cemmap

centre for microdata methods and practice

A nonparametric analysis of habits models

Ian Crawford

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP30/07

ESRC ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

A Nonparametric Analysis of Habits Models

Ian Crawford
Oxford University & IFS

October 2007

Abstract

This paper presents a nonparametric analysis of the canonical habits model. The approach is based on the combinatorial/revealed preference framework of Samuelson (1948), Houthakker (1950), Afriat (1967) and Varian (1982) and the extenstion and application of these ideas to intertemporal models in Browning (1989). It provides a simple finitely computable test of the model which does not require a parameterisation of the underlying (hypothesised) preferences. It also yields set identification of important features of the canonical habits model including the consumer's rate of time preference and the welfare effects of habit-formation. The ideas presented are illustrated using Spanish panel data.

Acknowledgements I am grateful to Martin Browning, Mette Christensen, Lars Nesheim, Stephen Nickell, Frans Spinnewyn and seminar particpants at Cambridge, Leuven, Oxford and York for helpful comments and to Neil Rickman who suggested looking at this topic in the first place. The usual disclaimer applies. Email: ian.crawford@economics.ox.ac.uk

1 Introduction

Habit-formation models have been used profitably to analyse a wide variety of micro and macroeconomic phenomena. Micro applications have, for example, included Becker and Murphy's (1988) classic study of the price-responsiveness of addictive activities, Meghir and Weber's (1996) work on intertemporal nonseparabilities and liquidity constraints and the explanation of asset-pricing anomalies such as the equity premium puzzle (Abel (1990), Campbell and Cochrane (1999), Constantinides (1990)). Macro-orientated studies have used habit-formation models to improve the ability of business cycle models to explain movements in asset prices (Jermann (1998), Boldrin *et al* (2001)), to investigate the idea that economic growth may cause savings rather than the other way around (Carroll *et al* (2000)) and to explain the finding that aggregate spending tends to have a gradual hump-shaped response to various shocks (Fuhrer (2000)).

Compared to the standard discounted utility model the principal feature of the habit-formation model is the relaxation of consumption independence. The implication of consumption independence in the standard discounted utility model is that preferences over consumption in one period are unaffected by consumption in another. Samuelson (1952) was evidently sceptical about this feature and noted that,

> "the amount of wine I drank yesterday and will drink tomorrow can be expected to have effects upon my today's indifference slope between wine and milk".

Similarly Koopmans (1960), who provided an axiomatic derivation of the discounted utility model, remarked that

"One cannot claim a high degree of realism for [consumption independence], because there is no clear reason why complementarity of goods could not extend over more than one time period".

This, in effect, is an argument against the time-separability of preferences in the discounted utility framework. The leading example of the kind of behaviour which will give rise to nonseparabilities is, perhaps, habit formation (Duesenberry (1952), Pollak (1970), Ryder and Heal (1973), Spinnewyn (1981), and others). In habit formation models[1] the consumption vector is typically partitioned into a group of consumption goods and a group of addictive/habit-forming goods, and the period-t instantaneous utility function is allowed to depend on both current consumption and lagged consumption of the habit-forming goods. The effects of habit formation on preferences over consumption profiles and consequent behaviour can be fairly general: i.e. depending on how much one has already consumed and whether current consumption increases or decreases future utility, habit formation can lead to preferences for increasing, decreasing or even non-monotonic consumption profiles.

Thus far the empirical literature on habits models has been, to my knowledge, entirely parametric. That is to say based on parametrisations of the Euler equation or consumption function (typically) and of the hypothesised underlying preference structure. The problem which an approach based on a statistical fit of a parametric model to the data is that any test of the theory must be a joint one conflating a test of the hypothesis of interest with a joint hypothesis regarding the functional form of the model plus a number of statistical/econometric auxiliary hypotheses. It therefore does not follow that a rejection of the econometric model necessarily implies a rejection of the hypothesis of interest. Similarly any model-based empirical identification of discount rates or the welfare effects of habit-formation equally rests on the parametrisation.

This paper asks: *what are the nonparametric empirical implications of the habits model?* In particular; are there restrictions involving only data on observables which can allow us to test the model's empirical validity and (granted this) to recover its features, i.e. welfare measures and discount rates? The path which is taken in this paper is based on the nonparametric-revealed preference approach developed in Samuelson (1948), Houthakker (1950), Afriat (1967) and Varian (1982) and the extension of these ideas to the life-cycle/permanent income version of the discounted utility model developed by Browning (1989) who showed how the constancy of the marginal utility of income across periods can be used to generate finite linear-programming type restrictions which only involve data on observables: discounted prices and quantities. These provide a simple yes/no test of exact, error free, consistency between the data and the theory. Despite the strength of the assumptions underlying the life cycle-permanent income model, Browning (1989) found that there were very strong theory-coherent regularities in the post war aggregate data sets for Canada, the US and the UK.

The benefits of this style of approach are well known[2]: it is designed to work using finite (even small) datasets, it requires only data on observables and it avoids the need to fit parametric (or indeed nonparametric) statistical models to the data. The downside is that empirical identification is necessarily weakened; although to the extent that precise identification might flow from parametric/statistical assumptions this may be no bad thing.

No nonparametric test (in the sense of Afriat (1967), Varian (1982) and Browning (1989)) of the habits model has yet been proposed or implemented. Whilst Kubler (2004) shows that nonparametric testing of very general intertemporal choice models is not possible, the canonical habits model is rather special: it is additive and breaks intertemporal separability in a fairly specific manner. This paper asks whether the habits model is nonparametrically testable on the basis of

[1]It is worth noting that habits models bear a great many formal parallels with models of consumer durables (habit-forming goods being goods which are psychologically durable rather than - or as well as - physically durable), models of adaptive consumer learning, reference-point models, models incorporating rational anticipation and models of consumer behaviour under rationing.

[2]See for example the motivation given in Varian (1982).

observables. It is shown that it is, and also that the proposed test is a rather straightforward one.

The plan of this paper is as follows. Section 2 presents necessary and sufficient empirical conditions for the habits model and describes the implementation of the test. Section 3 sets out the way in which measurement errors might be accommodated. Section 4 discusses the relationship between the test described here and other nonparametric tests in the literature (e.g. GARP and Browning's (1989) test of the life-cycle/permanent income model). Section 5 presents the empirical identification results for the model. Section 6 applies these ideas to a Spanish Panel dataset. Section 7 concludes.

2 Necessary and sufficient conditions for the habits model

Suppose that there are T observations indexed[3] $t = 1, ..., T$ on a consumer's demands over time $\{\mathbf{q}_t\}$ and the corresponding prices $\{\mathbf{p}_t\}$ and interest rate $\{i_t\}$ which they were facing. Let the commodity vector be partitioned into a group of consumption goods $\mathbf{q}_t^c$ and a vector of habit-forming goods $\mathbf{q}_t^a$ such that $\mathbf{q}_t = [\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}]'$. To develop the main ideas without the loss of a great deal of generality, the discussion will focus on the simplest case in which the effects of lagged consumption of the addictive goods only persist for one period. This is precisely the type of model considered in, for example, Becker, Grossman and Murphy (1994) *inter alia*, and it is used here to introduce notation and to fix ideas before considering the extension to a general lag structure. The discussion of this extension (which is straightforward) is postponed until the end of this section.

The first question is whether it is possible to find necessary and sufficient empirical conditions on the observable discounted price and quantity data under which these data are consistent with the short-memory habits model. To this end, consistency between the habits model and the data is defined as follows.

Definition 1. The time series of the interest rate, prices and quantities $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\tau}$ satisfies the one-lag habits model if there exists a concave, strictly increasing (utility) function $u(.)$ and positive constants λ and β such that for all $t \in \tau$

$$\beta^{t-1}\mathbf{D}_{\mathbf{q}_t^c} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) = \lambda \boldsymbol{\rho}_t^c$$

$$\beta^{t-1}\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) + \beta^t \mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_{t+1}^c, \mathbf{q}_{t+1}^a, \mathbf{q}_t^a\right) = \lambda \boldsymbol{\rho}_t^a$$

where $\rho_t^i = p_t^i / \prod_{s=2}^{s=t}(1 + i_s)$ denotes discounted prices.

This says that the data are consistent with the theory if there exists a well-behaved instantaneous utility function (defined over the consumption goods and the habit-forming goods plus the one-period lag of the habit-forming goods), the derivatives of which satisfy the first order conditions of optimising behaviour. If such a utility function exists, and we know what it is, then it means that we can simply plug it into the habits model, solve the model and precisely replicate the observed demand choices of the consumer. To put it another way, the theory and the data are consistent if there exists a well-behaved and suitably-conditioned utility function which can provide perfect within-sample fit of the consumption/demand data.

From Definition 1 is it clear that the first order conditions for the consumption goods are identical to those of the standard life cycle model. Those for the habit-forming goods are a little more complex because the current consumption of the habit forming goods affects future utility as well as current utility. Nevertheless, as first pointed out by Spinnewyn (1981), these condition can

[3] Denote the index set by $\tau = \{1, ..., T\}$.

be transformed into a form which is analogous to a no-habits model by defining suitable shadow discounted prices which summarise these welfare effects:

$$\boldsymbol{\rho}_t^{a,0} = \frac{\beta^{t-1}\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right)}{\lambda} \tag{1}$$

$$\boldsymbol{\rho}_t^{a,1} = \frac{\beta^{t-1}\mathbf{D}_{\mathbf{q}_{t-1}^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right)}{\lambda} \tag{2}$$

Expression (1) is the shadow discounted price of current consumption and measures the willingness-to-pay for current consumption of the habit-forming goods. Expression (2) is the shadow discounted price of past consumption and measures the willingness-to-pay for past consumption of the habit-forming goods. It is worth noting that the shadow discounted price of current consumption can be interpreted as the (observed) discounted price adjusted to account for the future welfare effects of current decisions. That is

$$\boldsymbol{\rho}_t^{a,0} = \boldsymbol{\rho}_t^a - \frac{\beta^t\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_{t+1}^c, \mathbf{q}_{t+1}^a, \mathbf{q}_t^a\right)}{\lambda} \tag{3}$$

The habits model also entails an intertemporal dependence between the shadow discounted prices given by

$$\boldsymbol{\rho}_t^a = \boldsymbol{\rho}_t^{a,0} + \boldsymbol{\rho}_{t+1}^{a,1} \tag{4}$$

Note that for *a priori* harmfully addictive goods past consumption reduces current utility and current consumption reduces future utility so $\boldsymbol{\rho}_t^{a,0} \geq \boldsymbol{\rho}_t^a$ and $\boldsymbol{\rho}_{t+1}^{a,1} \leq \mathbf{0}$. The empirical/behavioural implications of the short memory habits model are therefore driven by: *(i)* links between the derivatives of discounted utility with respect to future and past consumption of the habit-forming goods and the (unobservable) shadow discounted prices, and *(ii)* intertemporal links between the (unobservable) shadow discounted prices and the (observable) discounted prices. The aim then, is to turn these insights into testable empirical conditions involving only observables. The following result for the short-memory habits model can now be given:

Theorem 1. The following statements are equivalent:
(T) The time series of the interest rate, prices and quantities $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\tau}$ satisfies the one-lag habits model.
(R) There exist shadow discounted prices $\{\boldsymbol{\rho}_t^{a,r}\}_{t\in\tau}^{r=0,1}$ and a positive constant β such that

$$0 \leq \sum_{\forall s,t\,\in\sigma} \boldsymbol{\pi}_s'\left(\mathbf{x}_t - \mathbf{x}_s\right) \quad \forall\sigma\subseteq\tau \tag{R1}$$

$$0 = \boldsymbol{\rho}_t^a - \boldsymbol{\rho}_t^{a,0} - \boldsymbol{\rho}_{t+1}^{a,1} \quad \forall \quad t, t+1\in\tau \tag{R2}$$

where $\mathbf{x}_t = \left[\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}, \mathbf{q}_{t-1}^{a\prime}\right]'$ and $\boldsymbol{\pi}_t = \frac{1}{\beta^{t-1}}\left[\boldsymbol{\rho}_t^{c\prime}, \boldsymbol{\rho}_t^{a,0\prime}, \boldsymbol{\rho}_t^{a,1\prime}\right]'$ and $t\in\tau$, $\tau = \{2, ..., T\}$.

Proof. See the Appendix.■

Theorem 1 is an equivalence result. It says that if we can find suitable shadow prices and a discount rate such that restrictions (R1) and (R2) hold, then the data are consistent with the theory and there does indeed exist a well-behaved utility function which gives perfect within-sample rationalisation of the model. Conversely if we cannot find such shadow discounted prices and a discount rate then there does not exist any theory-consistent utility representation. Restriction (R1) is a cyclical monotonicity condition[4] which is an implication of the concavity of the instantaneous utility function and the constant marginal utility of lifetime wealth. This condition involves the shadow discounted prices discussed above. Restriction (R2) is the intertemporal link between the shadow prices.

[4] Rockafellar, (1970, Theorem 24.8)

The empirical test is thus a question of searching for shadow price vectors and a discount rate which satisfies the restriction in (R). These restrictions are non-linear in unknowns and look forbidding but are, in fact, computationally quite straightforward. The important feature to note is that, conditional on the discount rate, the restrictions are linear in unknowns. This means that, for any choice of discount rate, we can readily check for the existence or non-existence of feasible shadow prices in a finite number of steps using phase one of a linear programme. The issue is then simply one of conducting an arbitrarily fine one-dimensional grid search over a sensible range[5] for the discount rate and running a linear programming problem at each node[6].

To end this section consider a more general model in which consumption of the habit-forming goods persists for R periods the instantaneous utility function is given by

$$u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a, \mathbf{q}_{t-2}^a, ..., \mathbf{q}_{t-R}^a\right) \tag{5}$$

The definition of what it means for data to be consistent with the R-lag model and the corresponding necessary and sufficient conditions for theoretical consistency are given in the appendix (Definition R and Theorem R). Both are natural extensions of Definition 1 and Theorem 1. Once more the restrictions come in the form of a cyclical monotonicity condition and an intertemporal condition linking the shadow and spot prices of the habit-forming goods. However in this more general model the lag lengths involved in the consumption vectors are longer and the intertemporal links between shadow prices extend further. Otherwise the restrictions are multi-period analogies of those in Theorem 1. Once more the conditions are conditionally linear given a choice of discount rate and so can be checked by grid search on the discount rate with a linear programming step at each node.

Note that whilst one can test data against the habits model with R lags, there is an empirical limit for the number of lags at $R = T$. Indeed if the lag length equals the number of observations on the consumer then the habits model is untestable/unrejectable because, in essence, any empirically-relevant degree of time separability has been lost and the result of Kubler (2004) applies. Another way to think about this is that when the lag length equals the number of observations, one only has a single complete observation on the consumer - too few to check the dynamic consistency of a consumer's behaviour. Of course, when $R \geq T$ there are no complete observations at all.

3 Allowing for errors

The conditions in Theorems 1 and R are, like all non-parametric/revealed preference type tests, rather exacting in the sense that if either the consumer's or the data collector's "hand trembles" then the data may be inconsistent with the model even if the deviations induced are very small. One might be particularly concerned that measurement error could induce violations of the conditions even though the underlying true data are theory-consistent.

Let D^0 denote the observed dataset $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\tau}$ and let $\Delta(R)$ denote the set of all such datasets which are consistent with the R-lag model

$$\Delta(R) = \{D : D \text{ is consistent with the R-lag model}\} \tag{6}$$

Then a violation of the empirical conditions for the observed data simply means that the observed data lie outside the theoretically consistent range

$$D^0 \notin \Delta(R) \tag{7}$$

[5]Since $\beta = 1/(1+\delta)$, where δ is the consumer's rate of time preference $\delta \in [0, \infty] \Rightarrow \beta \in [0, 1]$.

[6]Note that with T data points we can only test the one-lag model using $T-1$ observations from $t = 2$ to T as one observation is lost per-lag as we construct the $\mathbf{x}$ vectors. Also note that whilst it may appear at first blush that there are $rows(\boldsymbol{\rho}_t^a)\times 2(T-1)+1$ free parameters in the problem ($\boldsymbol{\rho}_t^{a,0}$ and $\boldsymbol{\rho}_t^{a,1}$ for $t = 2, ..., T$ plus β), in fact the intertemporal restrictions mean that there are only $rows(\mathbf{p}_t^a)\times T+1$ because, given $\boldsymbol{\rho}_t^a$ is observed (R2) implies that once $\boldsymbol{\rho}_t^{a,0}$is found, then $\boldsymbol{\rho}_{t+1}^{a,1}$ is tied down as well.

However, suppose that the data are contaminated by measurement error. Specifically suppose that the relationship between the true data D^* and the observed data is

$$D^* = D^0 + E \tag{8}$$

where $E = \{v_t; \mathbf{e}_t^c, \mathbf{e}_t^a; \mathbf{u}_t^c, \mathbf{u}_t^a\}_{t\in\tau}$ represents measurement error which is classical by assumption[7]. Thus $D^* = \{i_t + v_t; \mathbf{p}_t^c + \mathbf{e}_t^c, \mathbf{p}_t^a + \mathbf{e}_t^a; \mathbf{q}_t^c + \mathbf{u}_t^c, \mathbf{q}_t^a + \mathbf{u}_t^a\}_{t\in\tau}$. In this case a test statistic for the null hypothesis that the true data satisfy the model can be based on the loss function

$$L(E) = \frac{vec(E)' \, vec(E)}{\sigma^2} \tag{9}$$

where σ^2 is the variance of the measurement error. This is distributed as a $\chi^2_{2K(T-1)+T}$ and the null hypothesis that the true data satisfy the model would be rejected if the test statistic exceeded C_α, the critical value at the α significance level of the chi-squared distribution. Since the true data are unobserved one can instead compute the minimum perturbation to the data such that the perturbed data satisfy the model, and use the calculated errors as the basis for making conservative inferences. Of course the variance of the measurement errors is typically unknown but Varian (1990) suggests calculating how big it would need to be in order to reject the null and then comparing this to one's prior beliefs on the likely size of these errors. Alternatively one may be able to estimate it from a parametric or nonparametric fit of the data, or from other data sources. This provides a basis for analysing the model in the presence of measurement errors.

4 The relationship with other nonparametric tests

A natural question concerns how the test proposed in the previous section relate to other nonparametric integrability tests? Specifically, how does it relate to Browning's (1989) test of the life-cycle model/strong rational expectations hypothesis (LCM) and the (Afriat (1967), Varian (1982)) Generalised Axiom of Revealed Preference (GARP). The first result to note is that whilst the test of the habits model does not imply LCM, and neither is it implied by the LCM conditions, nevertheless the habits model nests the life cycle model in the following sense.

Theorem 2. If the data $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t=1,\ldots,T}$ satisfies the R-lag model with $\boldsymbol{\rho}_t^{a,r} = \mathbf{0}$ for all $t \in \tau$ and $r \geq 1$ then the data also satisfies LCM.

The life cycle model/strong rational expectations hypothesis can therefore be regarded as a special case of the habits model in which discounted willingness to pay for past consumption is always zero. The test proposed here can, therefore, be easily adapted to provide a test of the life cycle model by adding the constraints that $\boldsymbol{\rho}_t^{a,r} = \mathbf{0}$ for $r \geq 1$ to those in Theorems 1 and R, in which case the test becomes identical to that proposed in Browning (1989) (augmented to allow for time discounting which Browning does not explicitly consider). As a corollary note that this further restriction also provides the link between the integrability condition for the habits model and GARP.

Corollary 1. If the data $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\tau}$ satisfies the R-lag model with $\boldsymbol{\rho}_t^{a,r} = \mathbf{0}$ for all $t \in \tau$ and $r \geq 1$ then the data also satisfies GARP.

[7] See for example, Varian (1985).

The intuition is straightforward: if the data satisfy the conditions for the habits model with this added restriction on the shadow prices then habit formation is effectively ruled out and they then satisfy the conditions for the life-cycle model. GARP only requires within-period efficiency in expenditure, the life-cycle model however requires more; it requires the efficient within-period *and* between-period allocation of expenditure. The LCM condition is therefore over sufficient for GARP.

5 Identification

Given a dataset which is consistent with the habits model, the question then arises as to whether it might be possible to identify parts of the model. In general the restrictions in Theorems 1 and R will only provide set identification in the sense that, if there exist shadow prices and discount rates which are consistent with the habits model at all, then the constraints define a set of admissible values. Consider the discount factor; its identification set is given by

$$B(R) = \left\{\beta : \{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\tau} \text{ satisfies the R-lag model}\right\} \tag{10}$$

The set of theory and data-consistent discount factors is not convex, a fact which stems from the non-linear nature of the restrictions implied by the model. Since the empirical test proceeds by means of a grid search on β this means that the empirical identification interval for β has "gaps" in it both at nodes where the conditions may be rejected and also between nodes at untested values of β. These second class of gaps can be made arbitrarily small by choosing a finer grid.

The other elements of principal interest in the habits model are the willingness-to-pay measures which capture the welfare effects of habit formation:

$$\frac{\mathbf{D}_{\mathbf{q}_{t-r}^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a, ..., \mathbf{q}_{t-R}^a\right)}{\lambda} = \left\{\frac{\boldsymbol{\rho}_t^{a,r}}{\beta^{t-1}}\right\}^{r=1,...,R} \tag{11}$$

Again, given that the data are theory-consistent there will be a set of combinations of shadow prices and discount factors which will be admissible under the restrictions. The empirical procedure will return theory-consistent combinations of these parameters and the identification set for the willingness-to-pay measures will be given by

$$P(R) = \left\{\left\{\frac{\boldsymbol{\rho}_t^{a,r}}{\beta^{t-1}}\right\}^{r=1,...,R} : \{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\tau} \text{ satisfies the R-lag model; } \beta \in B(R)\right\} \tag{12}$$

Once more this identification set will be non-convex. The non-convexity of $P(R)$ stems from the non-convexity of $B(R)$. However, once again conditional on β the set of willingness-to-pay parameters is convex. The issue of how best to represent the set-identified objects is addressed in the next section.

6 An empirical implementation: the Spanish Continuous Family Expenditure Survey (ECPF)

6.1 Data

The data used here to investigate the empirical implementation of the ideas outlined above is the Spanish Continuous Family Expenditure Survey (the *Encuesta Continua de Presupuestos Familiares* - ECPF). The ECPF is a quarterly budget survey of Spanish households which interviews about 3,200 households every quarter. These households are randomly rotated at a rate of 12.5% each quarter. Thus it is possible to follow a participating household for up to eight consecutive quarters.

This dataset is a much studied survey which has often been used for the analysis of intertemporal models and particularly, latterly, the analysis of habits models (for example, Carrasco, Labeaga and López-Salido, (2005), Browning and Collado (2001, 2005)). The data used here are drawn from the years 1985 to 1997 and are the selected sub-sample of couples with and without children, in which the husband is in full-time employment in a non-agricultural activity and the wife is out of the labour force (this is to minimised the effects of nonseparabilities between consumption demands and leisure which the empirical appplication does not otherwise allow for). The dataset consists of 21866 observations on 3134 households. The data record household non-durable expenditures and these are disaggregated into 18 commodity groups (details are in the Appendix). The discounted price data are calculated from published prices aggregated to correspond to the expenditure categories and the average interest rate on consumer loans. Note that whilst in these data we observe expenditures rather than quantities this poses no problems for the application as it is straightforward to adapt the tests etc. outlined above[8].

In what follows the focus is on the consumption of cigarettes - one of the most studied addictive goods (see for example Chaloupka (1991), Becker, Grossman and Murphy (1994), Labeaga (1999), *inter alia*). Following these studies it is sensible to look at the special version of the habits model in which the habit-forming good is bad for the consumer. That is, the version of the model in which past consumption reduces current utility.

6.2 Results

6.2.1 Test results

The empirical results begin with the analysis of the comparative performance of the Varian (1982) test of GARP, the Browning (1989) test of the life-cycle/strong rational expectations hypothesis and the habits model with one and two lags respectively corresponding to one and two quarters. Each test is run *independently* for each household in the sample. The data across households is not pooled at any point. This therefore allows for complete heterogeneity, of unrestricted form across households with respect to (i) whether or not their behaviour is theory-consistent and (ii) the form of their preferences (if their behaviour is rationalisable at all). The results are given in Table 1.

Table 1: Pass Rates

Sample	GARP	Life-cycle	Habits (1)	Habits (2)
All ($N = 3134$)	94.19%	1.88%	-	-
Non-smokers (23.8%)	93.16 %	2.28%	-	-
Smokers (76.2%)	94.51%	1.76%	43.05%	47.43%

Table 1 shows the pass rates for each test for the full sample, and the sub-samples of smokers and non-smokers. The first column shows the pass rate for the GARP test. Recall that this tests for consistency between the data and the atemporal consumer choice model in which each period's budget is parametric. The results indicate a high level of agreement between the theory and model with about 94% of the sample satisfying GARP and no significant difference between the smokers and non-smokers. The atemporal model out-performs both the life-cycle model and the habits model by a significant degree. However, this is not surprising in because, as discussed above, the empirical

[8] For example, the cyclical monotonicity condition for the LCM becomes

$$\sum_{s,t\in\sigma}\left(\frac{\boldsymbol{\rho}_t}{\boldsymbol{\rho}_s}\right)'\mathbf{z}_s-\mathbf{z}_t$$

where $\frac{\boldsymbol{\rho}_t}{\mathbf{v}_s}$ denotes element-by-element division of the price vectors and $\mathbf{z}_t$ and $\mathbf{z}_s$ denote the vectors of commodity expenditures.

requirements of the static demand model are much less stringent than those of the intertemporal model.

The next column reports the results of Browning's (1989) test of the life cycle model. It is found, in contrast to Browning's study of aggregate data series, that the life-cycle model is heavily rejected in these microdata. Less than 2 percent of the sample satisfy the conditions required and whilst the pass rates for non-smokers are higher than those for smokers they are still low. It would appear that the data are generally inconsistent with the strong rational expectations version of the life-cycle model.

The last two columns show the pass rates for the habits model. Recall that the life-cycle model can be regarded as a special case of the habits model in which the welfare effects of past consumption are restricted to be zero. Habits models are therefore less restrictive than the life-cycle model and the more lags which are allowed, the less restrictive they progressively become (until the number of lags equals the number of observations at which point they provide no testable restrictions). The pass rates for the habits models should therefore be no worse than those for the life-cycle model. The results in the table show that this is indeed the case with 43 percent of smokers' behaviour rationalisable by the one-lag version of the model. Whilst the empirical results show a far better degree of agreement between the habits model and the data that between the life-cycle model and the data, the pass rates are far below those of the GARP test. One possibility may be that it is necessary to admit a longer lag structure. In the final column of Table 1 the lags are extended from one quarter to two. The increase in the pass rate is not enormous. This is probably because the addictive properties of tobacco are such that the addictive stock dissipates within six months.

To investigate whether or not a household's consistency with one of these models is correlated with observables, pass/fail indicators for each household for each model (atemporal, life-cycle and both habits model) were regressed on a number of standard observable household characteristics[9]. The pseudo R^2 of the probits were all around the 1% to 2% level and almost none of the coefficients were individually significant. It appears that whether or not a household's behaviour is likely to be rationalisable with theory is not predictable on the basis of standard observables.

6.2.2 Allowing for Measurement Errors

For smoking households whose behaviour violates the habits model it is possible, using the ideas outlined in section 3, to perturb their observed behaviour so that they satisfy the model. As discussed above, the data we observe are expenditures on commodity groups which are collected in the ECPF, and corresponding price indices and a consumer interest rate series published by the *Instituto Nacional de Estadística*. Given that individual expenditures are recorded in the survey, but the prices and the interest rate are not, but rather are national time series data, is seems most likely that, if there is any measurement error, most of it is in the discounted prices. We are not aware of any specific studies about price dispersion (or variation in the interest rates at which different households may borrow) in Spain but there is plenty of evidence for it in the UK[10] and there is no reason to expect that Spain is much different. In view of this we perturb the discounted price data for each violating households by the minimum (Euclidean) distance necessary such that they then satisfy the model.

The results are hard to interpret directly since distance metrics like this depend on the units involved: in the abscence of an estimate of the variance of the measurement error it is hard to tell whether the necessary perturbation is big. To help with interpretation the distances have been translated into $R^2 \in [0, 1]$ values based on how much the perturbations contribute to the fitting of the nearest theory-consistent values compared to the observed discounted price[11]. If the observed

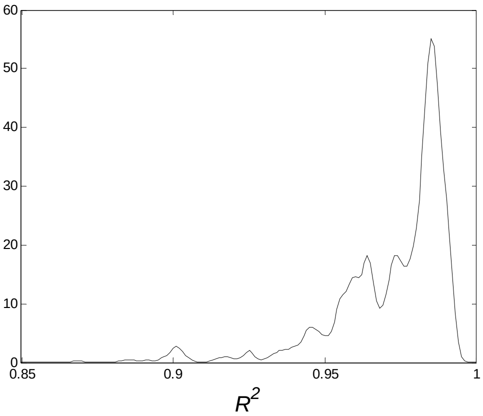
[9] These were {age of the head of household, the number of children in the household and dummy variables for highest educational qualification=university degree, highest educational qualification=high school, head's occupation = professional/managerial, head's occupation = skilled, homeowner, renter, drinker}.

[10] See for example Griffith and Leicester (2006).

[11] The model is $\rho_t^{i*} = \rho_t^i + e_t^i$ - see section 3.

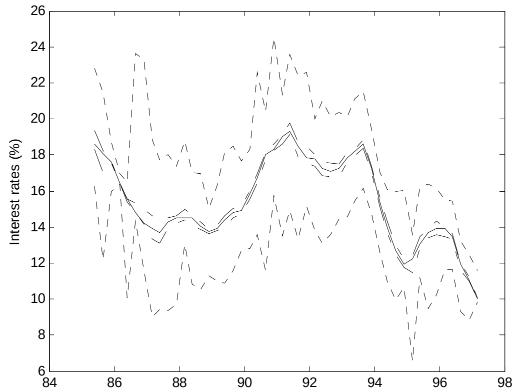
discounted prices did satisfy the model the minimum perturbation required would be zero and the R^2 would be 1. To the extent that the data violates the model and larger perturbations are required then $R^2 \to 0$. The interpretation of an R^2 around 1 is that the price data are "close" to passing. The calculation is carried out independently for each household which violates the model and for each value of the discount rate in the grid search. Figure 1 illustrates the density of the distribution of R^2 values. It shows a right-skewed distribution with 98% of the R^2 values greater than 0.9. The behaviour of many households appears to be close (on this measure) to rationalisable by the model - the required perturbation to the discounted prices is generally modest.

FIGURE 1: The density of the distribution of $R^2 : \rho_t^{i*} = \rho_t^i + e_t^i$



One interesting exercise is to take the perturbed (and now theoretically-consistent) discounted prices and to recover from them the implied interest rate which would rationalise the data. This allows for the fact that the aggregate discounted price data uses the average rate of interest on consumer loans whilst in reality different households might vary widely in the cost of borrowing which they face. Figure 2 shows the time series of quantiles of the interest rate which emerges from this.

FIGURE 2: Rationalisable interest rates



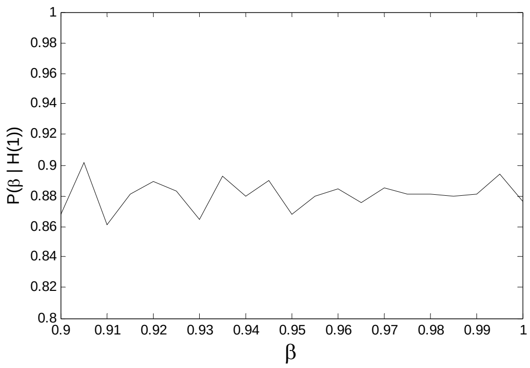
The solid line in the middle of the figure is the average consumer interest rate. The dashed lines surrounding this solid line show quantiles of the distribution of interest rates which are theory-consistent (i.e those which are implied by the minimum perturbation to the discounted price data needed to rationalise the data with the model). The outer bounds are the 2.5th and 97.5th percentiles of the theory-consistent distribution; i.e. 95% of the distribution of theory-consistent rates are within these bounds which can be seen to be fairly wide - typically ±4 to ±5 percentage points of the average rate. However, the distribution is quite "peaky": the inner dashed lines show the 90-10 interval and indicate that 80% of the data are very close (less than about ±0.5 percentage points) to the average rate.

To summarise: it appears that only reasonably modest adjustments to the prices and interest rates faced by individual households are required to rationalise the data.

6.2.3 Identification results

Discount rates In the tests conducted above the grid search with respect to the discount rate was carried out, for every individual household, over the range $[0.9, 1]$ using grid points spaced at 0.005. As noted in section 5 the identification region for the discount rate is non-convex and may contain gaps at values which were inconsistent with observed behaviour and the model. For each household which satisfies the restrictions implied by the one-lag model a set of admissible discount rates were recovered. Figure 3 illustrates the probability that each discount rate in the range examined is rationalisable conditional on *some* discount rate being appropriate. To put it another way the height of the line records the proportion of times each value of the discount rate was rationalisable over the sample of rationalisable household. The line is more or less flat. One way of interpreting this is that, if a household's behaviour is theory consistent at all, then there is little to choose between different discount rates (at least over the range studied).

FIGURE 3: Theory-consistent consumer discount rates



The welfare effects of habit-formation There are two main practical difficulties with set-identification results discussed above. The first is that when there is more than one or at most two, features of the model which are of interest, finding a suitable way of illustrating/representing the set becomes hard. It would be convenient to pick a point within the set which is in some way meaningful, or representative. This leads to the second problem. This is that whilst the economic theory and the empirical restrictions they imply tie down the set of theoretically feasible parameter values (or else indicate that such a set is empty) this set exhausts the (nonparametric) empirical implications of the theory. The theory cannot, therefore, provide any guidance as to which particular elements within the set of feasible alternatives are "best". The familiar idea of "fit" cannot help because for all feasible values there exists a corresponding well-behaved utility function which exactly fits the data.

To make further progress one needs to supply a loss function and provide a motivation for it or to try to recover bounds on the values of the welfare measures. This is computationally reasonably straightforward: given a discount rate for which there exist feasible shadow price values (determined by the initial grid-search), using equations (3) and(4) one can see that one can simply minimise $\left(\boldsymbol{\rho}_t^a - \boldsymbol{\rho}_t^{a,0}\right)$ for each t in phase two of the linear programming problem and thereby find the minimum values of $\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_{t+1}^c, \mathbf{q}_{t+1}^a, \mathbf{q}_t^a\right)/\lambda$ in each period. This will provide bounds on the welfare effects of past consumption on utility in each period. Note that these bounds will generally be mutually inconsistent as they will have been found by independent minimisation (for all welfare effects simultaneously to achieve their bounds may not be a feasible solution). Nevertheless these bounds are useful and robust pieces of information about the consumer's preferences and allow us to make statements of the kind: the absolute value of the welfare effects of habit-formation are *at least* x.

FIGURE 4: The density of the distribution of welfare effects

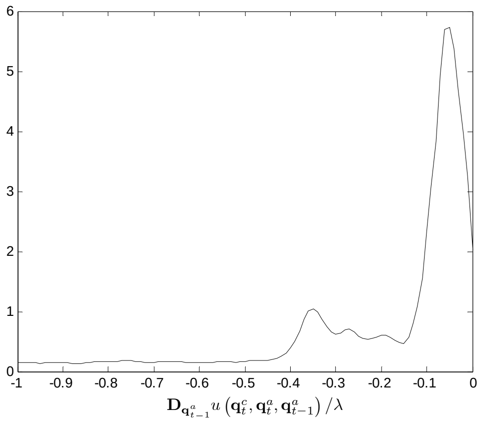


Figure 3 shows the density of the distribution of welfare effects for households whose behaviour is consistent with the short memory model. There is a long lower tail of the distribution which has been truncated in the figure. A mode is evident around -0.075 with another smaller mass point around -0.35. The median sample value is -0.1495. This is a standard willingness to pay measure (marginal utility of the object of interest relative to the marginal utility of income). The interpretation is therefore that on average amongst households whose behaviour is rationalisable with the model, the negative marginal effects of tobacco habit-formation are at least about one seventh of those of a marginal decrease in income.

7 Conclusions

This paper has derived general combinatorial/nonparametric conditions for the canonical habits model. The test proposed in the paper is shown to be computationally straightforward and set identification results for certain features of the model are available. The ideas outlined have been applied to a Spanish panel dataset in which it appears that, while the life-cycle model is very heavily rejected, for smokers at least the addition of habit formation to the discounted utility model improves the rationalisability of the data considerably. For households who were non-rationalisable it was shown that rather modest adjustments to prices and interest rates were sufficient to bring their expenditure behaviour in line with the model. For households whose behaviour is captured by the habits model set identification of discount rates and the welfare effects of tobacco-habit formation was illustrated.

Appendix

A. Proofs

Proof of Theorem 1.

$(T) \Rightarrow (R)$: Definition 1 and the definitions of the shadow discounted prices in (**??**) and (3) imply (4) which is restriction (R2). Together they imply

$$\mathbf{D}u\left(\mathbf{x}_t\right)' = \lambda \boldsymbol{\pi}_t' \tag{P1}$$

where $\boldsymbol{\pi}_t = 1/\left(\beta^{t-1}\right)\left[\boldsymbol{\rho}_t^{c\prime}, \boldsymbol{\rho}_t^{a,0\prime}, \boldsymbol{\rho}_t^{a,1\prime}\right]$ and $\mathbf{x}_t = \left[\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}, \mathbf{q}_{t-1}^{a\prime}\right]'$. The concavity of the instantaneous utility function $u\left(\mathbf{x}_t\right)$ means

$$u\left(\mathbf{x}_s\right) - u\left(\mathbf{x}_t\right) \leq +\mathbf{D}u\left(\mathbf{x}_t\right)'\left(\mathbf{x}_s - \mathbf{x}_t\right) \qquad \forall\ t, s \in \tau \tag{P2}$$

Therefore concavity (P2) and optimising behaviour (P1) together imply that

$$u\left(\mathbf{x}_s\right) \leq u\left(\mathbf{x}_t\right) + \lambda \boldsymbol{\pi}_t'\left(\mathbf{x}_s - \mathbf{x}_t\right) \qquad \forall\ t, s \in \tau \tag{P3}$$

Now consider *any* subset of observations from τ and denote this subset by σ. Then summing across all observations within the subset gives

$$0 \leq \sum_{\forall s,t\ \in \sigma} \boldsymbol{\pi}_s'\left(\mathbf{x}_t - \mathbf{x}_s\right) \qquad \forall \sigma \subseteq \tau \tag{P4}$$

which is restriction (R1).

$(R) \Rightarrow (T)$: Restriction (R1) is a cyclical monotonicity condition (Rockafellar, 1970, Theorem 24.8). Cyclical monotonicity for the data $\left\{\boldsymbol{\pi}_t, \mathbf{x}_t\right\}_{t\in\tau}$ and the definition of $\boldsymbol{\pi}_t$ implies that there exists a concave, strictly increasing (utility) function $u\left(.\right)$ and positive constant λ such that

$$\mathbf{D}_{\mathbf{q}_t^c} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) = \lambda \frac{1}{\beta^{t-1}} \boldsymbol{\rho}_t^c \tag{P5}$$

$$\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) = \lambda \frac{1}{\beta^{t-1}} \boldsymbol{\rho}_t^{a,0} \tag{P6}$$

$$\mathbf{D}_{\mathbf{q}_{t-1}^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) = \lambda \frac{1}{\beta^{t-1}} \boldsymbol{\rho}_t^{a,1} \tag{P7}$$

for all $t \in \tau$. Combining (P7) and restriction (R2) gives

$$\frac{1}{\lambda}\beta^{t-1}\mathbf{D}_{\mathbf{q}_{t-1}^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) = \boldsymbol{\rho}_{t-1}^a - \boldsymbol{\rho}_{t-1}^{a,0} \tag{P8}$$

Backdating (P6) (which must hold for all t) substitute for $\boldsymbol{\rho}_{t-1}^{a,0}$ to rewrite (P8) as

$$\beta^{t-2}\mathbf{D}_{\mathbf{q}_{t-1}^a} u\left(\mathbf{q}_{t-1}^c, \mathbf{q}_{t-1}^a, \mathbf{q}_{t-2}^a\right) + \beta^{t-1}\mathbf{D}_{\mathbf{q}_{t-1}^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) = \lambda \boldsymbol{\rho}_{t-1}^a \tag{P9}$$

which can then be updated to show

$$\beta^{t-1}\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) + \beta^{t}\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_{t+1}^c, \mathbf{q}_{t+1}^a, \mathbf{q}_t^a\right) = \lambda \boldsymbol{\rho}_{t-1}^a \tag{P10}$$

■

Proof of Theorem 2. Consider, without loss of generality the test for the one-lag habits model. If $\boldsymbol{\rho}_t^{a,1} = 0$ for all $t \in \tau$ then, from (R2) in Proposition 1 $\boldsymbol{\rho}_t^{a} = \boldsymbol{\rho}_t^{a,0}$ for all $t \in \tau$. Therefore $\boldsymbol{\pi}_t = \frac{1}{\beta^{t-1}} \left[\boldsymbol{\rho}_t^{c\prime}, \boldsymbol{\rho}_t^{a\prime}, \mathbf{0}'\right]'$, where $\mathbf{0}$ is a vector of zeros of appropriate length, and $\mathbf{x}_t = \left[\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}, \mathbf{q}_{t-1}^{a\prime}\right]'$ .Substituting into R(1) in Proposition 1 we have

$$0 \leq \sum_{\forall s,t \in \sigma} \frac{1}{\beta^{s-1}} \left[\boldsymbol{\rho}_s^{c\prime}, \boldsymbol{\rho}_s^{a\prime}, \mathbf{0}'\right]' \left(\left[\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}, \mathbf{q}_{t-1}^{a\prime}\right]_t' - \left[\mathbf{q}_s^{c\prime}, \mathbf{q}_s^{a\prime}, \mathbf{q}_{s-1}^{a\prime}\right]'\right) \qquad \forall \sigma \subseteq \tau$$

or equivalently

$$0 \leq \sum_{\forall s,t \in \sigma} \frac{1}{\beta^{s-1}} \boldsymbol{\rho}_s' \left(\mathbf{q}_t - \mathbf{q}_s\right) \qquad \forall \sigma \subseteq \tau$$

which is condition for the life-cycle model in Browning (Definition 1 and Proposition 1, (1989)) extended to allow for $\beta \neq 1$.■

Proof of Corollary 1. This follows immediately from Theorem 2 and Browning ((1989), Proposition 2). This is because GARP is a condition under which the data are consistent with a model of stable preferences and weak intertemporal separability. The conditions for SREH are stronger: additively time-separable and stable preferences.■

The R-lag habits model

Definition R. The time series of the interest rate, prices and quantities $\{i_t, \boldsymbol{\rho}_t^c, \boldsymbol{\rho}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t=1,\dots,T}$ satisfies the R-lag habits model if there exists a concave, strictly increasing (utility) function $u(.)$ and positive constants λ and β such that

$$\beta^{t-1} \mathbf{D}_{\mathbf{q}_t^c} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a, ..., \mathbf{q}_{t-R}^a\right) = \lambda \boldsymbol{\rho}_t^c$$

$$\beta^{t-1} \mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a, ..., \mathbf{q}_{t-R}^a\right) + \sum_{r=1}^{R} \beta^{k-1} \mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_k^c, \mathbf{q}_k^a, \mathbf{q}_{k-1}^a, ..., \mathbf{q}_{k-R}^a\right) = \lambda \boldsymbol{\rho}_t^a$$

where $k \equiv t + r$ for all $t \in \{R+1, ..., T\}$

Theorem R. The following statements are equivalent:
(T) The time series of the interest rate, prices and quantities $\{i_t, \boldsymbol{\rho}_t^c, \boldsymbol{\rho}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t=\tau}$ satisfies the R-lag model.
(R) There exist shadow prices $\{\boldsymbol{\rho}_t^{a,r}\}_{t\in\tau}^{r=0,\dots,R}$ and a positive constant β such that

$$0 \quad \leq \quad \sum_{\forall s,t \in \sigma} \boldsymbol{\pi}_s' \left(\mathbf{x}_t - \mathbf{x}_s\right) \qquad \forall \sigma \subseteq \tau \tag{R1}$$

$$0 \quad = \quad \boldsymbol{\rho}_{t-R}^a - \sum_{i=0}^{R} \boldsymbol{\rho}_{t-i}^{a,R-i} \quad \forall \quad t \in \tau \tag{R2}$$

where $\mathbf{x}_t = \left[\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}, \mathbf{q}_{t-1}^{a\prime}, ..., , \mathbf{q}_{t-R}^{a\prime}\right]'$ and $\boldsymbol{\pi}_t = \frac{1}{\beta^{t-1}} \left[\boldsymbol{\rho}_t^{c\prime}, \boldsymbol{\rho}_t^{a,0\prime}, \boldsymbol{\rho}_t^{a,1\prime}, ..., \boldsymbol{\rho}_t^{a,R\prime}\right]'$ and $t \in \{R+1, ..., T\}$.

Proof of Theorem R. The proof is analogous toTheorem 1 by induction on R.■

B. Variable definitions

The commoditiy groups are as follows: Food and non-alcoholic drinks at home; Alcohol; Tobacco; Clothing and footwear; House rent (includes imputed rent); Energy at home (heating by electricity); Services at home (heating not electricity, water, furniture repair); Non-durables at home (cleaning products); Nondurable medicines; Medical services; Transportation; Petrol; Leisure (cinema, theatre, clubs for sports); Education; Personal services; Personal non-durables (toothpaste, soap); Personal small items (combs, jewellery, lighters); Restaurants and bars.

References

[1] Abel, A. (1990), "Asset prices under habit formation and catching up with the Jones." *American Economic Review*, 80, 38-42.

[2] Afriat, S.N. (1967),"The construction of a utility function from expenditure data", *International Economic Review,* 8, 76-77.

[3] Becker, G. and K. M. Murphy. (1988), "A theory of rational addiction." *Journal of Political Economy*, 96:4, 675-701.

[4] Becker, G. M. Grossman and K. Murphy, (1994), "An empirical analysis of cigarette addiction" American Economic Review, 84(3) 396-418.

[5] Bewley, S. (1977), "The permanent income hypothesis: A theoretical formulation", *Journal of Economic Theory*, 16, 252-292

[6] Böhm-Bawerk, E. von. (1889), *Capital and interest.* South Holland: Libertarian Press.

[7] Boldrin, M., L. Christiano, and J. Fisher. (2001), "Habit persistence, asset returns, and the business cycle." *American Economic Review*, 91, pp. 149-66.

[8] Browning, M. (1989), "A nonparametric test of the life-cycle rational expections hypothesis" *International Economic Review*, 30(4), 979-992

[9] Browning M. and M. D. Collado, (2001) "The response of expenditures to anticipated income changes: panel data estimates," *American Economic Review*, 91(3), 681-692.

[10] Browning M. and M. D. Collado, (2004). "Habits and heterogeneity in demands: a panel data analysis," CAM Working Papers 2004-18, University of Copenhagen.

[11] Campbell, J. and J. Cochrane. (1999), "By force of habit: A consumption-based explanation of aggregate stock market behavior" *Journal of Political Economy*, 107, 205-51.

[12] Carrasco R., J-M. Labeaga and J. López-Salido, (2005), "Consumption and habits: Evidence from panel data," *Economic Journal*, 115(500), 144-165.

[13] Carroll, C, J. Overland, and D. Weil (2000), "Saving and growth with habit formation." *American Economic Review*, 90, 341-55.

[14] Chaloupka, F., (1991). "Rational addictive behavior and cigarette smoking". *Journal of Political Economy* 99, 722-742.

[15] Constantinides, G. (1990), "Habit formation: A resolution of the equity premium puzzle." *Journal of Political Economy*, 98:3, 519-43.

[16] Duesenberry, J. (1952), *Income, saving, and the theory of consumer behavior.* Cambridge, MA: Harvard University Press.

[17] Diewert, W.E. (1973), "Afriat and revealed preference theory", *Review of Economic Studies*, 40, 419-426.

[18] Fisher, I. (1930). *The Theory of Interest.* New York: MacMillan.

[19] Fuhrer, J. (2000), "Habit formation in consumption and its implications for monetray-policy models." *American Economic Review*, 90, 367-90.

[20] Griffith, R. and A. Leicester (2006), "Do the poor pay more: evidence from the UK", *Institute for Fiscal Studies*, mimeo.

[21] Hall, R. (1978), Stochastic Implications of the Life Cycle - Permanent Income Hypothesis: Theory and Evidence", *Journal of Political Economy*, 86, 971-987.

[22] Houthakker, H. S. (1950), "Revealed preference and the utility function", *Economica*, May, 159-174.

[23] Jermann, U. (1998), "Asset pricing in production economies." *Journal of Monetary Economics*, 41, 257-75.

[24] Koopmans, T. C. (1960), "Stationary ordinal utility and impatience". *Econometrica*, 28, 287-309.

[25] Kubler, F, (2004), "Is intertemporal choice theory testable?" *Journal of Mathematical Economics* 40, 177-189

[26] Labeaga, J-M, (1999), "A double-hurdle rational addiction model with heterogeneity: Estimating the demand for tobacco" *Journal of Econometrics* 93 49-72

[27] Meghir, C. and G. Weber, (1996), "Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis Using a U.S. Consumption Panel," *Econometrica,* 64(5), 1151-1181.

[28] Neary, J. P. and K. Roberts (1980), "The theory of household behaviour under rationing" *European Economic Review*, 13, 25-42.

[29] Pollak, R. (1970), "Habit formation and dynamic demand functions." *Journal of Political Economy*, 78:4, 745-63.

[30] Rae, J. (1834), *The Sociological Theory of Capital* (reprint of original 1834 edition). London:Macmillan.

[31] Ryder, H. and G. Heal. (1973), "Optimal growth with intertemporally dependent preferences." *Review of Economic Studies*, 40, 1-33.

[32] Samuelson, P. (1937), "A note on measurement of utility", *Review of Economic Studies*, 4, pp.155-61.

[33] Samuelson, P. (1952), "Probability, utility, and the independence axiom". *Econometrica*, 20:4, 670-78.

[34] Samuelson, P. (1948), "Consumption theory in terms of revealed preference", *Economica*, November, 243-253.

[35] Spinnewyn, Frans (1981), "Rational habit formation", *European Economic Review*, 15, 91-109.

[36] Varian, H. (1982), "The nonparametric approach to demand analysis", *Econometrica*, 50, 945-973.

[37] Varian, H. (1990), "Goodness-of-fit in optimizing models," *Journal of Econometrics*, 46(1-2), 125-140