Prices versus Preferences:

Taste Change and Revealed Preference

Abi Adams (Oxford and IFS), Richard Blundell (UCL and IFS),

Martin Browning (Oxford and IFS), and Ian Crawford (Oxford and IFS)

December 2013, this draft February 2015*

Abstract

A systematic approach for incorporating taste variation into a revealed preference framework for heterogeneous consumers is developed. We create a new methodology that enables the recovery of the minimal variation in tastes that are required to rationalise observed choice patterns. This approach is used to examine the extent to which changes in tobacco consumption have been driven by price changes or by taste changes, and whether the significance of these two channels varies across socioeconomic groups. A censored quantile approach is used to allow for unobserved heterogeneity and censoring of consumption. Statistically significant educational differences in the marginal willingness to pay for tobacco are recovered. More highly educated cohorts are found to have experienced a greater shift in their effective tastes away from tobacco.

*We would like to thank Laurens Cherchye, Dennis Kristensen, Arthur Lewbel and participants at the June 2014 CAM workshop, November 2014 cemmap workshop and the AEA January 2015 session on consumer demand for helpful comments. The authors also gratefully acknowledge financial support from the Economic and Social Research Council Centre for the Microeconomic Analysis of Public Policy at IFS and the European Research Council grant MicroConLab. All errors are ours.

1 Introduction

Structural empirical research on consumer behaviour is typically based upon the idea of choice-revealed preference: consumers choose what they prefer from the options available to them and thereby reveal their preferences through repeated observations of their choices from different budget sets. Simple techniques can then be used to recover a consumer's preferences from data on their choices using methods developed by Samuelson (1948), Houthakker (1953), Afriat (1967) and Varian (1982). However, classical revealed preference methods can only be applied when preferences are *stable*. If preferences change during the period of observation then these methods cannot solve the inverse problem.

In this paper we develop a systematic approach to incorporating taste variation into a revealed preference framework. We create a methodology that enables the recovery of the minimal variation in tastes that are required to rationalise the observed choice patterns. We represent taste heterogeneity as a linear perturbation to a heterogeneous base utility function, much as in McFadden and Fosgerau (2012) and Brown and Matzkin (1998). Under this specification, taste change can be interpreted as the shift in the marginal utility of a good for each individual. Our theorems show this specification is not at all restrictive. We derive inequalities that are a simple extension of Afriat (1967) such that when they hold there exists a well-behaved base utility function and a series of taste shifters that perfectly rationalise observed behaviour. We then show, under mild assumptions on the characteristics of available choice data, that we can always find a pattern of taste shifters on a single good that are sufficient to rationalise any finite time series of prices and quantities.

Our first results are derived as if we observed an individual consumer for a finite set of time periods in which prices, and possibly income, change. We also consider the case of repeated samples over time from a fixed population of individuals where we do not necessarily observe the same individual in each sample. This is analogous to our empirical setting where we follow the same birth cohort in repeated cross-sections. In this case we consider shifts in the *distribution of demands* that are consistent with Revealed Preference (RP). In particular, we make a rank invariance assumption on unobserved heterogeneity and taste perturbations that allow us to use censored quantile regression to recover individual demands and the minimal taste changes required to rationalise the observed distribution of choice behaviour.

We apply this approach to analyse the preferences for a good where there is strong *prima facie* evidence that tastes have changed: tobacco products. In particular, we ask how much of the fall in tobacco consumption is due to a rise in the relative price of tobacco and how much needs to be attributed to taste changes? We also consider how tastes evolve across different socio-economic strata, asking the question: Does education matter? The approach is implemented on household consumer expenditure survey data using RP inequality conditions on the censored conditional quantile demand functions for tobacco. We extend the analysis to allow for the non-separability of tobacco consumption with alcohol consumption.

The objective is to understand taste change and to inform policy on the balance between information/health campaigns with tax reform. Governments have a limited set of levers should they wish to influence household consumption patterns. These include quantity constraints, price changes through taxes and subsidies and information programs. The relative efficacy of the different modes is important for designing public policy. The approach in this paper allows us to consider the extent to which changes in tobacco consumption are due to price changes and how much is due to preference change.

This paper proceeds as follows. Sections 2 and 3 outline our theoretical framework and derive the necessary and sufficient conditions under which observed behaviour and our model of taste change are consistent. Section 4 develops a quadratic programming methodology that can be applied to recover the minimal amount of interpersonally comparable taste variation that is necessary to rationalise choice behaviour. Section 5 introduces the data used for our empirical investigation of tobacco consumption in the UK and discusses the construction of quantity sequences for the psuedo birth-cohorts that we draw from the UK Family Expenditure Survey using censored quantile regression methods. Section 6 applies our method to rationalise the changes in tobacco consumption occurring in the U.K. since 1980. Finally, Section 7 concludes our analysis and considers the implications of our findings for government anti-smoking policy moving forward.

2 Theoretical framework

Consider a consumer who selects a quantity vector $\mathbf{q} \in \mathbb{R}_+^K$ at time t to maximise their time-dependent utility function:

$$u(\mathbf{q}; \boldsymbol{\alpha}_t) = v(\mathbf{q}) + \boldsymbol{\alpha}_t' \mathbf{q} \tag{1}$$

subject to $\mathbf{p'q} = x$, where $\mathbf{p} \in \mathbb{R}^K_{++}$ is an exogenous price vector and x is total expenditure. It is assumed that $u(\mathbf{q}, \boldsymbol{\alpha}_t)$ is locally nonsatiated and concave conditional on $\boldsymbol{\alpha}_t \in \mathbb{R}^K$, where $\boldsymbol{\alpha}_t$ is a vector of marginal utility perturbations that indexes the consumer's tastes at time t. This utility function therefore consists of two components: a set of "base" preferences given by $v(\mathbf{q})$, and a time-varying part given by $\boldsymbol{\alpha}'_t\mathbf{q}$.[1] Our use of marginal utility shifters to capture taste changes follows the random utility approach of Brown and Matzkin (1998) and McFadden and Fosgerau (2012).

In this most general case, in which the marginal utility of all goods is potentially subject to arbitrary observation-to-observation changes it is clear that choice patterns are little-restricted by the model. Indeed, given any dataset D made up of a sequence of price-quantity observations for a consumer $D = \{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ it is clear that we can always find a sequence of taste-shifters $\{\boldsymbol{\alpha}_t\}_{t=1,...,T}$ that can rationalise the model. To see this, consider the first order conditions

$$\nabla u(\mathbf{q}_t) = \nabla v(\mathbf{q}_t) + \boldsymbol{\alpha}_t \leq \lambda_t \mathbf{p}_t$$

This can be rewritten in terms of the base preferences and shadow prices as

$$\nabla v(\mathbf{q}_t) \leq \lambda_t \tilde{\mathbf{p}}_t$$

where we use the shadow prices

$$\tilde{\mathbf{p}}_t = \mathbf{p}_t - \frac{\boldsymbol{\alpha}_t}{\lambda_t}$$

where $\boldsymbol{\alpha}_t/\lambda_t$ represents the innovations in willingness-to-pay or *'taste-wedge'* for every good in every period. The behaviour generated by the model

$$\max_{\mathbf{q}} v(\mathbf{q}) + \boldsymbol{\alpha}'_t\mathbf{q} \text{ subject to } \mathbf{p}'_t\mathbf{q} = x_t$$

is therefore identical to the behaviour generated by the model

$$\max_{\mathbf{q}} v(\mathbf{q}) \text{ subject to } \tilde{\mathbf{p}}'_t\mathbf{q} = \tilde{x}_t$$

where preferences are not subject to taste change, but where the prices and budget are replaced by their virtual counterparts.[2] Thus the question of whether there exist rationalising taste-shifters is equivalent to the question of

[1] The effect of the taste change parameters $\boldsymbol{\alpha}_t$ on consumer demand is not invariant to transformations of $v(\mathbf{q})$ and so further analysis is conditioned upon a given cardinal representation of the "base" preferences.

[2] The concept of a virtual budget was first suggested by Rothbarth (1941) and Neary and Roberts (1980) to develop the theory of choice behaviour under rationing.

whether we can always find shadow prices which can rationalise the observed quantity data. Varian (1988, Theorem 1) shows that this is indeed the case; indeed there will typically be multiple suitable shadow values consistent with any finite dataset.

This model of taste change, which allows for adjustments in the willingness-to-pay for every good in every period, introduces as many free parameters as there are observations and so is clearly extremely permissive - even given its apparently restrictive additive form. Instead, suppose that we have prior grounds to believe that the most significant taste changes which have taken place have been confined to a single good denoted good 1. For example, in our empirical application, we suppose that good 1 represents tobacco products- a good for which there are strong grounds for believing that tastes have changed significantly given increased awareness of its ill health effects. With the restriction of taste change to a single good, then $\boldsymbol{\alpha}_t = [\alpha_t^1, 0, ..., 0]'$, yielding the following temporal series of utility functions:[3]

$$u(\mathbf{q}; \alpha_t^1) = v(\mathbf{q}) + \alpha_t^1 q^1 \tag{2}$$

Taste changes thus enter the basic utility maximisation framework in a more restricted manner. Specifically, the additive-linear specification for taste perturbations implies that the marginal rate of substitution between any of the other goods $j, k \in \{2, ..., K\}$ is invariant to taste instability on good-1. Another implication of this functional form is that preferences will obey the single crossing property.

Definition (Milgrom and Shannon, 1994) A utility function $u(\mathbf{q}; \alpha_t^1)$ satisfies the single crossing property in $(\mathbf{q}; \alpha^1)$ if for $\mathbf{q}' > \mathbf{q}''$ and $\alpha^{1\prime} > \alpha^{1\prime\prime}$

$$u(\mathbf{q}'; \alpha^{1\prime\prime}) \geq u(\mathbf{q}''; \alpha^{1\prime\prime}) \text{ implies } u(\mathbf{q}'; \alpha^{1\prime}) \geq u(\mathbf{q}''; \alpha^{1\prime}) \tag{3}$$

This condition can be interpreted as stating that for any $\mathbf{q}' > \mathbf{q}''$, the function $f(\alpha^1) = u(\mathbf{q}'; \alpha^1) - u(\mathbf{q}''; \alpha^1)$ crosses the horizontal axis only once, from negative to positive, as α^1 increases. The single-crossing property means that there is an unambiguous ranking of tastes for the good of interest - the MRS with respect to the good of interest given α_t^1 is always high (lower) in one period than it is in another at every point in commodity space. Another basic

[3] This assumes intertemporal separability. Whether this is appropriate for the study of tobacco depends on the definition of the time period and the frequency of consumption observations. In our empirical application, this frequency is two years — there is evidence that this is sufficient time for habits to weaken (see, for example, Cuzbek and Johal (2010)).

result that follows from single crossing is that $\text{argmax}_{\mathbf{q}} u(\mathbf{q}; \alpha^1)$ increases with α^1 (see Milgrom and Shannon (1994)). We will use this feature of the model to motivate the use of quantile regression methods in the empirical section of this paper.

2.1 Empirical conditions

We are interested in establishing whether a consumer's observed choice behaviour, $D = \{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$, could have been generated by taste change on a single good. Rationalisation of D by our model is defined as follows.

Definition 2 A consumer's choice behaviour, $D = \{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$, can be 'good-1 taste rationalised' by the base utility function $v(\mathbf{q})$ and a temporal series of additive linear perturbations to the marginal utility of good-1, $\{\alpha_t^1\}_{t=1,...,T}$, if

$$v(\mathbf{q}) + \alpha_t^1 q^1 \leq v(\mathbf{q}_t) + \alpha_t^1 q_t^1 \tag{4}$$

for all $\mathbf{q}$ such that:

$$\mathbf{p}_t' \mathbf{q} \leq \mathbf{p}_t' \mathbf{q}_t \tag{5}$$

In words, D can be rationalised by the model if there exists a time-invariant base utility function, $v(\mathbf{q})$ and a series of perturbations to the marginal utility of good-1 such that observed choices are weakly preferred to all feasible alternatives. The empirical conditions, involving only observables, that are equivalent to a rationalisation of D by our theoretical model are given in Theorem 1.

Theorem 1. The following statements are equivalent:

1. Observed choice behaviour, $D = \{\mathbf{p}_t, \mathbf{q}_t^i\}_{t=1,...,T}$, can be good-1 taste rationalised.

2. One can find sets $\{v_t\}_{t=1,...,T}$, $\{\alpha_t^1\}_{t=1,...,T}$ and $\{\lambda_t\}_{t=1,...,T}$ with $\lambda_t > 0$ for all $t = 1, ..., T$, such that there exists a non-empty solution set to the following revealed preference inequalities:

$$\begin{aligned} v_s - v_t + \alpha_t^1 (q_s^1 - q_t^1) &\leq \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t) \\ \alpha_t^1 &\leq \lambda_t p_t^1 \end{aligned} \tag{6}$$

Theorem 1 is implied by optimising behaviour within the theoretical framework. If there exists a non-empty solution set to the inequalities defined by Theorem 1, then there exists a well-behaved base utility function and a series of taste shifters on good-1 that perfectly rationalise observed behaviour. The variables referred to by the revealed preference inequalities in part (2) of Theorem 1 have natural interpretations. The numbers $\{u_t\}_{t=1,..,T}$ and $\{\lambda_t\}_{t=1,..,T}$ can be interpreted respectively as measures of the level of baseline utility and the marginal utility of income at observed demands. The α_t^1 values can be interpreted as the marginal utility perturbation to good-1 relative to that dictated by base utility at observed demands since we can set $\alpha_t^1 = 0$ for all t.

Theorem 1 is an extension to the equivalence result originally derived by Afriat (1967) for the utility maximisation model with a time-invariant utility function. Imposing $\alpha_t^1 = 0$ for all $t = 1, .., T$ returns the standard Afriat inequalities. If there is no intertemporal variation in good-1, $q_t^1 = q_s^1$ for all $t, s \in 1, ..., T$, then we immediately have the standard Afriat conditions.[4]

Given solutions for condition (2) in Theorem 1, we can construct a rationalising utility function at each observation ($u(\mathbf{q}_t; \alpha_t^1)$) and also examine counterfactuals such as at $u(\mathbf{q}_t; \alpha_s^1)$. This indicates the utility which would be derived from consuming the period t bundle, but with one's the tastes from another period. For example, comparing $u(\mathbf{q}_t; \alpha_t^1)$ with $u(\mathbf{q}_t; \alpha_s^1)$ gives

$$u(\mathbf{q}_t; \alpha_t^1) - u(\mathbf{q}_t; \alpha_s^1) = \left(\alpha_t^1 - \alpha_s^1\right) q_t^1$$

whilst the shadow price of consumption of good 1 in period t given period s tastes is

$$u_1(\mathbf{q}_t; \alpha_s^1) = \lambda_t \left[p_t^1 - \frac{\left(\alpha_t^1 - \alpha_s^1\right)}{\lambda_t} \right]$$

Note that the shadow price of consumption in period t, given tastes in period s, may be negative if the taste-change term $\left(\alpha_t^1 - \alpha_s^1\right)$ is large enough. For example, if the consumer's taste for tobacco changes negatively between some earlier period $t-1$ and a later period t such that $\alpha_{t-1}^1 - \alpha_t^1$ is sufficiently positive then it is possible that

$$u_1(\mathbf{q}_{t-1}; \alpha_t^1) = \lambda_{t-1} \left[p_{t-1}^1 - \frac{\left(\alpha_{t-1}^1 - \alpha_t^1\right)}{\lambda_{t-1}} \right] < 0$$

The interpretation of this is that the consumer in period t would need to be paid to smoke as much as they did back in period $t-1$.

[4] In what follows we assume that we assume that we observe period-to-period variation in good-k such that $q_t^k \neq q_s^k$ for all $t, s = 1, ..., T$.

Solutions to Theorem 1 also enable us to construct the virtual prices at which the individual with preferences given by the base utility function would have purchased the bundle of goods purchased at t with taste for tobacco α_t^1

$$\tilde{p}_t^1 = p_t^1 - \frac{\alpha_t^1}{\lambda_t}$$

Interpreting taste change as an evolution of virtual prices supports the interpretation of information programmes as supplementary tax and incomes policies. For example, programmes designed to cultivate a negative taste for tobacco can be thought as levying a 'taste-tax' on the good because they manifest themselves in a rise in the virtual price for tobacco: $\tilde{p}_t^1 > p_t^1$ as $\alpha_t^1 < 0$ for a negative taste perturbation. Given the virtual price characterisation, variation in α^1 is more easily interpreted as a change in the marginal willingness to pay for good-1. The magnitude of the change in the marginal willingness to pay is captured by the term α_t^1/λ_t. This is useful because there is no clear behavioural interpretation of the magnitude of α_t^1 since its value depends upon the cardinal representation of base preferences.

2.2 Rationalisability

Surprisingly, under mild assumptions[5] on D, observed behaviour can *always* be explained by our simple model; that is, one can find sets of base-utility numbers, $\{v_t\}_{t=1,...T}$, marginal utilities of income $\{\lambda_t\}_{t=1,...,T}$ and taste perturbations on a single good $\{\alpha_t^1\}_{t=1,...,T}$ that rationalise observed choice behaviour.

Theorem 2 Given $q_t^1 \neq q_s^1$ for all t and s, any data set D can be good-1 taste rationalised.

Note that quantity variation is sufficient, but not necessary, for D to be good-1 taste rationalised. Let $\mathbf{p}_t^{\neg 1}$ denote the price vector for all goods except good-1, $\mathbf{p}_t^{\neg 1} = [p_t^2, ..., p_t^K]$, and define $\mathbf{q}_t^{\neg 1}$ analogously. For subsets $S_t \subseteq \{1, ..., T\}$ within which $q_t^1 = q$ for all $t \in S_t$, if the choice set $\{\mathbf{p}_t^{\neg 1}, \mathbf{q}_t^{\neg 1i}\}_{t \in S_t}$ satisfies the Generalised Axiom of Revealed Preference (GARP), then D will be rationalisable by our framework despite the violation of perfect variation in good-1. A leading example of this is when an agent does not buy good 1 in more than one period ($q_s^1 = q_t^1 = 0$ for some $s \neq t$).

Theorem 2 is closely related to Varian's (1988) Theorem 1, in which it is proved that the standard utility

[5] Our assumption that we observe period-to-period variations in quantities such that $q_t^k \neq q_s^k$ for all $t, s = 1, ..., T$ is important here.

maximisation model is virtually emptied of empirical content if the price of at least one good is not observed. In such circumstances, one can hypothesize that the unobserved prices take on values high enough that expenditures on goods with unobserved prices dominate all other revealed preference comparisons. The virtual budget characterisation of taste change makes clear the connection between Theorem 2 and Varian's result: tastes for good-1 could always decline to the extent that the virtual prices required to support observed bundles are high enough to prevent an intersection of the virtual budget hyperplanes.

2.3 Recoverability

The revealed preference inequalities associated with our theoretical framework can be used to recover the set of *minimal* perturbations to the marginal utility of good-1 that will rationalise observed behaviour. This set is always non-empty if D satisfies perfect variation with respect to good-1. We here outline an easy to implement quadratic programming procedure that enables the recovery of the minimal marginal utility perturbations on good-1 that are necessary to rationalise a data set D.

Recovery of α^1 The minimal squared perturbations to the marginal utility of good-1 relative to preferences in period 1 that are necessary to good-1 rationalise observed choice behaviour $D = \{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$ are identified as the unique solution set $\{\alpha_t^1\}_{t=1,\ldots,T}$ to the following quadratic programme.

$$\min_{\{v_t,\lambda_t,\alpha_t\}_{t=1,\ldots,T}} \sum_{t=1}^{T} \left(\alpha_t^1\right)^2 \tag{7}$$

subject to:

1. The revealed preference inequalities:

$$\begin{aligned} v_s - v_t + \alpha_t^1(q_s^1 - q_t^1) &\leq \lambda_t \mathbf{p}_t'(\mathbf{q}_s - \mathbf{q}_t) \\ \alpha_t^1 &\leq \lambda_t p_t^1 \end{aligned} \tag{8}$$

2. The normalisation conditions:

$$\begin{aligned} v_1 &= \delta \quad \text{(an arbitrary constant)} \\ \lambda_1 &= \beta \quad \text{(a strictly positive constant)} \\ \alpha_1^1 &= 0 \end{aligned} \tag{9}$$

for $s, t = 1, ..., T$.

Minimizing the sum of squared α^1's subject to the set of revealed preference inequalities ensures that the recovered pattern of minimal taste perturbations are sufficient to rationalise observed choice behaviour. The normalisation conditions are required because the quadratic programming problem is ill-posed in their absence. This is due to the ordinality of the utility function. Let $\{\bar{v}_t, \bar{\lambda}_t, \bar{\alpha}_t^1\}_{t=1,...,T}$ represent some feasible solution to the rationalisation constraints. As the set represents a feasible solution, the following inequalities hold for all $s, t \in \{1, ..., T\}$:

$$\bar{v}_s - \bar{v}_t + \bar{\alpha}_t(q_s^1 - q_t^1) \geq \bar{\lambda}_s \mathbf{p}_s'(\mathbf{q}_s - \mathbf{q}_t) \tag{10}$$

However, the following set of inequalities is also feasible:

$$\beta(\bar{v}_s + \delta) - \beta(\bar{v}_t + \delta) + \beta\bar{\alpha}_t^1(q_s^1 - q_t^1) \geq \beta\bar{\lambda}_s \mathbf{p}_s'(\mathbf{q}_s - \mathbf{q}_t) \tag{11}$$

for $\delta \in (-\infty, \infty)$ and $\beta > 0$. Thus, without a location and scale normalisation, there exist an infinite number of feasible sets of utility numbers that can be associated with a given set of feasible taste shifters. Without loss of generality, we impose $\alpha_1^1 = 0$, which allows us to interpret $\{\alpha_t^1\}_{t=2,...,T}$ as the minimal rationalising marginal utility perturbations to good-1 relative to preferences at $t = 1$. We also impose the scale normalisation $\lambda_1 = \beta$ to ensure that the output of the quadratic programming procedure is scaled sensibly.

3 Interpersonal unobserved heterogeneity

In conventional revealed preference analyses, recovery exercises are typically conducted on an individual-by-individual basis. However, in many settings panel data is not available (or its time-dimension is short). Data must be pooled across different individuals over time to make statements about taste changes. There are then two dimensions of unobserved preference heterogeneity to consider: heterogeneity in a given individual's tastes over time (i.e. $\alpha_{it} \neq \alpha_{is}$ for an individual i and time periods t and s) and heterogeneity in tastes *across* individuals at any given time (i.e. $\alpha_{it} \neq \alpha_{jt}$ for individuals i and j). Heterogeneity in tastes across individuals gives rise to a distribution of demands at any given budget regardless of the existence of intertemporal taste instability. In this data environment, the aim is to recover the minimal changes in tastes such that changes in the distribution of demands are consistent with revealed preference.

Drawing on Blundell, Kristensen and Matzkin (2014), we restrict the dimensionality of the demand system to $K = 2$ and assume that individual demands (in any given time period) are monotonic in scalar unobserved heterogeneity that is distributed $U(0,1)$. That is, we develop the specification of individual preferences in (2) to take the form:

$$\begin{aligned} u_i(q^1, q^0; \alpha_{it}) &= v_i(q^1, q^0, \varepsilon_i) + \alpha_{it}^1 q^1 \\ &= v(q^1, q^0) + w(q^1, \varepsilon_i) + \alpha_{it}^1 q^1 \\ &= v(q^1, x - p^1 q^1) + \varepsilon_i q^1 + \alpha_{it}^1 q^1 \end{aligned} \tag{12}$$

where ε_i represents time-invariant interpersonal taste heterogeneity and α_{it}^1 gives the perturbation to marginal utility experienced by i at t, as in (2).

Let $d^1(\mathbf{p}, x, \varepsilon_i + \alpha_{it}^1)$ give the demand associated with individual taste heterogeneity $\varepsilon_i + \alpha_{it}^1$. With time-invariant tastes ($\alpha_{it}^1 = 0$ for all i, t), our assumptions guarantee that quantile demands, conditional on $(\mathbf{p}, x)$, identify individual demands, i.e.

$$\begin{aligned} d^1(\mathbf{p}, x, \varepsilon_i) &= \max_{q^1} v(q^1, x - p^1 q^1) + \varepsilon_i q^1 \\ &= F^{-1}_{q^1|\mathbf{p},x}(\varepsilon_i) \end{aligned} \tag{13}$$

where $\varepsilon_i \in [0,1]$ and $\alpha_{it}^1 = 0$ for all i, t.

To equate quantile and individual demands, given the time varying component to unobserved preference heterogeneity, additional assumptions must be made. Specifically, taste evolution cannot alter the ranking of individuals in the q^1 distribution conditional on $(\mathbf{p}, x)$. That is, if

$$d^1(\mathbf{p}_s, x_s, \varepsilon_i + \alpha_{is}^1) < Q_{q^1|\mathbf{p}_s, x_s}(\tau) \tag{14}$$

then

$$d^1(\mathbf{p}_t, x_t, \varepsilon_i + \alpha_{it}^1) < Q_{q^1|\mathbf{p}_t, x_t}(\tau) \tag{15}$$

This assumption is strong but not incredible, and motivated in our empirical application in Section 4.

3.1 Recovery

The quadratic programming procedure of Section 2 can be applied to recover the minimal taste perturbations necessary for consistency with the RP inequalities at each quantile. However, to meaningfully compare the magnitude of minimal α^1 across quantiles, the procedure must be applied simultaneously to the full of observations $\{D^i\}^{i=1,\ldots,\tau}$, where i is a particular quantile (individual) of interest. Thus, rather than apply the method to data on a single individual, $D^i = \{\mathbf{p}_t, \mathbf{q}^i_t\}_{t=1,\ldots,T}$, one recovers $\{\alpha^{1i}_t\}^{i=1,\ldots,\tau}_{t=1,\ldots,T}$ sufficient to rationalise the pooled data set $\{D^i\}^{i=1,\ldots,\tau}$. Running the recovery exercise independently on each time series of quantile demands would not guarantee commonality of base preferences, i.e. that $v_i(q^1, q^0) = v_j(q^1, q^0) = v(q^1, q^0)$,[6] precluding comparisons of the magnitude of α^1 parameters across quantiles given that these parameters represent marginal utility perturbations and their magnitude is meaningless independent of some ordinal representation of base preferences. Here the normalisation of tastes will be to some quantile i in some period t, allowing us to interpret recovered α^1_{it}'s as the minimum taste shifts relative to i's tastes at t that are necessary to rationalise observed behaviour.

4 Rationalising tobacco consumption

In this application, we apply the new revealed preference methodology to recover the minimal taste perturbations for tobacco that are required to rationalise the consumptions patterns of a set of psuedo-cohorts that were stratified by education level and the level of tobacco consumption.

4.1 The data

We use repeated cross-section data drawn from the U.K. Family Expenditure Survey (FES) with biennial periodicity between 1980 and 2000.[7] The FES records detailed expenditure and demographic information for 7,000 randomly selected households each year. From this data set we construct two groups of individuals that are stratified by

[6] Differences in the base utility function across individuals could lead to an interpersonal violation of single-crossing and preclude one from making global statements about taste difference for good-1 across individuals.

[7] We restrict observations to biennial periodicity to reduce the salience of habit formation in tobacco for our method and to limit the computational burden of the procedure.

education level and who were aged between 25 and 35 years old in 1980.[8] The "low" education group, L, is formed from those individuals who left school no older than the legal minimum, 15 years old at the time. Those staying in school past this age comprise the "high education" group, H. Appendix B provides further details and summary statistics.

We are primarily concerned with choice over a tobacco aggregate and a nondurable commodity aggregate, i.e. $K = 2$. Appendix B provides a detailed list of goods that are classed as nondurables in our data set. Total expenditure is defined as all spending on these goods. Price indices are constructed for the tobacco and nondurable aggregates using the sub-indices of the U.K. Retail Price Index. As displayed in Figure 1, the relative of price of tobacco relative to nondurables rose significantly between 1980 and 2000. The rise was relatively smooth, apart from a brief stagnation in the late 1980s that was caused by a reduction in growth of the absolute price of tobacco rather than an increase in the growth rate of nondurable prices.

Figure 1. Relative price of tobacco

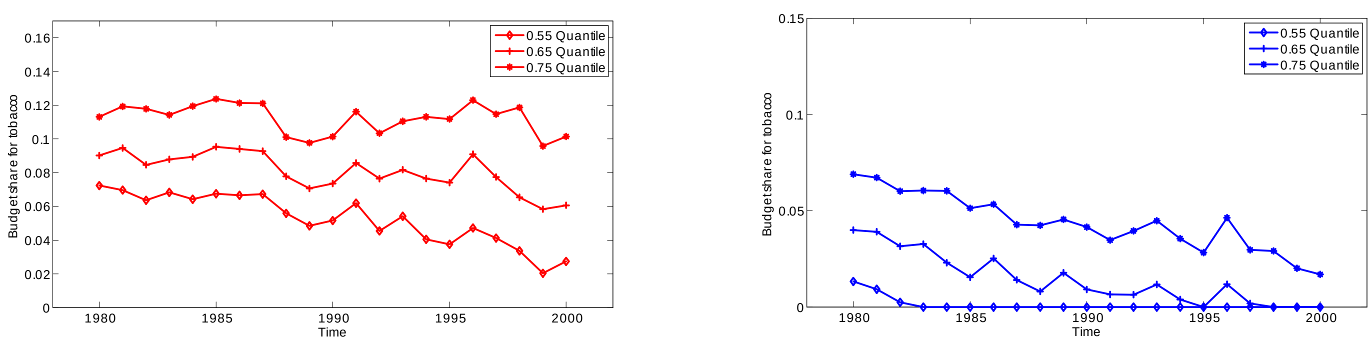
Over the same period, there is evidence to suggest that individuals acquired more information about the ill-effects of smoking, perhaps causing tastes for tobacco to change. The Health and Lifestyle Survey 1984 and the Office for National Statistics Omnibus Survey 1996 both asked questions about the connection between smoking and poor

[8] We select this birth cohort because less than 5% of smokers start smoking after they reach their 25^{th} birthday (Office for National Statistics, 2012). The assumption that the population of smokers is stable is then relatively mild.

health. Between these years, there was a large rise in awareness of the link between smoking and heart disease over the period; in 1984, only 25.8% of the sample believed that smoking caused heart disease, while by 1996, 80.6% recognised the link.

In the following sections, we will rationalise the changes in tobacco consumption at three different quantiles of the budget share for tobacco distribution, $\boldsymbol{\tau} = \{0.55, 0.65, 0.75\}$, for each education group. We refer to the 0.55-quantile as "light smokers", to the 0.65-quantile as "moderate smokers" and we refer to the 0.75-quantile as "heavy smokers". We thus recover taste changes for six different cohorts. Figure 2 displays the budget shares for tobacco at the relevant quantiles of both education cohorts between 1980 and 2000. At each quantile, the high education cohort devotes a smaller proportion of their budget to tobacco than the low education cohort. In fact, within the high education cohort, the 0.55 quantile of the budget share for tobacco distribution falls to zero in 1983 and remains nil for the the rest of the period considered, while the 0.65 quantile hits zero by the end of the period. Tobacco consumption remains strictly positive in all periods for the quantiles of the low education cohort that are considered.

Figure 2. Budget share for Tobacco



(a) Low Education (b) High Education

4.2 Controlling for budget variation

We are interested in recovering the contribution of taste change to the trends in the distribution of tobacco consumption over time and between different cohorts. However, budget variation in the data hinders our progress somewhat. Income growth in survey data can prevent the intersection of budget hyperplanes, which weakens ones ability to detect violations of stable preferences, see Varian (1983). To make our methodology as powerful as possible, we evaluate each cohort's demands at a series of expenditure levels that are constructed in a similar manner to Blundell, Browning and Crawford's (2003) "Sequential Maximum Power" (SMP) paths. Intuitively, the SMP method controls for budget variation to provide a choice set that maximises the chance of detecting violations of a time-invariant utility function in observational data.

The distribution of demands will change in response to price and income changes even when preferences are stable and consistent with the Revealed Preference inequalities. We recover the minimal α_t taste change perturbations that are required to rationalise the distribution of choice behaviour. We maintain the assumption of rank invariant shifts in the distribution of demands. Along the SMP path that starts at an education group's demand in 1980, and which then continues sequentially over time, we select the demand that is just weakly preferred to the SMP demand in the previous period. Specifically, we recover taste changes for quantile demands of each education cohort at the budgets $\{\mathbf{p}_t, \tilde{x}_t^{ed}\}_{t=1,..,T}$, where the SMP expenditure levels $\{\tilde{x}_t^{ed}\}_{t=1,..,T}$ are given as:[9]

$$x_1^{ed} = Q_{x_1}(0.5|ed) \tag{16}$$

and

$$\widetilde{x}_t^{ed} = \mathbf{p}_t' \mathbf{q}_{t-1}^{ed} \tag{17}$$

for $t = 2, ..., T$ and $ed = \{L, H\}$, where $\mathbf{q}_t^{ed} = \mathbf{q}^{ed}(\mathbf{p}_t, \widetilde{x}_t^{ed})$.[10] We can abstract from the complicating issues caused by transitivity in the construction of the SMP path that are examined in Blundell *et al.* (2015) because we consider a two-good demand system. As first highlighted by Rose (1958), transitivity has no empirical content for a 2-good demand system.

4.3 Censored quantile demands

As in Blundell, Kristensen and Matzkin (2014) we assume demands are monotonic in scalar unobserved heterogeneity so that conditional quantiles, conditional on income and price regime, identify individual demands. To allow for the censoring in the observed demand data we use censored quantile regression methods in the spirit of Chernozhukov and Hong (2002) and Chernozhukov, Fernandez-Val and Kowalski (2010). Censored quantile regression models are a natural choice given the theoretical framework: with single crossing, the quantile rank of an individual's α parameter

[9] This reflects an assumption that individuals with the same education level face the same relative prices in each period.

[10] Please note that demands for light, moderate and heavy smokers within each education group are recovered at the same expenditure level, but that the SMP expenditure level differs between education groups.

identifies the quantile rank of their tobacco demand. The latent budget share for tobacco, $w_i^\star$ is left censored at zero: $w_i = max\{0, w_i^\star\}$. We follow Blundell and Powell (2007) and Imbens and Newey (2009), using a quantile control function approach to correct for the endogeneity of total expenditure.

More specifically, the following triangular system of quantile equations determines the τ_i^{th} quantile of household i's budget share of tobacco, w_i, at each price regime:

$$
\begin{aligned}
w_i &= max(0, w_i^\star) \\
w_i^\star &= Q_{w_i^\star}(\tau_i | x_i, \mathbf{z}_i, v_i, ed_i) \\
x_i &= Q_{x_i}(v_i | z_i, m_i, ed_i)
\end{aligned}
\tag{18}
$$

where

$$
\begin{aligned}
\tau_i &\sim U(0,1) | x_i, \mathbf{z}_i, v_i, ed_i, m_i \\
v_i &\sim U(0,1) | z_i, m_i, ed_i
\end{aligned}
\tag{19}
$$

and x_i is log total expenditure (on nondurables and tobacco), $\mathbf{z}_i$ is a vector of household characteristics, v_i is an unobserved latent variable that is included to account for the possible endogeneity of x_i, and $ed_i \in \{L, H\}$ denotes individual i's education cohort membership. m_i, the log of disposable income, is our excluded instrument that allows us to recover control function v_i from the conditional quantile of x given (z_i, m_i, ed_i).

As discussed in Section 3, the model imposes a rank invariance restriction on the budget share for tobacco distribution within education groups; taste evolution and price effects cannot alter the ranking of individuals within the same education group in the budget share for tobacco distribution. Therefore, for any uncensored quantiles of the budget share distribution within a particular education group, τ and τ', with $\tau > \tau'$, then $\alpha_t^\tau > \alpha_t^{\tau'}$ for all $t = 1, ..., T$. In this application, we only consider individuals who are old enough to have made an initial decision to smoke and high profile anti-smoking campaigns often are not targeted in a way that would cause one to expect significant re-ranking within education cohorts.

5 Estimation

The model developed above is used to recover conditional cohort demands at a set of quantiles of the tobacco distribution $\boldsymbol{\tau} = \{0.55, 0.65, 0.75\}$, for each price regime and education group at the SMP expenditure levels. In the implementation, $[x_i, z_i] = [log(x), log(x)^2, oecd]$, where *oecd* is the OECD demographic index.[11]

To estimate this model, we extend the semiparametric estimator of Chernozhukov, Fernandez-Val and Kowalski (2010) to allow for unrestricted heterogeneity between education groups.

$$\begin{aligned} Q_{w_i^*}(\tau_i | x_i, \mathbf{z}_i, v_i, ed_i) &= \mathbf{X}_i' \boldsymbol{\beta}^{ed_i}(\tau_i) \\ \mathbf{X}_i &= f(x_i, \mathbf{z}_i, v_i) \end{aligned} \tag{20}$$

Assuming a random sample of N observations at each price regime $\boldsymbol{\beta}^{ed}(\tau)$ is estimated as:

$$\boldsymbol{\beta}^{ed}(\tau) = arg \min_{\beta} \sum_{i \in S^{ed}} \rho_\tau (w_i - \hat{\mathbf{X}}_i' \boldsymbol{\beta}) \tag{21}$$

where ρ_τ is the standard asymmetric absolute loss function of Koenker and Bassett (1978). S^{ed} denotes the set of observations on individuals i for which $ed_i = ed$ and $Pr(w_i > 0) > \delta$, i.e. the subset of individuals in education group *ed* for which the probability of censoring is negligible and a linear functional form for the conditional quantile is justified.

$\mathbf{X}_i$ is replaced with $\hat{\mathbf{X}}_i = f(x_i, \mathbf{z}_i, \hat{v}_i)$, where $\hat{v}_i$ is an estimator of v_i, because the true value of v_i is unobserved. An additive linear model allows v to be easily recovered from the cumulative distribution function of the least squares residuals:

$$\begin{aligned} Q_x(\nu | z, m, ed) &= \mathbf{Z}' \boldsymbol{\pi}^{ed} + Q_v(\nu | ed) \\ \widehat{v}_i &= Q_v^{-1}(x_i - \mathbf{Z}_i' \widehat{\boldsymbol{\pi}}^{ed} | ed) \end{aligned} \tag{22}$$

We refer the reader to Chernozhukov, Fernandez-Val and Kowalski (2010) for a more detailed exposition, including the construction of S^{ed}.

5.1 Implementation

The quadratic programming procedure of Section 2 is applied to $\{\mathbf{p}_t, \mathbf{q}_t^{ed,\tau}\}_{t=1,\dots,T}$, with $ed = \{L, H\}$ and $\tau = \{0.55, 0.65, 0.75\}$ and $\mathbf{q}_t^{ed,\tau}$ estimated as above. Application to the pooled choice set of all quantiles and education groups imposes a common base utility function for all. Taste changes are normalised relative to the heavy smoking, low education group in 1980, $\alpha_1^{L,0.55} = 0$. Implementation of the method is a computationally intensive procedure. To limit the computational burden of the empirical exercise, observations are restricted to biennial periodicity.

[11] When predicting the demands, we use the SMP incomes, $oecd = 0.67$ for a single person household and v is set to 0.

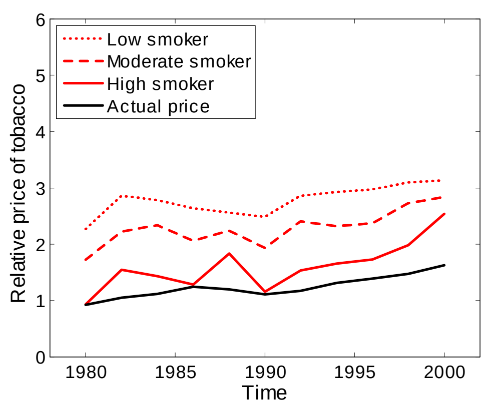
The procedure is bootstrapped to address the issue of sampling variation in estimated quantity sequences. Observations are randomly drawn with replacement within each education-time cell 1000 times and quantile demands are estimated on each resampled set of observations.

Minimal taste change is estimated for these sets of perturbed quantities, allowing us to construct a simultaneous 95% confidence interval on taste perturbations for each cohort.

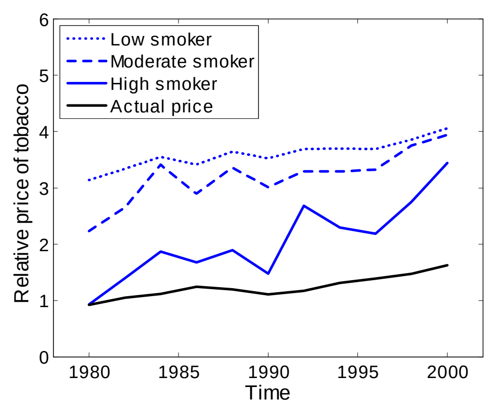
6 Results

We find significant differences in the virtual price trajectories along the SMP total expenditure path across cohorts. Figures 3 and 4 depict the minimum virtual prices and the minimal taste wedge that are necessary to rationalise the choice behaviour of the low and high education cohorts, given the normalisation of taste shifters relative to the heavy-smoking low-education cohort in 1980. In addition to recovered virtual prices, the observed relative price for tobacco is depicted. This is the trajectory that each cohort's virtual price would follow in the absence of intercohort and intertemporal taste heterogeneity. This represents the change in the marginal willingness to pay for tobacco relative to base tastes along the SMP expenditure paths.

Figure 3. Virtual price for tobacco

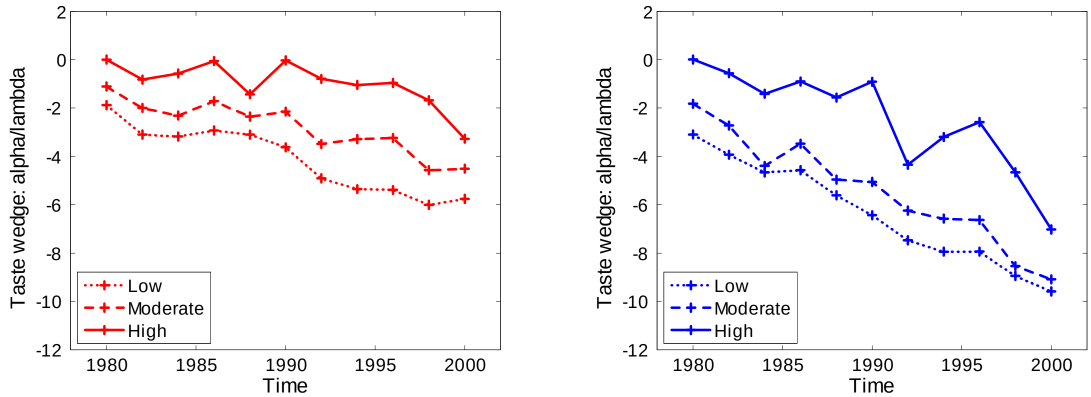


(a) Low Education



(b) High Education

Figure 4. Taste wedge for tobacco



(a) Low Education

(b) High Education

Some degree of taste variation is necessary to rationalise the behaviour of every cohort for the period 1980-2000; all virtual price trajectories are significantly different from the price that is observed in reality. Figures 3 and 4 display the intracohort heterogeneity in virtual prices and the taste wedge. We here suppress the 95% confidence intervals to allow for an uncluttered overview of intracohort trends but they are given in later Figures. Unsurprisingly, the virtual price of heavier smoking cohorts is lower compared to that of lighter smoking cohorts. The marginal willingness to pay for tobacco, captured by the taste wedge, is, therefore, lower for lighter smoking cohorts compared to that dictated by the tastes of the low-education heavy-smokers in 1980.

Turning now to compare the virtual price trajectories across education groups, Figure 5 shows the evolution of the taste wedge by smoking group. We find statistically significant heterogeneity across education cohorts for light and moderate smokers along their respective quantity sequences. This is highlighted by reference to panels (b) and (c), which illustrate that the confidence intervals on the taste wedge trajectories for low and high education cohorts are disjoint at light and moderate smoking quantiles.

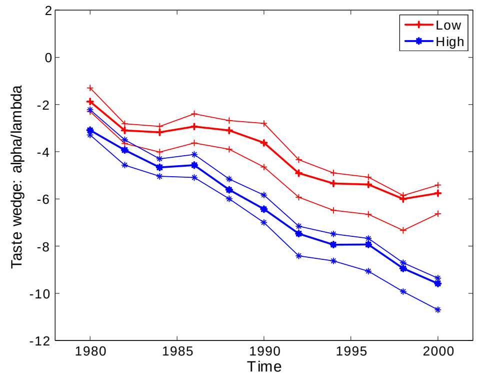
The finding of significant educational differences in tastes is coherent with independent evidence on the differences in the propensity of the education cohorts to link smoking and heart disease: in 1984 for example, 45.2% of highly educated individuals and 39.8% of less educated individuals believed that smoking caused heart disease; by 1996, 83.3% of highly educated individuals and 72.4% of less educated individuals recognised the link.[12] Thus, the percentage point gap in responses by education level rose from 5.7 to 10.9 over the decade. Such patterns are

[12] Health and Lifestyle Survey 1984 and the Office for National Statistics Omnibus Survey 1996

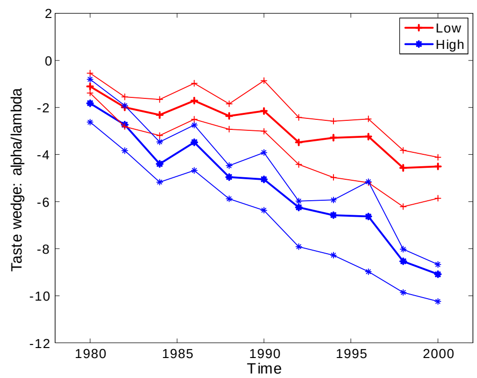
suggestive of a link between information acquisition and taste change for tobacco, although we leave a formal test of this hypothesis to later work.

However, we find that educational differences in virtual prices are not significant for heavy smokers; the 95% confidence intervals over the minimal virtual prices trajectories overlap for the low and high education cohorts at the 0.75-quantile. It should also be noted that there are a number of periods in which the taste wedge of the low-education heavy-smoking group is not significantly different from zero. These findings suggest that the taste for tobacco is more robust amongst heavier smokers, which accords with biological evidence that nicotine addiction has a genetic basis. Particular gene sequences (especially variants in chromosome 15) have been found to be associated with, among other factors, the number of cigarettes smoked per day and nicotine intake (see, for example, Mansfelder (2000), Berrettini (2008), Keskitalo (2009)). Given that one's genome does not change except by random mutation, one would expect tastes grounded in genetic factors to be relatively time invariant as they appear to be in the case of our heavy-smokers.

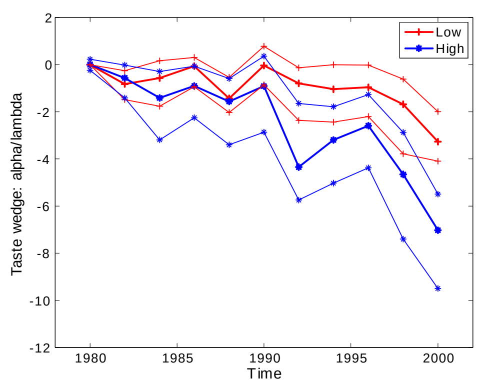
Figure 5. Taste wedge



(a) Light smokers



(b) Moderate smokers



(c) Heavy smokers

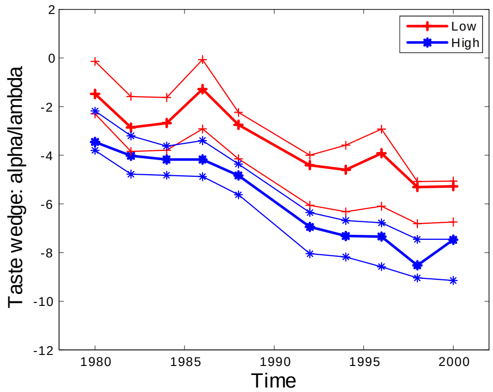
7 Conditional taste change

The estimation strategy we used above implicitly assumes that tobacco is weakly separable from all other non-durables, including alcohol. To examine the robustness of our findings to this assumption, and to explore whether additional patterns emerge from the data once this condition is relaxed, we follow the approach of Browning and Meghir (1998) and re-run our quadratic programming procedure on quantile demands that are estimated conditional on alcohol consumption. Specifically, we partition the set of observations that comprise each education group into "light" and "heavy" drinkers depending on whether an individual consumes below or above the median budget share for alcohol in that time period. The quantile regression method outlined in the previous section is then applied to estimate demands within each education-alcohol-time cell at the expenditures along the same SMP paths that were calculated previously. Base preferences are normalised relative to those of the heavy smoking, heavy drinking, low education cohort in 1980.

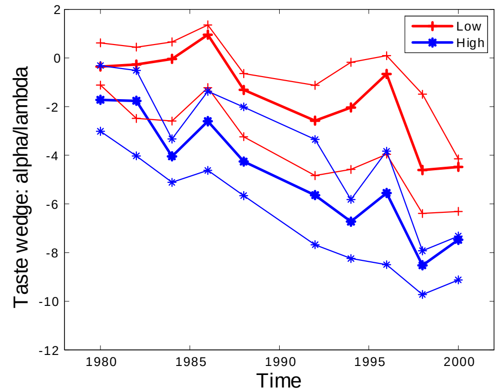
Taste shifters are only estimated for moderate and heavy smokers that are drawn from our education-alcohol cohorts. The requirement of perfect intertemporal variability of good-1 is violated for demands on the "high education"-"light smoking"-"light drinking" cohort; for most of the period considered the estimated budget share on tobacco for this cohort is zero. We thus do not calculate taste changes for light smokers. This ensures that there will always exist a set of comparable, choice-rationalising taste shifters that can be estimated in a reasonable amount of time. However, please note that the fact that light smokers are not included at this stage implies that the magnitude of estimated taste shifters are not comparable to the unconditional quantile results of the previous section.

We recognise that the rank invariance assumption, which is imposed by the quantile regression model that is used to estimate SMP demands, is stronger in this context of conditional demands. It amounts to a no re-ranking requirement on the joint distribution of tobacco-alcohol group budget shares. We cannot determine how strong this assumption is because we do not have access to panel data for the period. However, research suggests a robust, if modest, positive correlation between alcohol and smoking consumption that is replicated across many studies, which lends some support to our strategy (Bobo and Husten, 2000; Falk, Yi and Hiller-Sturmhofel, 2006).

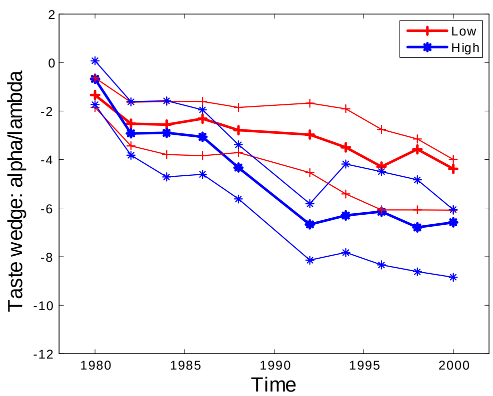
Figure 6: Conditional demands- Taste wedge



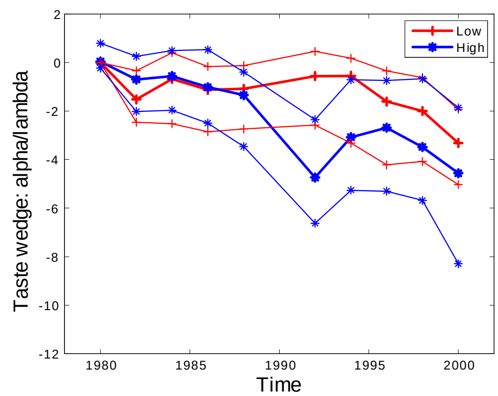
Moderate Smoker - Light Drinker



Heavy Smoker - Light Drinker



Moderate smoker - Heavy Drinker



Heavy Smoker - Heavy Drinker

The main themes arising from our earlier results are robust to conditioning of demands upon alcohol consumption. We find our low education cohorts have typically experienced less taste change than high education cohorts in each smoking level-drinking level cell. However, it is only for the moderate smoking- low drinking cohort that educational differences in tastes are statistically significant; in this case, the 95% confidence intervals on virtual prices and the taste wedge are disjoint from 1984 onwards. The marginal willingness to pay evolved very little for the heavy-smoking heavy-drinking cohorts, regardless of education level. This finding is consistent with government and health practitioner reports that note low smoking cessation rates amongst heavy drinkers (Dollar *et al.*, 2009). It also motivates the use of further restrictions on the joint consumption of alcohol and tobacco (e.g. bans on smoking in pubs and bars). Such restrictions would exploit the complementarity between alcohol and tobacco to reduce smoking amongst the heavy-smoking heavy-drinking group given the persistence in their tastes for tobacco.[13]

[13] Although see Adda and Cornaglia (2010) on the additional effects of such restrictions.

8 Conclusions

This paper has provided a theoretical and empirical framework for characterising taste change. We have uncovered a surprising non-identification result: observational data sets on a K-dimensional demand system can always be rationalised by taste change on a single good in a nonparametric setting. Our theoretical results were used to develop a quadratic programming procedure to recover the minimal intertemporal (and interpersonal) taste heterogeneity required to rationalise observed choices. The approach we have developed establishes that we can almost always rationalise a data set when allowing for taste change in a single good. We have shown that these conditions are equivalent to solving a "latent virtual price" problem.

A censored quantile approach was used to allow for unobserved heterogeneity and censoring in the application of our approach to the consumption of tobacco in the UK over the period 1980 to 2000 where we might expect large shifts in demand due to taste change. Non-separability between tobacco and alcohol consumption was incorporated using a conditional (quantile) demand analysis. We focussed on a single birth cohort aged between 25 and 35 years old in 1980 and allowed for examined different education groups. The estimation results suggested that systematic taste change was required for all groups in our expenditure survey data. In no case could the fall in tobacco consumption over the twenty year period be rationalised solely by the rise in the relative price of tobacco. A series of negative perturbations to the marginal utility of tobacco were found to be sufficient to rationalise the trends in tobacco consumption.

Statistically significant educational differences in the marginal willingness to pay for tobacco were recovered; more highly educated cohorts experienced a greater shift in their effective tastes away from tobacco. We find virtual prices and the taste wedge are disjoint across education groups for all cohorts except for the "heavy smoking"-"heavy drinking" group. Education is largely irrelevant for explaining the evolution of virtual prices amongst heavy smokers. This might suggest diminished differences in smoking behaviour by education group in the future.

References

[1] Jerome Adda and Francesa Cornaglia (2010), "The Effect of Bans and Taxes on Passive Smoking", *American Economic Journal: Applied Economics*, 2(1), 1-32.

[2] Sydney Afriat (1967), "The Construction of Utility Functions from Expenditure Data", *International Economic Review*, 8(1), 67-77.

[3] Sydney Afriat (1977), *The Price Index*, London: Cambridge University Press.

[4] Laura Blow and Ian Crawford (1999), "Valuing Quality", *IFS Working Paper Series W99/21*, The Institute for Fiscal Studies.

[5] Richard Blundell, Martin Browning, Laurens Cherchye, Ian Crawford, Bram De Rock and Frederic Vermeulen (2015), “Sharp for SARP: Nonparametric bounds on the behavioural and welfare effects of price changes”, *American Economic Journal: Microeconomics*, 7(1): 43–60

[6] Richard Blundell, Martin Browning and Ian Crawford (2003), "Nonparametric Engel Curves and Revealed Preference", *Econometrica*, 71(1), 205-240.

[7] Richard Blundell, Martin Browning and Ian Crawford (2008), "Best Nonparametric Bounds on Demand Responses", *Econometrica*, 76(6), 1227-1262.

[8] Richard Blundell, Dennis Kristensen and Rosa Matzkin (2014) “Bounding Quantile Demand Functions using Revealed Preference Inequalities”, *Journal of Econometrics*, 117, 112-127.

[9] Richard Blundell and James Powell (2007), "Censored Regression Quantiles with Endogenous Regressors", *Journal of Econometrics*, 141(1), 65-83.

[10] Janet Bobo and Corinne Husten (2000), "Sociocultural Influences on Smoking and Drinking", *Alcohol Research and Health*, 24(4), 225-232.

[11] Donald Brown and Rosa Matzkin (1998), "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand", *Cowles Foundation Discussion Papers 1175*, Cowles Foundation, Yale University.

[12] Martin Browning (1989), "A Nonparametric Test of the Life-Cycle Rational Expectations Hypothesis", *International Economic Review*, 30(4) 979-992.

[13] Martin Browning and Costas Meghir (1991), "The Effects of Male and Female Labor Supply on Commodity Demands", *Econometrica*, 59(4), 925-951.

[14] Victor Chernozhukov and Han Hong (2002), "Three-Step Censored Quantile Regression and Extramarital Affairs", *Journal of the American Statistical Association*, 40(459), 872-882.

[15] Victor Chernozhukov, Iván Fernández-Val and Amanda Kowalski (2010), "Quantile Regression with Censoring and Endogeneity", *Mimeo*, Massachusetts Institute for Technology.

[16] Ian Crawford (2010), "Habits Revealed", *Review of Economic Studies*, 77(4), 1382-1402.

[17] Department of Health (2011), *White Paper: Healthy Lives, Healthy People: Our Strategy for Public Health in England*, London: TSO.

[18] W. Erwin Diewert (1973), "Afriat and Revealed Preference Theory", *The Review of Economic Studies*, 40(3), 419-425.

[19] Charles Manski (2007), *Identification for Prediction and Decision*, Cambridge, MA: Harvard University Press.

[20] Daniel McFadden and Mogens Fosgerau (2012), "A Theory of the Perturbed Consumer with General Budgets", *NBER Working Papers 17953*, National Bureau of Economic Research, Inc.

[21] Paul Milgrom and Chris Shannon (1994), "Monotone Comparative Statics", *Econometrica*, 62(1), 157-180.

[22] Christopher Murray and Alan Lopez (1997), "Global Mortality, Disability, and the Contribution of Risk Factors: Global Burden of Disease Study", *The Lancet*, 349(9063), 1436-1442.

[23] Peter Neary and Kevin Roberts (1980), "The Theory of Household Behaviour Under Rationing", *European Economic Review*, 13(1), 25-42.

[24] James Powell (1984), "Least Absolute Deviations Estimation for the Censored Regression Model", *Journal of Econometrics*, 25(3), 303-325.

[25] Hugh Rose (1958), "Consistency of Preference: The Two-Commodity Case", *Review of Economic Studies*, 25(2), 124-125.

[26] Hal Varian (1982), "The Nonparametric Approach to Demand Analysis", *Econometrica*, 50(4), 945-973.

[27] Hal Varian (1983), "Non-Parametric Tests of Consumer Behaviour", *Review of Economic Studies*, 50(1), 99-110.

[28] Hal Varian (1988), "Revealed Preference with a Subset of Goods", *Journal of Economic Theory*, 46(1), 179-185.

[29] World Health Organisation (2012), "Mortality Attributable to Tobacco", *WHO Global Report Series*.

Appendix A: Proofs

Theorem 1. The following statements are equivalent:

1. Observed choice behaviour, $D = \{\mathbf{p}_t, \mathbf{q}_t^i\}_{t=1,...,T}$, can be good-1 taste rationalised.
2. One can find sets $\{v_t\}_{t=1,...,T}$, $\{\alpha_t^1\}_{t=1,...,T}$ and $\{\lambda_t\}_{t=1,...,T}$ with $\lambda_t > 0$ for all $t = 1, ..., T$, such that there exists a non-empty solution set to the following revealed preference inequalities:

$$\begin{aligned} v_s - v_t + \alpha_t^1(q_s^1 - q_t^1) &\leq \lambda_t \mathbf{p}_t'(\mathbf{q}_s - \mathbf{q}_t) \\ \alpha_t^1 &\leq \lambda_t p_t^1 \end{aligned} \tag{23}$$

Proof:
Necessity: Let us consider the case where our data set has been generated by the model. Observed choices are then the solution to the following optimisation problem:

$$\max_{\{\mathbf{q}_t\}_{t=1,...,T}} v(\mathbf{q}_t) + \alpha_t^1 q_t^1$$

subject to

$$\mathbf{p}_t'\mathbf{q}_t \leq x_t$$

An optimal interior solution to the problem must satisfy:

$$\begin{aligned} \nabla_{q_t^1} v(\mathbf{q}_t) + \alpha_t^1 &= p_t^1 \\ \nabla_{\mathbf{q}_t^{-1}} v(\mathbf{q}_t) &= \mathbf{p}_t^{-1} \end{aligned}$$

Given a particular level of the taste shifter, α_t^1, concavity of the utility function implies:

$$u(\mathbf{q}_t, \alpha_t^1) + \nabla_{\mathbf{q}_t} u(\mathbf{q}_t, \alpha_t^1)'(\mathbf{q}_s - \mathbf{q}_t) \quad \geq \quad u(\mathbf{q}_s, \alpha_t^1)$$

Substituting the first order conditions into the concavity condition and rearranging gives:

$$v(\mathbf{q}_s) - v(\mathbf{q}_t) + \alpha_t^1(q_s^1 - q_t^1) \leq \lambda_t \mathbf{p}_t'(\mathbf{q}_s - \mathbf{q}_t)$$

Letting $v_t = v(\mathbf{q}_t)$, returns the first set of inequalities.

The second set of inequalities are required for the base utility function to be strictly increasing in $\mathbf{q}$. From the first order conditions,

$$\alpha_t^1 \quad > \quad \lambda_t p_t^{k,i}$$

would imply

$$\nabla_{q_t^1} v(\mathbf{q}_t) \quad < \quad 0$$

violating monotonicity.
Sufficiency: The concavity condition associated with the taste-varying utility function, $u(\mathbf{q}, \alpha_t^1)$ implies the existence of T overestimates of the utility of some bundle $\mathbf{q}$:

$$\begin{aligned} v(\mathbf{q}) &\leq v_t + \lambda_t \mathbf{p}_t'(\mathbf{q} - \mathbf{q}_t) - \alpha_t^1(q^1 - q_t^1) \\ v(\mathbf{q}) &\leq v_t + \lambda_t \widetilde{\mathbf{p}}_t'(\mathbf{q} - \mathbf{q}_t) \end{aligned}$$

where $\widetilde{p}_t^1 = p_t^1 - \alpha_t^1/\lambda_t^i$ and $\widetilde{\mathbf{p}}_t^{\neg 1} = \mathbf{p}$. A piecewise linear utility function can be derived from the lower envelope of the hyperplanes defined by these T overestimates:

$$v(\mathbf{q}) = \min_t \{v_t + \lambda_t \widetilde{\mathbf{p}}_t'(\mathbf{q} - \mathbf{q}_t)\}$$

The utility of any feasible consumption bundle cannot be strictly greater than that conferred by observed choices with the utility function defined as above. Consider an arbitrary feasible bundle, $\widehat{\mathbf{q}}$:

$$\mathbf{p}_t'\widehat{\mathbf{q}} \leq \mathbf{p}_t'\mathbf{q}_t$$

Given the definition of the base individual utility function:

$$v(\widehat{\mathbf{q}}) + \alpha_t^1 \widehat{q}^1 \leq v_t + \lambda_t \widetilde{\mathbf{p}}_t'(\widehat{\mathbf{q}} - \mathbf{q}_t) + \alpha_t^1 \widehat{q}_t^1$$

Noting that

$$\lambda_t \widetilde{\mathbf{p}}_t'(\widehat{\mathbf{q}} - \mathbf{q}_t) = \lambda_t \mathbf{p}_t(\widehat{\mathbf{q}} - \mathbf{q}_t) - \alpha_t^1(\widehat{q}^1 - q_t^1)$$

returns

$$\begin{aligned} v(\widehat{\mathbf{q}}) + \alpha_t^1 \widehat{q}^1 &\leq v_t + \lambda_t \mathbf{p}_t'(\widehat{\mathbf{q}} - \mathbf{q}) + \alpha_t^1 q_t^1 \\ &\leq u(\mathbf{q}_t, \alpha_t^1) + \lambda_t \mathbf{p}_t'(\widehat{\mathbf{q}} - \mathbf{q}_t) \end{aligned}$$

Further noting that

$$\begin{aligned} \mathbf{p}_t'\widehat{\mathbf{q}} &\leq \mathbf{p}_t'\mathbf{q}_t \\ \mathbf{p}_t'(\widehat{\mathbf{q}} - \mathbf{q}_t) &\leq 0 \end{aligned}$$

Implies that

$$v(\widehat{\mathbf{q}}) + \alpha_t^1 \widehat{q}^1 \leq v(\mathbf{q}_t) + \alpha_t^1 q_t^1$$

In words, any other feasible bundle yields weakly lower utility than $\mathbf{q}_t$. Therefore, we can always construct a utility function which taste rationalises the data set given that a non-empty solution set is associated with the inequalities of Theorem 1. ■

Theorem 2. Given $q_t^1 \neq q_s^1$ for all t and s, any data set D can be good-1 taste rationalised.

Proof.
The inequalities of Theorem 1 can be expressed in terms of virtual prices.

$$\begin{aligned} v_s - v_t + \alpha_t^1(q_s^1 - q_t^1) &\leq \lambda_t \mathbf{p}_t'(\mathbf{q}_s - \mathbf{q}_t) \\ v_s - v_t &\leq \lambda_t \widetilde{\mathbf{p}}_t'(\mathbf{q}_s - \mathbf{q}_t) \end{aligned}$$

where

$$\begin{aligned} \widetilde{\mathbf{p}}_t &= \begin{bmatrix} p_t^1 - \alpha_t^{1i}/\lambda_t^i \\ \mathbf{p}_t^{\neg 1} \end{bmatrix} \\ \widetilde{p}_t^1 &\geq 0 \end{aligned}$$

Varian (1982) proves that the following conditions are equivalent.

1. A data set $\{\widetilde{\mathbf{p}}_t, \mathbf{q}_t^i\}_{t=1,\ldots,T}$ satisfies GARP.
2. There exist numbers $\{v_t\}_{t=1,\ldots,T}$, $\{\alpha_t^1\}_{t=1,\ldots,T}$ and $\{\lambda_t\}_{t=1,\ldots,T}$ with $\lambda_t > 0$ for all $t = 1, \ldots, T$ such that the following "Afriat" inequalities hold.

$$v_s - v_t \leq \lambda_t \widetilde{\mathbf{p}}_t^{'}(\mathbf{q}_s - \mathbf{q}_t)$$

We first establish the existence of rationalising shadow prices $\widetilde{\mathbf{p}}_t$. We observe the data set $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$. Assume that good-1 exhibits perfect intertemporal variation, i.e. $q_t^1 \neq q_s^1$ for all $t \neq s$. We proceed by extending Theorem 1 from Varian (1988) to the current setting.
If $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ satisfies GARP, then the choice set satisfies the inequalities of Theorem 1 with:

$$\alpha_t^1 = \mathbf{0}$$

for $t = 1, ..., T$. If $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ fails GARP, then there exist periods s and t such that

$$\begin{aligned} \mathbf{p}_s'\mathbf{q}_s &\leq \mathbf{p}_s'\mathbf{q}_t \\ \mathbf{p}_t'\mathbf{q}_t &\leq \mathbf{p}_t'\mathbf{q}_s \end{aligned}$$

Given perfect intertemporal variation of good-1, there always exists a set of modifications to p_t^1 such that the GARP inequalities are satisfied. This result follows from Theorem 1 in Varian (1988), in which it is proved that, given perfect intertemporal variation for a good with a missing price, one can always find a price trajectory for this good such that the entire data set satisfies GARP. To demonstrate the relevance of Varian (1988) result in this context, let us consider what value p_t^1 would have to take on, if we were to conjecture that, once taste change is taken into account, the consumer prefers the bundle $\mathbf{q}_t$ to $\mathbf{q}_s$. This conjecture implies the need to prove the existence of a price $\widetilde{p}_t^1$ such that:

$$\begin{aligned} \mathbf{p}_t^{\neg 1\prime}\mathbf{q}_t^{\neg 1} + \widetilde{p}_t^1 q_t^1 &\geq \mathbf{p}_t^{\neg 1\prime}\mathbf{q}_s^{\neg 1} + \widetilde{p}_t^1 q_s^1 \\ \widetilde{p}_t^1 &\geq \frac{\mathbf{p}_t^{\neg 1\prime}(\mathbf{q}_s^{\neg 1} - \mathbf{q}_t^{\neg 1})}{q_t^1 - q_s^1} \end{aligned}$$

where $\mathbf{p}_t^{\neg 1}$ gives the price vector for all goods except good-1, $\mathbf{p}_t^{\neg 1} = [p_t^2, ..., p_t^K]$, and $\mathbf{q}_t^{\neg 1}$ is defined analogously. For each period t, define the lower bound on the virtual price of good 1 such that:

$$\widetilde{p}_t^1 > \max_{s \neq t} \left\{ \frac{\mathbf{p}_t^{\neg 1\prime}(\mathbf{q}_s^{\neg 1} - \mathbf{q}_t^{\neg 1})}{q_t^1 - q_s^1}, 1 \right\}$$

Further define the "taste adjusted direct revealed preferred relation", $\widetilde{\mathbb{R}}^0$. If $\widetilde{\mathbf{p}}_t'\mathbf{q}_t \geq \widetilde{\mathbf{p}}_t'\mathbf{q}_s$, then we conclude that $\mathbf{q}_t$ is directly revealed taste preferred to $\mathbf{q}_s$, or $\mathbf{q}_t\widetilde{\mathbb{R}}^0\mathbf{q}_s$. There are then two cases to consider:

1. $q_t^1 > q_s^1$: In this case, we must have that

$$\begin{aligned} \widetilde{p}_t^1(q_t^1 - q_s^1) &> \mathbf{p}_t^{\neg 1\prime}(\mathbf{q}_s^{\neg 1} - \mathbf{q}_t^{\neg 1}) \\ \widetilde{p}_t^1 q_t^1 + \mathbf{p}_t^{\neg 1\prime}\mathbf{q}_t^{\neg 1} &> \widetilde{p}_t^1 q_s^1 + \mathbf{p}_t^{\neg 1\prime}\mathbf{q}_s^{\neg 1} \\ \widetilde{\mathbf{p}}_t'\mathbf{q}_t &> \widetilde{\mathbf{p}}_t'\mathbf{q}_s \end{aligned}$$

   and set $\mathbf{q}_t\widetilde{\mathbb{R}}^0\mathbf{q}_s$.

2. $q_t^1 < q_s^1$. In this case, we must have that

$$\begin{aligned} \widetilde{p}_t^1(q_t^1 - q_s^1) &< \mathbf{p}_t^{\neg 1\prime}(\mathbf{q}_s^{\neg 1} - \mathbf{q}_t^{\neg 1}) \\ \widetilde{p}_t^1 q_t^1 + \mathbf{p}_t^{\neg 1\prime}\mathbf{q}_t^{\neg 1} &< \widetilde{p}_t^1 q_s^1 + \mathbf{p}_t^{\neg 1\prime}\mathbf{q}_s^{\neg 1} \\ \widetilde{\mathbf{p}}_t'\mathbf{q}_t &> \widetilde{\mathbf{p}}_t'\mathbf{q}_s \end{aligned}$$

   and thus it is not the case that $\mathbf{q}_t\widetilde{\mathbb{R}}^0\mathbf{q}_s$.

Therefore, one can determine the preference ordering of consumption bundles solely by reference to the quantity of good-1 consumed and set the taste adjusted price of good 1 to dominate the impact of revealed preference violations in the unadjusted data set. The choice set $\{\widetilde{\mathbf{p}}_t, \mathbf{q}_t\}_{t=1,...,T}$ then passes GARP. Given the equivalence of GARP and a non-empty feasible set to the standard Afriat inequalities, this result then implies that for for any element of the rationalising price set, $\{\widetilde{\mathbf{p}}_t\}_{t=1,...,T}$, there exist numbers $\{v_t\}_{t=1,...,T}$ and $\{\lambda_t\}_{t=1,...,T}$ with $\lambda_t > 0$ such that the following inequalities hold.

$$v_s - v_t \leq \lambda_t \widetilde{\mathbf{p}}_t'(\mathbf{q}_s - \mathbf{q}_t)$$

An element of the set of choice rationalising taste shifters associated with $\{\widetilde{\mathbf{p}}_t\}_{t=1,...,T}$ can then be constructed as:

$$\alpha_t^1 = \lambda_t(p_t^1 - \widetilde{p}_t^1)$$

for $t = 1, ..., T$. The fact that the set of rationalising $\widetilde{p}_t^1$ is unbounded above implies that the taste modification to virtual prices, or equivalently, the change in the marginal willingness to pay for good-1, α_t^1/λ_t, is unbounded below.■

Appendix B: Data

This section provides further details on the data used in our analysis. In the first part of our analysis, the nondurable aggregate is the union of the nondurables and alcohol groups below.

The tobacco, nondurable and alcohol good aggregates are constructed from the following subgroups.

1. **Tobacco**: Cigarettes; Other Tobacco.
2. **Nondurables**: Bread; Cereals; Biscuits; Beef; Lamb; Pork; Bacon; Poultry; Fish; Butter; Oils; Cheese; Eggs; Fresh Milk; Processed Milk; Tea; Coffee; Soft Drinks; Sugar; Sweets; Potatoes; Other Vegetables; Fruit; Other Food; Canteen; Other Snacks; Coal; Electric; Gas; Oil; Household consumables; Pet Care; Postage; Telephone; Domestic Services; Chemical Products; Petrol; Rail Fares; Bus Fares; Other Travel; Toys; Books; Entertainment.
3. **Alcohol**: Beer; Wine; Spirits.

Table B.1 gives the average and SMP expenditure levels for both education cohorts for the period considered and the number of observations per cohort in each year. The 5% and 95% confidence intervals on the SMP Expenditure levels, calculated by running the quantity estimation sequence on 1000 random samples drawn with replacement, are also given. Table B.2a. and B.2b. give the mean, median and SMP-path budget shares for tobacco in each period for each education group. The 95% confidence interval on the SMP budget shares are also given, calculated by randomly drawing observations with replacement within each education-time cell 1000 times and estimating quantile demands on each resampled set of observations at the SMP expenditure level.

Table B.1 Expenditure Levels

Time	Low Ed		High Ed	
	Mean Ex.	SMP Ex.	Mean Ex.	SMP Ex.
	(st. dev)	(95% CI)	(st. dev)	(95% CI)
1980	110.3	98.4	131.2	115.0
	(59.2)	(95,120)	(72.5)	(111,120)
1982	138.7	123.5	169.2	145.3
	(95.6)	(120,131)	(97.6)	(141,152)
1984	150.5	138.0	192.1	162.0
	(80.8)	(135,147)	(115.5)	(157,169)
1986	192.2	152.5	246.0	176.4
	(123.1)	(149,162)	(198.0)	(171,184)
1988	227.6	166.4	294.8	195.5
	(143.9)	(163,177)	(206.2)	(189, 204)
1990	283.8	192.9	371.6	229.1
	(189.0)	(187, 205)	(237.6)	(222, 239)
1992	307.8	222.9	432.6	278.8
	(187.0)	(220, 238)	(340.9)	(270, 290)
1994	307.4	245.9	420.7	288.9
	(206.6)	(242, 262)	(255.2)	(281, 302)
1996	327.5)	260.5	446.8	311.9
	(208.2)	(257, 278)	(278.4)	(302,325)
1998	346.5	287.9	501.2	343.0
	(220.8)	(283, 307)	(337.5)	(332,357)
2000	380.0	314.5	524.3	377.1
	(246.0)	(307, 333)	(354.6)	(364,390)

Table B.2a. Budget Share for Tobacco: Low Education

Time	Mean (st. dev)	Median	SMP 0.55 (95% CI)	SMP 0.65 (95% CI)	SMP 0.75 (95% CI)
1980	0.0762	0.0679	0.0667	0.0831	0.1070
	(0.07)		(0.06,0.09)	(0.08,0.11)	(0.11,0.13)
1982	0.0750	0.0565	0.0411	0.0766	0.0983
	(0.08)		(0.03,0.07)	(0.07,0.10)	(0.10,0.12)
1984	0.0765	0.0552	0.0481	0.0764	0.1074
	(0.08)		(0.04,0.07)	(0.07,0.10)	(0.10,0.14)
1986	0.0781	0.0563	0.0675	0.0947	0.1247
	(0.08)		(0.06,0.09)	(0.09,0.12)	(0.12,0.15)
1988	0.0678	0.0484	0.0600	0.0845	0.1022
	(0.08)		(0.05,0.08)	(0.08,0.1)	(0.10,0.13)
1990	0.0610	0.0436	0.0594	0.0813	0.1069
	(0.08)		(0.05,0.08)	(0.07,0.1)	(0.10,0.13)
1992	0.0603	0.0343	0.0468	0.0800	0.1165
	(0.08)		(0.03,0.08)	(0.08,0.11)	(0.10,0.13)
1994	0.0724	0.0309	0.0431	0.0872	0.1189
	(0.09)		(0.03,0.08)	(0.07,0.11)	(0.11,0.15)
1996	0.0778	0.0344	0.0433	0.0909	0.1256
	(0.09)		(0.03,0.07)	(0.07,0.12)	(0.12,0.16)
1998	0.0769	0.0132	0.0274	0.0673	0.1172
	(0.10)		(0.02,0.06)	(0.05,0.11)	(0.11,0.16)
2000	0.0748	0.0097	0.0295	0.0666	0.0938
	(0.10)		(0.01,0.05)	(0.04,0.10)	(0.09,0.13)

Table B.2b. Budget Share for Tobacco: High Education

Time	Mean (st. dev)	Median	SMP 0.55 (95% CI)	SMP 0.65 (95% CI)	SMP 0.75 (95% CI)
1980	0.0445	0.0029	0.0133	0.0363	0.708
	(0.06)		(0.01,0.03)	(0.03,0.06)	(0.06,0.09)
1982	0.0375	0.0000	0.0082	0.0337	0.0581
	(0.06)		(0.00,0.02)	(0.02,0.05)	(0.05,0.08)
1984	0.0392	0.0000	0.0017	0.0055	0.0491
	(0.06)		(0.00,0.01)	(0.00,0.04)	(0.04,0.08)
1986	0.0381	0.0000	0.0075	0.0300	0.0578
	(0.06)		(0.00,0.02)	(0.02,0.05)	(0.05,0.08)
1988	0.0330	0.0000	0.0000	0.0109	0.0512
	(0.06)		(0.00,0.01)	(0.00,0.03)	(0.04,0.07)
1990	0.0323	0.0000	0.0023	0.0128	0.0468
	(0.06)		(0.00,0.02)	(0.01,0.04)	(0.05,0.09)
1992	0.0320	0.0000	0.0039	0.0193	0.0391
	(0.06)		(0.00,0.01)	(0.01,0.03)	(0.03,0.06)
1994	0.0299	0.0000	0.0000	0.0174	0.0482
	(0.06)		(0.00,0.01)	(0.00,0.03)	(0.04,0.07)
1996	0.0363	0.0000	0.0039	0.0202	0.0533
	(0.07)		(0.00,0.01)	(0.00,0.04)	(0.05,0.07)
1998	0.0320	0.0000	0.0000	0.0048	0.0429
	(0.07)		(0.00,0.01)	(0.00,0.02)	(0.03,0.06)
2000	0.0314	0.0000	0.0000	0.0060	0.0252
	(0.07)		(0.00,0.01)	(0.00,0.01)	(0.01,0.05)