Bounding counterfactual demand with unobserved heterogeneity and endogenous expenditures

Laurens Cherchye*

Thomas Demuynck[†]

Bram De Rock[‡]

March 6, 2019

Abstract

We propose a method to predict rational counterfactual demand responses from repeated cross-sections. We derive bounds on the distribution of counterfactual demands that are consistent with the Weak Axiom of Revealed Preferences, without putting any restriction on the preference heterogeneity across consumers. In contrast to existing methods, we also allow for endogeneity of total expenditures. In addition, the method can readily incorporate restrictions on the income elasticities of the consumption goods, which further enhances its identifying power (i.e. tighter bounds). We illustrate our approach through an application to data drawn from the U.S. Consumer Expenditure Survey.

JEL classification: C14, D12

Keywords: Unobserved heterogeneity, WARP, endogenous expenditures, counterfactual demand

1 Introduction

We consider a setting in which the empirical analyst can only use repeated cross-sectional data on consumers' demands for a set of goods, with each cross-section corresponding to a different price regime. This is the type of data most frequently used in applied demand analysis. Having only a single consumption observation for each individual consumer, we are faced with the problem that individual preferences and, hence, counterfactual choices are not identified. Unless one is willing to impose some additional restrictions on the structure of the unobserved heterogeneity component or on the form of the underlying consumer preferences.

*Department of economics, University of Leuven. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven-kortrijk.be. Laurens Cherchye thanks the European Research Council (ERC; ERC-2013-CoG-614221), the Fund for Scientific Research - Flanders (FWO) and the Research Fund of the KU Leuven for financial support.

[†]ECARES, Université Libre de Bruxelles. Avenue F.D. Roosevelt, CP 114, B-1050 Brussels, Belgium. E-mail: thomas.demuynck@ulb.ac.be. Thomas Demuynck gratefully acknowledges FWO and FRNS for their financial support.

[‡]ECARES, Université Libre de Bruxelles and Department of Economics, University of Leuven. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges FWO and FRNS for their financial support.

Unobserved preference heterogeneity means that different consumers (e.g. households) behave differently even when taking account of all observable characteristics (like household composition, household age profile, marital status, education, etc.). As we discuss in more detail in Section 2, the literature has invoked several different assumptions to deal with unobserved preference heterogeneity in applied demand analysis. These assumptions can be broadly classified into three categories: (i) assumptions on the heterogeneity component, (ii) assumptions on the structure of consumer behavior (i.e. rationality assumptions) and (iii) independence assumptions regarding the data generating process.

Our contribution. We propose a novel method for counterfactual demand analysis while accounting for unobserved preference heterogeneity. Our method makes minimalistic prior assumptions. First of all, we impose no restrictions on the structure or distribution of the unobserved heterogeneity component. Under this condition, we will be able to identify bounds on counterfactual demand distributions associated with unobserved prices and expenditure levels. Next, a specific feature of our method is that it defines these demand bounds by exploiting the basic empirical implications of rational consumer behavior that are captured by the Weak Axiom of Revealed Preference (WARP). More specifically, this means that we do not impose transitivity of preferences and that we avoid using unverifiable functional structure for the consumer utilities.¹

Finally, while we maintain independence of unobserved preference heterogeneity on commodity prices, we do account for the possibility that preference heterogeneity and total expenditures are interdependent.² The assumption that preference heterogeneity does not depend on commodity prices is rather uncontroversial for the type of setting that we analyze (i.e. repeated cross-sections with fixed prices per cross-section). However, and more importantly, the fact that we account for endogenous total expenditures contrasts with most of the existing literature, which typically assumes (conditional) independence between preference heterogeneity and total expenditures. Moreover, in the applied demand literature that does account for endogeneous expenditures, the usual practice is to resolve this endogeneity issue by using instrumental variables. In this respect, gross earnings is popularly used as an instrument in empirical work.³ This instrument, however, is only valid if total earnings influence demand *only* through their effect on the total expenditures. For example, it also requires that the wage rate (which partially determines the gross earnings) is not correlated with consumers' preferences, and this is arguably a controversial assumption in many situations.

¹Basically, WARP imposes that preferences are acyclic for all chains of length two (see Section 3). As such, it does not impose transitivity of preferences (in contrast to the Strong/General Axiom of Revealed Preference).

²At this point, we remark that endogeneity of expenditures does not invalidate our WARP-based reasoning. For our arguments to hold, we basically need that expenditure levels and consumption quantities are not chosen simultaneously. More specifically, in that case the consumption behavior can be represented as if it corresponds to some "two step budgeting process": in the first step, the consumer decides on the total expenditure; in the second step, for the given expenditure allocated to the observed goods, the consumer will make the choices that optimize utility. The WARP conditions for rational consumer behavior then remain valid as long as preferences are independent from prices.

³See, for example, Blundell and Duncan (1998), Blundell, Browning, and Crawford (2003) and Blundell, Chen, and Kristensen (2007b).

More generally, it is notoriously difficult to find an instrument that influences total expenditures but does not indirectly influence (or is influenced by) the consumers' preferences.⁴

A main novel feature of our method is that it avoids assuming that expenditures are exogenous or that valid instruments are available. Our method does not assume a particular structural model describing the mechanics that drive the interdependence of total expenditures and preferences. In principle, there are a multitude of possible explanations of endogenous expenditures, which –in our opinion– makes it practically impossible to come up with a model that is fully comprehensive. Admittedly, the absence of such a structural model implies some limitations. Most notably, it prevents us from determining the distribution of endogenous expenditures in a counterfactual situation with new prices. As a “second best” option, in the current paper we choose to estimate the distribution of demands while keeping the new expenditure level fixed. In other words, the counterfactual puts the entire population on the same budget set, thus providing insight in the distribution of preferences for this specific counterfactual situation. This use of fixed budgets for evaluating counterfactual demand responses falls in line with existing work on the empirical analysis of consumer behavior with preference heterogeneity.

An important motivation for our study originates from a recent strand of literature that focuses on the issue of partial identification in treatment response models (see for example Manski (1997)). The starting point in this literature is what is known as the credibility problem. The credibility problem says that every empirical model faces a negative trade-off between the assumptions imposed on the model and the credibility of its outcomes: the stronger the assumptions imposed on the model, the less credible are the findings. To control this trade-off, we must start from a base model which imposes only minimal assumptions on the data generating process. Next, departing from this base scenario, additional assumptions are imposed. By comparing the results with and without these additional assumptions, their identifying power can be analyzed.

Our base model uses minimalistic (“uncontroversial”) assumptions: we only impose a very basic rationality constraint (WARP), we do not impose any restriction on the unobserved heterogeneity component, and we allow expenditure levels to be endogenous. We show that, in principle, this base model can define informative WARP-consistent bounds on counterfactual demand distributions, despite these very weak priors. These bounds provide a concrete insight into the identifying power of our minimal assumptions. However, not very surprisingly, these WARP-consistent bounds may often be quite wide in practice. Therefore, in a following step, we show that our method can readily incorporate specific restrictions regarding the income elasticities of the consumption goods. We believe that such restrictions may be fairly easy to define in many practical situations. For example, we may often reasonably assume (or learn from existing studies) that a particular commodity is typically a normal good (i.e. positive income elasticity) or a necessary good (i.e. income elasticity between zero and one). Our empirical application will show that such additional restrictions can substantially enhance the identifying power of our method in empirical analysis (i.e.

⁴One example of a valid instrument may be a lump sum transfer. Unfortunately, however, such transfers rarely occur in real life and are therefore difficult to use in practical demand analysis.

tighter bounds).

Structure of the paper. The remainder of the paper is organized as follows. Section 2 discusses the main related literature. Section 3 sets the stage by introducing some necessary definitions and notation. Section 4 shows how to obtain WARP-consistent bounds on demand distributions under our minimalistic priors. Section 5 explains how we can derive tighter bounds by restricting the income elasticities of the consumption goods. Section 6 presents an empirical illustration to data drawn from the U.S. Consumer Expenditure Survey (CEX). Finally, Section 7 summarizes and concludes.

2 Related literature

The literature has invoked three types of assumptions to deal with unobserved preference heterogeneity in applied demand analysis. The first category of assumptions concerns the unobserved preference heterogeneity component. The most widely used assumption is that the unobserved heterogeneity component is single-dimensional and enters the demand functions in an additively separable way. If, in addition, a mean independency assumption is imposed, then demand behavior can be analyzed by focusing on the mean demand function of all consumers in the population (e.g. Blundell, Browning, and Crawford (2003, 2007a, 2008), Blundell, Horowitz, and Parey (2012), Blundell et al. (2015)). However, as discussed in detail by Brown and Walker (1989) and Lewbel (2001), this approach is only valid if individual preferences satisfy a number of quite stringent requirements.

In response to this critique, a recent strand of literature (including Blundell, Kristensen, and Matzkin (2014), Blundell, Horowitz, and Parey (2017) and Hoderlein and Vanhems (2018)) has relaxed the additivity assumption, and used the weaker assumption that the unobserved heterogeneity component has a dimension that equals at most the number of goods minus one and, in addition, satisfies an invertibility requirement (basically stating that demand functions are invertible in the unobserved heterogeneity component). As shown by Matzkin (2003), see also Beckert and Blundell (2008) and Matzkin (2008), in a two-goods setting these assumptions identify each consumer with its own quantile in the demand distribution. Specifically, for two different price-income situations, the consumers situated on the same quantile of the demand distribution have the same preferences.

If we abstain from imposing such assumptions, the identification of individual preferences and demands must be abandoned. In such a case, the best we can do is identify or bound the distribution of preferences or demands over the heterogeneous population (e.g. Hoderlein (2011), Hoderlein and Stoye (2014), Dette, Hoderlein, and Neumeyer (2016), Hausman and Newey (2016), Kawaguchi (2017), Cosaert and Demuyne (2018) and Kitamura and Stoye (2018)). Notably, this still allows us, for example, to identify the median demand function in the population, but it will no longer be possible to recover the demand for the “median” consumer (in fact, this median consumer need no longer exist). As indicated in the Introduction, we will also follow this minimalistic approach in

the current paper, focusing on the recovery of (bounds on) demand distributions while imposing no restrictions on the structure or distribution of the unobserved heterogeneity component.

The second group of assumptions is of a statistical nature and imposes independence restrictions on the distribution of the unobserved heterogeneity component. Essentially, these restrictions impose that, when conditioning on observable consumer characteristics, the unobserved heterogeneity component is independent from the consumers' budget set (i.e. prices and income). In a cross-sectional setting, with fixed prices within each cross-section, this assumption requires that the distribution of preferences remains unchanged when considering subpopulations with different expenditure levels (conditional on other observable characteristics). As an implication, it also imposes that the distributions of preferences for rich and poor consumers are the same.

In many cases, however, total expenditures are not independent of preferences, which makes them endogenous. As indicated above, this endogeneity problem is typically resolved by using gross earnings as an instrumental variable. To the best of our knowledge, all papers in the literature either require such (conditional) independence between preferences and expenditures or at least the availability of valid instruments. A main novel feature of the method that we develop in the following sections is that it avoids assuming that expenditures are exogenous or that valid instruments are available.

The third group of assumptions imposes rationality constraints on consumers' behavior. Such rationality conditions allow us to predict the demand distribution in counterfactual price-income situations.⁵ We distinguish two types of rationality conditions that are used in the literature. The first type takes the form of differential conditions on consumer demands. Examples are the well-known Slutsky symmetry, Slutsky negativity, adding up and homogeneity conditions (e.g. Dette, Hoderlein, and Neumeyer (2016) and Blundell, Horowitz, and Parey (2017)). The second type of conditions is expressed in terms of revealed preference axioms. The most frequently used axioms are the Weak Axiom of Revealed Preference (WARP) (e.g. Blundell, Browning, and Crawford (2003, 2007a, 2008), Hoderlein and Stoye (2014) and Cosaert and Demuyne (2018)) and the Strong Axiom of Revealed Preference (SARP) (e.g. Kawaguchi (2017) and Kitamura and Stoye (2018)). In a population setting, combining these revealed preference axioms with the "independence of budgets" restriction that we discussed above gives the so-called stochastic axioms of revealed preference (see McFadden (2005) for an overview). Importantly, the validity of the budget independence assumption is again crucial to justify the application of these stochastic axioms in empirical demand analysis.

⁵In principle, when the data show sufficient price and income variation, rationality conditions can be omitted for conducting demand analysis. However, even in such a case the conditions can be useful to increase estimation precision.

3 Setting the stage

In the main part of the paper, we focus on a setting with two goods, with quantities $q \in \mathbb{R}_+$ and $z \in \mathbb{R}_+$. We normalize the price of good z to one and denote the price of good q by $p \in \mathbb{R}_{++}$. Our focus on a two-goods settings allows us to graphically illustrate our main theoretical results. Importantly, however, it is possible to generalize our analysis to settings with more than two goods (see Appendix A.1).

Set-up. We consider a probability space $(\mathcal{U}, \Omega, P)$ where $\mathcal{U}$ is the space of consumers, i.e. preference types. The set Ω a suitable σ -algebra and P the probability measure. A consumer is denoted by $u \in \mathcal{U}$. We will use the upper case letter U to represent a random element of the probability space. Every consumer $u \in \mathcal{U}$ has a demand function for good q , denoted by the function $q_u(p, x_u)$, where $x_u \in \mathbb{R}_+$ stands for total expenditures of consumer u . Thus, the demand of consumer u for good q is defined as a continuous function $q_u : \mathbb{R}_{++} \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that, $0 \leq q_u(p, x_u) \leq x_u/p$ (i.e. the maximum expenditure on good q equals x_u). Correspondingly, the demand of consumer u for good z is equal to $z_u(p, x_u) = x_u - p q_u(p, x_u)$.

As indicated above, we assume a repeated cross-sectional setting. Within each cross-section, the empirical analyst observes the demanded quantities and expenditure levels for a large number of consumers drawn i.i.d from $(\mathcal{U}, \Omega, P)$. The price is assumed to be fixed and identical for all consumers within the same cross-section, but prices may vary over cross-sections. Thus, we treat prices as non-stochastic, which also means that unobserved preference heterogeneity (captured by the random variable U) is distributed independently from the prices. If the distribution of consumer preferences remains stable over time (i.e. over the different cross-sections) and if each cross-section is randomly drawn from the same population, then this assumption will be satisfied. We denote the cross-sections by $\{1, \dots, t, \dots, T\}$ and we represent the price in cross-section t by p_t .

We assume a sufficiently large number of independent observations for each cross-section t . Then, we can identify the joint distribution of the random vector $(x_{t,U}, q_{t,U}, z_{t,U})$, with, $x_{t,U}$ the random total expenditures, $q_{t,U} = q_U(p_t, x_{t,U})$ the random demanded quantity of good q , and $z_{t,U} = z_U(p_t, x_{t,U})$ the random quantity of good z . Of course, these random variables satisfy the adding up condition, $x_{t,U} = p_t q_{t,U} + z_{t,U}$. For now, we will assume that for each cross-section t , we know the joint distribution of $(x_{t,U}, q_{t,U}, z_{t,U})$. In our empirical application we use Andrews and Soares (2010) to deal with small sample and estimation issues.

Importantly, we do not impose that the unobserved heterogeneity component is distributed independently of the expenditure level. For example, we do not require that $\Pr[q_U(p_t, x) \leq q | x_{t,U} = x] = \Pr[q_U(p_t, x) \leq q]$. Neither do we have an instrument I such that, $\Pr[q_U(p_t, x) \leq q | x_{t,U} = x, I = i] = \Pr[q_U(p_t, x) \leq q | I = i]$. As discussed in the Introduction, this constitutes a main difference between our method and the literature.

WARP restrictions. We use the Weak Axiom of Revealed Preference (WARP; Samuelson (1938)) as our condition of rational consumer behavior.⁶ To formally define the WARP restrictions, let us consider two income levels $x_{t,u}, x_{v,u} \in \mathbb{R}_+$ and two prices $p_t, p_v \in \mathbb{R}_{++}$. Consumer u satisfies WARP if for the associated consumption bundles $(q_u(p_t, x_{t,u}), z_u(p_t, x_{t,u})) \neq (q_u(p_v, x_{v,u}), z_u(p_v, x_{v,u}))$ we have

$$x_{v,u} \geq p_v q_u(p_t, x_{t,u}) + z_u(p_t, x_{t,u}) \Rightarrow x_{t,u} < p_t q_u(p_v, x_{v,u}) + z_u(p_v, x_{v,u}).$$

In words, if the bundle $(q_u(p_t, x_{t,u}), z_u(p_t, x_{t,u}))$ could be bought with income $x_{v,u}$ at price p_v , then the bundle $(q_u(p_v, x_{v,u}), z_u(p_v, x_{v,u}))$ should not have been strictly cheaper at income $x_{t,u}$ and price p_t .

4 Bounds on counterfactual demands

Consider some measurable set $S \in \Omega$. In what follows, we will use the notation

$$\Pr[S] = \Pr[U \in S] = \int_{\mathcal{U}} \mathbb{1}[u \in S] P(du),$$

where the indicator function $\mathbb{1}[\cdot]$ equals 1 if the expression between brackets is true and 0 otherwise. In words, the function $\Pr[S]$ gives the probability that the expression $U \in S$ holds for a randomly drawn consumer U . Or, putting it differently, it stands for the proportion of consumers that belong to the set S .

Counterfactual demand analysis wants to predict demand for some unobserved price p_0 and expenditure level x_0 . More specifically, we want to identify information on the distribution of the random variable $q_U(p_0, x_0)$. We address this question by defining upper and lower bounds on the value of $\Pr[q_U(p_0, x_0) \leq q_0]$ for any $q_0 \in]0, \frac{x_0}{p_0}[$. Obviously, for $z_0 = x_0 - p_0 q_0$, we have that $\Pr[q_U(p_0, x_0) \leq q_0] = \Pr[z_U(p_0, x_0) \geq z_0]$.

To define our bounds on $\Pr[q_U(p_0, x_0) \leq q_0]$ we consider the sets $\underline{T} = \{t \in T | p_t < p_0\}$ and $\overline{T} = \{t \in T | p_t > p_0\}$ (using that p_0 is different from any observed price), i.e. two mutually exclusive subsets of the observed cross-sections. In what follows, the set $\underline{T}$ will be used to define a WARP-consistent lower bound on $\Pr[q_U(p_0, x_0) \leq q_0]$, and the set $\overline{T}$ to define a WARP-consistent upper bound.

Lower WARP-consistent bounds. Our first result defines the tightest WARP-consistent lower bound on the value of $\Pr[q_U(p_0, x_0) \leq q_0]$. The proofs of our main results can be found in the Appendix.

⁶Interestingly, in a two goods setting such as ours, WARP is equivalent to SARP; see Rose (1958) for a proof. Cherchye, Demuyne, and De Rock (2018) provide a recent discussion on the empirical equivalence of WARP and SARP.

Theorem 1. *For all q_0 , we have that*

$$\max_{t \in \underline{T}} \ell_t(q_0) \leq \Pr[q_U(p_0, x_0) \leq q_0],$$

with

$$\ell_t(q_0) = \Pr[x_0 \geq p_0 q_{t,U} + z_{t,U} \text{ and } x_{t,U} \geq p_t q_0 + z_0].$$

This bound is the tightest WARP-consistent lower bound.

We illustrate the reasoning underlying the result in Theorem 1 in the left panel of Figure 1. The solid line depicts the budget line at income level x_0 and price level p_0 . Similarly, the dashed line gives a budget line for cross-section t with observed price $p_t < p_0$ (i.e. $t \in \underline{T}$). In particular, this dashed line corresponds to an expenditure level $x_0 + (p_t - p_0)q_0$ (i.e. the budget necessary to buy the bundle (q_0, z_0) at price level p_t). The solid and dashed budget lines intersect at the bundle (q_0, z_0) . For this situation, the lower WARP-consistent bound in Theorem 1 corresponds to the fraction of all observed bundles $(q_{t,u}, z_{t,u})$ in cross-section t that are situated in the shaded area of the figure, i.e. region (1). These are the bundles for which $x_0 \geq p_0 q_{t,u} + z_{t,u}$ (below the solid line, i.e. regions (1) and (4)) and $x_{t,u} \geq z_0 + p_t q_0$ (above the dashed budget line, i.e. regions (1) and (2)).

To grasp the intuition behind this lower bound, let us assume that some consumer u chooses a bundle $(q_{t,u}, z_{t,u})$ that lies in the shaded area (1) of the left panel of Figure 1. Then, consistency with WARP requires that, under the new prices p_0 and income x_0 , this consumer cannot consume $(q_u(p_0, x_0), z_u(p_0, x_0))$ on the line segment $[A, x_0/p_0]$ of the solid budget line (as this would imply a violation of the WARP condition in Section 3). As a result, this consumer must consume $(q_u(p_0, x_0), z_u(p_0, x_0))$ on the segment $[x_0, A]$. This implies a consumption of $q_u(p_0, x_0)$ below q_0 , which can be used to define a WARP-consistent lower bound on $\Pr[q_u(p_0, x_0) \leq q_0]$.

Next, we can use the same example to illustrate that Theorem 1 actually defines the tightest WARP-consistent lower bound. In particular, it is easy to see that, if some bundle $(q_{t,u}, z_{t,u})$ is not in the shaded area (1), then we can always draw an income expansion path for consumer u and price p_t that does not cross the shaded area.⁷ In this case, it is always possible to find an allocation $(q_u(p_0, x_0), z_u(p_0, x_0))$ on the solid budget line with $q_u(p_0, x_0) > q_0$ (i.e. the segment $[A, x_0/p_0]$), which implies no violation of WARP. As a result, the lower bound in Theorem 1 effectively produces the highest WARP-consistent lower bound that is possible for the set-up (with minimalistic priors) that we consider.

Upper WARP-consistent bounds. Following an analogous reasoning, we can define the tightest WARP-consistent upper bound on $\Pr[q_u(p_0, x_0) \leq q_0]$. In this case we make use of the observed prices in the set $\bar{T}$ that we defined above.

⁷For consumer u and price p_t , the income expansion path defines the coordinates $(q_u(p_t, x), z_u(p_t, x))$ that correspond to varying values of the expenditures x .

Figure 1: lower and upper bounds for $\Pr[q_u(p_0, x_0) \leq q_0]$

Theorem 2. For all q_0 , we have that

$$\Pr[q_U(p_0, x_0) \leq q_0] \leq \min_{t \in \bar{T}} h_t(q_0),$$

with

$$h_t(q_0) = \Pr \left[x_0 < z_{t,U} + p_0 q_{t,U}, \text{ or } x_{t,U} < z_0 + p_t q_0, \text{ or } (q_{t,U}, z_{t,U}) = (q_0, z_0) \right].$$

This bound is the tightest WARP-consistent upper bound.

We use the right panel of Figure 1 to illustrate this result. The solid line has the same interpretation as before, and the dashed line now corresponds to a cross-section t with $p_t > p_0$ (i.e. $t \in \bar{T}$). In this case, the upper WARP-consistent bound in Theorem 2 equals the fraction of observed bundles $(q_{t,u}, z_{t,u})$ with $x_0 < z_{t,u} + p_0 q_{t,u}$ or $x_{t,u} < z_0 + p_t q_0$ or $(q_{t,u}, z_{t,u}) = (q_0, z_0)$, which are situated in the shaded area in the right panel of Figure 1, i.e. regions (1), (2) and (4). To see the intuition that this defines an upper bound on $\Pr[q_U(p_0, x_0) \leq q_0]$, we note that this upper bound is actually the complement of the lower bound on the probability $\Pr[z_U(p_0, x_0) < z_0]$, which is approximately equal to $\Pr[z_U(p_0, x_0) \leq z_0]$. Using this, the argument follows directly from the lower bound argument that we gave above (but now applied to $\Pr[z_U(p_0, x_0) \leq z_0]$ instead of $\Pr[q_U(p_0, x_0) \leq q_0]$). The fact that the upper bound also includes the agents for which $(q_{t,u}, z_{t,u}) = (q_0, z_0)$ follows from the fact that this choice is always WARP-consistent with a choice of (q_0, z_0) on the counterfactual budget set (p_0, x_0) , meaning that $q_u(p_0, x_0) = q_0$.

Following a similar logic as before, we can also verify that the shaded area in the right panel of the figure effectively defines the lowest WARP-consistent upper bound that we can define. Specifically, for any consumer u in this shaded area (1), (2) and (4), it is possible to draw an income expansion path (for prices p_t) that does not go through the non-shaded area (3). For such an expansion path, the corresponding bundle on the budget line for (p_0, x_0) will have q below q_0 , which guarantees consistency with WARP.

For the above examples, our argument that Theorems 1 and 2 define tightest WARP-consistent bounds uses that income expansion paths have not been restricted so far. This suggests that we may improve the identifying power of our method (i.e. obtain tighter demand bounds) by imposing specific structure on the shape of income expansion paths. We explore this further in the following section. In particular, we will consider shape restrictions on income expansion paths that take the form of restricted income elasticities for the consumption goods.

5 Restrictions on income elasticities

From our graphical example in the previous section, it should be clear that, if the cross-sectional prices p_t ($t \in T$) and the counterfactual price p_0 are close to each other, then the WARP-consistent bounds on the probability $\Pr[q_U(p_0, x_0) \leq q_0]$ will be close to the trivial bounds of zero and one. As a result, we can expect the bounds defined in Theorems 1 and 2 to be little informative if there is little cross-sectional price variation.

In such a case, it can be useful to add structure. As indicated above, an obvious strategy is to impose shape restrictions on the income expansion paths. More specifically, let us consider again the situation in Figure 1 and now assume that we do know the exact income expansion path for each consumer u . Then, if $p_t < p_0$ (i.e. $t \in \underline{T}$), the fraction of all expansion paths (over the set of consumers $\mathcal{U}$) that go through the shaded area (1) in the left panel of Figure 1 will define a lower bound on $\Pr[q_U(p_0, x_0) \leq q_0]$. Similarly, if $p_t > p_0$ (i.e. $t \in \overline{T}$), then the fraction of all paths that do not enter the non-shaded area (3) in the right panel of Figure 2 will define an upper bound on $\Pr[q_U(p_0, x_0) \leq q_0]$.

Of course, in practice we typically do not observe the expansion paths. However, in many cases it makes sense to place some restrictions on them. For example, we may often reasonably assume that goods are normal (i.e. positive income elasticities). In some cases, the empirical analyst may actually be able to define even more stringent bounds on the income elasticities of particular consumption goods (e.g. based on existing empirical studies). In what follows, we will consider these situations in more detail.

Normal goods. If q and z are both normal goods, then the income expansion paths will be non-decreasing in x . For the moment, assume that we know the expansion path $q_u(p_t, \cdot)$ of consumer u . Let the intersection of this expansion path with the budget (p_0, x_0) occur at the bundle (q, z) . Blundell, Browning, and Crawford (2008) call these bundles the *intersection demands*. It can be shown (see Blundell, Browning, and Crawford (2008) and Lemma 2 in Appendix A.2) that the counterfactual demand $q_u(p_0, x_0)$ will be less than q_0 if and only if this intersection demand q is less than or equal to q_0 , i.e. the intersection demand (q, z) lies on the segment $[x_0, A]$. A direct implication of this is that, under normality of the goods q and z , we only need information on the intersection demands to define our WARP-consistent bounds. However, we cannot directly use this result in practice as, for consumer u and cross-section t , we only observe the consumption

Figure 2: Lower and upper demand bounds for normal goods

quantities $q_{t,u} = q_u(p_t, x_{t,u})$ and $z_{t,u} = z_u(p_t, x_{t,u})$ associated with the expenditure level $x_{t,u}$. We do not observe the expansion paths $q_u(p_t, x)$ for values $x \neq x_{t,u}$. In a sense, the structure of this problem is similar to the one of a treatment response model in which $q_u(p_t, \cdot)$ is the outcome function and $x_{t,u}$ the treatment (i.e. we only observe the outcome $q_u(p_t, x)$ for the treatment $x = x_{t,u}$).

However, we can still identify the fraction of consumers for which all possible expansion paths have an intersection demand less than or equal to q_0 . This gives the following bounds.

Theorem 3. *Assume that the goods q and z are both normal. For all q_0 , we have that*

$$\max_{t \in \underline{T}} \ell_t(q_0) \leq \Pr [q_U(p_0, x_0) \leq q_0] \leq \min_{t \in \overline{T}} h_t(q_0),$$

with

$$\ell_t(q_0) = \Pr \left[q_{t,U} \leq q_0 \text{ and } z_{t,U} \geq z_0 \right] \text{ and } h_t(q_0) = \Pr \left[q_{t,U} \leq q_0 \text{ or } z_{t,U} \geq z_0 \right].$$

These bounds are the tightest WARP-consistent bounds.

This result is illustrated in Figure 2. Let us first consider the left panel, which presents the lower bound on $\Pr [q_U(p_0, x_0) \leq q_0]$. From the reasoning above, we know that this lower bound is defined by the fraction of expansion paths that go through the segment $[x_0, A]$. Given that both goods are normal, these expansion paths are everywhere increasing. Therefore, any expansion path associated with some observed consumption bundle in the shaded region necessarily crosses the line segment $[x_0, A]$.

The right panel of Figure 2 illustrates the WARP-consistent upper bound that is defined in Theorem 3. In particular, for every consumer u that is in the shaded region, it is possible to construct a monotonically increasing expansion path that crosses the line segment $[x_0, A]$, meaning that this consumer's behavior (as summarized by the expansion path) is informative to construct the upper bound. Comparing the shaded regions in Figures 1 and 2 clearly reveals that the assumption of normal goods can considerably enhance the identifying power of our method.

Additional restrictions. Let us then turn to the case where the goods q and z are again normal, but now the empirical analyst can use even more information on possible values for the income elasticities of the consumption goods. In particular, we consider parameters $\eta_\ell, \eta_h, \delta_\ell$ and δ_h such that,

$$\begin{aligned} 0 \leq \eta_\ell &\leq \frac{\partial q_u(p_t, x)}{\partial x} \frac{x}{q_u(p_t, x)} \leq \eta_h, \\ 0 \leq \delta_\ell &\leq \frac{\partial z_u(p_t, x)}{\partial x} \frac{x}{z_u(p_t, x)} \leq \delta_h. \end{aligned}$$

Obviously, normality of q and z can be imposed by setting η_ℓ and δ_ℓ equal to zero, with η_h and δ_h arbitrarily high. We consider elasticity bounds that are independent of t and income x . In principle, we can easily allow for different $\eta_\ell, \delta_\ell, \eta_h$ and δ_h associated with different cross-sections or income levels. We choose not to do so here to facilitate our discussion.

The next lemma uses the information on $\eta_\ell, \delta_\ell, \eta_h$ and δ_h to bound the counterfactual values of $q_u(p_t, x)$ and $z_u(p_t, x)$.

Lemma 1. *Assume that the goods q and z are both normal with income elasticities in the intervals $[\eta_\ell, \eta_h]$ and $[\delta_\ell, \delta_h]$. For any consumer $u \in \mathcal{U}$, we have that, for $x \leq x_{t,u}$,*

$$\begin{aligned} q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_h} &\leq q_u(p_t, x) \leq q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_\ell}, \\ z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_h} &\leq z_u(p_t, x) \leq z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_\ell}, \end{aligned}$$

and, for $x \geq x_{t,u}$,

$$\begin{aligned} q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_\ell} &\leq q_u(p_t, x) \leq q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_h}, \\ z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_\ell} &\leq z_u(p_t, x) \leq z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_h}. \end{aligned}$$

As we will explain below, this result allows us to define the following WARP-consistent bounds on $\Pr[q_u(p_0, x_0) \leq q_0]$.

Theorem 4. *Assume that the goods q and z are both normal with income elasticities in the intervals $[\eta_\ell, \eta_h]$ and $[\delta_\ell, \delta_h]$. For all q_0 , we have that*

$$\max_{t \in \underline{T}} \ell_t(q_0) \leq \Pr(q_U(p_0, x_0) \leq q_0) \leq \min_{t \in \overline{T}} h_t(q_0),$$

with

$$\begin{aligned} \ell_t(q_0) = & \Pr \left[\left(q_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\eta_h} \leq q_0 \text{ or } z_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\delta_\ell} \geq z_0 \right) \text{ and } p_t q_0 + z_0 \geq x_{t,U} \right] \\ & + \Pr \left[\left(q_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\eta_\ell} \leq q_0 \text{ or } z_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\delta_h} \geq z_0 \right) \text{ and } p_t q_0 + z_0 \leq x_{t,U} \right], \end{aligned}$$

and

$$\begin{aligned} h_t(q_0) = & \Pr \left[q_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\eta_\ell} \leq q_0 \text{ and } z_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\delta_h} \geq z_0 \text{ and } p_t q_0 + z_0 \geq x_{t,U} \right] \\ & + \Pr \left[q_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\eta_h} \leq q_0 \text{ and } z_{t,U} \left(\frac{p_t q_0 + z_0}{x_{t,U}} \right)^{\delta_\ell} \geq z_0 \text{ and } p_t q_0 + z_0 \leq x_{t,U} \right]. \end{aligned}$$

These bounds are the tightest WARP-consistent bounds.

Let us clarify the reasoning behind this result for the first term of $\ell_t(q_0)$, which defines the lower bound on $\Pr[q_U(p_0, x_0) \leq q_0]$. The explanation for the other term of $\ell_t(q_0)$ and for $h_t(q_0)$ proceeds analogously. We have $p_t < p_0$ (i.e. $t \in \underline{T}$) with $x_{t,u} \leq p_t q_0 + z_0$ for consumer u , i.e. the bundle $(q_{t,u}, z_{t,u})$ is below the dashed budget line in the left panel of Figure 2. Following our above reasoning, we want to know whether the expansion path through this bundle crosses the counterfactual budget line through the segment $[x_0, A]$, which corresponds to $q_u(p_t, p_t q + z) \leq q_0$ for the intersection demand (q, z) or, equivalently, $z_u(p_t, p_t q + z) \geq z_0$. From the figure, it is clear to see that if both goods are normal (i.e. expansion paths are non-decreasing), then the expansion path will cross the segment $[x_0, A]$ of the solid budget line if and only if it crosses the segment $[B, A]$ of the dashed budget line, corresponding to the budget at prices p_t and income $p_t q_0 + z_0$. As such, we get the condition $q_u(p_t, p_t q_0 + z_0) \leq q_0$ or, equivalently, $z_u(p_t, p_t q_0 + z_0) \geq z_0$. Because $x_{t,u} \leq p_t q_0 + z_0$, Lemma 1 implies

$$q_u(p_t, p_t q_0 + z_0) \leq q_{t,u} \left(\frac{p_t q_0 + z_0}{x_{t,u}} \right)^{\eta_h},$$

and, thus, we have $q_u(p_t, p_t q_0 + z_0) \leq q_0$ if $q_{t,u} \left(\frac{p_t q_0 + z_0}{x_{t,u}} \right)^{\eta_h} \leq q_0$.

Similarly, Lemma 1 also tells us that

$$z_{t,u} \left(\frac{p_t q_0 + z_0}{x_{t,u}} \right)^{\delta_\ell} \leq z_u(p_t, p_t q_0 + z_0),$$

which implies $z_u(p_t, p_t q_0 + z_0) \geq z_0$ if $z_0 \leq z_{t,u} \left(\frac{p_t q_0 + z_0}{x_{t,u}} \right)^{\delta_\ell}$.

Taken together, for consumer u (with $x_{t,u} \leq p_t q_0 + z_0$) to be informative for our WARP-based lower bound, we need

$$q_{t,u} \left(\frac{p_t q_0 + z_0}{x_{t,u}} \right)^{\eta_h} \leq q_0 \text{ or } z_0 \leq z_{t,u} \left(\frac{p_t q_0 + z_0}{x_{t,u}} \right)^{\delta_\ell},$$

Figure 3: Illustration bounds if q is income inelastic

which is the condition that appears in Theorem 4.

To conclude, Figure 3 provides an example for the case where q is a necessary good ($\eta_\ell = 0$ and $\eta_h = 1$) and z a normal good ($\delta_\ell = 0$ and $\delta_h = \infty$). Then, Theorem 4 tells us that

$$\begin{aligned} \ell_t(q_0) &= \Pr \left[\frac{z_0}{q_0} \leq \frac{z_{t,U}}{q_{t,U}} \text{ and } p_t q_0 + z_0 \geq x_{t,U} \right] + \Pr \left[q_{t,U} \leq q_0 \text{ and } p_t q_0 + z_0 \leq x_{t,U} \right], \text{ and} \\ h_t(q_0) &= \Pr \left[q_{t,U} \leq q_0 \text{ and } p_t q_0 + z_0 \geq x_{t,U} \right] + \Pr \left[\frac{z_{t,U}}{q_{t,U}} \geq \frac{z_0}{q_0} \text{ and } p_t q_0 + z_0 \leq x_{t,U} \right], \end{aligned}$$

which defines the shaded areas in Figure 3. Comparing these areas with the ones in Figure 2 clearly shows the identifying power of using specific elasticity bounds for the consumption goods.

6 Application

We demonstrate the practical usefulness of our method by means of an empirical illustration using data drawn from the Consumer Expenditure Survey (CEX). We have cross-sectional household consumption data for U.S. households from 1994 until 2007 (14 years), which are obtained from the diary survey contained in the CEX. We assume that the distribution of preferences is stable over the whole period. In principle, it is fairly easy to relax this assumption, by dividing our total time period into subperiods and computing bounds on counterfactual demands for the different subperiods. By construction, such an exercise will generate demand bounds that are wider than the bounds obtained under the stronger assumption of preference stability over the full period of analysis. For compactness, and to focus our discussion, we will not report on demand bounds based on subperiods in what follows.

Further, our (partial) identification results are based on a finite collection of moment inequalities. An active and rapidly growing literature has recently focused on dealing with statistical inference for this type of set identified problems.⁸ See the supplementary material, including our

⁸See, among others, Imbens and Manski (2004), Chernozhukov, Hong, and Tamer (2007), Stoye (2009), Bugni (2010) and Chernozhukov, Lee, and Rosen (2013) for various approaches to deal with statistical inference for models where parameters are defined through moment inequalities.

data and programs, for more details on the bootstrap procedure set out by Andrews and Soares (2010) that we use to compute the confidence intervals in our empirical exercise.

Set-up. Following our theoretical discussion, we focus on a setting with two goods, i.e. food (good q) and other (nondurable) goods and services (good z). Analyzing price responsiveness of food consumption relative to other goods and services can be interesting, for example, when debating a possible change of the indirect taxes (VAT) on food (which affects the relative price of food relative to other consumption). More generally, our analysis shows the possible use of our method to anticipate demand effects of rising or falling food prices (e.g. because of political intervention or changed conditions on the food world market). In our following analysis, we will also consider these price effects for the four quartiles of the observed expenditure distribution, so investigating whether these effects vary over income groups.

Our consumers are households, for which we deflate total expenditures by OECD equivalence scales to correct for household size.⁹ We control for observable characteristics by restricting our sample to households (i) that have completed the two-week diary, (ii) that are not living in student housing, (iii) that are vehicle owners, (iv) where both members work at least 17 hours, (v) in which both members are not self-employed, (vi) in which the age of the reference person is at least 21, and (vii) that consist of a husband, a wife and possibly children. By conditioning on housing status and vehicle ownership, we also mitigate the potential impact of our implicit assumption of separability between durable and non-durable consumption.

Given that the CEX diary data have been collected for only two weeks per household, we deseasonalized the original data by using a dummy regression approach before annualizing them. Next, we also removed a number of outlier observations.¹⁰ We can use an average of 2163 observations (households) per cross-section, with a minimum of 1775 in 1994 and a maximum of 2379 in 1995. The full data set is available from the authors upon request. Figure 4 plots the evolution of the relative price of food (left axis) and the mean consumption share of food (right axis) over the considered time period. This shows that there is not so much price variation, but there is quite some variation in the average share of food consumption. The latter motivates the empirical relevance of our exercise, while the former suggests that we will need to impose extra assumptions in order to obtain meaningful bounds.

Counterfactual demand distribution. As a first exercise, we study the demand distribution for a specific counterfactual price-income scenario. In particular, we set the new price p_0 equal to the mean of the cross-sectional prices p_t over all periods, and we use as reference expenditure level

⁹In principle, we could explicitly account for differences in household size by adopting a so-called collective model of household consumption (after Chiappori (1988)). See Cherchye, De Rock, and Vermeulen (2007) and Cherchye, De Rock, Lewbel, and Vermeulen (2015) for associated axioms of revealed preference. We see the extension of our method towards collective consumption models as an interesting avenue for follow-up research.

¹⁰In particular, we removed households for which rescaled total expenditures or expenditure shares are not within 3 standard deviations from the mean, households for which rescaled total expenditures are among the 5 percent lowest or 5 percent highest expenditures, and households for which expenditure shares are very close to 0.

Figure 4: Evolution of relative price and consumption share of food

x_0 the median household expenditures in 2007. For our sample of households, we compute bounds on the probability $\Pr[q_U(p_0, x_0) \leq q_0]$ for different values of q_0 . Specifically, we consider alternative budget shares of food $\frac{p_0 q_0}{x_0}$ for the given p_0 and x_0 .

Table 1 contains our main results. It reports bounds and 95% confidence intervals (CIs) for the upper and lower bounds on $\Pr[q_U(p_0, x_0) \leq q_0]$ when the food share $\frac{p_0 q_0}{x_0}$ varies between 0.05 and 0.8. The confidence intervals are based on 10,000 bootstrap samples. We consider three types of income elasticity restrictions: (i) no restrictions (i.e. the basic scenario of Section 4), (ii) both goods are normal (i.e. positive elasticities), and (iii) food is a necessary good (i.e. income elasticity between zero and one) and other consumption is normal. We believe these scenarios involve very reasonable assumptions for the goods under study.

A few interesting observations emerge from Table 1. If we do not restrict the income elasticities of the goods, our method does not really deliver informative bounds for the current sample of households. In this case, the lower and upper bounds on $\Pr[q_U(p_0, x_0) \leq q_0]$ are everywhere very close (or even equal) to the trivial bounds of, respectively, zero and one. As indicated above, the intuition is that our data set contains too little relative price variation to yield empirically useful results without additional assumptions.

Importantly, however, the bounds do become substantially more informative when we put restrictions on the shape of the income expansion paths. Both the normality assumption and the assumption that food is a necessary good contribute substantially to the identifying power of our method. In fact, for many values of $\frac{p_0 q_0}{x_0}$ our bounds on the counterfactual demand distribution become remarkably tight under the mere (and in our opinion fairly weak) prior that food is a necessary good while the demand for other goods and services is normal. This clearly demonstrates the empirical usefulness of our method. This conclusion is all the more attractive given that, apart from WARP as a minimal condition of rational consumer behavior, we impose no structure on the unobserved heterogeneity component and do not assume independence between preference heterogeneity and total expenditures.

Table 1: Bounds on $\Pr[q_U(p_0, x_0) \leq q_0]$

$\frac{p_0 q_0}{x_0}$		no elasticity bounds		two normal goods		food as necessity	
		lb	ub	lb	ub	lb	ub
0.05	bounds	0.0000	0.9949	0.0113	0.3515	0.0339	0.0428
	95% CI	0.0000	0.9949	0.0083	0.3755	0.0261	0.0516
0.10	bounds	0.0000	0.9958	0.0488	0.4524	0.0913	0.1658
	95% CI	0.0000	0.9958	0.0393	0.4746	0.0794	0.1869
0.15	bound	0.0007	0.9968	0.1101	0.5859	0.1950	0.3761
	95% CI	0.0001	0.9983	0.0974	0.6097	0.1757	0.4006
0.20	bounds	0.0008	0.9977	0.1996	0.7239	0.3401	0.5977
	95% CI	0.0002	0.9992	0.1863	0.7454	0.3177	0.6188
0.25	bounds	0.0013	0.9983	0.3111	0.8354	0.4827	0.7728
	95% CI	0.0005	0.9997	0.2906	0.8531	0.4633	0.7862
0.30	bounds	0.0025	0.9993	0.4086	0.9132	0.6207	0.8783
	95% CI	0.0012	0.9999	0.3896	0.9250	0.5991	0.892
0.35	bounds	0.0042	1.0000	0.4897	0.9574	0.7285	0.9368
	95% CI	0.0023	1.0000	0.4728	0.9632	0.7123	0.9458
0.40	bounds	0.0059	1.0000	0.5634	0.9777	0.8113	0.9687
	95% CI	0.0035	1.0000	0.5459	0.9833	0.7961	0.9757
0.45	bounds	0.0066	1.0000	0.6229	0.9873	0.8714	0.9819
	95% CI	0.0042	1.0000	0.6052	0.9921	0.8590	0.9887
0.50	bounds	0.0084	1.0000	0.6813	0.9928	0.9147	0.9861
	95% CI	0.0057	1.0000	0.6594	0.9968	0.9035	0.9938
0.55	bounds	0.0094	1.0000	0.7301	0.9958	0.9449	0.9928
	95% CI	0.0067	1.0000	0.7090	0.9988	0.9353	0.9977
0.60	bounds	0.0113	1.0000	0.7804	0.9981	0.9676	0.997
	95% CI	0.0077	1.0000	0.7606	0.9997	0.9581	0.9993
0.65	bounds	0.0156	1.0000	0.8233	0.9988	0.9798	0.9982
	95% CI	0.0107	1.0000	0.8054	0.9999	0.9721	0.9998
0.70	bounds	0.0164	1.0000	0.861	0.9994	0.9874	0.9988
	95% CI	0.0124	1.0000	0.8462	0.9999	0.9810	0.9999
0.75	bounds	0.0195	1.0000	0.9006	0.9994	0.9924	0.9994
	95% CI	0.0152	1.0000	0.8896	0.9999	0.9885	0.9999
0.80	bounds	0.021	1.0000	0.9475	1.0000	0.9963	1.0000
	95% CI	0.0162	1.0000	0.9357	1.0000	0.9933	1.0000

Bounds on quantiles. We next consider the demand effects of changing food prices for different quantiles of the demand distribution. We maintain the assumption that food is a necessary good and remaining consumption is normal. For food, we will also consider the additional scenario in which the income elasticity is bounded below by 0.5, i.e. $[\eta_\ell, \eta_h] = [0.5, 1]$. Like before, this will give us insight into the identifying power of more stringent elasticity bounds.

Before presenting our empirical results, we recall that, for a fixed price p_0 and expenditure level x_0 , the π th quantile of the distribution of $q_U(p_0, x_0)$ is given by the value $q(\pi)$ that defines

$$\pi = \Pr(q_U(p_0, x_0) \leq q(\pi)).$$

In order to construct bounds on the value of $q(\pi)$, we first parametrize the upper and lower bounds in terms of q_0 , i.e.

$$lb(q_0) = \max_{t \in \underline{T}} \ell_t(q_0) \leq \Pr(q_U(p_0, x_0) \leq q_0) \leq \min_{t \in \overline{T}} h_t(q_0) = ub(q_0),$$

It is easy to show that both $lb(q_0)$ and $ub(q_0)$ are monotone in q_0 . Now, define the functions $q_\ell(\pi)$ and $q_h(\pi)$ such that $\pi = lb(q_h(\pi))$ and $\pi = ub(q_\ell(\pi))$. Then, we have that

$$\begin{aligned} \Pr(q_U(p_0, x_0) \leq q(\pi)) &= \pi = ub(q_\ell(\pi)) \geq \Pr(q_U(p_0, x_0) \leq q_\ell(\pi)), \\ \Pr(q_U(p_0, x_0) \leq q(\pi)) &= \pi = lb(q_h(\pi)) \leq \Pr(q_U(p_0, x_0) \leq q_h(\pi)). \end{aligned}$$

As such, we see that $q_\ell(\pi) \leq q(\pi) \leq q_h(\pi)$.

Figures 5 and 6 present these lower and upper bounds for the two scenarios that we described above, with the elasticity intervals $[\eta_\ell, \eta_h] = [0, 1]$ and $[\eta_\ell, \eta_h] = [0.5, 1]$, respectively. Like before, we set x_0 equal to the median income in 2007, but now we vary the price of food p_0 (measured as a percentage change from the mean price level in our sample). The figures show bounds for the median ($\pi = 0.5$), the 25th percentile ($\pi = 0.25$) and the 75th percentile ($\pi = 0.75$) of the counterfactual distribution of the food share $pq(\pi)/x_0$.

We find that the quantile bounds are fairly tight for price values close to 1 (i.e. prices close to the mean food price in our sample), while they become uninformatively wide for price values around 0.9 or 1.1 (i.e. away from the bulk of observed prices). As expected, we also observe that the budget share for the 25th percentile of the counterfactual demand distribution is significantly below that for the 75th percentile. In Figure 5, there is no very pronounced difference between the bounds for the median quantile demand and the 25th or 75th percentile demand. By contrast, in Figure 6 we do see a clear difference, with the median demand situated distinctively between the 25th and 75th percentile. Generally, we find that the quantile bounds are substantially tighter in Figure 6 than in Figure 5, which again demonstrates the enhancing power effect of imposing specific restrictions on the income elasticity of food.

Figure 5: Quantiles bounds for $[\eta_\ell, \eta_h] = [0, 1]$ (food as necessary good)

Different income groups. As a final exercise, we analyze whether demand effects vary over income groups. To this end, we stratified our sample into four groups according to their quartile in the observed distribution of total expenditures (as a proxy for disposable income). For each of these subgroups, we redid our previous exercise but now we (only) looked at the median budget share of food as a function of varying food prices (again measured as a percentage change from the mean price level in our sample). In doing so, we fix the expenditure x_0 at the subgroup’s median level in 2007. Finally, in contrast to our previous exercises, we no longer impose that food should be a necessary good. Specifically, we use the elasticity interval $[\eta_\ell, \eta_h] = [0, 1.5]$, which means that food may well be a luxury (but not “too luxury”) good for some consumers.

Figure 7 summarizes our results. First, and not really surprisingly, we see clear differences in the (median) budget shares of the different quartiles. In general, the lower the income (quartile), the higher the budget share of food. Also, for each quartile we find that the budget share of food increases with its price. The extent to which the budget share increases is roughly the same for the different income groups; from the figure, we cannot identify an obvious difference in price responsiveness over subgroups. Also, within income subgroups, we do not see a lot of variation in the price responses when varying the food price, at least not for our sample of households. A notable observation is that the bounds in Figure 7 are again fairly tight, even when using a weaker prior than before (i.e. elasticity interval $[\eta_\ell, \eta_h] = [0, 1.5]$).

Summarizing, our different exercises clearly suggest the empirical usefulness of our method to address alternative questions related to counterfactual demand analysis. The method does allow for informative demand predictions under minimalistic prior assumptions. Our results also indicate that it may often be useful to impose specific restrictions on the income elasticities of the consumption goods (e.g. regarding the normal, necessary or luxury nature of goods). This makes our method particularly attractive for applied demand analysis in settings where such assumptions are fairly easy to justify.

Figure 6: Quantiles bounds for $[\eta_\ell, \eta_h] = [0.5, 1]$

7 Conclusion

We have proposed a novel method to derive WARP-consistent bounds on the distribution of demands for counterfactual income and price situations. Our method does not impose any condition on the unobserved heterogeneity component, and it does not assume (conditional) independence between total expenditures and unobserved preference heterogeneity. In addition, the method can readily incorporate restrictions on the income elasticities of the consumption goods, which further enhances its identifying power (i.e. tighter bounds). The possibility to include this type of restrictions is a useful feature from an applied point of view, as assumptions on income elasticities are often fairly easy to justify in real life applications (e.g. based on existing empirical work). More generally, our analysis fits within an interesting research avenue that combines more standard (often (semi)parametric) demand analysis, imposing specific restrictions on consumer demand behavior, with rationality constraints that are obtained from revealed preference theory.

The method is easy to implement and yields informative bounds on demand, which we have demonstrated through an empirical illustration to data drawn from the U.S. Consumer Expenditure Survey (CEX). In this application, we focused on the responsiveness of household demand for food relative to other (nondurable) goods and services as a function of food prices. First, we have shown that our method can deliver meaningfully tight bounds on the counterfactual demand distribution for a fixed new price. Next, we have illustrated that the method can yield informative results on the demand effects of changing prices, for specific quantiles of the demand distribution. Finally, we considered the median demand effect of changing prices for the four quartiles of the observed expenditure distribution in our sample. This last exercise showed the usefulness of our method to investigate differential demand effects over (observable) groups of consumers (in our case income groups).

In our empirical illustration we restricted our attention to demand responses due to price

Figure 7: Median demand bounds for income quartiles and $[\eta_\ell, \eta_h] = [0, 1.5]$

changes. In doing so, we have assumed expenditure levels that are fixed. As discussed in the Introduction, one may further investigate the impact of endogenous expenditures by considering some empirically specified distribution of (endogenous) expenditure levels (which, for example, may vary depending on the consumer type). Next, counterfactual demand responses are essential ingredients for welfare analyses that are based on cost of living indices (or, more generally, money metric welfare concepts), which are often studied in empirical policy evaluations. In our opinion, using our method for evaluating the welfare effects of counterfactual policies constitutes another fruitful avenue for follow-up research. Finally, as we have emphasized above, our method allows us to conduct counterfactual demand analyses without requiring appropriate instruments to deal with endogenous expenditures. Alternatively, one may actually use our framework to evaluate the validity of potential instruments. By construction, the (bounds on) counterfactual demand responses obtained through our revealed preference method must include the ones generated by more structured (semi)parametric methods that make use of specific (valid) instruments.

References

- Andrews, D. W. K., Soares, G., 2010. Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* 78, 119–157.
- Beckert, W., Blundell, R., 2008. Heterogeneity and the non-parametric analysis of consumer choice: Conditions for invertibility. *Review of Economic Studies* 75, 1069–1080.
- Blundell, R., Browning, M., Cherchye, L., Crawford, I., De Rock, B., Vermeulen, F., 2015. Sharp for SARP: Nonparametric bounds on the behavioral and welfare effects of price changes. *American Economic Journal: Microeconomics* 7, 43–60.

- Blundell, R., Browning, M., Crawford, I., 2003. Nonparametric engel curves and revealed preference. *Econometrica* 71, 205–240.
- Blundell, R., Browning, M., Crawford, I., 2007a. Improving revealed preference bounds on demand responses. *International Economic Review* 48, 1227–1244.
- Blundell, R., Browning, M., Crawford, I., 2008. Best nonparametric bounds on demand responses. *Econometrica* 76, 1227–1262.
- Blundell, R., Chen, X., Kristensen, D., 2007b. Semi-nonparametric IV estimation of shape-invariant engel curves. *Econometrica* 75, 1613–1669.
- Blundell, R., Duncan, A., 1998. Kernel regression in empirical micro-economics. *Journal of Human Resources* 33, 62–87.
- Blundell, R., Horowitz, J., Parey, M., 2017. Nonparametric estimation of a nonseparable demand function under the Slutsky inequality restriction. *Review of Economics and Statistics* 99, 291–304.
- Blundell, R., Horowitz, J. L., Parey, M., 2012. Measuring the price responsiveness of gasoline demand: Economic shape restrictions and nonparametric demand estimation. *Quantitative Economics* 3, 29–51.
- Blundell, R., Kristensen, D., Matzkin, R., 2014. Bounding quantile demand functions using revealed preference inequalities. *Journal of Econometrics* 179, 112–127.
- Brown, B. W., Walker, M. B., 1989. The random utility hypothesis and inference in demand systems. *Econometrica* 57, 815–829.
- Bugni, F. A., 2010. Bootstrap inference in partially identified models defined by moment inequalities: Coverage of the identified set. *Econometrica* 78, 735–753.
- Cherchye, L., De Rock, B., Lewbel, A., Vermeulen, F., 2015. Sharing rule identification for general collective consumption models. *Econometrica* 83, 2001–20041.
- Cherchye, L., De Rock, B., Vermeulen, F., 2007. The collective model of household consumption: a nonparametric characterization. *Econometrica* 75, 553–574.
- Cherchye, L., Demuynck, T., De Rock, B., 2018. Transitivity of preferences: when does it matter? *Theoretical Economics* 13, 1043–1076.
- Chernozhukov, V., Hong, H., Tamer, E., 2007. Estimation and confidence regions for parameter sets in econometric models. *Econometrica* 75, 1243–1284.
- Chernozhukov, V., Lee, S., Rosen, A. M., 2013. Intersection bounds: Estimation and inference. *Econometrica* 81, 667–737.

- Chiappori, P.-A., 1988. Rational household labor supply. *Econometrica* 56, 63–89.
- Cosaert, S., Demuynck, T., 2018. Nonparametric welfare and demand analysis with unobserved individual heterogeneity. *Review of Economic and Statistics* 100, 348–361.
- Dette, H., Hoderlein, S., Neumeyer, N., 2016. Testing multivariate economic restrictions using quantiles: The example of Slutsky negative semidefiniteness. *Journal of Econometrics* 191, 129–144.
- Hausman, J. A., Newey, W. K., 2016. Individual heterogeneity and average welfare. *Econometrica* 84, 1225–1248.
- Hoderlein, S., 2011. How many consumers are rational. *Journal of Econometrics* 164, 294–309.
- Hoderlein, S., Stoye, J., 2014. Revealed preference in a heterogeneous population. *Review of Economics and Statistics* 96, 197–213.
- Hoderlein, S., Vanhems, A., 2018. Estimating the distribution of welfare effects using quantiles. *Journal of Applied Econometrics* 33, 52–72.
- Imbens, G. W., Manski, C. F., 2004. Confidence interval for partially identified parameters. *Econometrica* 72, 1845–1857.
- Kawaguchi, K., 2017. Testing rationality without restricting heterogeneity. *Journal of Econometrics* 197, 153–171.
- Kitamura, Y., Stoye, J., 2018. Nonparametric analysis of random utility models. *Econometrica* 86, 1883–1909.
- Lewbel, A., 2001. Demand systems with and without errors. *American Economic Review* 91, 611–618.
- Manski, C. F., 1997. Monotone treatment response. *Econometrica* 65, 1311–1334.
- Matzkin, R. L., 2003. Nonparametric estimation of nonadditive random functions. *Econometrica* 71, 1339–1375.
- Matzkin, R. L., 2008. Identification in nonparametric simultaneous equations models. *Econometrica* 76, 945–978.
- McFadden, D. L., 2005. Revealed stochastic preference: a synthesis. *Economic Theory* 26, 245–264.
- Rose, H., 1958. Consistency of preference: The two-commodity case. *Review of Economic Studies* 25, 124–125.
- Samuelson, P. A., 1938. A note on the pure theory of consumer’s behavior. *Economica* 5, 61–71.

Stoye, J., 2009. More on confidence intervals for partially identified parameters. *Econometrica* 77, 1299–1315.

A Appendix

A.1 General setting with n goods

In this appendix, we derive the sharpest WARP-consistent bounds for a general setting with n goods. Consumption bundles are denoted by $\mathbf{q} \in \mathbb{R}_+^n$ and price vectors by $\mathbf{p} \in \mathbb{R}_{++}^n$. Individual heterogeneity is captured by a random variable U with values $u \in \mathcal{U}$. Every consumer with heterogeneity u has WARP-consistent demand functions $\mathbf{q}_u(\mathbf{p}, x)$, which give the desired consumption bundle given prices $\mathbf{p}$ and income x . For a given cross-section t , we observe the distribution of $(\mathbf{p}_t, \mathbf{q}_{t,U}, x_{t,U})$ where $x_{t,U} = \mathbf{p}_t \mathbf{q}_{t,U}$.

A.1.1 Bounds on counterfactual demands

Consider a counterfactual price $\mathbf{p}_0 \in \mathbb{R}_{++}^n$ and income level $x_0 > 0$. We denote the budget hyperplane associated with the prices $\mathbf{p}_0$ and income x_0 by $B(\mathbf{p}_0, x_0) = \{\mathbf{q} \in \mathbb{R}_+^n : \mathbf{p}_0 \mathbf{q} = x_0\}$. Let $A \subset B(\mathbf{p}_0, x_0)$ and define $\neg A = B(\mathbf{p}_0, x_0) \setminus A$ to be the complement of A relative to the budget hyperplane $B(\mathbf{p}_0, x_0)$. We are interested in obtaining bounds on the proportion of consumers u for which $\mathbf{q}_u(\mathbf{p}_0, x_0) \in A$, i.e. $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$.

If consumer u is WARP-consistent, then the the following condition must be satisfied:

$$x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u} \Rightarrow x_{t,u} < \mathbf{p}_t \mathbf{q}_u(\mathbf{p}_0, x_0) \text{ or } \mathbf{q}_{t,u} = \mathbf{q}_u(\mathbf{p}_0, x_0).$$

We begin by computing an upper bound on $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$. To do so, we exploit this WARP condition and identify all consumers u for which this WARP condition does not exclude the possibility that $\mathbf{q}_u(\mathbf{p}_0, x_0) \in A$. These are the consumers u for which either (i) $x_0 < \mathbf{p}_0 \mathbf{q}_{t,u}$ or (ii) $x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u}$ and (either $\mathbf{q}_{t,u} \in A$ or there is at least one $\mathbf{q} \in A$ for which $x_{t,u} < \mathbf{p}_t \mathbf{q}$). We gather these consumers in the set $E_t(A)$:

$$\begin{aligned} E_t(A) &= \left\{ u \in \mathcal{U} : \begin{array}{l} x_0 < \mathbf{p}_0 \mathbf{q}_{t,u}, \text{ or} \\ (x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u} \text{ and } (\mathbf{q}_{t,u} \in A \text{ or } \exists \mathbf{q} \in A : x_{t,u} < \mathbf{p}_t \mathbf{q})) \end{array} \right\} \\ &= \left\{ u \in \mathcal{U} : \begin{array}{l} x_0 < \mathbf{p}_0 \mathbf{q}_{t,u}, \text{ or} \\ (x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u} \text{ and } (\mathbf{q}_{t,u} \in A \text{ or } x_{t,u} < \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q})) \end{array} \right\} \\ &= \left\{ u \in \mathcal{U} : x_0 < \mathbf{p}_0 \mathbf{q}_{t,u} \text{ or } \mathbf{q}_{t,u} \in A \text{ or } x_{t,u} < \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q} \right\}. \end{aligned}$$

The construction of the set $E_t(A)$ is such that, for all WARP-consistent consumers in $E_t(A)$, it is not possible to exclude that $\mathbf{q}_u(\mathbf{p}_0, x_0) \in A$. Taking the minimum across all cross-sections then

obtains the following tightest WARP-consistent upper bound:

$$\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A) \leq \min_{t \in \{1, \dots, T\}} \Pr(U \in E_t(A)).$$

We next turn to the construction of a lower bound on $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$. We have

$$\begin{aligned} \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \notin A) &= \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in \neg A) \leq \Pr(U \in E_t(\neg A)) \\ \Leftrightarrow 1 - \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \notin A) &\geq 1 - \Pr(U \in E_t(\neg A)) \\ \Leftrightarrow 1 - \Pr(U \in E_t(\neg A)) &\leq \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A). \end{aligned}$$

The lower bound can be rewritten in the following way:

$$\begin{aligned} 1 - \Pr(U \in E_t(\neg A)) &= 1 - \Pr\left(x_0 < \mathbf{p}_0 \mathbf{q}_{t,U} \text{ or } \mathbf{q}_{t,u} \in \neg A \text{ or } x_{t,U} < \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q}\right) \\ &= \Pr\left(x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,U} \text{ and } \mathbf{q}_{t,u} \notin \neg A \text{ and } x_{t,U} \geq \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q}\right). \end{aligned}$$

In order to see that this is indeed the tightest WARP-consistent lower bound for cross-section t , let us show that this probability equals the proportion of consumers $u \in \mathcal{U}$ for which we are absolutely sure that WARP implies that $\mathbf{q}_u(\mathbf{p}_0, x_0) \in A$. This equals the proportion of consumers u such that, for all $\mathbf{q} \in \neg A$, the set $(\mathbf{p}_t, \mathbf{q}_{t,u}, \mathbf{p}_0, \mathbf{q})$ must be a WARP violation. These are the consumers that belong to the following set:

$$\begin{aligned} F_t(A) &= \{u \in \mathcal{U} : x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u} \text{ and } \forall \mathbf{q} \in \neg A : \mathbf{q} \neq \mathbf{q}_{t,u} \text{ and } x_{t,u} \geq \mathbf{p}_t \mathbf{q}\} \\ &= \left\{u \in \mathcal{U} : x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u} \text{ and } \mathbf{q}_{t,u} \notin \neg A \text{ and } x_{t,u} \geq \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q}\right\}. \end{aligned}$$

Taking the maximum across all cross-sections then obtains the following tightest WARP-consistent lower bound:

$$\max_{t \in \{1, \dots, T\}} \Pr(U \in F_t(A)) \leq \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A).$$

We summarize our findings in the following theorem.

Theorem 5. *We have that*

$$\max_{t \in \{1, \dots, T\}} \Pr(U \in F_t(A)) \leq \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A) \leq \min_{t \in \{1, \dots, T\}} \Pr(U \in E_t(A)),$$

with

$$F_t(A) = \left\{ u \in \mathcal{U} : x_0 \geq \mathbf{p}_0 \mathbf{q}_{t,u} \text{ and } \mathbf{q}_{t,u} \notin \neg A \text{ and } x_{t,u} \geq \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q} \right\},$$

and

$$E_t(A) = \left\{ u \in \mathcal{U} : x_0 < \mathbf{p}_0 \mathbf{q}_{t,u} \text{ or } \mathbf{q}_{t,u} \in A \text{ or } x_{t,u} < \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q} \right\}.$$

These bounds are the tightest WARP-consistent bounds.

A.1.2 Expansion path bounds

For a consumer $u \in \mathcal{U}$, and a given price vector $\mathbf{p}_t$, the expansion path $\mathbf{q}_u(\mathbf{p}_t, x)$ gives the demanded bundles of consumer u as function of income x .¹¹ As indicated in the main text, we only focus on a setting where all goods are normal, which implies that all expansion paths are non-decreasing.

We only observe bundles $\mathbf{q}_{t,u}$ for a finite set of observations and consumers, and thus not the full expansion paths. This implies that for each consumer u and each observation t we need to consider several possibilities for their expansion paths. For a cross-section t and a consumer u with observation $(\mathbf{q}_{t,u}, x_{t,u})$, we denote by $\mathcal{F}_{t,u}$ a possible collection of non-decreasing expansion paths $\mathbf{q}_u(\mathbf{p}_t, \cdot)$ of consumer u that satisfy $\mathbf{q}_u(\mathbf{p}_t, x_{t,u}) = \mathbf{q}_{t,u}$. The exact definition of $\mathcal{F}_{t,u}$ will depend on the model under consideration. In the weakest setting, $\mathcal{F}_{t,u}$ contains all increasing expansion paths satisfying $\mathbf{q}_u(\mathbf{p}_t, x_{t,u}) = \mathbf{q}_{t,u}$. In the strongest setting, $\mathcal{F}_{t,u}$ consists of exactly one non-decreasing expansion path through $\mathbf{q}_{t,u}$. In intermediate settings, $\mathcal{F}_{t,u}$ will contain a subset of expansion paths (e.g., all expansion paths with income elasticities of the goods belonging to some bounded set).

For a given expenditure level $x \in \mathbb{R}_+$, consider the set of possible bundles with expenditure x that agree with some expansion path in $\mathcal{F}_{t,u}$. We collect these vectors in the set $Q(\mathcal{F}_{t,u}, x)$, which gives

$$Q(\mathcal{F}_{t,u}, x) = \bigcup_{\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}} \mathbf{q}_u(\mathbf{p}_t, x).$$

We make the mild assumption that $Q(\mathcal{F}_{t,u}, x)$ is lower hemi-continuous in $x \in \mathbb{R}_+$.¹² This is satisfied as soon as income expansion paths are continuous.

The following lemma is similar in spirit to the reasoning in Blundell, Browning, and Crawford (2003, 2007a, 2008). It states that it suffices to focus on a set of bundles instead of a collection of expansion paths.

Lemma 2. *Let $\mathbf{q} \in B(\mathbf{p}_0, x_0)$. There exists an expansion path $\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}$ such that, for all expenditure levels $x \in \mathbb{R}_+$, $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \mathbf{q}_u(\mathbf{p}_t, x))$ is WARP-consistent if and only if there exists a*

¹¹Formally, an expansion path is a function $\mathbf{q}_u(\mathbf{p}_t, \cdot)$ from $\mathbb{R}_+$ to $\mathbb{R}_+^n$ such that, for all $x \in \mathbb{R}_+$, $\mathbf{p}_t \mathbf{q}_u(\mathbf{p}_t, x) = x$.

¹²This means that if we have a sequence $x_n \rightarrow x$ and $\mathbf{q} \in Q(\mathcal{F}_{t,u}, x)$, then there is a number N and a sequence $\mathbf{q}_n \rightarrow \mathbf{q}$ such that $\mathbf{q}_n \in Q(\mathcal{F}_{t,u}, x_n)$ for all $n \geq N$.

$\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$ such that $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \tilde{\mathbf{q}})$ is WARP-consistent.

Proof. The necessity trivially follows from setting $x = \mathbf{p}_t \mathbf{q}$ for the specific expansion path. To show sufficiency, let us assume the statement is not true. That is, for all expansion paths $\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}$, there is an income level $x \in \mathbb{R}_+$ such that $(\mathbf{p}_0, \mathbf{q}, \mathbf{p}_t, \mathbf{q}_u(\mathbf{p}_t, x))$ is WARP-violating, and there also exists a bundle $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$ such that $(\mathbf{p}_0, \mathbf{q}, \mathbf{p}_t, \tilde{\mathbf{q}})$ is WARP-consistent.

Take an expansion path $\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}$ such that $\mathbf{q}_u(\mathbf{p}_t, \mathbf{p}_t \mathbf{q}) = \tilde{\mathbf{q}}$. This is always possible as $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$. Then there should be an x such that $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \mathbf{q}_u(\mathbf{p}_t, x))$ is WARP-violating, i.e.

$$x_0 \geq \mathbf{p}_0 \mathbf{q}_u(\mathbf{p}_t, x), x \geq \mathbf{p}_t \mathbf{q} \text{ and } \mathbf{q}_u(\mathbf{p}_t, x) \neq \mathbf{q}.$$

Since $\mathbf{q}_u(\mathbf{p}_t, \mathbf{p}_t \mathbf{q}) = \tilde{\mathbf{q}}$, we have that $\mathbf{p}_t \tilde{\mathbf{q}} = \mathbf{p}_t \mathbf{q}$ and $x_0 \geq \mathbf{p}_0 \tilde{\mathbf{q}}$. This implies that $(\mathbf{p}_0, \mathbf{q}, \mathbf{p}_t, \tilde{\mathbf{q}})$ gives a WARP violation, unless $\mathbf{q} = \tilde{\mathbf{q}}$. However, since all expansion paths are non-decreasing, the second inequality then implies $\mathbf{q}_u(\mathbf{p}_t, x) \geq \mathbf{q}_u(\mathbf{p}_t, \mathbf{p}_t \mathbf{q}) = \tilde{\mathbf{q}} = \mathbf{q}$. As such $x_0 = \mathbf{p}_0 \mathbf{q} \geq \mathbf{p}_0 \mathbf{q}_u(\mathbf{p}_t, x)$ obtains that $\mathbf{q} = \mathbf{q}_u(\mathbf{p}_t, x)$. This contradiction finishes the proof. $\square$

As before, we take a counterfactual price $\mathbf{p}_0$ and income level x_0 . Let $A \subset B(\mathbf{p}_0, x_0)$ and define again $\neg A = B(\mathbf{p}_0, x_0) \setminus A$. We need to obtain the tightest WARP-consistent lower and upper bounds on the probability $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$, while including information on the sets of feasible expansion paths $\mathcal{F}_{t,u}$.

Following a similar argument as in Appendix A.1.1, we begin by constructing the tightest upper bound on the probability $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$ by identifying all consumers $u \in \mathcal{U}$ such that, for at least one expansion path $\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}$ and one bundle $\mathbf{q} \in A$, the sets $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \mathbf{q}_u(\mathbf{p}_t, x))$ are WARP-consistent for all $x \in \mathbb{R}_+$. Lemma 2 states that this is equivalent to requiring that, for at least one bundle $\mathbf{q} \in A$ and at least one bundle $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$, the set $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \tilde{\mathbf{q}})$ is WARP-consistent. This gives the set

$$E_t(A) = \{u \in \mathcal{U} : \exists \mathbf{q} \in A, \exists \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}) \text{ such that } (\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \tilde{\mathbf{q}}) \text{ is WARP-consistent}\}.$$

As argued in Appendix A.1.1, $\Pr(U \in E_t(A))$ is the tightest WARP-consistent upper bound on $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$ given cross-section t . Therefore,

$$\Pr(\mathbf{q}_U(\mathbf{p}_t, x_0) \in A) \leq \min_{t \in T} \Pr(U \in E_t(A))$$

is the tightest upper bound across all observations. For $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$, we have $\mathbf{p}_t \tilde{\mathbf{q}} = \mathbf{p}_t \mathbf{q}$ by definition. As such WARP-consistency of $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \tilde{\mathbf{q}})$ is equivalent to $\mathbf{p}_0 \mathbf{q} = x_0 < \mathbf{p}_0 \tilde{\mathbf{q}}$ or $\tilde{\mathbf{q}} = \mathbf{q}$.

This allows the following reformulation

$$\begin{aligned}
E_t(A) &= \{u \in \mathcal{U} : \exists \mathbf{q} \in A, \exists \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}), x_0 < \mathbf{p}_0 \tilde{\mathbf{q}} \text{ or } \tilde{\mathbf{q}} = \mathbf{q}\} \\
&= \left\{ u \in \mathcal{U} : \exists \mathbf{q} \in A, x_0 < \sup_{\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ or } \mathbf{q} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}) \right\} \\
&= \left\{ u \in \mathcal{U} : \left(x_0 < \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}) \text{ and } \mathbf{q} \in A \right) \text{ or } \exists x \in \mathbb{R}_+, A \cap Q(\mathcal{F}_{t,u}, x) \neq \emptyset \right\}.
\end{aligned}$$

The following lemma states that it suffices to focus on $x_t^* = \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q}$ to characterize $E_t(A)$.

Lemma 3. *Let $x_t^* = \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q}$. Then $v \equiv \sup_{\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})} \mathbf{p}_0 \tilde{\mathbf{q}}$ s.t. $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$ and $\mathbf{q} \in A$ has the same value as $w \equiv \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}}$ s.t. $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^*)$.*

Proof. To show $v \geq w$, take any $\varepsilon > 0$ and let $\tilde{\mathbf{q}}^* \in Q(\mathcal{F}_{t,u}, x_t^*)$ be such that $w - \varepsilon < \mathbf{p}_0 \tilde{\mathbf{q}}^*$. Consider a sequence $\mathbf{q}_n^* \in A$ such that $\mathbf{p}_t \mathbf{q}_n^* \rightarrow x_t^*$. Since $\tilde{\mathbf{q}}^* \in Q(\mathcal{F}_{t,u}, x_t^*)$, the lower hemi-continuity of $Q(\mathcal{F}_{t,u}, \cdot)$ implies that there exists a N and a sequence $\tilde{\mathbf{q}}_n^* \rightarrow \tilde{\mathbf{q}}^*$ such that, $\tilde{\mathbf{q}}_n^* \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}_n^*)$ for all $n \geq N$. For all $n \geq N$, we have $\mathbf{q}_n^* \in A$ and $\tilde{\mathbf{q}}_n^* \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}_n^*)$, which implies $v \geq \mathbf{p}_0 \tilde{\mathbf{q}}_n^*$. Taking limits then gives $v \geq \mathbf{p}_0 \tilde{\mathbf{q}}^* > w - \varepsilon$. This conclusion holds for any $\varepsilon > 0$, which obtains $v \geq w$.

Conversely, to show $v \leq w$, take any $\varepsilon > 0$ and let $\mathbf{q}^* \in A$ and $\tilde{\mathbf{q}}^* \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}^*)$ such that $v - \varepsilon < \mathbf{p}_0 \tilde{\mathbf{q}}^*$. Since $\mathbf{p}_t \mathbf{q}^* \leq x_t^*$ and expansion paths are non-decreasing, we have, for any $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}^*)$, that there exists a $\tilde{\mathbf{q}}' \in Q(\mathcal{F}_{t,u}, x_t^*)$ such that $\tilde{\mathbf{q}}' \geq \tilde{\mathbf{q}}$. For $\tilde{\mathbf{q}}^*$ we therefore obtain that there exists $\tilde{\mathbf{q}}' \in Q(\mathcal{F}_{t,u}, x_t^*)$ such that

$$v < \mathbf{p}_0 \tilde{\mathbf{q}}^* + \varepsilon \leq \mathbf{p}_0 \tilde{\mathbf{q}}' + \varepsilon \leq w + \varepsilon.$$

This conclusion holds for any $\varepsilon > 0$, which obtains $v \leq w$. □

Lemma 3 allows for the following simplest reformulation of $E_t(A)$:

$$E_t(A) = \left\{ u \in \mathcal{U} : \left(x_0 < \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^*) \right) \text{ or } \exists x \in \mathbb{R}_+, A \cap Q(\mathcal{F}_{t,u}, x) \neq \emptyset \right\},$$

with $x_t^* = \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q}$.

The construction of a lower bound on $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$ again proceeds similarly as in Appendix A.1.1:

$$\begin{aligned}
\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \notin A) &= \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in \neg A) \leq \Pr(U \in E_t(\neg A)) \\
&\Leftrightarrow 1 - \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A) \geq 1 - \Pr(U \in E_t(\neg A)) \\
&\Leftrightarrow 1 - \Pr(U \in E_t(\neg A)) \leq \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A).
\end{aligned}$$

Using the above we have

$$E_t(\neg A) = \left\{ u \in \mathcal{U} : \begin{array}{l} x_0 < \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^{**}) \text{ or} \\ \exists x \in \mathbb{R}_+, \neg A \cap Q(\mathcal{F}_{t,u}, x) \neq \emptyset \end{array} \right\},$$

with $x_t^{**} = \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q}$. As such

$$1 - \Pr(U \in E_t(\neg A)) = \Pr \left(\begin{array}{l} x_0 \geq \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^{**}) \text{ and} \\ \forall x \in \mathbb{R}_+, \neg A \cap Q(\mathcal{F}_{t,u}, x) = \emptyset \end{array} \right).$$

The intuition that this defines for a given cross-section the tightest WARP-consistent lower bound is similar to Appendix A.1.1. Consider all consumers $u \in \mathcal{U}$ such that for all $\mathbf{q} \in B(\mathbf{p}_0, x_0)$, the set $(\mathbf{p}_t, \mathbf{q}_u(\mathbf{p}_t, x); \mathbf{p}_0, \mathbf{q})$ is WARP-consistent for all $\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}$ and all $x \in \mathbb{R}_+$ only if $\mathbf{q} \in A$. This equals the fraction of consumers such that, for all $\mathbf{q} \in \neg A$ and all expansion paths $\mathbf{q}_u(\mathbf{p}_t, \cdot) \in \mathcal{F}_{t,u}$, $(\mathbf{p}_t, \mathbf{q}_u(\mathbf{p}_t, x); \mathbf{p}_0, \mathbf{q})$ is WARP-violating for at least one value of $x \in \mathbb{R}_+$. Given Lemma 2 it is sufficient to show that, for all $\mathbf{q} \in \neg A$ and all $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$, the set $(\mathbf{p}_0, \mathbf{q}; \mathbf{p}_t, \tilde{\mathbf{q}})$ is WARP-violating. This gives the following set:

$$F_t(A) = \{u \in \mathcal{U} : \forall \mathbf{q} \in \neg A, \forall \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}), (\mathbf{p}_0, \mathbf{q}, \mathbf{p}_t, \tilde{\mathbf{q}}) \text{ is WARP-violating}\}.$$

As $\tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})$, we know that $\mathbf{p}_t \tilde{\mathbf{q}} = \mathbf{p}_t \mathbf{q}$ and $(\mathbf{p}_0, \mathbf{q}, \mathbf{p}_t, \tilde{\mathbf{q}})$ is therefore WARP-violating if and only if $\mathbf{p}_0 \mathbf{q} = x_0 \geq \mathbf{p}_0 \tilde{\mathbf{q}}$ and $\mathbf{q} \neq \tilde{\mathbf{q}}$. This allows the following reformulation

$$\begin{aligned} F_t(A) &= \{u \in \mathcal{U} : \forall \mathbf{q} \in \neg A, \forall \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}), x_0 \geq \mathbf{p}_0 \tilde{\mathbf{q}} \text{ and } \mathbf{q} \neq \tilde{\mathbf{q}}\} \\ &= \left\{ u \in \mathcal{U} : \forall \mathbf{q} \in \neg A, x_0 \geq \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}) \text{ and } \mathbf{q} \notin Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q}) \right\} \\ &= \left\{ u \in \mathcal{U} : \begin{array}{l} (x_0 \geq \sup_{\mathbf{q}, \tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \mathbf{q} \in \neg A \text{ and } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, \mathbf{p}_t \mathbf{q})) \\ \text{and } (\forall x \in \mathbb{R}_+, \neg A \cap Q(\mathcal{F}_{t,u}, x) = \emptyset) \end{array} \right\} \\ &= \left\{ u \in \mathcal{U} : \left(x_0 \geq \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^{**}) \right) \text{ and } (\forall x \in \mathbb{R}_+, \neg A \cap Q(\mathcal{F}_{t,u}, x) = \emptyset) \right\}, \end{aligned}$$

with $x_t^{**} = \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q}$.

Summarizing, we obtain the following result.

Theorem 6. *Assume that all goods are normal. We have that*

$$\max_{t \in \{1, \dots, T\}} \Pr(U \in F_t(A)) \leq \Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A) \leq \min_{t \in \{1, \dots, T\}} \Pr(U \in E_t(A)),$$

where

$$E_t(A) = \left\{ u \in \mathcal{U} : \left(x_0 < \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^*) \right) \text{ or } \exists x \in \mathbb{R}_+, A \cap Q(\mathcal{F}_{t,u}, x) \neq \emptyset \right\},$$

with $x_t^* = \sup_{\mathbf{q} \in A} \mathbf{p}_t \mathbf{q}$ and

$$F_t(A) = \left\{ u \in \mathcal{U} : \left(x_0 \geq \sup_{\tilde{\mathbf{q}}} \mathbf{p}_0 \tilde{\mathbf{q}} \text{ s.t. } \tilde{\mathbf{q}} \in Q(\mathcal{F}_{t,u}, x_t^{**}) \right) \text{ and } (\forall x \in \mathbb{R}_+, \neg A \cap Q(\mathcal{F}_{t,u}, x) = \emptyset) \right\},$$

with $x_t^{**} = \sup_{\mathbf{q} \in \neg A} \mathbf{p}_t \mathbf{q}$. These bounds are the tightest WARP-consistent bounds.

A.2 Proofs of Theorems 1-4 and Lemma 1

In what follows, we consider the setting with two goods, which we treat as a special case of the case with n goods discussed in Appendix A.1. Using our notation of the main text, we denote the two quantities consumed by q and z (i.e. $\mathbf{q} = (q, z)$ when using our notation of Appendix A.1). We normalize the price of z at 1, and we let p represent the normalized price of q . To formally link our set-up in the main text to the one in Appendix A.1, we let $A = \{(q, z) \in \mathbb{R}_+^2 : p_0 q + z = x_0 \text{ and } q \in [0, q_0]\}$, which is a subset of $B(p_0, x_0)$. We are interested in $Pr(q_U(p_0, x_0) \leq q_0)$ for any $q_0 \in]0, \frac{x_0}{p_0}[$.

A.2.1 Proof of Theorem 1

In Appendix A.1, we showed that we can obtain the lower bound on $Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$ by computing the probability of the event that U belongs to

$$F_t(A) = \left\{ u \in \mathcal{U} : x_0 \geq p_0 q_{t,u} + z_{t,u} \text{ and } (q_{t,u}, z_{t,u}) \notin \neg A \text{ and } x_{t,u} \geq \sup_{(q,z) \in \neg A} p_t q + z \right\}.$$

We can derive the following two results for the cases $p_t > p_0$ and $p_t < p_0$, respectively.

Lemma 4. *If $p_t > p_0$ then $F_t(A) = \emptyset$.*

Proof. Towards a contradiction, assume that $p_t > p_0$ and $u \in F_t(A)$. By substituting out z , we see that $x_{t,u} \geq \sup_{(q,z) \in \neg A} p_t q + z$ is equivalent to

$$\begin{aligned} x_{t,u} &\geq \sup p_t q + z \text{ s.t. } (q, z) \in \neg A \\ &= \sup (p_t - p_0)q + x_0 \text{ with } q \in]q_0, x_0/p_0] \\ &= (p_t - p_0) \frac{x_0}{p_0} + x_0 \\ &= p_t \frac{x_0}{p_0}. \end{aligned}$$

Together with $x_0 \geq p_0 q_{t,u} + z_{t,u}$ this implies

$$\frac{p_0}{p_t} x_{t,u} = \frac{p_0}{p_t} (p_t q_{t,u} + z_{t,u}) = p_0 q_{t,u} + \frac{p_0}{p_t} z_{t,u} \geq x_0 \geq p_0 q_{t,u} + z_{t,u},$$

which gives $\frac{p_0}{p_t} z_{t,u} \geq z_{t,u}$. This inequality can only hold if $z_{t,u} = 0$, and thus $q_{t,u} = \frac{x_{t,u}}{p_t}$. But then we obtain $x_0 \geq p_0 q_{t,u} = p_0 \frac{x_{t,u}}{p_t} \geq x_0$, and we can conclude that $q_{t,u} = \frac{x_{t,u}}{p_t} = \frac{x_0}{p_0}$. This implies that $q_{t,u} > q_0$, since $q_0 < \frac{x_0}{p_0}$. This yields a contradiction with $(q_{t,u}, z_{t,u}) \notin \neg A$. $\square$

Lemma 5. *If $p_t < p_0$ then $\Pr(U \in F_t(A)) = \Pr(x_0 \geq p_0 q_{t,U} + z_{t,U} \text{ and } x_{t,U} \geq p_t q_0 + z_0)$.*

Proof. We have that

$$\begin{aligned} & \Pr(x_0 \geq p_0 q_{t,U} + z_{t,U} \text{ and } x_{t,U} \geq p_t q_0 + z_0) \\ &= \Pr(x_0 \geq p_0 q_{t,U} + z_{t,U} \text{ and } (q_{t,U}, z_{t,U}) \notin \neg A \text{ and } x_{t,U} \geq \sup_{(q,z) \in \neg A} p_t q + z) \\ &+ \Pr(x_0 \geq p_0 q_{t,U} + z_{t,U} \text{ and } (q_{t,U}, z_{t,U}) \in \neg A \text{ and } x_{t,U} \geq \sup_{(q,z) \in \neg A} p_t q + z). \end{aligned}$$

It thus suffices to show that the last term equals zero. Similar to before, if $p_t < p_0$, then (by substituting out z)

$$\begin{aligned} x_{t,u} &\geq \sup p_t q + z \text{ s.t. } (q, z) \in \neg A \\ &= \sup (p_t - p_0)q + x_0 \text{ with } q \in]q_0, \frac{x_0}{p_0}] \\ &= (p_t - p_0)q_0 + x_0. \end{aligned}$$

Next, substituting out $z_{t,u}$ in $x_0 \geq p_0 q_{t,u} + z_{t,u}$ obtains $x_0 \geq (p_0 - p_t)q_{t,u} + x_{t,u}$. These two inequalities taken together then imply

$$\begin{aligned} x_0 &\geq (p_0 - p_t)q_{t,u} + x_{t,u} \geq (p_0 - p_t)q_{t,u} + (p_t - p_0)q_0 + x_0 \\ &= (p_0 - p_t)(q_{t,u} - q_0) + x_0. \end{aligned}$$

Given that $p_0 > p_t$, we thus obtain that we need $q_0 \geq q_{t,u}$. This implies that it can never be that $(q_{t,U}, z_{t,U}) \in \neg A$. $\square$

Lemmata 4 and 5 define the tightest lower bound for all cross-sections t . Taking the maximum over all the available cross-sections gives us the lower bound in Theorem 1.

A.2.2 Proof of Theorem 2

As shown in Appendix A.1, we can obtain the upper bound on $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$ by computing the probability of the event that U belongs to

$$E_t(A) = \left\{ u \in \mathcal{U} : x_0 < p_0 q_{t,u} + z_{t,u} \text{ or } (q_{t,u}, z_{t,u}) \in A \text{ or } x_{t,u} < \sup_{(q,z) \in A} p_t q + z \right\}.$$

In this case, we can derive the following two results for $p_t < p_0$ and $p_t > p_0$.

Lemma 6. *If $p_t < p_0$ then $E_t(A) = \mathcal{U}$.*

Proof. Towards a contradiction, let $p_t < p_0$ and assume that $u \notin E_t(A)$. That is

$$x_0 \geq p_0 q_{t,u} + z_{t,u} \text{ and } (q_{t,u}, z_{t,u}) \notin A \text{ and } x_{t,u} \geq \sup_{(q,z) \in A} p_t q + z.$$

Similar to before, if $p_t < p_0$, then (by substituting out z)

$$\begin{aligned} x_{t,u} &\geq \sup p_t q + z \text{ s.t. } (q, z) \in A \\ &= \sup (p_t - p_0)q + x_0 \text{ s.t. } q \in [0, q_0] \\ &= x_0. \end{aligned}$$

Substituting out $z_{t,u}$ then implies $x_{t,u} \geq x_0 \geq p_0 q_{t,u} + z_{t,u} = (p_0 - p_t)q_{t,u} + x_{t,u} \geq x_{t,u}$, since $p_t < p_0$. As such, all inequalities must be equalities and, therefore, $q_{t,u} = 0$ and $x_{t,u} = x_0$. This implies $(q_{t,u}, z_{t,u}) = (0, x_0)$ belongs to A , which is a contradiction. $\square$

Lemma 7. *If $p_t > p_0$ then*

$$\Pr(U \in E_t(A)) = \Pr(x_0 < p_0 q_{t,U} + z_{t,U} \text{ or } x_{t,U} < p_t q_0 + z_0 \text{ or } (q_{t,U}, z_{t,U}) = (q_0, z_0)).$$

Proof. Similar to before, if $p_t > p_0$, then (by substituting out z)

$$\begin{aligned} x_{t,u} &< \sup p_t q + z \text{ s.t. } (q, z) \in A \\ &= \sup (p_t - p_0)q + x_0 \text{ s.t. } q \in [0, q_0] \\ &= (p_t - p_0)q_0 + x_0 = p_t q_0 + z_0. \end{aligned}$$

If $(q_{t,u}, z_{t,u}) \in A$, then we obtain again

$$\begin{aligned} x_{t,u} &= p_t q_{t,u} + z_{t,u} = (p_t - p_0)q_{t,u} + x_0 \\ &\leq (p_t - p_0)q_0 + x_0 = p_t q_0 + z_0, \end{aligned}$$

since $p_t > p_0$. As such we can conclude that

$$\Pr(U \in E_t(A)) \leq \Pr(x_0 < p_0 q_{t,U} + z_{t,U} \text{ or } x_{t,U} < p_t q_0 + z_0 \text{ or } (q_{t,U}, z_{t,U}) = (q_0, z_0)).$$

Conversely, if $(q_{t,u}, z_{t,u}) = (q_0, z_0)$ then $(q_{t,u}, z_{t,u}) \in A$. As such, we also obtain

$$\Pr(U \in E_t(A)) \geq \Pr(x_0 < p_0 q_{t,U} + z_{t,U} \text{ or } x_{t,U} < p_t q_0 + z_0 \text{ or } (q_{t,U}, z_{t,U}) = (q_0, z_0)).$$

□

Lemmata 6 and 7 define the tightest upper bound for all cross-sections t . Taking the minimum over all the available cross-sections gives us the upper bound in Theorem 2.

A.2.3 Lemmata for the proofs of Theorems 3 and 4

To prove Theorem 4, we need to allow for further restricting the sets $Q(\mathcal{F}_{t,u}, x)$ that we introduced in Appendix A.1.2.

Assumption 1. *There exist two non-decreasing functions $\underline{q}_{t,u}(x) \leq \overline{q}_{t,u}(x)$ such that*

$$Q(\mathcal{F}_{t,u}, x) = \left\{ (q, z) \in \mathbb{R}_+^2 : \begin{array}{l} p_t q + z = x, \\ \underline{q}_{t,u}(x) \leq q \leq \overline{q}_{t,u}(x). \end{array} \right\}.$$

In words, $\underline{q}_{t,u}(x)$ and $\overline{q}_{t,u}(x)$ define lower and upper bounds on the expansion path for the first good. Note that adding these restrictions does not interfere with our arguments in Appendix A.1.2.

Under our maintained assumption that expansion paths are non-decreasing (i.e. all goods are normal), the functions $\underline{q}_{t,u}(x)$ and $q \leq \overline{q}_{t,u}(x)$ cannot be any function (see, for example, Lemma 13 for a characterization of the case where $\mathcal{F}_{t,u}$ includes all non-decreasing expansion paths). In particular, the following lemma will be useful below.

Lemma 8. *Assume all goods are normal. If $x > y$, then $\underline{q}_{t,u}(x) \leq \underline{q}_{t,u}(y) + \frac{x-y}{p_t}$.*

Proof. Consider an expansion path $q_u(p_t, y)$ that goes through the lower bound $\underline{q}_{t,u}(y)$. Given that all goods are normal, the increase in good q is bounded above by $(x - y)/p_t$. As such,

$$\underline{q}_{t,u}(x) \leq q_u(p_t, x) \leq q_u(p_t, y) + \frac{x-y}{p_t} = \underline{q}_{t,u}(y) + \frac{x-y}{p_t}.$$

□

From Theorem 6 we know that to construct our upper bound on $\Pr(\mathbf{q}_U(\mathbf{p}_0, x_0) \in A)$, we need to compute the probability of the event that U belongs to

$$E_t(A) = \left\{ u \in \mathcal{U} : \begin{array}{l} x_0 < \sup_{\tilde{q}, \tilde{z}} p_0 \tilde{q} + \tilde{z} \text{ s.t. } (\tilde{q}, \tilde{z}) \in Q(\mathcal{F}_{t,u}, x_t^*) \\ \text{or } \exists x \in \mathbb{R}_+, A \cap Q(\mathcal{F}_{t,u}, x) \neq \emptyset \end{array} \right\},$$

with $x_t^* = \sup_{(q,z) \in A} p_t q + z$.

The first condition of $E_t(A)$ can be rewritten as

$$\begin{aligned} x_0 &< \sup_{\tilde{q}, \tilde{z}} p_0 \tilde{q} + \tilde{z} & \text{s.t. } \overline{q_{t,u}}(x_t^*) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^*) \\ & & p_t \tilde{q} + \tilde{z} &= x_t^* \\ \Leftrightarrow \quad x_0 &< \sup (p_0 - p_t) \tilde{q} + x_t^* & \text{s.t. } \overline{q_{t,u}}(x_t^*) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^*). \end{aligned}$$

The following lemma shows that $\Pr(U \in E_t(A)) = 1$ when $p_t < p_0$.

Lemma 9. *If $p_t < p_0$ then $E_t(A) = \mathcal{U}$.*

Proof. If $p_t < p_0$, then (by substituting out z)

$$\begin{aligned} x_t^* &= \sup p_t q + z & \text{s.t. } (q, z) &\in A \\ &= \sup (p_t - p_0) q + x_0 & \text{s.t. } q &\in [0, q_0] \\ &= x_0. \end{aligned}$$

As such,

$$\begin{aligned} x_0 &< \sup (p_0 - p_t) \tilde{q} + x_t^* & \text{s.t. } \overline{q_{t,u}}(x_t^*) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^*) \\ &= \sup (p_0 - p_t) \tilde{q} + x_0 & \text{s.t. } \overline{q_{t,u}}(x_0) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_0) \\ &= (p_0 - p_t) \overline{q_{t,u}}(x_0) + x_0. \end{aligned}$$

This inequality is satisfied if and only if $\overline{q_{t,u}}(x_0) > 0$. If $\overline{q_{t,u}}(x_0) = 0$, then $(q, z) = (0, x_0)$ both belongs to A and $Q(\mathcal{F}_{t,u}, x_0) = 0$, which implies that the second condition in $E_t(A)$ is satisfied. $\square$

The following lemma considers the opposite scenario.

Lemma 10. *If $p_t > p_0$ then $\Pr(U \in E_t(A)) = \Pr(q_0 \geq \underline{q_{t,U}}(p_t q_0 + z_0))$.*

Proof. We first show that $\Pr(U \in E_t(A)) \leq \Pr(q_0 \geq \underline{q_{t,U}}(p_t q_0 + z_0))$ for $p_t > p_0$. By substituting out z we obtain

$$\begin{aligned} x_t^* &= \sup p_t q + z & \text{s.t. } (q, z) &\in A \\ &= \sup (p_t - p_0) q + x_0 & \text{s.t. } q &\in [0, q_0] \\ &= (p_t - p_0) q_0 + x_0 \\ &= p_t q_0 + z_0. \end{aligned}$$

As such,

$$\begin{aligned}
x_0 &< \sup(p_0 - p_t)\tilde{q} + x_t^* && \text{s.t. } \overline{q_{t,u}}(x_t^*) \geq \tilde{q} \geq \underline{q_{t,u}}(x_t^*) \\
&= \sup(p_0 - p_t)\tilde{q} + (p_t - p_0)q_0 + x_0 && \text{s.t. } \overline{q_{t,u}}(p_t q_0 + z_0) \geq \tilde{q} \geq \underline{q_{t,u}}(p_t q_0 + z_0) \\
&= \sup(p_0 - p_t)(\tilde{q} - q_0) + x_0 && \text{s.t. } \overline{q_{t,u}}(p_t q_0 + z_0) \geq \tilde{q} \geq \underline{q_{t,u}}(p_t q_0 + z_0) \\
&= (p_0 - p_t)(\underline{q_{t,u}}(p_t q_0 + z_0) - q_0) + x_0.
\end{aligned}$$

This inequality is satisfied if and only if $q_0 > \underline{q_{t,u}}(p_t q_0 + z_0)$.

Next, we consider the second condition defining $E_t(A)$. Assume there is an $x \in \mathbb{R}_+$ such that $(q, z) \in Q(\mathcal{F}_{t,u}, x)$ and $(q, z) \in A$. This implies

$$\underline{q_{t,u}}(x) \leq q \leq \overline{q_{t,u}}(x), p_t q + z = x, p_0 q + z = x_0 \text{ and } q \in [0, q_0],$$

which in turn obtains

$$\underline{q_{t,u}}(p_t q + z) \leq q \text{ and } p_t q + z = (p_t - p_0)q + x_0 \leq (p_t - p_0)q_0 + x_0 = p_t q_0 + z_0.$$

Using Lemma 8, we then have

$$\begin{aligned}
\underline{q_{t,u}}(p_t q_0 + z_0) &\leq \underline{q_{t,u}}(p_t q + z) + \frac{p_t q_0 + z_0 - p_t q - z}{p_t} \\
&\leq q + q_0 - q + \frac{1}{p_t}(z_0 - z) \\
&\leq q_0,
\end{aligned}$$

where the last inequality follows from the fact that $(q, z) \in A$ and $q \leq q_0$ implies $z \geq z_0$. Taking together, we can conclude $\Pr(U \in E_t(A)) \leq \Pr(q_0 \geq \underline{q_{t,u}}(p_t q_0 + z_0))$.

Conversely, we show $\Pr(U \in E_t(A)) \geq \Pr(q_0 \geq \underline{q_{t,u}}(p_t q_0 + z_0))$ for $p_t > p_0$. Above we proved that the first condition defining $E_t(A)$ is satisfied if $q_0 > \underline{q_{t,u}}(p_t q_0 + z_0)$. If $q_0 = \underline{q_{t,u}}(p_t q_0 + z_0)$, then per construction (q_0, z_0) belongs to both $Q(\mathcal{F}_{t,u}; p_t q_0 + z_0)$ and A . As such, $A \cap Q(\mathcal{F}_{t,u}; p_0 q_0 + z_0) \neq \emptyset$, which shows that the second condition of $E_t(A)$ is satisfied. $\square$

Finally, we turn to computing the lower bound on $\Pr(\mathbf{q}_U(\mathbf{p}_t, x_0) \in A)$. From Theorem 6 we know that this requires to compute the probability of the event that U belongs to

$$F_t(A) = \left\{ u \in \mathcal{U} : \begin{array}{l} x_0 \geq \sup_{\tilde{q}, \tilde{z}} p_0 \tilde{q} + \tilde{z} \text{ s.t. } (\tilde{q}, \tilde{z}) \in Q(\mathcal{F}_{t,u}, x_t^{**}) \\ \text{and } \forall x \in \mathbb{R}_+, \neg A \cap Q(\mathcal{F}_{t,u}, x) = \emptyset \end{array} \right\},$$

with $x_t^{**} = \sup_{(q,z) \in \neg A} p_t q + z$. The first condition of $F_t(A)$ can be rewritten as

$$\begin{aligned} x_0 &\geq \sup_{\tilde{q}, \tilde{z}} p_0 \tilde{q} + \tilde{z} & \text{s.t. } p_t \tilde{q} + \tilde{z} &= x_t^{**} \\ & & \overline{q_{t,u}}(x_t^{**}) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^{**}) \\ \Leftrightarrow \quad x_0 &\geq \sup_{\tilde{q}, \tilde{z}} (p_0 - p_t) \tilde{q} + x_t^{**} & \text{s.t. } \overline{q_{t,u}}(x_t^{**}) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^{**}). \end{aligned}$$

The following lemma shows that $\Pr(U \in F_t(A)) = 0$ when $p_t > p_0$.

Lemma 11. *If $p_t > p_0$ then $F_t(A) = \emptyset$.*

Proof. Towards a contradiction, assume that $p_t > p_0$ and $u \in F_t(A)$. By substituting out z we obtain

$$\begin{aligned} x_t^{**} &= \sup p_t q + z & \text{s.t. } (q, z) &\in \neg A \\ &= \sup (p_t - p_0) q + x_0 & \text{s.t. } q &\in]q_0, x_0/p_0] \\ &= (p_t - p_0) \frac{x_0}{p_0} + x_0 \\ &= \frac{p_t}{p_0} x_0. \end{aligned}$$

As such,

$$\begin{aligned} x_0 &\geq \sup_{\tilde{q}, \tilde{z}} (p_0 - p_t) \tilde{q} + x_t^{**} & \text{s.t. } \overline{q_{t,u}}(x_t^{**}) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^{**}) \\ &= \sup (p_0 - p_t) \tilde{q} + \frac{p_t}{p_0} x_0 & \text{s.t. } \overline{q_{t,u}}\left(\frac{p_t}{p_0} x_0\right) &\geq \tilde{q} \geq \underline{q_{t,u}}\left(\frac{p_t}{p_0} x_0\right) \\ &= \sup (p_t - p_0) \left(\frac{x_0}{p_0} - \tilde{q} \right) + x_0 & \text{s.t. } \overline{q_{t,u}}\left(\frac{p_t}{p_0} x_0\right) &\geq \tilde{q} \geq \underline{q_{t,u}}\left(\frac{p_t}{p_0} x_0\right) \\ &= (p_t - p_0) \left(\frac{x_0}{p_0} - \underline{q_{t,u}}(p_t x_0 / p_0) \right) + x_0. \end{aligned}$$

This inequality is satisfied if and only if $x_0/p_0 \leq \underline{q_{t,u}}(p_t x_0 / p_0)$. However, this implies that $(x_0/p_0, 0)$ belongs to both $\mathcal{Q}(\mathcal{F}_{t,u}, \frac{p_t}{p_0} x_0)$ and $\neg A$, which contradicts with the second defining condition of $F_t(A)$. $\square$

The following lemma considers the opposite scenario.

Lemma 12. *If $p_t < p_0$ then $\Pr(U \in F_t(A)) = \Pr(q_0 \geq \overline{q_{t,U}}(p_t q_0 + z_0))$.*

Proof. We first show that $\Pr(U \in F_t(A)) \geq \Pr(q_0 \geq \overline{q_{t,U}}(p_t q_0 + z_0))$ for $p_t < p_0$. By substituting

out z we obtain

$$\begin{aligned}
x_t^{**} &= \sup p_t q + z & \text{s.t. } (q, z) &\in \neg A \\
&= \sup (p_t - p_0)q + x_0 & \text{s.t. } q &\in]q_0, x_0/p_0] \\
&= (p_t - p_0)q_0 + x_0 \\
&= p_t q_0 + z_0.
\end{aligned}$$

As such,

$$\begin{aligned}
x_0 &\geq \sup_{\tilde{q}, \tilde{z}} (p_0 - p_t)\tilde{q} + x_t^{**} & \text{s.t. } \overline{q_{t,u}}(x_t^{**}) &\geq \tilde{q} \geq \underline{q_{t,u}}(x_t^{**}) \\
&= \sup (p_0 - p_t)\tilde{q} + p_t q_0 + z_0 & \text{s.t. } \overline{q_{t,u}}(p_t q_0 + z_0) &\geq \tilde{q} \geq \underline{q_{t,u}}(p_t q_0 + z_0) \\
&= \sup (p_0 - p_t)\tilde{q} + (p_t - p_0)q_0 + x_0 & \text{s.t. } \overline{q_{t,u}}(p_t q_0 + z_0) &\geq \tilde{q} \geq \underline{q_{t,u}}(p_t q_0 + z_0) \\
&= \sup (p_0 - p_t)(\tilde{q} - q_0) + x_0 & \text{s.t. } \overline{q_{t,u}}(p_t q_0 + z_0) &\geq \tilde{q} \geq \underline{q_{t,u}}(p_t q_0 + z_0) \\
&= (p_0 - p_t)(\overline{q_{t,u}}(p_t q_0 + z_0) - q_0) + x_0.
\end{aligned}$$

This inequality is satisfied if and only if $\overline{q_{t,u}}(p_t q_0 + z_0) \leq q_0$. We can conclude that $\Pr(U \in F_t(A)) \leq \Pr(q_0 \geq \overline{q_{t,u}}(p_t q_0 + z_0))$.

To prove the reverse inequality, it suffices to show that, if $q_0 \geq \overline{q_{t,u}}(p_t q_0 + z_0)$, then $\forall x \in \mathbb{R}_+ : \neg A \cap Q(\mathcal{F}_{t,u}, x) = \emptyset$. Towards a contradiction, assume that there exists a $(q, z) \in \neg A \cap Q(\mathcal{F}_{t,u}, x)$. This implies

$$p_0 q + z = x_0, q \in]q_0, x_0/p_0], p_t q + z = x, \text{ and } \underline{q_{t,u}}(x) \leq q \leq \overline{q_{t,u}}(x),$$

which in turn obtains

$$\overline{q_{t,u}}(p_t q + z) \geq q \text{ and } p_t q_0 + z_0 = (p_t - p_0)q_0 + x_0 \geq (p_t - p_0)q + x_0 = p_t q + z.$$

As such, $\overline{q_{t,u}}(p_t q_0 + z_0) = \overline{q_{t,u}}(p_t q + z) \geq q > q_0$, which is a contradiction. This allows us to conclude $\Pr(q_0 \geq \overline{q_{t,u}}(p_t q_0 + z_0)) \leq \Pr(U \in F_t(A))$. $\square$

The following theorem summarizes our results in Lemmata 9-12.

Theorem 7. *Assume that all goods are normal. We have that*

$$\max_{t \in \underline{T}} \Pr(q_0 \geq \overline{q_{t,u}}(p_t q_0 + z_0)) \leq \Pr(q_U(p_0, x_0) \leq q_0) \leq \min_{t \in \overline{T}} \Pr(q_0 \geq \underline{q_{t,u}}(p_t q_0 + z_0)).$$

These bounds are the tightest WARP-consistent bounds.

A.2.4 Proof of Theorem 3

To prove Theorem 3 we need to consider the setting where $\mathcal{F}_{t,u}$ consists of non-decreasing expansion paths. This complies with the following characterization of the functions $\underline{q}_{t,u}(x)$ and $\overline{q}_{t,u}(x)$.

Lemma 13. *Let $\mathcal{F}_{t,u}$ be the set of non-decreasing expansion paths. Then, for all $x < x_{t,u}$,*

$$\underline{q}_{t,u}(x) = \max\{0, (x - z_{t,u})/p_t\} \text{ and } \overline{q}_{t,u}(x) = \min\{q_{t,u}, x/p_t\},$$

and, for all $x \geq x_{t,u}$,

$$\underline{q}_{t,u}(x) = q_{t,u} \text{ and } \overline{q}_{t,u}(x) = (x - z_{t,u})/p_t.$$

Proof. For income $x_{t,u}$ we observe the point $(q_{t,u}, z_{t,u})$ on the income expansion path. Given that expansion paths are non-decreasing this defines lower and upper bounds on q for alternative income levels x .

If $x < x_{t,u}$, then q must be smaller than the observed $q_{t,u}$ and smaller than the maximal available budget that can be spend on q . This defines $\overline{q}_{t,u}(x)$. Next, q must always be positive. Moreover, given that both goods are normal, the consumer cannot consume more than $(x_{t,u} - x)/p_t$ goods less of the first good. Thus,

$$q_t - \frac{x_{t,u} - x}{p_t} = \frac{x - z_{t,u}}{p_t}$$

is also a lower bound. This defines $\underline{q}_{t,u}(x)$.

Finally, if $x \geq x_{t,u}$, then $q_{t,u}$ is a lower bound. Similar to above, the consumer cannot consume more than $(x - x_{t,u})/p_t$ goods more of the first good. Thus,

$$q_t + \frac{x - x_{t,u}}{p_t} = \frac{x - z_{t,u}}{p_t}$$

is an upper bound. This defines $\overline{q}_{t,u}(x)$. □

Combining Lemma 13 and Theorem 7 gives the following lower bound on $\Pr(q_U(p_0, x_0) \leq q_0)$:

$$\max_{t \in \overline{T}} (\Pr[p_t q_0 + z_0 < x_{t,U} \text{ and } q_0 \geq q_{t,U}] + \Pr[p_t q_0 + z_0 \geq x_{t,U} \text{ and } z_{t,U} \geq z_0]),$$

which equals $\Pr[z_{t,U} \geq z_0 \text{ and } q_0 \geq q_{t,U}]$ because $x_{t,u} = p_t q_{t,u} + z_{t,u}$.

Similarly, we obtain the following upper bound on $\Pr(q_U(p_0, x_0) \leq q_0)$:

$$\min_{t \in \overline{T}} (\Pr[p_t q_0 + z_0 < x_{t,U} \text{ and } z_{t,U} \geq z_0] + \Pr[p_t q_0 + z_0 \geq x_{t,U} \text{ and } q_0 \geq q_{t,U}]),$$

which is equal to $\Pr[z_{t,U} \geq z_0 \text{ or } q_0 \geq q_{t,U}]$. This proves Theorem 3.

A.2.5 Proof of Theorem 4

To prove Theorem 4 we need to consider the setting where $\mathcal{F}_{t,u}$ consists of non-decreasing expansion paths that satisfy the extra bounds on the income elasticities given by

$$\begin{aligned} 0 \leq \eta_\ell &\leq \frac{\partial \ln(q_u(p_t, x))}{\partial \ln(x)} \leq \eta_h, \\ 0 \leq \delta_\ell &\leq \frac{\partial \ln(z_u(p_t, x))}{\partial \ln(x)} \leq \delta_h. \end{aligned}$$

This complies with the following characterization of the functions $\underline{q}_{t,u}(x)$ and $\overline{q}_{t,u}(x)$.

Lemma 14. *Let $\mathcal{F}_{t,u}$ be the set non-decreasing expansion paths that satisfy*

$$\begin{aligned} 0 \leq \eta_\ell &\leq \frac{\partial \ln(q_u(p_t, x))}{\partial \ln(x)} \leq \eta_h, \\ 0 \leq \delta_\ell &\leq \frac{\partial \ln(z_u(p_t, x))}{\partial \ln(x)} \leq \delta_h. \end{aligned}$$

Then, for all $x < x_{t,u}$,

$$\begin{aligned} \underline{q}_{t,u}(x) &= \max \left\{ q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_h}, \frac{x - z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_\ell}}{p_t} \right\} \text{ and} \\ \overline{q}_{t,u}(x) &= \min \left\{ q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_\ell}, \frac{x - z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_h}}{p_t} \right\}, \end{aligned}$$

and, for all $x \geq x_{t,u}$,

$$\begin{aligned} \underline{q}_{t,u}(x) &= \max \left\{ q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_\ell}, \frac{x - z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_h}}{p_t} \right\} \text{ and} \\ \overline{q}_{t,u}(x) &= \min \left\{ q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_h}, \frac{x - z_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\delta_\ell}}{p_t} \right\}. \end{aligned}$$

Proof. We obtain this result by using exactly the same reasoning as in Lemma 13 in combination with Lemma 1. □

Combining Lemma 14 and Theorem 7 then proves Theorem 4.

A.2.6 Proof of Lemma 1

Take the case where $x \leq x_{t,u}$. Then,

$$\begin{aligned} \ln(q_{t,u}) - \ln(q_u(p_t, x)) &= \int_{\ln(x)}^{\ln(x_{t,u})} \frac{\partial \ln(q_u(p_t, z))}{\partial \ln(z)} \, d \ln(z) \\ &\leq \int_{\ln(x)}^{\ln(x_{t,u})} \eta_h \, d \ln(z) \\ &= \eta_h [\ln(x_{t,u}) - \ln(x)]. \end{aligned}$$

This gives $q_u(p_t, x) \geq q_{t,u} \left(\frac{x}{x_{t,u}} \right)^{\eta_h}$. The other inequalities in Lemma 1 are obtained in a readily similar fashion.