root MSE	0.293	0.270
Observations	161	161

***,** significant at 1%, 5%, respectively.

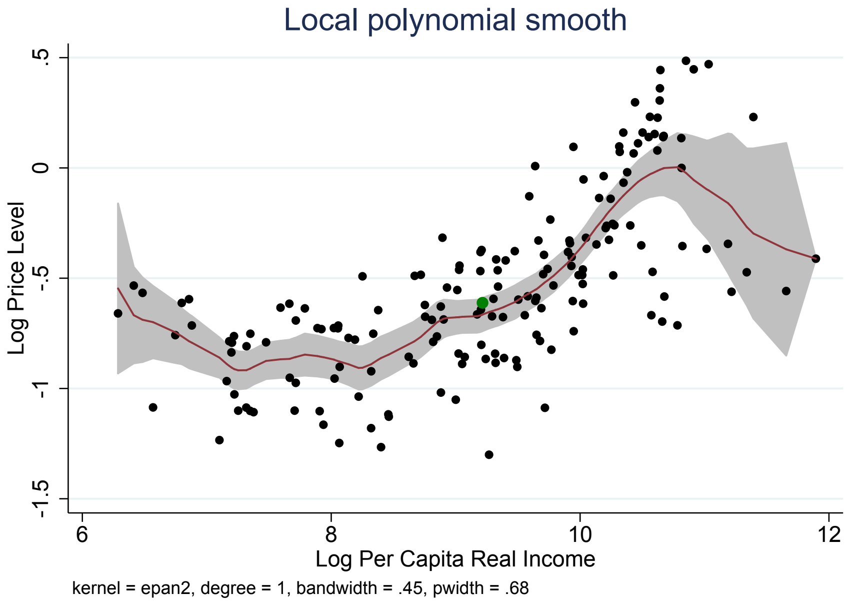
Note: $GDPP_i$ stands for the country i's per capita real income. $\hat{PL}_{China}$ is the estimated price level of China. % misalignment is calculated as the difference between the log of the observed price level of China, which is -0.612, and the log of $\hat{PL}_{China}$ if the difference is statistically significant.

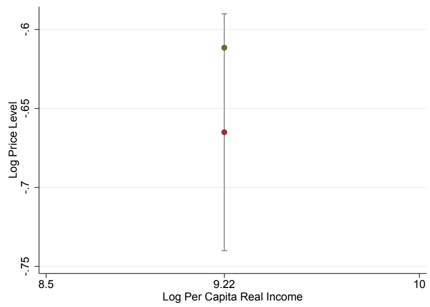
Table B.6 Model Selection Criteria Excluding Oil Exporting Economies

Model	AIC	BIC
Log Linear	63.441	69.604
Log Quadratic	37.692	46.936

Note: AIC stands for Akaike information criterion, and BIC stands for Bayesian information criterion. Both criteria suggest the functional form with the lowest score as the preferred functional form.

Fig. B.1 Nonparametric Estimation Results with Income Data





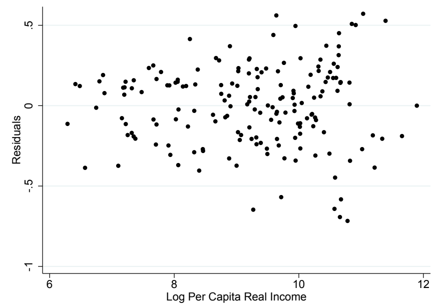
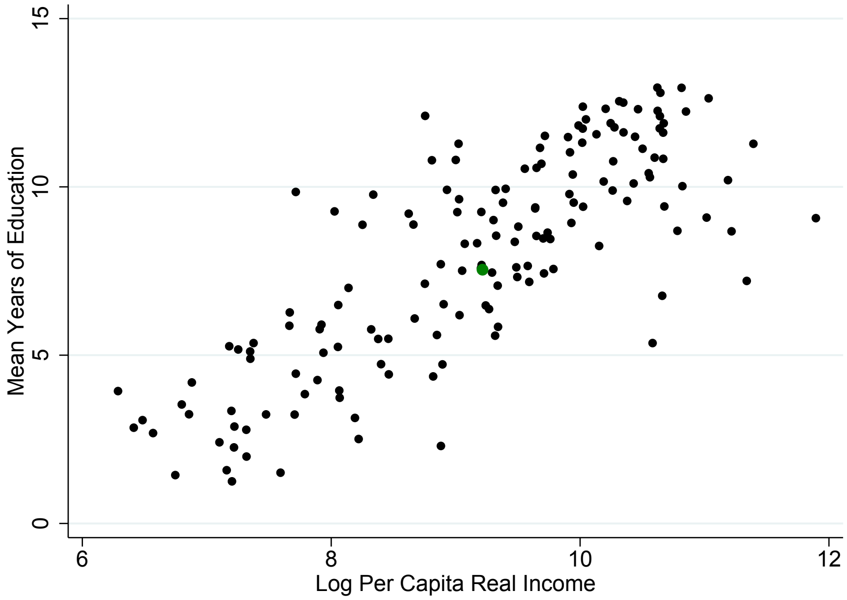


Fig. B.2 Nonparametric - China

Fig. B.3 Nonparametric - Residuals

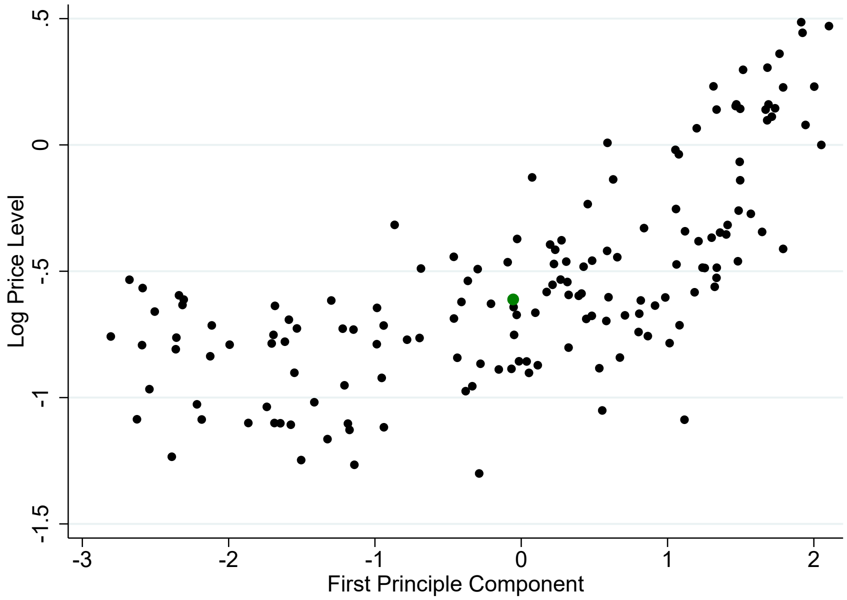
Note: Figure B.1 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure B.2 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure B.3 plots the residuals from the nonparametric estimation against the regressor.

Fig. B.4 Educational Attainment and Per Capita Real Income



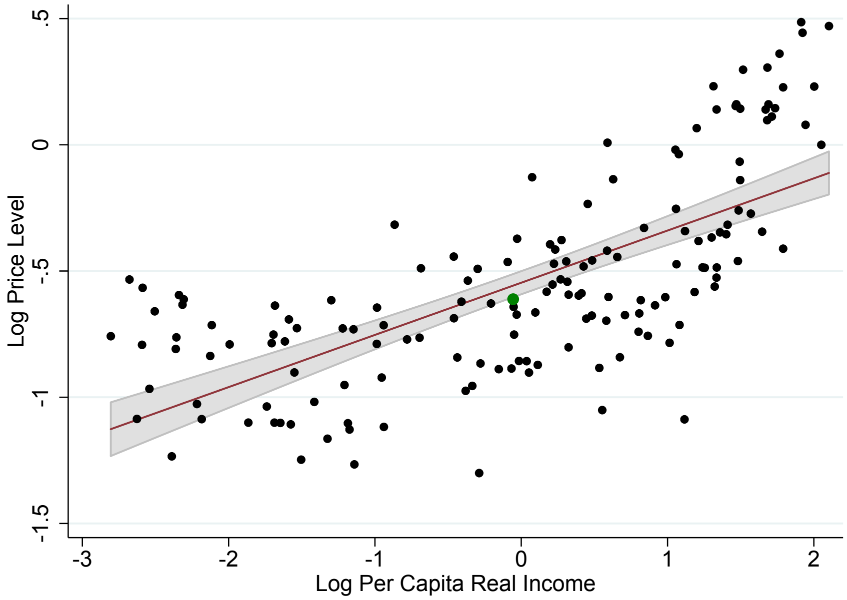
Note: The mean years of education is plotted on the vertical axis, and log per capita real income is plotted on the horizontal axis. China is represented by a green data point.

Fig. B.5 Price Levels and First Principle Component



Note: Log price levels are plotted on the vertical axis, and values of the first principle component of per capita real incomes and mean years of education are plotted on the horizontal axis. China is represented by a green data point.

Fig. B.6 OLS Estimation Controlling Sectoral Labor Mobility - Log Linear Functional Form



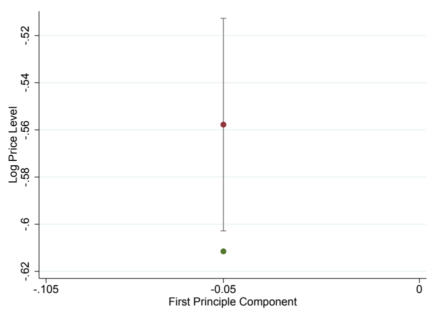


Fig. B.7 Log Linear - China

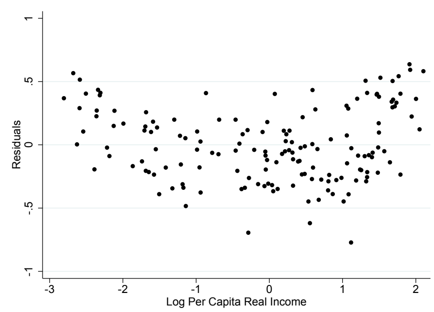
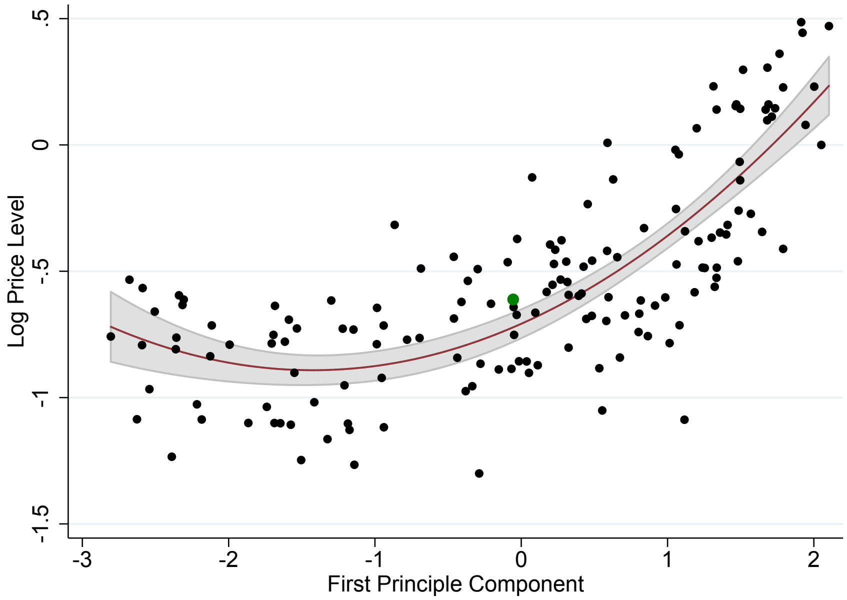
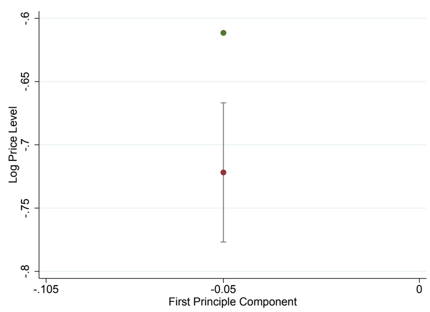


Fig. B.8 Log Linear - Residuals

Note: Figure B.6 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure B.7 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure B.8 plots the residuals from the OLS estimation using the linear functional form against the regressor.

Fig. B.9 OLS Estimation Controlling Sectoral Labor Mobility - Log Quadratic Functional Form





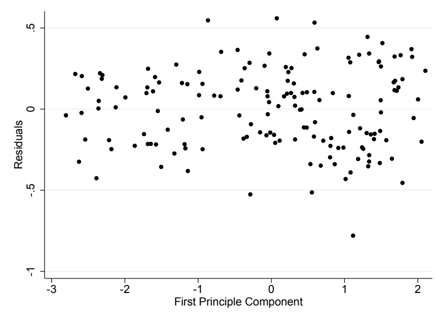
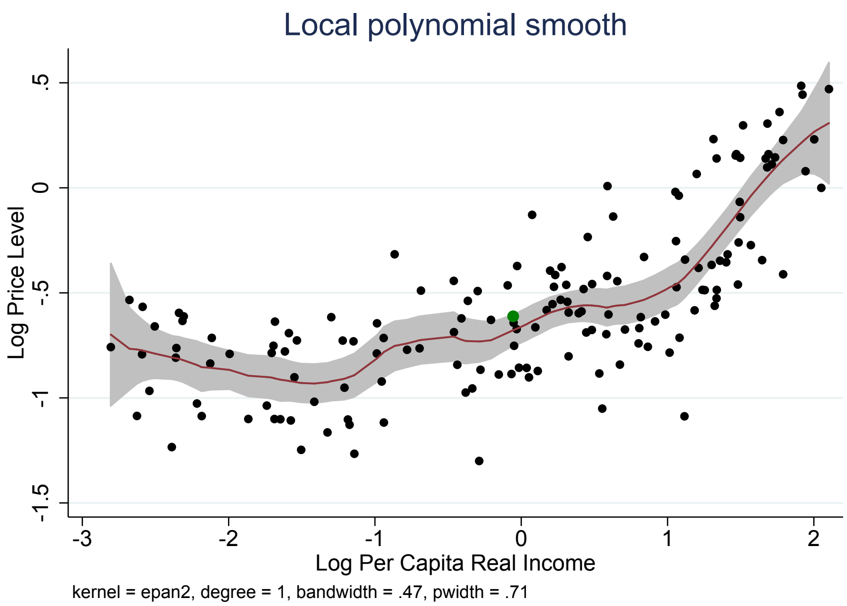


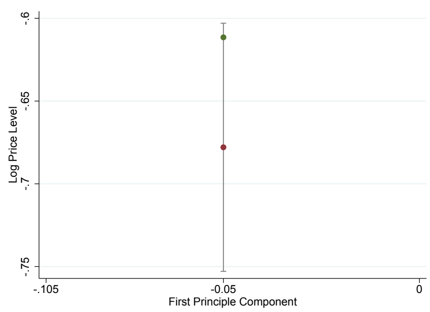
Fig. B.10 Log Quadratic - China

Fig. B.11 Log Quadratic - Residuals

Note: Figure B.9 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure B.10 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure B.11 plots the residuals from the OLS estimation using the quadratic functional form against the regressor.

Fig. B.12 Nonparametric Estimation Controlling Sectoral Labor Mobility





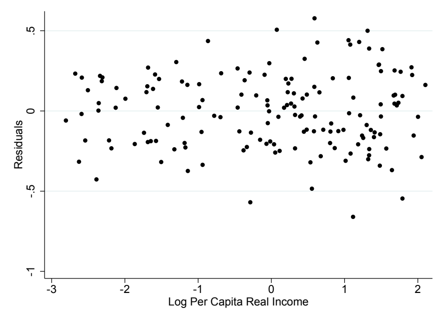


Fig. B.13 Nonparametric Estimation - China **Fig. B.14** Nonparametric Estimation - Residuals

Note: Figure B.12 shows the estimated price levels in red, and the 95% confidence interval in gray. The observed price level of China is represented by the green data point. Figure B.13 emphasizes the observed price level of China, the estimated price level of China, and the confidence interval. Figure B.14 plots the residuals from the nonparametric estimation against the regressor.

Appendix C
Appendix to Chapter 3

C.1 Common Multilateral Indices

Gini-Éltetö-Köves-Szulc (GEKS)

The GEKS method, as used in the International Comparison Program (ICP) of the World Bank, is a multilateral conversion to symmetric bilateral Fisher indices. It is a geometric mean of all possible bilateral comparisons taking each country as the base:

$$S_{h,f}^{GEKS} = \left(\prod_{c=1}^{C} \Pi_{h,c} \Pi_{c,f} \right)^{1/C}.$$

Consistency of the Fisher index as an approximation to the behavior of an optimizing individual when preferences are homogeneous and homothetic does not extend to the GEKS. A special case where GEKS is also consistent, and is equal to the Fisher index, is when such preferences are represented by a homogeneous quadratic utility function (Neary, 2004).

Geary-Khamis (GK)

The GK method computes purchasing power parities as the ratio of the value of a fixed basket of goods at hypothetical world prices and at observed local prices. The world prices of G goods, and the exchange rates for C countries are the solutions to the following system of equations:

$$S_{h,f}^{GK} = \frac{\mathbf{p}_0 \mathbf{Q}_h^h}{\mathbf{p}_h \mathbf{Q}_h^h}$$
$$p_0^g = \frac{\sum_{c=1}^{C} GK_c p_c^g q_c^{c,g}}{\sum_{c=1}^{C} q_c^{c,g}}$$

where $\mathbf{Q}_h^h$ is a vector of aggregate quantities for country h, p_c^g and $q_c^{c,g}$ are price and quantity of commodity g in country c, and $\mathbf{p}_0$ is a vector of world prices.

Solving simultaneously for p_c^g and GK_c, one can calculate each country's purchasing power parity and the world prices. Thus, purchasing power parities for bilateral and multilateral comparisons are calculated simultaneously. GK method is consistent with the economic theory if preferences are Leontief, i.e., if there is no substitution away from more expensive goods, or vice versa.

Geary-Allen International Accounts (GAIA)

Neary (2004) suggests a GK-variant formula that allows for substitution between consumption goods. The preferences for a reference consumer who mimics world consumption patterns are estimated using pooled country-level data for preferences that exhibit generalized linearity. Using these preferences, hypothetical consumption levels in world prices are estimated at a fixed price vector of world prices.

Given a reference consumer who minimizes an expenditure function $e(\mathbf{p}, u)$, price indices S_h^{GAIA} are calculated as the solution to the following system of equations:

$$S_h^{GAIA} = \frac{e(\mathbf{p}_0, u^h)}{e(\mathbf{p}_h, u^h)}$$
$$p_0^g = \frac{\sum_{c=1}^{C} S_c^{GAIA} p_c^g q_c^{c,g}}{\sum_{c=1}^{C} q_0^{c,g}}$$

where $q_0^{c,g} = \partial e(\mathbf{p}_0, u^h) \big/ \partial p^g$ by Shephard's Lemma. The corresponding Geary-Allen indices of real incomes are the ratios of the minimum expenditure levels of the reference consumer in the living standards of each country:

$$\mathcal{Q}_{h,f}^{GAIA} = \frac{e(\mathbf{p}_0, u^f)}{e(\mathbf{p}_0, u^h)}.$$

Therefore, contrary to the previous methods, GAIA is an economic approach to price comparisons by building a link between how an individual responds to price differentials. However, it evaluates purchasing power parities for a single type of preferences and at a single income level. Hence, it is consistent with optimizing individual behavior if preferences are homogeneous and homothetic. Moreover, use of a reference world price vector and homogeneous preferences make GAIA a symmetric and transitive index.

C.2 Proofs

Consistency with Index Axioms

For $\Pi^L_{h,f} = \Pi^L(\mathbf{p}_f, \mathbf{p}_h, \mathbf{Q}^h_h)$:

1. $\partial \Pi^L_{f,h}/\partial p^g_h > 0$ and $\partial \Pi^L_{f,h}/\partial p^g_h < 0$ follows from the definition of the index.
2. If the away prices are $\alpha \mathbf{p}_f$, then

$$\Pi^L(\alpha \mathbf{p}_f, \mathbf{p}_h, \mathbf{Q}^h_h) = \frac{\alpha \mathbf{p}^T_f \mathbf{Q}^h_h}{\mathbf{p}^T_h \mathbf{Q}^h_h}.$$

   Therefore, $\Pi^L(\alpha \mathbf{p}_f, \mathbf{p}_h, \mathbf{Q}^h_h) = \alpha \Pi^L(\mathbf{p}_f, \mathbf{p}_h, \mathbf{Q}^h_h)$.
3. Follows from the definition.
4. If away prices are $\alpha \mathbf{p}_f$ and home prices are $\alpha \mathbf{p}_h$, then

$$\Pi^L(\alpha \mathbf{p}_f, \alpha \mathbf{p}_h, \mathbf{Q}^h_h) = \frac{\alpha \mathbf{p}^T_f \sum^{I_h}_{i_h} \mathbf{Q}^h_h}{\alpha \mathbf{p}^T_h \sum^{I_h}_{i_h} \mathbf{Q}^h_h},$$

   and $\Pi^L(\alpha \mathbf{p}_f, \alpha \mathbf{p}_h, \mathbf{Q}^h_h) = \Pi^L(\mathbf{p}_f, \mathbf{p}_h, \mathbf{Q}^h_h)$.
5. By multiplying units of goods with a diagonal matrix Γ with elements γ^g, we have

$$\Pi^L(\Gamma \mathbf{p}_f, \Gamma \mathbf{p}_h, \Gamma^{-1} \mathbf{q}_h) = \frac{\sum^{I^h}_{i_h=1} \sum^G_{g=1} \gamma^g p^g_f q^{i^h,g}_h / \gamma^g}{\sum^{I^h}_{i_h=1} \sum^G_{g=1} \gamma^g p^g_h q^{i^h,g}_h / \gamma^g}.$$

   Then $\Pi^L(\Gamma \mathbf{p}_f, \Gamma \mathbf{p}_h, \Gamma^{-1} \mathbf{Q}^h_h) = \Pi^L(\mathbf{p}_f, \mathbf{p}_h, \mathbf{q}_h)$.

   For $\Pi^{CU}_{h,f}$ and $\Pi^{CS}_{h,f}$:

1. Both properties follow from the properties of an expenditure function with respect to prices and definitions of these indices.
2. Follows from the homogeneity of an expenditure function in all prices.
3. For axiom 3, we evaluate each index separately:

   a. $\Pi^{CU}_{h,f}$: If the foreign price vector is equal to the home price vector, i.e., $\mathbf{p}_f = \mathbf{p}_h = \mathbf{p}$, then by definition, $CV^{i_h} = e^{i_h}(\mathbf{p}, u^{i_h}) - e^{i_h}(\mathbf{p}, u^{i_h}) = 0, \forall i$. Then,

$$\Pi^{CU}_{h,f} = 1 + \frac{\sum^{I_h}_{i_h} CV^{i_h}}{\sum^{I_h}_{i_h} m^{i_h}_h} = 1$$

   b. $\Pi^{CS}_{h,f}$: If $\mathbf{p}_h = \mathbf{p}_f = \mathbf{p}$, then

$$B^h(\mathbf{p}, \mathbf{m}^h_h) = B^h(\mathbf{p}, \Pi^{CS}_{h,f} \mathbf{m}^h_h)$$

   holds only if $\Pi^{CS}_{h,f} = 1$.

4. Follows from the homogeneity of an expenditure function in all prices.
5. Follows from the definition of an expenditure function.

Proof of Theorem 1

By definition of $\Pi^L_{h,f}$, social cost of society h in prices of society f is

$$M_f^{h,L} = \sum_{i_h=1}^{I_h} \pi_{h,f}^{i_h,L} m_h^{i_h}.$$

By definition of $\Pi^{CU}_{h,f}$, social cost is

$$\begin{aligned} M_f^{h,CU} &= \sum_{i_h=1}^{I_h} e^{i_h}\left(\mathbf{p}_f, v^{i_h}(\mathbf{p}_h, m_h^{i_h})\right) \\ &= \sum_{i_h=1}^{I_h} e^{i_h}\left(\mathbf{p}_f, u_h^{i_h}\right) \\ &= \sum_{i_h=1}^{I_h} \pi_{h,f}^{i_h,K} m_h^{i_h} \end{aligned}$$

From lemma 1 and monotonicity of individual utilities in individual costs, $\pi_{h,f}^{i_h,L} \geq \pi_{h,f}^{i_h,K}$ holds for all i_h. Therefore, $M_f^{h,L} \geq M_f^{h,CU}$, which trivially implies $\Pi^L_{h,f} \geq \Pi^{CU}_{h,f}$.

Proof of Theorem 2

We can decompose the change in the utility level of an individual i with respect to prices and budget as

$$\mathrm{d}u^{i_h} = \sum_{g=1}^{G} \frac{\partial v^{i_h}(\mathbf{p}_h, m_h^{i_h})}{\partial p_g} \mathrm{d}p_g + \frac{\partial v^{i_h}(\mathbf{p}_h, m^{i_h})}{\partial m} \mathrm{d}m^{i_h}.$$

If preferences are identical and homothetic, then $v^{i_h}(\mathbf{p}_h, m^{i_h}) = \alpha(\mathbf{p}_h) m^{i_h}, \forall i_h$, where $\alpha(\cdot)$ is a homogeneous function. At constant utilities, i.e., when $\mathrm{d}u^{i_h} = 0$ for all individuals i_h,

$$\mathrm{d}m^{i_h} = -\left(m^{i_h} \sum_{g=1}^{G} \frac{\partial \alpha(\mathbf{p}_h)}{\partial p_g} \mathrm{d}p_g \right) \Big/ \alpha(\mathbf{p}_h)$$

holds, which implies $\mathrm{d}m^{i_h} = \gamma m^{i_h}, \forall i_h$. Hence, constant utility changes in budgets are proportionate to actual budget levels. Therefore, social cost level in a foreign price vector $\mathbf{p}_f$ is $M^h_f = (1+\gamma)M^h_h$, which implies $\Pi^{CV}_{h,f} = \Pi^{CS}_{h,f} = 1+\gamma$.

Proof of Proposition 1

S^{CGW} is symmetric if $S^{CGW}_{h,f} \times S^{CGW}_{f,h} = 1$. By definition,

$$S^{CGW}_{h,f} = \frac{\sum_{c=1}^{C} \Pi_{c,f} M^c_c}{\sum_{c=1}^{C} \Pi_{c,h} M^c_c}$$

and

$$S^{CGW}_{f,h} = \frac{\sum_{c=1}^{C} \Pi_{c,h} M^c_c}{\sum_{c=1}^{C} \Pi_{c,f} M^c_c}$$

where $\Pi = \left\{\Pi^L, \Pi^{CU}, \Pi^{CS}\right\}$. Then, $S^{CGW}_{h,f} \times S^{CGW}_{f,h} = 1$ holds.

For transitivity, $S^{CGW}_{h,k} \times S^{CGW}_{k,f} = S^{CGW}_{k,f}$. By definition,

$$\begin{aligned} S^{CGW}_{h,k} \times S^{CGW}_{k,f} &= \frac{\sum_{c=1}^{C} \Pi_{c,k} M^c_c}{\sum_{c=1}^{C} \Pi_{c,h} M^c_c} \times \frac{\sum_{c=1}^{C} \Pi_{c,f} M^c_c}{\sum_{c=1}^{C} \Pi_{c,k} M^c_c} \\ &= \frac{\sum_{c=1}^{C} \Pi_{c,f} M^c_c}{\sum_{c=1}^{C} \Pi_{c,h} M^c_c}. \end{aligned}$$

Therefore S^{CGW} is transitive.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org

Abramovsky L, Attanasio O, Phillips D (2015) Value Added Tax Policy and The Case for Uniformity: Empirical Evidence from Mexico. IFS Working Paper W15/08, Institute for Fiscal Studies

Aguiar M, Bils M (2015) Has Consumption Inequality Mirrored Income Inequality? American Economic Review 105(9):2725–2756

Almås I (2012) International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food. American Economic Review 102(2):1093–1117

Altman EI, Marco G, Varetto F (1994) Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience). Journal of Banking and Finance 18(3):505–529

Artuc E, Lederman D, Porto G (2015) A Mapping of Labor Mobility Costs in the Developing World. Journal of International Economics 95(1):28–41

Asian Development Bank (2014) Purchasing Power Parities and Real Expenditures. Asian Development Bank, Mandaluyong City, Philippines

Athey S, Imbens GW (2017) The State of Applied Econometrics: Causality and Policy Evaluation. Journal of Economic Perspectives 31(2):3–32

Bajari P, Nekipelov D, Ryan SP, Yang M (2015) Machine Learning Methods for Demand Estimation. American Economic Review: Papers & Proceedings 105(5):481–485

Balassa B (1964) The Purchasing-Power Parity Doctrine: A Reappraisal. Journal of Political Economy 72(6):584–596

Balk BM (2008) Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference. Cambridge University Press

Banks J, Blundell R, Lewbel A (1997) Quadratic Engel Curves and Consumer Demand. The Review of Economics and Statistics 79(4):527–539

Barnett WA, Seck O (2008) Rotterdam Model versus Almost Ideal Demand System: Will the Best Specification Please Stand Up? Journal of Applied Econometrics 23:795–824

Barnett WA, Serletis A (2008) Consumer Preferences and Demand Systems. Journal of Econometrics 147:210–224

Barro RJ, Lee JW (2013) A New Data Set of Educational Attainment in the World, 1950-2010. Journal of Development Economics 104:184–198

Barron AR (1994) Approximation and Estimation Bounds for Artificial Neural Networks. Machine Learning 14(1):115–133

Bénassy-Quéré A, Duran-Vigneron P, Lahréche-Revil A, Mignon V (2004) Burden Sharing and Exchange-Rate Misalignments within the Group of Twenty. Working Papers 2004-13, CEPII research center

Blundell R (1988) Consumer Behavior: Theory and Empirical Evidence - A Survey. The Economic Journal 98(389):16–65

Blundell R, Robin JM (1999) Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems. Journal of Applied Econometrics 14(3):209–232

Blundell R, Horowitz JL, Parey M (2012) Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation. Quantitative Economics 3(1):29–51

Blundell R, Horowitz J, Prey M (2016) Nonparametric Estimation of a Nonseparable Demand Function under the Slutsky Inequality Restriction. The Review of Economics and Statistics 99(2):291–304

Cardell NS, Joerding WH, Li Y (1995) Symmetry Constraints for Feedforward Network Models of Gradient Systems. IEEE Transactions on Neural Networks 6(5):1249–1254

Cavallo A (2017) Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers. American Economic Review 107(1):283–303

Cavallo A, Rigobon R (2016) The Billion Prices Project: Using Online Prices for Measurement and Research. Journal of Economic Perspectives 30(2):151–178

Cavallo A, Diewert WE, Feenstra RC, Inklaar R, Timmer MP (2018) Using Online Prices for Measuring Real Consumption Across Countries. NBER Working Paper 24292, National Bureau of Economic Research

Cheung YW, Chinn M, Fujii E (2007) The Overvaluation of Renminbi Undervaluation. Journal of International Money and Finance 26(5):762–785

Cheung YW, Chinn M, Nong X (2017) Estimating Currency Misalignment Using the Penn Effect: It is not As Simple As It Looks. International Finance 20(3):222–242

Choromanska A, Henaff M, Mathieu M, Arous GB, LeCun Y (2015) The Loss Surfaces of Multilayer Networks. Journal of Machine Learning Research 38:192–204

Christensen LR, Jorgensen DW, J LL (1975) Trancendental Logarithmic Utility Functions. American Economic Review 65(3):367–383

Cline WR, Williamson J (2008) Estimates of the Equilibrium Exchange Rate of the Renminbi: Is There a Consensus and If Not, Why Not? In: Debating China's Exchange Rate Policy, Peterson Institute for International Economics, Washington, DC, pp. 131–154

Cline WR, Williamson J (2010) Notes on Equilibrium Exchange Rates: January 2010. Policy Brief 10-2, Peterson Institute for International Economics, Washington, DC

Coudert V, Couharde C (2009) Currency Misalignments and Exchange Rate Regimes in Emerging and Developing Countries. Review of International Economics 17(1):121–136

Crossley TF, Pendakur K (2010) The Common-Scaling Social Cost-of-Living Index. Journal of Business and Economic Statistics 28(4):523–538

Cui Y (2013) How Is the RMB Exchange Rate Misaligned? A Recent Application of Behavioral Equilibrium Exchange Rate (BEER) to China. Journal of East Asian Economic Integration 17(3):281–310

Cybenko G (1989) Approximation by Superpositions of a Sigmoidal Function. Mathematics of Control, Signals, and Systems 2(4):303–314

d'Aspermont C, Gevers L (1977) Equity and the Informational Basis of Collective Choice. Review of Economic Studies 44(2):199–209

Deaton A (1997) The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. The Johns Hopkins University Press, Baltimore, Maryland

Deaton A, Aten B (2017) Trying to Understand the PPPs in ICP 2011: Why Are the Results So Different? American Economic Journal: Macroeconomics 9(1):243–264

Deaton A, Heston A (2010) Understanding PPPs and PPP-Based National Accounts. American Economic Journal: Macroeconomics 2(4):1–35

Deaton A, Muellbauer J (1980a) An Almost Ideal Demand System. American Economic Review 70(3):312–326

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

Diewert WE (1976) Exact and Superlative Index Numbers. Journal of Econometrics 4(2):115–145

Diewert WE, Nakamura AO (eds) (1993) Essays in Index Number Theory Volume 1, Contributions to Economic Analysis, vol 217. Emerald Group Publishing Limited

Diewert WE, Avriel M, Zang I (1981) Nine Kinds of Quasiconcavity and Concavity. Journal of Economic Theory 25(3):397–420

Donaldson D, Pendakur K (2012) Index-Number Tests and the Common-Scaling Social Cost-of-Living Index. Social Choice and Welfare 38(3):407–429

Eichhorn W (1978) Functional Equations in Economics. Addison-Wesley, Reading, MA

Eichhorn W, Voeller J (1976) Theory of the Price Index: Fisher's Test Approach and Generalizations, Lecture Notes in Economics and Mathematical Systems, vol 140. Springer-Verlag, Berlin, Heidelberg

Elbadawi I, Gallant AR, Souza G (1983) An Elasticity can be Estimated Consistently without a Priori Knowledge of Functional Form. Econometrica 51(6):1731–1751

Feenstra RC, Ma H, Neary PJ, Prasada Rao DS (2013) Who Shrunk China? Puzzles in the Measurement of Real GDP. The Economic Journal 123(573):1100–1129

Feenstra RC, Inklaar R, Timmer MP (2015a) The Next Generation of the Penn World Table. American Economic Review 105(10):3150–3182

Feenstra RC, Inklaar R, Timmer MP (2015b) What is new in PWT 8.1? URL `http://www.rug.nl/ggdc/docs/what_is_new_in_pwt_81.pdf`

Fleurbaey M, Gaulier G (2009) International Comparisons of Living Standards by Equivalent Incomes. The Scandinavian Journal of Economics 111(3):597–624

Frankel J (2006) On the Yuan: The Choice between Adjustment under a Fixed Exchange Rate and Adjustment under a Flexible Rate. CESifo Economic Studies 52(2):246–275

Frisch R (1933) Editor's Note. Econometrica 1(1):1–4

Gallant AR (1981) On the Bias in Flexible Functional Forms and an Essentially Unbiased Form. Journal of Econometrics 15(2):211–245

Gallant AR, Golub HG (1984) Imposing Curvature Restrictions on Flexible Functional Forms. Journal of Econometrics 26(3):295–321

Gallant AR, White H (1988) There Exists a Neural Network That Does Not Make Avoidable Mistakes. In: IEEE 1988 International Conference on Neural Networks, vol 1, pp. 657–664

Geary RC (1958) A Note on Comparisons of Exchange Rates and Purchasing Power Between Countries. Journal of the Royal Statistical Society 121(1):97–99

Geman S, Bienenstock E, Doursat R (1992) Neural Networks and the Bias/Variance Dilemma. Neural Computation 4(1):1–58

Gijbels I, Fan J (1996) Local Polynomial Modelling and Its Applications. Chapman and Hall, London

Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press

Greene WH (2012) Econometric Analysis, 7th edn. Pearson, New Jersey

Haag BR, Hoderlein S, Pendakur K (2009) Testing and Imposing Slutsky Symmetry in Nonparametric Demand Systems. Journal of Econometrics 153(1):33–50

Hassan F (2016) The Price of Development: The Penn-Balassa-Samuelson Effect Revisited. Journal of International Economics 102:291–309

Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edn. Springer

Hausman JA, Newey WK (1995) Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss. Econometrica 63(6):1445–1476

Hausman JA, Newey WK (2016) Individual Heterogeneity and Average Welfare. Econometrica 84(3):1225–1248

Hill RJ, Syed I (2010) Improving International Comparisons of Real Output: The ICP 2005 Benchmark and its Implications for China. Discussion Papers 2010-25, School of Economics, The University of New South Wales

Hornik K (1991) Approximation Capabilities of Multilayer Feedforward Networks. Neural Networks 4(2):251–257

Hornik K, Stinchcombe M, White H (1990) Universal Approximation of an Unknown Mapping and Its Derivatives Using Multilayer Feedforward Networks. Neural Networks 3:551–560

Hurwicz L, Uzawa H (1971) On Integrability of Demand Functions. In: Chipman JS (ed) Preferences, Utility, and Demand, chap 6, pp. 114–148

IMF (2011) People's Republic of China: 2011 Article IV Consultation. IMF Country Report 11/192, International Monetary Fund, Washington, DC

Inklaar R, Prasada Rao DS (2017) Cross-Country Income Levels over Time: Did the Developing World Suddenly Become Much Richer? American Economic Journal: Macroeconomics 9(1):265–290

Jean N, Burke M, Xie M, Davis WM, Lobell DB, Ermon S (2016) Combining Satellite Imagery and Machine Learning to Predict Poverty. Science 353(6301):790–794

Joerding W, Li Y (1994) Global Estimation of Feedforward Neural Networks with a Priori Constraints. Computational Economics 7:73–87

Joerding WH, Meador JL (1991) Encoding A Priori Information in Feedforward Networks. Neural Networks 4:847–856

Joliffe IT (2002) Principal Component Analysis, 2nd edn. Springer, New York

Jones E, Oliphant T, Peterson P, et al. (2001) SciPy: Open Source Scientific Tools For Python

Kaufmann D, Kraay A, Mastruzzi M (2010) The Worldwide Governance Indicators: A Summary of Methodology, Data and Analytical Issues. Policy Research Working Paper 5430, World Bank

Kessler M, Subramanian A (2014) Is the Renminbi Still Undervalued? Not According to New PPP Estimates. RealTime Economic Issues Watch, Peterson Institute for International Economics

King MA (1983) Welfare Analysis of Tax Reforms Using Household Data. Journal of Public Economics 21:183–214

Kingma DP, Ba J (2014) Adam: A Method for Stochastic Optimization. CoRR abs/1412.6980

Konüs AA (1939) The Problem of the True Index of the Cost of Living. Econometrica 7(1):10–29

Korhonen I, Juurikkala T (2009) Equilibrium exchange rates in oil-exporting countries. Journal of Economics and Finance 33(1):71–79

Kuan CM, Liu T (1995) Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks. Journal of Applied Econometrics 10(4):347–364

Kuan CM, White H (1994) Artificial Neural Networks: An Econometric Perspective. Econometric Reviews 13(1):1–91

Lau LJ (1978) Testing and Imposing Monotonicity, Convexity and Quasi-Convexity Constraints. In: Fuss M, McFadden DL (eds) Production Economics: A Dual Approach to Theory and Application, vol I: The Theory of Production, Amsterdam: North-Holland, pp. 409–445

LeCun Y, Bengio Y, Hinton GE (2015) Deep Learning. Nature 521:436–444

Lee J, Milesi-Ferretti GM, Ostry J, Prati A, Ricci LA (2008) Exchange Rate Assessments: CGER Methodologies. Occasional Paper 261, International Monetary Fund

Lee JJ (2006) Persistent Wage Differential and Its Implications on the Balassa-Samuelson Hypothesis. Applied Economics Letters 12(10):643–648

Lewbel A (1991) The Rank of Demand Systems: Theory and Nonparametric Estimation. Econometrica 59(3):711–730

Lewbel A (1997) Consumer Demand Systems and Household Equivalence Scales. In: Pesaran MH, Schmidt P (eds) Handbook of Applied Econometrics, Blackwell Handbooks in Economics, vol II: Microeconomics, Wiley-Blackwell, pp. 167–201

Luenberger DG, Ye Y (2016) Linear and Nonlinear Programming, 4th edn. Springer Science+Business Media, LLC

McAleer M, Medeiros MC, Slottje D (2008) A Neural Network Demand System with Heteroskedastic Errors. Journal of Econometrics 147(4):359–371

McCarthy P (2013) Extrapolating PPPs and Comparing ICP Benchmark Results. In: Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program - ICP, The World Bank, Washington, DC, pp. 473–506

Meyer BD, Sullivan JX (2011) Viewpoint: Further results on measuring the well-being of the poor using income and consumption. Canadian Journal of Economics 44(1):52–87

Morrison WM, Labonte M (2013) China's Currency Policy: An Analysis of the Economic Issues. CRS Report for Congress RS21625, Congressional Research Service

Mullainathan S, Spiess J (2017) Machine Learning: An Applied Econometric Approach. Journal of Economic Perspectives 31(2):87–106

Neary JP (2004) Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income. American Economic Review 94(5):1411–1428

Nesterov Y (2004) Introductory Lectures on Convex Optimization: A Basic Course. Applied Optimization, Kluwer Academic Publ., Boston, Dordrecht, London

Nicholson JL (1975) Whose Cost of Living? Journal of the Royal Statistical Society 138(4):540–542

Pollak RA (1980) Group Cost-of-Living Indices. American Economic Review 70(2):273–278

Prais SJ (1959) Whose Cost of Living? The Review of Economic Studies 26(2):126–134

Rao DSP (2016) Welfare Comparisons with Heterogeneous Prices, Consumption, and Preferences. In: Adler MD, Fleurbaey M (eds) The Oxford Handbook of Well-Being and Public Policy, Oxford University Press, pp. 785–820

Rios LM, Sahinidis NV (2013) Derivative-free Optimization: A Review of Algorithms and Comparison of Software Implementations. Journal of Global Optimization 56(3):1247–1293

Rodrik D (2008) The Real Exchange Rate and Economic Growth. Brookings Papers on Economic Activity pp. 365–412

Rogoff K (1996) The Purchasing Power Parity Puzzle. Journal of Economic Literature 34(2):647–668

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning Representations by Back-propagation Errors. Nature 323(6088):533–536

Samuelson PA (1964) Theoretical Notes on Trade Problems. The Review of Economics and Statistics 46(2):145–154

Samuelson PA (1974) Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory. Journal of Economic Literature 12(4):1255–1289

Samuelson PA (1994) Facets of Balassa–Samuelson Thirty Years Later. Review of International Economics 2(3):201–226

Samuelson PA, Swamy S (1974) Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis. American Economic Review 64(4):566–593

Schmillen A (2013) Are Wages Equal Across Sectors of Production? A Panel Data Analysis for Tradable and Non-tradable Goods. Economics of Transition 21(4):655–682

Sonoda S, Murata N (2017) Neural Network with Unbounded Activations is Universal Approximator. Applied and Computational Harmonic Analysis 43(2):233–268

Stone R (1954) Linear Expenditure Systems and Demand Analysis : An Application to the Pattern of British Demand. The Economic Journal 64(255):511–527

Strauss J, Ferris ME (1996) The Role of Nontraded and Traded Wages in the Productivity Differential Model. Southern Economic Journal 63(2):327–338

Subramanian A (2010) New PPP-Based Estimates of Renminbi Undervaluation and Policy Implications. Policy Brief PB10-8, Peterson Institute for International Economics

Sutskever I, Martens J, Dahl G, Hinton GE (2011) On the Importance of Initialization and Momentum in Deep Learning. In: ICML'2011

Swanson NR, White H (1997) A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks. The Review of Economics and Statistics 79(4):540–550

The United States Energy Information Administration (2011) Exports of Crude Oil Including Lease Condensate. URL `http://www.eia.gov/beta/international/`

Theil H (1965) The Information Approach to Demand Analysis. Econometrica 33(1):67–87

UNDP (2014) Sustaining Human Progress: Reducing Vulnerability and Building Resilience. Human Development Report 2014, United Nations Development Program

US Department of the Treasury Office of International Affairs (2014) Report to congress on international economic and exchange rate policies. Tech. rep.

US Department of the Treasury Office of International Affairs (2015) Report to congress on international economic and exchange rate policies. Tech. rep.

White H (1990) Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings. Neural Networks 3:535–549

World Bank (2013) Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program - ICP. World Bank, Washington, DC

World Bank (2014) Purchasing Power Parities and Real Expenditures of World Economies: Summary of Results and Findings of the 2011 International Comparison Program. The World Bank, Washington, DC

Yatchew A (2003) Semiparametric Regression for the Applied Econometrician. Cambridge University Press, Cambridge

Zhang Z, Chen L (2014) A New Assessment of the Chinese RMB Exchange Rate. China Economic Review 30:113–122