Habits Revealed

Ian Crawford
Department of Economics, University of Oxford & IFS

February 25, 2009

Abstract

Models which incorporate habits have been shown many times and in many contexts to be useful in both macroeconomics and microeconomics. This paper sets out necessary and sufficient empirical conditions for the canonical rational intrinsic habits model in the revealed preference/nonparametric tradition of Samuelson (1948), Houthakker (1950), Afriat (1967) and Browning (1989). This allows an assessment of habits models which is free from the confounding effects of a choice of functional form. The conditions in the paper are shown to be computationally straightforward and to yield set identification for certain features of the model. The ideas outlined are applied to a microeconomic panel dataset. The addition of habit formation to the discounted utility model is shown to improve the rationalisability of the microdata considerably. Even if habit formation is rejected by the data it is shown that modest and plausible allowance for heterogeneity in prices and interest rates is sufficient to bring consumption behaviour into line with the theory. Theory-consistent discount rates and welfare measures revealed by the data are presented.

Acknowledgements I am grateful to the Editor Enrique Sentana and to two referees whose comments and suggestions have greatly improved the paper. I am also grateful to Richard Blundell, Martin Browning, Chris Carroll, Mette Christensen, John Hey, Richard Mash, Stephen Nickell, Frans Spinnewyn and seminar participants at Cambridge, IFS/cemmap, Leuven, Oxford and York for helpful comments on eariler versions and to Neil Rickman who suggested looking at this topic in the first place. Remaining errors are mine. Email: ian.crawford@economics.ox.ac.uk

1 Introduction

Models which allow for various kinds of habit formation have been used profitably to analyse a wide variety of both microeconomic and macroeconomic issues. Microeconomic 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 non-separabilities 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 tastes 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[1], 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 and the leading example of the kind of phenomenon 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 commodities are typically partitioned into a set of consumption goods and a set of habit-forming goods, and the instantaneous utility/felicity 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 generally been based on parametric estimates of the Euler equation or consumption function[2] and of the hypothesised underlying preference structure. The problem which an approach based on a statistical fit of a 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 a number of statistical/econometric auxiliary hypotheses. Similarly any model-based empirical identification of preferences will rest on these assumptions/choices.

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? The path which is taken in this paper is based on the revealed preference approach developed in Samuelson (1948), Houthakker (1950), Afriat (1967) and Varian (1982) and the extension of these ideas to the perfect foresight 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

[1] See Rozen (2008) for a recent axiomatic characterisation of rational intrinsic habits models.

[2] Although a recent paper by Chen and Ludvigson (2008) estimate a semiparametric asset pricing model which incorporates habit-formation.

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[3]: 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 price 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 perfect foresight habits model has yet been proposed or implemented. Whilst Kubler (2004) shows that nonparametric testing of general nonseparable 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 observables. It is shown, using ideas akin to those from the rationing literature (Neary and Robert (1980) and Spinnewyn (1981)) that habits models are testable, 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 rational intrinsic habits model, describes the implementation of these conditions and sets out the way in which measurement errors might be accommodated. It also discusses the identification of preference features and the relationship between the test described here and other nonparametric tests in the literature. Section 3 described the results of the application of these ideas to a microeconomic panel dataset. Section 4 concludes.

2 Characterising the habits model

2.1 Necessary and Sufficient Conditions

Suppose we have T observations indexed by t on a consumer's demands over time $\{\mathbf{q}_t\}$ and the corresponding prices $\{\mathbf{p}_t\}$ and interest rate $\{i_t\}$. Let the commodity vector be partitioned into a group of consumption goods $\mathbf{q}_t^c$ and a group of goods which are thought to be habit-forming $\mathbf{q}_t^a$ 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 initially focus on the simplest case in which the effects of lagged consumption of the addictive goods only persist for one period. The discussion of this extension (which is straightforward) is postponed until the end of this section.

The model of interest is

$$\max_{\mathbf{q}_t^c, \mathbf{q}_t^a} \sum_{t=1}^{\infty} \beta^{t-1} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) \text{ subject to } \sum_{t=1}^{\infty t-1} \left(\boldsymbol{\rho}_t^{c\prime} \mathbf{q}_t^a + \boldsymbol{\rho}_t^{a\prime} \mathbf{q}_t^a\right) = A_0 \text{ and } \mathbf{q}_0^a = \mathbf{q}^a$$

where $\rho_t^i = p_t^i / \prod_{s=2}^{s=t} (1 + i_s)$ denotes discounted prices and $\beta = 1/(1+\delta)$ where $\delta \in [0, \infty)$ is the consumer's rate of time preference. It is assumed that the instantaneous utility (felicity) function u is locally non-satiated differentiable and concave. Thus the model studied in this paper is the canonical version of the rational intrinsic habits model considered by Ryder and

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

Heal (1973), Boyer (1978, 1983), Spinnewyn (1981), Iannaccone (1986), Becker and Murphy (1988) and Becker Grossman and Murphy (1994) *inter alia.* Other variations/extensions of habits-like models have been put forward in the literature. These include models in which consumers are myopic (Pollak (1970)), discount rates that depend on prior consumption (Shi & Epstein (1993)), extrinsic/keeping-up-with-the-Joneses habits (Abel (1990) and Campbell and Cochrane (1999)) and reference point models which incorporate ideas from prospect theory and in which instantaneous utility/felicity is S-shaped (Loewenstein and Prelec (1992), Camerer and Loewenstein (2004)). The investigation of revealed preference conditions for these models is left for future work.

As Frederick *et al* (2002, p.396) point out, although the kind of standard rational habit formation model considered here is often said to induce a preference for an increasing consumption profile, in fact they are much more flexible and can also allow for preferences for decreasing or even non-monotonic consumption profiles (see also Rozen (2008) for an axiomatic basis for this property). This depends on various factors such as the level of the initial habits stock and whether current consumption raises or lowers future utility - in other words, whether the habit-forming good is good for you or not. Becker, Grossman and Murphy (1994), for example, employ precisely this rational intrinisic habits model in their study of cigarette addiction and use it to allow for the fact that current consumption can reduce future utility.

The first question is whether it is possible to find necessary and sufficient empirical conditions on observables under which the data are consistent with the model. To this end, consistency between the habits model and the data is defined as follows.

Definition 1. The data $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\{2,\ldots,T\}}$ satisfy the one-lag habits model if there exists a locally non-satiated, differentiable and concave utility/felicity function $u(.)$ and positive constants λ and β such that

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

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

This says that the data are consistent with the theory if there exists a well-behaved instantaneous utility/felicity 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 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 perfect foresight model. Those for the habit-forming goods are a little more complex because current consumption affects future utility as well as current utility. In the case of *a priori* harmfully addictive goods the discounted effect of current consumption on next period's utility is negative, but in general the model simply allows this term to be non-zero (see Becker, Grossman and Murphy (1994, p. 398) for a discussion of this point). Despite this complication this condition can be transformed into a form which is analogous to a no-habits model by defining suitable shadow discounted prices which account for these welfare effects[4]:

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

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

Expression (1) is the shadow discounted price of current consumption and measures the discounted willingness-to-pay for current consumption of the habit-forming goods. Expression (2) is the shadow discounted price of past consumption and measures the discounted 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, using Definition 1,

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

Given (1), (2) and (3) the habits model entails an intertemporal dependence between the shadow discounted prices:

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

The empirical/behavioural implications of the 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 can now be given:

Theorem 1. The following statements are equivalent:
(T) The data $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\{2,...,T\}}$ satisfy the one-lag habits model.
(R) There exist shadow discounted prices $\{\boldsymbol{\rho}_t^{a,r}\}_{t\in\{2,...,T\}}^{r=0,1}$ and a positive constant β such that

$$0 \leq \sum_{\forall s,t\, \in\sigma} \boldsymbol{\pi}_s' \left(\mathbf{x}_t - \mathbf{x}_s\right) \qquad \forall\sigma \subseteq \{2,...,T\} \tag{R1}$$

$$0 = \boldsymbol{\rho}_t^a - \boldsymbol{\rho}_t^{a,0} - \boldsymbol{\rho}_{t+1}^{a,1} \quad \forall \quad t, t+1 \in \{2,...,T\} \tag{R2}$$

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

Proof. See the Appendix.■

Theorem 1 is an equivalence result. It says that if one 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 data. Conversely if such shadow discounted prices and a discount

[4] As pointed out by Spinnewyn (1981). See also Neary and Roberts (1980).

rate cannot be found then there does not exist any theory-consistent utility representation. Restriction (R1) is a condition which is an implication of the concavity of the instantaneous utility function[5] 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 discussed above.

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. This means that, for any choice of discount rate, the existence or non-existence of feasible shadow prices can be readily determined in a finite number of steps using phase one of a (simplex method) linear programme. The issue is then simply one of conducting an arbitrarily fine grid search for the discount rate and running a linear programming problem at each node.

Theorem 1 shows that there are necessary and sufficient empirical conditions on observables which can be used to form a simple nonparametric test for rationalisability with the standard habits model. Like all tests in the revealed preference tradition the result is a straightforward yes/no. If the data satisfy the restrictions then they are consistent with the model and there exists a well behaved instantaneous utility/felicity function which satisfies the conditions in Definition 1[6] precisely. If the data do not satisfy the conditions then no such function exists. Whilst this seems very black and white, and indeed it is, it is worth exploring a little further.

Firstly, like all revealed preference tests, this test tells us whether the data and the model are mutually consistent, but it does not tell us whether the model is *right*. This is because, in general, but especially so in the case of revealed preference theory where predictions are set-valued, more than one model might explain/fit any given dataset and they cannot all be right. Indeed this is explored further below in the empirical work where some households are shown to be rationalisable by three different economic models (at least). So even if the data for a household are consistent with the habits model, the model might still be incorrect. Equally, if the model is rejected the test alone cannot tell us the reasons why. A related point is that the model might appear to be data-consistent, but this may turn out to be for the wrong reason. For example, while the empirical conditions take the level of aggregation over goods and periods in the data as given and ask if such data can be rationalised by a model defined over corresponding aggregates, the true model of household behaviour might be defined over differently aggregated data. The decision period for the household, for instance, might be different from the periodicity of the observed data. The smoothness in consumption which comes from time-aggregation, for example, can look a lot like habit persistence in a discrete-time model[7] and could potentially mean that the data are consistent with the model even though the model is not the true representation of a household's behaviour. While the investigator's ability to experiment with disaggregation is ultimately limited by the granularity of the data at hand, the sensitivity of the results to different aggregation schemes can easily be investigated within the framework developed in this paper by running the test whilst varying the aggregation over goods and periods.

To end this section consider a more general model in which consumption of the habit-forming goods persists for R periods[8] the instantaneous utility function is given by

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

[5] Rockafellar, (1970, Theorem 24.8)

[6] A suitable function can be constructed using the algorithm provided by Varian (1982).

[7] See Heaton (1993) for a discussion of this identification issue.

[8] It is assumed that the number of lags is strictly fewer than the number of observations. If this is not the case then obviously the habits model is untestable/unrejectable.

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. In other respects the restrictions are multi-period analogues of those in Theorem 1.

2.2 Allowing for errors

The conditions described above, like all revealed preference type tests, are 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. A useful framework within which one can address the effects of measurement errors on these kinds of tests has been suggested by Varian (1985) and this section briefly discusses how the habits model fits into this approach.

Let D^0 denote the observed data 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 set

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

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 $E = \{v_t; \mathbf{e}_t^c, \mathbf{e}_t^a; \mathbf{u}_t^c, \mathbf{u}_t^a\}_{t\in\{R+1,\ldots,T\}}$ represents measurement error which is classical by assumption[9]. Thus $D^* = \{i_t + v_t; \mathbf{p}_t^c + \mathbf{e}_t^c, \mathbf{p}_t^a + \mathbf{e}_t^a; \mathbf{q}_t^c + \mathbf{u}_t^c, \mathbf{q}_t^a + \mathbf{u}_t^a\}_{t\in\{R+1,\ldots,T\}}$. In this case a test statistic for the null hypothesis that the true data satisfy the model can be based on the loss function

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

where σ^2 is the variance of the measurement error. The statistic L is distributed as a chi-squared[10]. 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 (1985) 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.

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

[10] With degrees of freedom equal to the number of data points being perturbed.

2.3 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 and the Afriat (1967), Varian (1982) Generalised Axiom of Revealed Preference (GARP) test. The first result to note is that whilst the conditions for the habits model neither imply, nor are implied by those of the life-cycle model, nevertheless the test for habits nests the life cycle model in the following sense.

Theorem 2. If the data $\{i_t, \mathbf{p}_t^c, \mathbf{p}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\{R+1,\ldots,T\}}$ satisfies the R-lag model with $\boldsymbol{\rho}_t^{a,r} = \mathbf{0}$ and $r \geq 1$ then the data satisfies the conditions for the life-cycle model/strong rational expectations hypothesis.

This says that the life cycle model/strong rational expectations hypothesis can 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).

Similarly there is no general connection between rationalisability with the habits model and rationalisabilty with GARP. It is possible for data which satisfy the habits model to violate GARP and it is equally possible to construct data which satisfy GARP but which violate the habits model. However as a corollary to Theorem 2 we can note the following.

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

The intuition is straightforward: if the data satisfy the conditions for the habits model with the added zero-restriction on the shadow prices then they satisfy the conditions for the strong rational expectations life-cycle model (Theorem 2). If this is the case then they must also satisfy GARP because whilst GARP only requires within-period efficiency in expenditure the life-cycle model requires more; it requires efficient within-period *and* between-period allocation of expenditure. The conditions for the life-cycle model are therefore over-sufficient for GARP.

2.4 Identification

Given data which is consistent with the habits model, the question then arises as to whether it might be possible to identify features of the model. In general the empirical restrictions I have described will only allow identification of a *set* of admissible values for the discount rate and the shadow prices. The test procedure can be interpreted as simply determining whether or not this identification set is empty. The set of preference parameters which are data-consistent is not convex, a fact which stems from the non-linear nature of the restrictions implied by the habits model. In particular this means that the empirical identification interval for the discount rate has "gaps" in it and it is therefore possible for a certain value for the discount rate which cannot rationalise the data to lie between two values which can. This is why it is necessary to conduct an arbitrarily fine grid search over discount rates. The set of values of discount rates which satisfy the model can then simply be recorded.

The other elements of principal interest in the habits model are the relative willingness-to-pay measures (shadow prices) which capture the welfare effects of habit formation. 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 and the empirical procedure identifies theory-consistent combinations of these parameters. The only difficulty which arises is how best to represent the resulting set. Conditional on a choice of discount rate the set of willingness-to-pay parameters is convex and it is a straightforward problem to determine bounds on the welfare measures given a feasible set of starting values from the linear program. This is discussed further in the empirical application below.

3 Testing habits in microeconomic panel data

This section investigates the ideas discussed above using a household level panel dataset. The empirical results are organised as follows. They begin with an investigation of the consumption of tobacco - one of the most studied habit-forming goods (see for example Chaloupka (1991), Becker, Grossman and Murphy (1994), Labeaga (1999), *inter alia*). The performance of the one-lag habits model is looked at in comparison to the standard static utility maximisation model and the life-cycle model. The theory consistent set of discount rates is described for those households whose behaviour is rationalisable. For households which cannot be reconciled with the short memory habits model the effects of (i) allowing for measurement errors (ii) extending the lag length and (iii) allowing for habit formation in other goods is considered in turn. The impact on the rationalisability of data by the theory is shown in each case.

3.1 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, 2004)). 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 application 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 14 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 (these data and the issues they raise are further discussed below).

3.2 Results

3.2.1 Habits in Tobacco Consumption

The ECPF indicates that 76% of the sample households have positive expenditures on tobacco. Taking positive expenditures to be the indicator of smoking what follows concentrates on this sub-sample of 2388 households. The empirical results begin with the analysis of the comparative performance of the Varian (1982) test of GARP[11], the Browning (1989) test of the life-cycle/strong rational expectations hypothesis and the habits model with one lag on tobacco consumption[12]. It is important to note that each test is run *independently* using the data for each household in the sample, one at a time. The data across households are 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 (provided that their behaviour is rationalisable). It should also be noted that these results treat the household as a unitary entity and abstracts from issues to do with collective household behaviour[13]. Finally note that observations $\{3, .., T\}$ are used in the calculations for all of the models. Obviously, with one lag one can only use a maximum of $T - 1$ observations in the test of the habits model so it is important for comparability to truncate the data used for the life-cycle model and GARP tests to cover the same data points. The truncation by two periods is to make the results in this table comparable with those in Table 3 (below) which in due course considers the extension to 2 lags. The results are given in Table 1.

TABLE 1: Rationalisability results

Test:	Static u-max	Life-cycle	Habits (1 Lag)
Pass Rates:	0.972	0.046	0.244

The first column shows the pass rate for the GARP test. Recall that this tests for consistency between the data and the canonical static 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 97% of the sample satisfying GARP. The static model out-performs both the life-cycle model and the habits model by a significant degree.

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 5 percent of the sample satisfy the conditions required. It would appear that the data are generally inconsistent with the strong rational expectations version of the life-cycle model.

[11] The test is Varian's (1982) standard GARP test, which takes the total expenditures in a given period as fixed and does not require any connection across periods - except for stable preferences over a fixed set of commodities.

[12] Following the literature (e.g. Becker, Grossman and Murphy (1994)) 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. This boils down to adding the restriction that $\rho_t^{a,1} \leq 0$ to the conditions described in Theorem 1. The grid search for β was over the range $[0.95, 1]$ with a spacing of 0.005. Whilst it would be perfectly feasible to extend the range to the whole of $[0, 1]$ this range was chosen as reasonable given the quarterly frequency of the data.

[13] For a discussion of the issues raised by collective models of households in a revealed-preference framework see Cherchye *et al* (2007).

The last column shows the pass rate for the one lag 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 fixed at zero. Habits models relax this 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 and that the performance of the habits model is substantially better than that of the simple life-cycle model with about a quarter of smokers' behaviour rationalisable by the one-lag version of the model. Whilst the empirical results show a far better agreement between the habits model and the data than between the life-cycle model and the data, the pass rates are far below those of the GARP test.

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 (static, life-cycle and habits model) were regressed on a number of standard observable household characteristics and also the number of times the household is observed in the data. The pseudo R^2 of the probits were all low and virtually none of the coefficients were individually significant other than the variable measuring the number of times the household was observed which was strongly negative as one would expect[14]. 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 demographics.

The results in Table 1 seem to indicate that the static utility maximisation model performs very well, compared to the life-cycle model and the habits model. One potential explanation for this could be that if there is not much variation in nominal prices, then GARP will indeed perform well because, in the extreme case, the budget lines associated with different observations on a household would never cross and GARP *could* never be rejected for that household. To try to get some further understanding of these issues I have conducted a power analysis for each test. Power is, of course, a measure of $\Pr(\text{Rejecting } H_0 \mid H_0 \text{ is false})$ and the calculation of any power measure requires a specified alternative hypothesis. I use the approach developed by Bronars (1987) which adopts Becker's (1962) idea of uniform random behaviour as a general alternative hypothesis to optimising behaviour. The null hypothesis is, in turn: static utility maximisation, the life-cycle model and the one lag habits-in-tobacco model.

TABLE 2: Rationalisability and Power

Test:	Static u-max	Life-cycle	Habits (1 Lag)
Pass Rate	0.972	0.046	0.244
Power	0.039	0.969	0.754

The power results are reported in Table 2 which also recaps the pass rates from Table 1. We can see that while the pass rate of the GARP test was about 97% the power of the test is indeed very low. Conversely, while the pass rates of the two intertemporal models were more

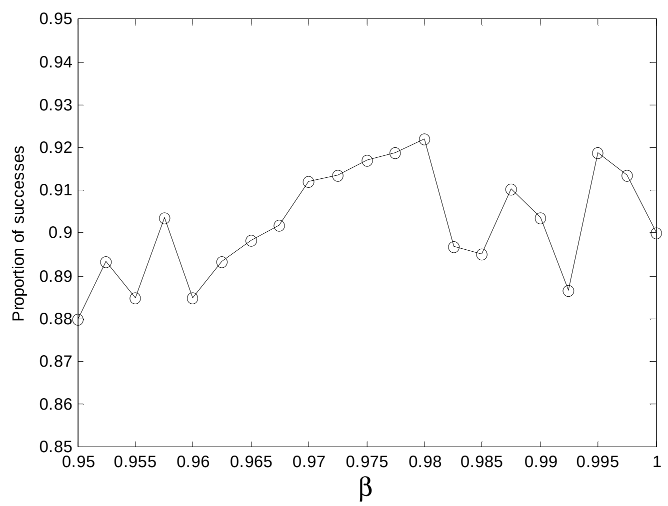
[14] The controls were {age of the head of household, the number of children in the household and dummy variables for highest educational qualification=university degree, highest educational qualification=high school, head's occupation = professional/managerial, head's occupation = skilled, homeowner, renter, drinker, number of times the household is observed in the data}. The probability of failing any RP-type test is weakly increasing in the number times one observes the agent. I am grateful to an anonymous referee for suggesting this regression. Details of these probits are available from the author.

modest the power calculations show that these tests were actually much more demanding than the GARP test. We see too that since the life-cycle model is a special case of the habits model the power of the life-cycle test is correspondingly higher than that of the habits model.

The comparison of the pass rates and power measures for the different tests raises the question of which model we might prefer[15]. The static utility maximisation model certainly performs very well compared to the alternatives in terms of pass rates. However the power measure indicates that the two intertemporal models are much more empirically demanding than the static model. How should we weight the impressive pass rate against the low power? A framework for thinking about these issues is provided by Selten (1991) who suggests that predictive models should be compared on the basis of their power-adjusted pass rates. Specifically Selten provides an axiomatic framework which cardinally identifies the difference between the pass rate and one minus the power as a suitable basis upon which to compare models. If we apply Selten's approach to these results the evident success of the static model is much undermined.

3.2.2 Preferences

FIGURE 1: Theory-consistent consumer discount rates



For each household which was rationalisable by the one-lag model, a set of admissible discount rates was recorded from the grid search. Figure 1 illustrates the probability that each discount rate in the range examined is rationalisable conditional on *some* discount rate being appropriate – i.e. the height of the line records the proportion of times each value of the discount rate was rationalisable over the sample of rationalisable household. If the line had reached 1 at any point that would mean that that value of β was acceptable for all of the rationalisable households, a value of 0.5 would have meant that that value of β was rationalisable with the data for only half of these households. The line slopes marginally upwards and is not smooth (due to the non-convex nature of the identification set). The lowest success rate was for $\beta = 0.95$ and the highest was for $\beta = 0.98$. Given these are quarterly data it is reasonably pleasing that higher values of β are somewhat more easily rationalised than lower ones. Nevertheless it is important to bear the vertical scale in mind when interpreting

[15] I'm grateful to a referee for raising this issue.

this figure. The range of variation is narrow (around 4 percentage points) so that the line is, in fact, rather 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 2: Theory-consistent rates of substitution

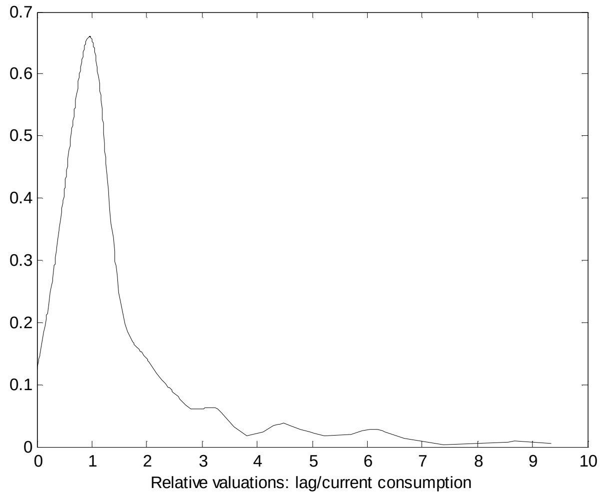


Figure 2 shows the distribution of absolute rates of substitution between current and lagged consumption of the habit-forming good amongst households whose behaviour is consistent with the short memory model. That is, it shows the distribution of ratios[16] of the absolute values of shadow prices in expressions (1) and (2):

$$\frac{\rho_t^{a,0}}{\rho_t^{a,1}} = \frac{D_{q_t^a} u\left(\mathbf{q}_t^c, q_t^a, q_{t-1}^a\right)}{D_{q_{t-1}^a} u\left(\mathbf{q}_t^c, q_t^a, q_{t-1}^a\right)}$$

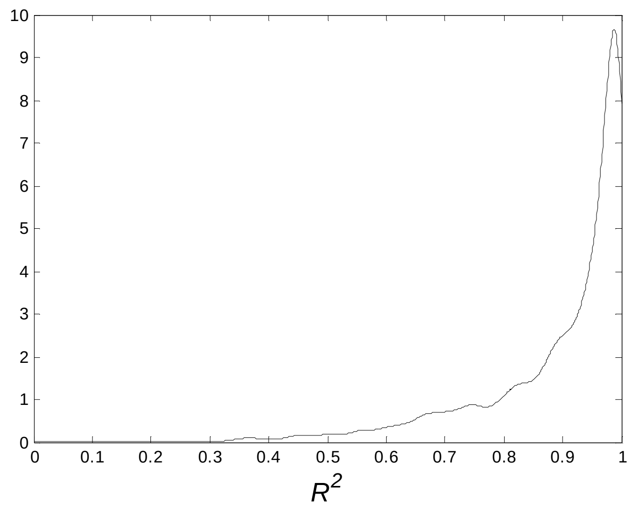
This measures how much the consumer is willing to pay for current consumption relative to lagged consumption and measures the relative importance of the habit-formation component. A ratio which is less than one indicates that the lagged effect outweighs the current effect and that the habitual element is rather strong. Conversely rates of substitution greater than one would indicate relatively weak effects from habits. In the limit, of course as the habit effect drops to zero this measure would tend to infinity. It appears that, for tobacco, habits are important; the median value is 1.03 indicating that the habit is about as important as current consumption on average, and in fact 13% of the sample have relative valuations which are less than 0.5 indicating that, for them, the habitual element is twice as important as current consumption. Nevertheless, as is seen from the figure there is a long right hand tail (which is actually truncated) and for half of the sample current consumption out-weighs lagged consumption whilst for about one quarter of household the effect of current consumption are twice as big as the lagged effect.

[16] Note that whilst the testing procedure only requires that feasible values are found for the shadow prices, the calculations here involve maximising and minimising $\rho_t^{a,0}/\rho_t^{a,1}$ at each observation to find bounds on the range of rates of substitution. The procedure `fmincon` in MatLab 7.3.0 was used to implement this taking the feasible values from the LP as starting values. The figure is a kernel density calculated by placing Gaussian kernels at the mid points of the bounds for each household.

3.2.3 Allowing for Measurement Errors

For households whose behaviour violates the habits model it is possible, using the ideas outlined in section 2.2, to perturb the data so that they satisfy the model. As discussed above, the data are composed of expenditures on disaggregated 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 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. I am 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[17] and there is no obvious reason to expect that Spain is much different. In view of this the discounted price data for each violating household has been individually perturbed by the minimum distance necessary such that they then satisfy the model.

FIGURE 3: The density of the distribution of $R^2 : \rho_t^i = \rho_t^{i*} + e_t^i$



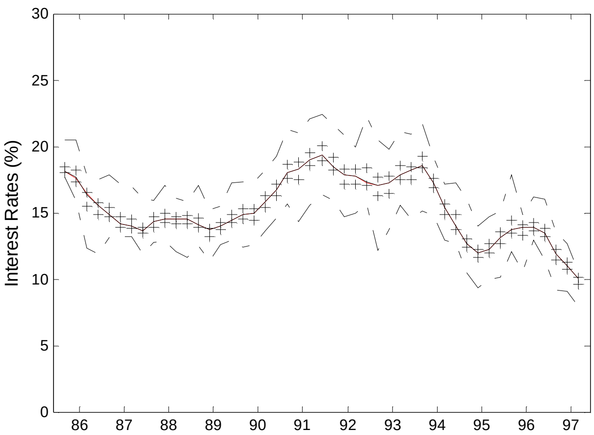
Raw distance measures generally depend on the units involved so in the absence of an estimate of the variance of the measurement error it is hard to tell whether the necessary perturbation is big. To help with interpretation the distances have been translated into R^2-type values based on how much the perturbations contribute to the fitting of the nearest theory-consistent values compared to the observed discounted price. If the observed 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 of R^2 is carried out independently for each household. Figure 3 illustrates the density of the distribution of R^2 values. It shows a right-skewed distribution with 90% of the R^2 values greater than 0.7. The behaviour of many households appears to be reasonably close (on this measure) to rationalisable by the model.

Even if the observed and perturbed discounted prices are close by an R^2 measure it is still not completely clear whether the difference is economically significant. One interesting

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

exercise is to take the perturbed (and now theoretically-consistent) discounted prices and to recover from them the implied household-specific interest rate which rationalises the data. This allows for the fact that the aggregate discounted price data is constructed using the published aggregate average rate of interest on consumer loans whilst in reality different households might vary widely in the cost of borrowing which they face. Figure 4 shows the time series of certain quantiles of the interest rate distribution which emerges.

FIGURE 4: Rationalisable interest rates and the observed interest rate



What appears to be a single solid line in the middle of the figure is, in fact two lines: they are the published consumer interest rate along with the median of the rationalisable (perturbed) interest rate distribution. However, it is impossible to tell them apart because they are virtually identical. The crosses indicate the quartiles of the rationalisable interest rate distribution. These are on average less than ± 0.5 percentage points of the observed interest rate over the period. In other words, for half of those households whose behaviour cannot be reconciled with the observed data an adjustment of only half of one percentage point in the interest rate they face is sufficient to be able to rationalise them. The outer dashed lines are the 90th and 10th percentiles of the rationalisable interest rate distribution. On average these lie within ± 2.5 percentage points from the observed interest rate over the period. Together these results seem to indicate that the distribution of theory-consistent interest rates is quite "peaky". The implication one might draw from Figures 3 and 4 is that it appears that only reasonably modest adjustments to the prices and interest rates faced by households are required to rationalise the data.

3.2.4 Adding Lags

The results in Table 1 showed that a habits model with one lag in tobacco consumption did something to improve the rationalisability of the data with theory compared to the unadorned life-cycle model. However, the proportion of the data which satisfied the habits model was still far lower than that for the static utility maximisation model. Possible explanations were discussed above and it was shown that fairly small household-specific adjustments to prices and interest rates could reconcile the data with the model. This approach essentially imposes the model by altering the data where necessary. An alternative is to increase the number of lags. To explore this the number of lags in tobacco consumption was increase to two. Table 4 shows the results.

Table 3: Rationalisability results, adding lags

Test:	Static u-max	Life-cycle	Habits (1 Lag)	**Habits (2 Lags)**
Pass Rate	0.972	0.046	0.244	**0.913**
Power	0.039	0.969	0.754	**0.202**

The results in the first three columns recap those shown in Table 1 and are there for comparison. The final column shows what happens when the lag length is increased to two periods[18]. Increasing the lag length in tobacco consumption increases the pass rate substantially. Note that as with Table 1 all of these results are based on the same observations (3 to T) for each household in order that the results are comparable. The increase in the ability of the habits model to rationalise the data is not, therefore, to do with the 2-lag model simply being tested against fewer observations than the alternative models. The implications appear to be significant: allowing for just two lags in the consumption of a single, plausibly habit forming, good can improve the agreement between the theory and the data from a situation in which almost none of observed behaviour is rationalisable (as with the strong rational expectations, life-cycle model) to a point at which the vast majority of the data are rationalisable. However it is important to note that the increased flexibility of the model (essentially adding a free time-varying parameter by relaxing the assumption that the shadow price of the two-period habit is zero) has a large negative effect on the power of the test compared to the one-lag version. This drop in power is further exacerbated if the lags are extended to three periods. Allowing for a lag of three quarters means that, compared to Table 3 which uses observations 3 to T, it is necessary to drop a period from the test and whilst this does improve the pass rate somewhat to around 95% (recall that the only useful predictor of the outcome of the test for a single lag was the number of observations involved and that this was strongly negative) it also significantly reduces the power of the test (again partly due the introduction of free parameters, but also this time, because of the reduced number of observations used).

3.2.5 Adding Goods

So far the empirical work has concentrated on tobacco; that is, the consumption vector has 14 disaggregated commodity groups in it, but amongst them only tobacco is allowed to have lagged effects. Whilst tobacco is the classic habit-forming good in the literature, there is no reason to suppose that complementarities between consumption in different periods do not exist for other goods as well. Leaving aside alcohol and gambling which are almost as frequently investigated as tobacco, a far from exhaustive list of other commodities which have been looked at in the literature on habit-formation is heterogeneous enough to include milk[19], coffee[20], cinema[21] and religious practice[22]. With this in mind the next set of results looks at the effects of allowing for habits in all spending categories on the rationalisability of the data. As was the case with adding lags, one can think of this as a relaxation of zero-constraints on shadow prices.

[18] Note that habits models with a one period lag is nested by the two period lag version (See Theorem 2). Hence if a household's behaviour can be explained by a short memory habits model then it has to be explainable by a longer memory model too.

[19] Auld and Grootendorst (2004).

[20] Olekalns and Bardsley (1996).

[21] Cameron (1999).

[22] Iannaccone (1990).

Table 4: Rationalisability results, adding goods

Test:	Static u-max	Life-cycle	Habits (Tobacco)	**Habits (All goods)**
Pass Rate	0.972	0.046	0.244	**0.985**
Power	0.039	0.969	0.754	**0.091**

The results in the first three columns in Table 4 again recap those shown in Table 1. The final column shows the degree of rationalisability between the data and a short memory habits model in which all goods are allowed to be habit forming (note too that intertemporal cross-complementarities can exits between different commodity groups). Once again, the number of observations involved in the test of each model is the same. Allowing for one-period habits in other goods increases the rationalisability of behaviour as expected and now over 98% of the data are theory consistent. However once more the power calculation serves as an important *caveat*; whilst in the more parsimonious model in which habit formation is restricted to just one commodity the power of the test was good (0.754) if habit formation in all goods is admitted then the power of the test drops markedly (although not as low as the GARP test for the static model). Finally, if one allows for two period lags in tobacco consumption (which Table 2 suggested is a good idea) and single period lags in all other goods one finds 99.87% agreement between the data and theory. That is to say, all of the data except for those relating to just three households are perfectly rationalisable with the habits model.

4 Conclusions

Gorman (1967) claimed that "It is commonplace that choices depend on tastes and tastes on past choices" and since then habits model have been shown, many times and in many contexts, to be useful in both macroeconomics and microeconomics. The literature suggests that habits models often fit the data well and provide insights into various economic issues which might otherwise prove resistant to straightforward explanation.

This paper has derived general empirical conditions for the standard intrinsic habits model in the revealed preference tradition of Samuelson (1948), Houthakker (1950), Afriat (1967) and Browning (1989). This allows, for the first time, an assessment of habits models which is nonparametric and therefore free from the confounding effects of a choice of functional form. The conditions in the paper are shown to be computationally straightforward and to yield set identification results for certain features of the model.

The ideas outlined have been applied to a microeconomic panel dataset. It appears that, in contrast to the results for aggregate data found by Browning (1989), the strong rational expectations version of the life-cycle model is heavily rejected. However, the addition of habit formation to the discounted utility model was shown to improve the rationalisability of the microdata considerably - virtually to the point where one hundred percent of the data are perfectly rationalisable if one allows intertemporal complementarities for many goods. Nevertheless it is important to recognise that by allowing more pervasive lags into the model, the power of the restrictions is much weakened. When habit-formation is rejected it was shown that rather modest and plausible allowance for heterogeneity in prices and interest rates was sufficient to bring consumption behaviour in line with the theory. Theory-consistent discount rates and welfare measures revealed by the data were presented. Overall, it appears that habits models are capable of explaining longitudinal household behaviour reasonably well.

Appendix

A. Proofs

Proof of Theorem 1.

$(T) \Rightarrow (R)$: Definition 1 and the definitions of the shadow discounted prices in (1) and (2) imply (4) which is restriction (R2). Together they imply

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

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 and differentiability of the instantaneous utility function $u(\mathbf{x}_t)$ means

$$u(\mathbf{x}_s) - u(\mathbf{x}_t) \leq +\mathbf{D}u(\mathbf{x}_t)'(\mathbf{x}_s - \mathbf{x}_t) \quad \forall\ t, s \in t \in \{2, ..., T\} \tag{P2}$$

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

$$u(\mathbf{x}_s) \leq u(\mathbf{x}_t) + \lambda \boldsymbol{\pi}_t'(\mathbf{x}_s - \mathbf{x}_t) \quad \forall\ t, s \in t \in \{2, ..., T\} \tag{P3}$$

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

$$0 \leq \sum_{\forall s,t\ \in \sigma} \boldsymbol{\pi}_s'(\mathbf{x}_t - \mathbf{x}_s) \quad \forall \sigma \subseteq t \in \{2, ..., T\} \tag{P4}$$

which is restriction (R1).

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

$$\begin{aligned}
\mathbf{D}_{\mathbf{q}_t^c} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) &= \lambda \frac{1}{\beta^{t-1}} \boldsymbol{\rho}_t^c && \text{(P5)}\\
\mathbf{D}_{\mathbf{q}_t^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) &= \lambda \frac{1}{\beta^{t-1}} \boldsymbol{\rho}_t^{a,0} && \text{(P6)}\\
\mathbf{D}_{\mathbf{q}_{t-1}^a} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right) &= \lambda \frac{1}{\beta^{t-1}} \boldsymbol{\rho}_t^{a,1} && \text{(P7)}
\end{aligned}$$

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 $\boldsymbol{\rho}_t^{a,1} = \mathbf{0}$ then, from (R2) in Proposition 1 $\boldsymbol{\rho}_t^{a} = \boldsymbol{\rho}_t^{a,0}$. 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

$$0 \leq \sum_{\forall s,t\ \in\sigma} \frac{1}{\beta^{s-1}}\left[\boldsymbol{\rho}_s^{c\prime}, \boldsymbol{\rho}_s^{a\prime}, \mathbf{0}'\right]\left(\left[\mathbf{q}_t^{c\prime}, \mathbf{q}_t^{a\prime}, \mathbf{q}_{t-1}^{a\prime}\right]_t' - \left[\mathbf{q}_s^{c\prime}, \mathbf{q}_s^{a\prime}, \mathbf{q}_{s-1}^{a\prime}\right]'\right) \qquad \forall\sigma \subseteq t \in \{2,...,T\}$$

or equivalently

$$0 \leq \sum_{\forall s,t\ \in\sigma} \frac{1}{\beta^{s-1}}\boldsymbol{\rho}_s'\left(\mathbf{q}_t - \mathbf{q}_s\right) \qquad \forall\sigma \subseteq t \in \{2,...,T\}$$

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). ■

B. The R-lag habits model

Definition R. The data $\{i_t, \boldsymbol{\rho}_t^c, \boldsymbol{\rho}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\{R+1,...,T\}}$ satisfy the R-lag habits model if there exists a locally non-satiated, differentiable and concave 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 $k \equiv t + r$.

Theorem R. The following statements are equivalent:
(T) The data $\{i_t, \boldsymbol{\rho}_t^c, \boldsymbol{\rho}_t^a; \mathbf{q}_t^c, \mathbf{q}_t^a\}_{t\in\{R+1,...,T\}}$ satisfy the R-lag model.
(R) There exist shadow prices $\{\boldsymbol{\rho}_t^{a,r}\}_{t\in\{R+1,...,T\}}^{r=0,...,R}$ and a positive constant β such that

$$0 \ \leq \ \sum_{\forall s,t\ \in\sigma} \boldsymbol{\pi}_s'\left(\mathbf{x}_t - \mathbf{x}_s\right) \qquad \forall\sigma \subseteq t \in \{R+1,...,T\} \tag{R1}$$

$$0 \ = \ \boldsymbol{\rho}_{t-R}^a - \sum_{i=0}^{R}\boldsymbol{\rho}_{t-i}^{a,R-i} \quad \forall \quad t \in \{R+1,...,T\} \tag{R2}$$

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 $\boldsymbol{\pi}_t = \frac{1}{\beta^{t-1}}\left[\boldsymbol{\rho}_t^{c\prime}, \boldsymbol{\rho}_t^{a,0\prime}, \boldsymbol{\rho}_t^{a,1\prime}, ..., \boldsymbol{\rho}_t^{a,R\prime}\right]'$.

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

C. Variable definitions

The commoditiy groups are as follows: Food and non-alcoholic drinks at home; Alcohol; Tobacco; 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); Personal services; Personal non-durables (toothpaste, soap); Restaurants and bars.

The ECPF data are collected quarterly. There are eight one-week survey periods within each quarter. Participating households are surveyed in the same week of each successive quarter (e.g. always the third, or always the first). Adult household members complete expenditure diaries in which they record their spending during the survey week. These data are supplemented with retrospective recall questionnaires and the data grossed-up to quarterly values. See Browning and Collado (2001) for a description of the data collection process.

References

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

[2] Auld, M. C. and P. Grootendorst, (2004) "An empirical analysis of milk addiction" *Journal of Health Economics* 23(6), 1117-1133

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

[4] Baltagi, B. H. and J. M. Griffin (2002) "Rational addiction to alcohol: panel data analysis of liquor consumption" *Health Economics,* 11(6), 485-491

[5] Becker, G., (1962), ""Irrational Behavior and Economic Theory", *Journal of Political Economy*, **70**, 1-13.

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

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

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

[9] Bronars, S.G., (1987): "The Power of Nonparametric Tests of Preference Maximization," *Econometrica*, **55** (3), 693–698.

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

[11] 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.

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

[13] Cameron, S.,(1999), "Rational addiction and the demand for cinema" *Applied Economics Letters*, 6(9), 617-620

[14] 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.

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

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

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

[18] Cherchye L, B. De Rock and F. Vermeulen (2007), "The collective model of household consumption: a nonparametric characterization", *Econometrica*, 75(2), p 553-574.

[19] Chen, X., and S. Ludvigson (2004), "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Behavior," NBER Working Paper.

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

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

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

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

[24] Frederick, S., Loewenstein, G. and E. O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review", *Journal of Economic Literature*, 40(2), 351-401

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

[26] Gorman, W.M. (1967) "Tastes, Habits and Choices", *International Economic Review*, 8(2), 218-222

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

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

[29] Heaton, J., (1993) "The interaction between time-nonseparable preferences and aggregation" *Econometrica*, 61(2), 353-385

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

[31] Iannaccone , L. R. (1990), "Religious Practice: A Human Capital Approach" *Journal for the Scientific Study of Religion*, 29(3), 297-314.

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

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

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

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

[36] 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.

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

[38] Olekalns, N and P. Bardsley (1996), “Rational Addiction to Caffeine: An Analysis of Coffee Consumption” *Journal of Political Economy*, 104(5),1100-1104.

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

[40] Rozen, K (2008), “Foundations of Intrinsic Habit Formation” mimeo

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

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

[43] Selten, R. (1991), “Properties of a measure of predictive success” *Mathematical Social Sciences* 21,153-167

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

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

[46] Varian, H. (1985), “Nonparametric analysis of optimizing behavior with measurement error”, *Journal of Econometrics*, **30**, 445–458.