ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Blow, Laura; Browning, Martin; Crawford, Ian

Working Paper

Never mind the hyperbolics: Nonparametric analysis of time-inconsistent preferences

IFS Working Papers, No. W14/17

Provided in Cooperation with:
Institute for Fiscal Studies (IFS), London

Suggested Citation: Blow, Laura; Browning, Martin; Crawford, Ian (2014) : Never mind the hyperbolics: Nonparametric analysis of time-inconsistent preferences, IFS Working Papers, No. W14/17, Institute for Fiscal Studies (IFS), London, https://doi.org/10.1920/wp.ifs.2014.1416

This Version is available at:
https://hdl.handle.net/10419/101334

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Institute for Fiscal Studies

Never mind the hyperbolics: nonparametric analysis of time-inconsistent preferences

IFS Working Paper W14/17

Laura Blow
Martin Browning
Ian Crawford

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The Institute for Fiscal Studies (IFS) is an independent research institute whose remit is to carry out rigorous economic research into public policy and to disseminate the findings of this research. IFS receives generous support from the Economic and Social Research Council, in particular via the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP). The content of our working papers is the work of their authors and does not necessarily represent the views of IFS research staff or affiliates.

Never Mind the Hyperbolics: Nonparametric Analysis of Time-Inconsistent Preferences

Laura Blow*, Martin Browning† & Ian Crawford‡

Abstract

We investigate necessary and sufficient nonparametric conditions for the quasi-hyperbolic consumer. These turn out to be quite tractable. We investigate the performance of this model compared to the standard exponential discounting model using consumer panel data.

Key Words: Quasi-hyperbolic discounting, revealed preference.

JEL Classification: C43, D11.

Acknowledgements: We are very grateful for financial support from the European Research Council grant ERC-2009-StG-240910-ROMETA, the Leverhulme Trust grant F/08 519/E and the ESRC Centre for Microeconomic Analysis of Public Policy at IFS (grant RES-544-28-5001). We thank seminar audiences at Oxford, University College London, Universitat Pompeu Fabra and Bristol for their patience and helpful comments.

1 Introduction

What are the necessary and sufficient nonparametric/revealed preference empirical conditions for quasi-hyperbolic consumption behaviour? This paper explores this in the elementary, choice-revealed preference tradition of Samuelson (1948), Houthakker (1950) and Afriat (1967). We describe these conditions for the leading forms of the model and discuss how to implement them and also how to make sensible empirical comparisons with alternative models including the standard exponential discounting model. We also provide an empirical application using a large, nationally representative panel household consumption survey.

In recent years quasi-hyperbolic discounting has begun to prove itself a major challenge to the traditional exponential discounting framework. The standard assumption of exponential discounting with its constant discount rate is, of course, a parsimonious one since it allows

*Institute for Fiscal Studies. 7 Ridgmount Street, London, WC1E 7AE. E-mail: laura_b@ifs.org.uk.

†Nuffield College and University of Oxford, Department of Economics, Manor Road Building, Manor Road, Oxford OX1 3UQ, United Kingdom. E-mail: martin.browning@economics.ox.ac.uk.

‡IFS, Nuffield College and University of Oxford, Department of Economics, Manor Road Building, Manor Road, Oxford OX1 3UQ, United Kingdom. E-mail: ian.crawford@economics.ox.ac.uk.

a person's time preference to be summarised as a single parameter and a useful one too since it makes the strong prediction that the consumer's intertemporal preferences are time-consistent. However a certain amount of evidence has accrued which indicates that people often do not behave in a time-consistent manner.[1] Indeed Samuelson (1937), when he introduced the exponential discounted utility model, in the same breath disavowed it on precisely these grounds:

> "our equations hold only for an individual who ... will still make at each instant the same decision with respect to expenditure as he would have, if at the beginning of the period he were to decide on his expenditure for the whole period. But the fact that this is so is in itself, a presumption that individuals do not behave in terms of our functions." [2]

Despite Samuelson's manifest misgivings, the exponential form was too useful to be ignored and was quickly adopted as the standard model of intertemporal consumer behaviour. This state of affairs lasted until the late 1990's when economists began first to question it and then to abandon it in favour of the quasi-hyperbolic discounting model[3,4]. This process is (pungently) described in Rubinstein (2003). The model has now been widely adopted and applied, it would seem with some success, to describe a range of phenomena from the role of illiquid assets as commitments[5], the excess sensitivity of consumption to income and the retirement savings puzzle[6], the simultaneous holding by households of high pre-retirement wealth, low liquid assets and high credit-card debt[7], labour supply and welfare program participation[8], procrastination in a number of contexts[9], addiction/habit formation[10], information acquisition[11] and the Phillips curve[12]. This is a far from exhaustive list of applications but it serves to give a sense of the breadth of topics to which the model has been fruitfully applied. The emerging consensus seems to be that, compared the the exponential model, quasi-hyperbolic discounting better fits the stylised facts regarding individuals' intertemporal behaviour and the experimental evidence. Indeed, Frederick *et al* (2001, p361) conclude that "the collective evidence ... seems overwhelmingly to support hyperbolic discounting".

[1] See Frederick *et al* (2002) for a survey of the empirical literature.

[2] Samuelson (1937, p160)

[3] Samuelson foreshadowed precisely this in his original article, writing: "Actually, however, as the individual moves along in time there is a sort of perspective phenomenon in that his view of the future in relation to his instantaneous time position remains invariant, rather than his evaluation of any particular year (e.g. 1940). This relativity effect is expressed in the behaviour of men who make irrevocable trusts, in the taking out of life insurance as a compulsory savings measure, etc." (Samuelson, (1937, p. 160)).

[4] Strotz (1955-1956) considered non-exponential discounting and Phelps and Pollak (1968), and then Elster (1979) studied the, now firmly established, (β, δ) form. Laibson (1994,1997,1998) Harris and Laibson (2001) in particular have analysed the implications of this form extensively.

[5] Laibson (1997)

[6] Laibson (1998)

[7] Angeletos et al. (2001)

[8] Fang and Silverman (2009).

[9] Fischer (1999) and O'Donoghue and Rabin (1999, 2001)

[10] O'Donoghue and Rabin (2000), Gruber and Koszegi (2000), and Carrillo (1999).

[11] Carrillo and Mariotti (2000) and Benabou and Tirole (2000).

[12] Graham and Snower (2007).

In this paper we study quasi-hyperbolic consumption behaviour from a revealed preference perspective. When researchers are armed with a theory about what the individual may be maximising they typically proceed in one of two ways. The first is the structural econometric approach in which the implications of the theory are described in terms of the properties of certain key functions (compensated demand curves, Euler equations and the like). Since these functions are never observed directly they need to be estimated from the data so the econometrician appends a statistical structure to the economic model in order to account for the fact that the economic theory, as expressed through the structural equations, does not perfectly fit the data. This extra structure generally entails statistical assumptions regarding the joint distribution of unobserved explanatory variables and unobservable error terms. The art of structural modelling thus mainly lies in getting this statistical aspect right, because the source and the properties of these unobserved and unobservable factors can have a critical impact on the estimation results (see for example, McElroy (1987), Brown and Walker (1989), Lewbel (2001)).

The alternative is empirical revealed preference analysis. This is the approach which we adopt. Rather than describing the implications of the theory in terms of estimated structural equations, empirical revealed preference uses systems of inequalities which depend neither on the form of structural functions nor on unobservables. Statistical error terms and special assumptions about the functional form of the economic model may be added but it is not an essential requirement of the approach. The classic example is the Generalised Axiom of Revealed Preference (GARP) which is the necessary and sufficient condition for the standard utility maximisation model with competitive linear pricing. GARP is a simple system of inequalities (or an equivalent linear program) which only involves the observed prices and choices and which is "nonparametric" in the sense that it assumes nothing about the functional form of preferences. By Afriat's Theorem[13] if the data satisfy GARP, then they are consistent with the hypothesis that they were generated by a rational consumer with well-behaved preferences. If they are not consistent with GARP then the data are not rationalisable by the standard neoclassical utility maximisation model and the model, or perhaps the data, need modification. Thus GARP exhausts the empirical implications of the standard consumer choice model.

In this paper we are, in essence, trying to see whether there is a GARP-like condition for the quasi-hyperbolic consumption model. The answer is a non-obvious - or at least was to us. This is because empirical revealed preference conditions usually exploit some consistency property in the individual's behaviour which is implied by the model. The quasi-hyperbolic model, however, explicitly implies behaviour which is *inconsistent* in some respects. The answer is also of interest: if a revealed preference characterisation does not exist then the model is not falsifiable in the same way that, for example, the exponential discounting model is (see Browning (1989) for a version of Afriat's Theorem which applies to the exponential model); if such a characterisation does exist then this is an important supportive result for one of the leading models in the Behavioural canon.

[13] Afriat (1967) , Diewert (1973) and Varian (1982).

We present necessary and sufficient conditions for the quasi-hyperbolic consumption[14] model and show that the associated empirical tests are rather straightforward. We also show a number of interesting related results which describe, in general terms, what can be learned about the quasi-hyperbolic model from the data without making strong identifying assumptions. We show that whilst it is possible to distinguish the sophisticated quasi-hyperbolic consumer from his exponential cousin, it is not possible to disentangle the hyperbolic individual's two discount rates. In contrast, we also show that using observables alone one cannot distinguish between the naive quasi-hyperbolic consumer and an exponential consumer with a very high discount rate - the same behaviour is rationalisable by both models. We then show how our results relate to those of Afriat (1967), Browning (1989) and others. We also investigate how the quasi-hyperbolic consumption model performs compared to the exponential model using a large, consumption panel data set. In this respect it is worth noting that it is often claimed that the hyperbolic model outperforms the exponential model empirically (see many of the papers cited above). We investigate this and it is indeed the case. But of course it could hardly be otherwise: the quasi-hyperbolic model is more general than the exponential model because, compared to the exponential model it contains a free parameter. It therefore necessarily fits the data no worse than the standard model. Merely claiming that the hyperbolic model is better in terms of fit is not a rational reason to prefer it. In this paper we consider this problem in some detail and suggest a different way of comparing models using their empirical informational content.

2 Necessary and Sufficient Conditions for the Quasi-hyperbolic Consumption Model

2.1 The self-aware individual

We are interested in the empirical implications of the quasi-hyperbolic discounting model for a finite dataset of interest rates, spot prices and purchases of K goods $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t=1,\ldots,T}$ for a mortal, self-aware individual who has no ability to commit to consumption plans. We begin by considering the perfect-foresight case and discuss the extension to uncertainty (which is straightforward) in due course. We consider what we take to be a standard version of the

[14] We focus on the consumption model for a number of reasons. The first that consumption behaviour is intrinsically important in both macro and microeconomics (we take this as read). The second is that we would like to develop methods which can be applied to readily-available real world data. The strongest evidence in favour of hyperbolic discounting behaviour often comes from the lab (Thaler (1981), Benzion, Rapoport and Yagil (1989), Kirby and Herrnstein (1996)) but the methodological status of lab studies in economics is far from uncontroversial (see for example Rubinstein (2001) and Levitt and List (2007)). We are therefore interested in investigating hyperbolic discounting using data from standard, widely-available real world datasets on consumption behaviour of the sort which have been used for many years to estimate Euler equations like the Consumer's Expenditure Survey (CEX) in the US or the Family Expenditure Survey (FES) in the UK. Thirdly, and perhaps most importantly, consumption behaviour is a good focus of study for our purposes because the data are often particularly abundant. Time-inconsistency may be an important feature of households' decisions to invest in a house or an individual's decision to investment in education, but such decisions are made rather infrequently. Consumption decisions, on the other hand, are made all the time by households and so provide an excellent data-rich context for studying intertemporal models.

hyperbolic consumption model in which the individual is modelled as a composite of temporal selves indexed by their respective periods of control, $t = 0, 1, 2, ..., T$, over a consumption decision. During their period of control, self t observes all past consumption decisions, and the current level of financial wealth A_t and chooses a consumption bundle for period t, such that

$$\mathbf{p}_t' \mathbf{c}_t + S_t = Y_t + A_t$$

Savings/dis-savings are denoted S_t. Self $t+1$ then inherits wealth equal to

$$A_{t+1} = (1 + r_{t+1}) S_t$$

The game continues, with self $t+1$ in control. We assume that $A_T = 0$. The payoff for the t'th player of this game is

$$U(\mathbf{c}_t, \mathbf{c}_{t+1}, ..., \mathbf{c}_T) = u(\mathbf{c}_t) + \beta \sum_{i=1}^{T-t} \delta^i u(\mathbf{c}_{t+i})$$

where u is a concave, continuous and differentiable instantaneous utility function. Optimising behaviour in this model can be characterised using the first order condition on the equilibrium path (see Harris and Laibson,(2001)). The following is a slight generalisation of their result which allows for multiple consumption goods - which will feature in our application.

Lemma 1 *(Harris and Laibson (2001)). On the equilibrium path:*

$$\frac{\partial u}{\partial c_t^k} = \delta^{-t} \lambda \rho_t^k \prod_{i=1}^{t} \left[1 - (1-\beta) \sum_{k=1}^{K} \left(p_i^k \frac{\partial c_i^k}{\partial A_i} \right) \right]^{-1} \qquad \forall k$$

and the corresponding Euler equation is

$$\frac{\partial u}{\partial c_t^k} = \delta (1 + r_{t+1}) \frac{p_t^k}{p_{t+1}^k} \left[1 - (1-\beta) \sum_{k=1}^{K} \left(p_{t+1}^k \frac{\partial c_{t+1}^k}{\partial A_{t+1}} \right) \right] \frac{\partial u}{\partial c_{t+1}^k} \qquad \forall k$$

where $\rho_t^k = p_t^k / \prod_{i=1}^{t} (1 + r_i)$ denotes the discounted price of the kth good, λ is a strictly positive constant (the marginal utility of wealth in period 0), β and δ lie in the interval $[0, 1]$ and $\partial c_t^k / \partial A_t$ is the marginal propensity to consume out of current wealth for the kth good.

Proof. See the Appendix for this and all subsequent proofs. ■

Note that this first order condition and Euler equation simplify to the standard exponential discounting case when $\beta = 1$. Both expressions are slight generalisations of those in Harris and Laibson (2001)[15] and readily reduce to those in Harris and Laibson (2001) if we consider

[15]Note that in deriving the first order condition and the Euler equation we follow the heuristic method of Harris and Laibson (2001) - that is we simply make the necessary assumptions regarding the smoothness/differentiability of demands. As Harris and Laibson (2001) show, the same results can be derived under less benign conditions. Readers are referred to their paper for this derivation.

a single consumption good, no inflation and a fixed interest rate. The object

$$\sum_{k=1}^{K}\left(p_{t+1}^{k}\frac{\partial c_{t+1}^{k}}{\partial A_{t+1}}\right)$$

is the period $t+1$ marginal propensity to spend out of wealth and, assuming that demands are normal, it lies between zero and one. Thus we can rewrite the hyperbolic Euler equation as

$$\frac{\partial u}{\partial c_t^k}=\delta^*\left(1+r_{t+1}\right)\frac{p_t^k}{p_{t+1}^k}\frac{\partial u}{\partial c_{t+1}^k}$$

where

$$\delta^*=\delta\left[1-\sum_{k=1}^{K}\left(p_{t+1}^{k}\frac{\partial c_{t+1}^{k}}{\partial A_{t+1}}\right)\right]+\delta\beta\left[\sum_{k=1}^{K}\left(p_{t+1}^{k}\frac{\partial c_{t+1}^{k}}{\partial A_{t+1}}\right)\right]$$

is sometimes termed the "effective discount factor". This is a weighted average of the short-run discount factor $\delta\beta$, and the long-run discount factor δ. This captures the conflict between the short-term and the long-term self because what the future self does matters to the current self. The larger the future self's marginal propensity to spend the more important the short-run discount factor and the current marginal utility is lower relative to the future than it would otherwise be in the standard case. This is due to the tendency of the future self to over-consume, as far as the current self is concerned, and thus to reduce that period's marginal utility of consumption relative to subsequent periods.

Following Browning (1989) we define what it means for the model to rationalise the data as follows.

Definition 1 *The self-aware quasi-hyperbolic discounting model rationalises the data* $\{r_t,\mathbf{p}_t,\mathbf{c}_t\}_{t\in\tau}$ *if there exists a locally non-satiated, differentiable and concave instantaneous utility (felicity) function* $u\left(.\right)$ *and constants* $\lambda>0$, $\beta\in[0,1]$, $\delta\in[0,1]$ *and* $\{\mu_t\in[0,1]\}_{t\in\tau}$*such that*

$$\frac{\partial u}{\partial c_t^k}=\lambda\frac{1}{\delta^t}\rho_t^k\prod_{i=1}^{t}\frac{1}{\left[1-\left(1-\beta\right)\mu_i\right]}\qquad\qquad\forall k$$

where $\rho_t^k=p_t^k/\prod_{i=1}^{t}\left(1+r_i\right)$ *and* $\mu_i=\sum_{k=1}^{K}\left(p_i^k\frac{\partial c_i^k}{\partial A_i}\right)$ *denotes the marginal propensity to spend in period* i.

This says that the data are consistent with the theory if there exists a well-behaved instantaneous utility/felicity function, the derivatives of which satisfy the hyperbolic first order conditions (or equivalently the hyperbolic Euler equation) on the equilibrium path. If such a utility function exists, and we know what it is, then it means that we can simply plug it into the model and precisely replicate the observed 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. Note that to allow for easy comparison this definition simplifies with $\beta=1$ to the rationalisability definition given

by Browning (1989) for the exponential discounting model.

We can now give the following result.

Proposition 1 *The following statements are equivalent.*
(1) The data $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$ *satisfy the self-aware quasi-hyperbolic discounting model.*
(2) There exist constants $\delta \in [0,1]$ $\{\psi_t \geq 0\}_{t\in\tau}$ *such that*

$$0 \leq \sum_{\forall s,t\,\in\sigma} \frac{\psi_t}{\delta^t} \boldsymbol{\rho}_t' \left(\mathbf{c}_s - \mathbf{c}_t\right) \qquad \forall \sigma \subseteq \tau \tag{RS1}$$

$$1 \leq \psi_{t-1} \leq \psi_t \qquad \forall t \tag{RS2}$$

where $\rho_t^k = p_t^k / \prod_{i=1}^t (1 + r_i)$.

The interpretation of ψ_t in the proposition is $\psi_t = \prod_{i=1}^t (1/[1-(1-\beta)\mu_i])$. Proposition 1 is an equivalence result. It says that if one can find a suitable path of increasing constants and a δ discount rate such that restrictions (RS1) and (RS2) 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 constants and a δ discount rate do not exist then there is no theory-consistent utility representation. Restriction (RS1) is a cyclical monotonicity condition which is an implication of the concavity of the instantaneous utility function[16] and the Euler equation on the equilibrium path for this consumer. Restriction (RS2) is related to the way in which the marginal utility of consumption evolves over time for the hyperbolic consumer.

The empirical problem is thus a question of determining whether a set of increasing constants and a δ discount rate which satisfy (RS1) and (RS2) can be found. These restrictions are computationally quite straightforward. The important feature to note is that the restrictions are linear conditional on δ. This means that, for any choice of δ discount rate, the existence or non-existence of the ψ_t terms can be efficiently 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 δ discount rate and running a linear programming problem to check the conditions at each node.

It is useful to show how Proposition 1 relates to some important existing results so we give two corollaries of Proposition 1 below.

Corollary 1 *If condition (RS1) holds, then the data also satisfy the Generalised Axiom of Revealed Preference (Afriat (1967), Diewert (1973), Varian (1982)).*

Corollary 2 *If* $\psi_t = \psi_{t+1}$ *in condition (RS1) then the data satisfy the conditions for the exponential model described by Browning (1989).*

Corollary 1 says that the hyperbolic discounter satisfies GARP. That is to say, regardless of how the matter of the allocation of a budget to each period is handled, he acts as if he is

[16] Rockafellar, (1970, Theorem 24.8)

maximising a stable well-behaved instantaneous utility function within each period. This is essentially due to the stability of within-period tastes and intertemporal separability of the hyperbolic consumer's preferences which is necessary and sufficient for the second stage of two-stage budgeting (Gorman (1959)).

The second corollary concerns the relationship with the exponential model. It shows that the conditions in Browning (1989) for the exponential model can be regarded as a special case of the conditions for the current model. This result is useful computationally if one wants to compare the predictive performance of the models (see below).

Another useful aspect of the connection with the results in Browning (1989) is that it is relatively simple to see how to extend Proposition 1 to the case where there is uncertainty. Under certainty, (RS1) uses the fact that marginal utility of wealth, λ_t, evolves according to

$$\lambda_t = \frac{1}{\delta\left(1+r_t\right)} \frac{\lambda_{t-1}}{\left[1-\left(1-\beta\right)\mu_t\right]} = \lambda \frac{\psi_t}{\delta^t} \prod_{i=1}^{t} \frac{1}{\left(1+r_i\right)}$$

and so λ can be cancelled out. As is well known, under uncertainty the consumer is modelled as adopting expected utility as the objective function conditional on the information available at time t. The key empirical implication of this version of the model in the exponential discounting case is that, in contrast to the perfect foresight case where it is flat, the marginal utility of lifetime income follows a martingale $E[\lambda_t|\lambda_{t-1}, ..., \lambda_1, \lambda] = \frac{1}{\delta(1+r_t)}\lambda_{t-1}$ (Hall (1978)). With exponential discounting $\psi_t = 1\,\forall t$ but its presence in (RS1) means it will pick up the stochastic part of marginal utility, and hence the stochastic version of Corollary 2 has ψ_t following a martingale: $E[\psi_t|\psi_{t-1}, ..., \psi_1] = \psi_{t-1}$. Thus under uncertainty the test of exponential discounting in Browning (1989) is simply the condition (RS1) plus $E[\psi_t|\psi_{t-1}, ..., \psi_1] = \psi_{t-1}$. Likewise to allow for uncertainty in Proposition 1 (i.e. with hyperbolic discounting) we use

$$E[\left(1-\left(1-\beta\right)\mu_t\right)\lambda_t|\lambda_{t-1}, ..., \lambda_1, \lambda] = \frac{1}{\delta\left(1+r_t\right)}\lambda_{t-1}$$
$$\Longrightarrow E[\lambda_t] = \frac{1}{E[\left(1-\left(1-\beta\right)\mu_t\right)]}\left[\frac{1}{\delta\left(1+r_t\right)}\lambda_{t-1} + \left(1-\beta\right)Cov\left(\mu_t, \lambda_t\right)\right]$$

(removing the conditional notation for simplicity). Since $E[\left(1-\left(1-\beta\right)\mu_t\right)] \leq 1$ this implies

$$E[\lambda_t|\lambda_{t-1}, ..., \lambda_1, \lambda] \geq \frac{1}{\delta\left(1+r_t\right)}\lambda_{t-1}$$

as long as $\left(1-\beta\right)Cov\left(\mu_t, \lambda_t\right)$ is not large and negative. There is no reason to expect a relationship between the marginal utility of wealth and how a shock to wealth is allocated across remaining periods: for example, in the standard esponential discounting model (with discount rate equal to interest rate for simplicity) a transitory shock to wealth is divided equally among remaining periods of life and so μ_t and λ_t are independent. Thus to allow for uncertainty for the hyperbolic discounting case we just require that (RS2) is replaced by the condition that ψ_t follows a sub-martingale (again ψ_t picks up the stochastic element of

marginal utility):

$$E[\psi_t|\psi_{t-1}, ..., \psi_1] \geq \psi_{t-1} \qquad \forall t. \tag{RS2'}$$

A further interesting result is that it does not appear to be possible separately to learn much about the the additional discount factor which is the distinguishing feature of the hyperbolic consumption model in the sense that if the data satisfies the model then all that it is possible to establish about β is that it is less than 1.

Corollary 3 *If the data satisfies conditions (RS1) and (RS2) then it is consistent with any* $\beta \in (0, 1)$.

This result would seem to indicate that, nonparametrically at least and with consumption data of the type we focus on, whilst it is possible to establish that a consumer has quasi-hyperbolic preferences, it is not in general possible to say anything more about their β discount factor. This result does not mean that parametric assumptions cannot provide means of delivering β but it does show that those parametric assumptions are likely to be crucial.

2.2 The naive individual

In this section we consider the case of an individual who is a hyperbolic discounter but wrongly assumes that his future self will simply fall into line with the consumption plan which he maps out. The literature (e.g. O'Donoghue and Rabin (1999)) describes this as naivety and means by this that the current self knows themselves to be a hyperbolic discounter with an inclination for immediate gratification, but believes that future selves do not have present-biased preferences but will behave as exponential discounters with $\beta = 1$. That is, in period t they maximise

$$u(\mathbf{c}_t) + \beta \sum_{i=1}^{T-t} \delta^i u(\mathbf{c}_{t+i})$$

but believe that in periods $\tau = t + 1...T - 1$ their future selves will maximise

$$u(\mathbf{c}_\tau) + \sum_{i=1}^{T-\tau} \delta^i u(\mathbf{c}_{\tau+i})$$

The implications for the first order condition and Euler equations of individuals who behave in this way are given in the next Lemma. This shows that the Euler equation, for example, for this type of individual is structurally very similar to that of an exponential discounter; the difference is that the combination of the two discount rates produces an effective discount rate which is higher than the standard case with the implication that such an individual will run down their assets faster than the pure exponential discounter to the extent that $\beta < 1$.

Lemma 2 *On the equilibrium path:*

$$\frac{\partial u}{\partial c_t^k} = (\beta\delta)^{-t} \lambda \rho_t^k \qquad \forall k$$

and the corresponding Euler equation is

$$\frac{\partial u}{\partial c_t^k} = \beta\delta\left(1 + r_{t+1}\right)\frac{p_t^k}{p_{t+1}^k}\frac{\partial u}{\partial c_{t+1}^k} \qquad \forall k$$

where $\rho_t^k = p_t^k / \prod_{i=1}^t (1 + r_i)$ *denotes the discounted price of the* k*th good,* λ *is a strictly positive constant,*β *and* δ *lie in the interval* $[0, 1]$.

We are interested in the conditions under which observations on the interest rate, spot prices and consumption decisions may be rationalisable by the naive version of the hyperbolic model. Paralleling Definition 1 we describe the requirements for rationalisability below. Once more this reduces to the case of exponential discounting considered in Browning (1989) when $\beta = 1$.

Definition 2 *The quasi-hyperbolic discounting model with naive beliefs rationalises the data* $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$ *if there exists a locally non-satiated, differentiable and concave instantaneous utility (felicity) function* $u(.)$ *and constants* $\lambda > 0$, $\beta \in [0, 1]$, $\delta \in [0, 1]$ *such that*

$$\frac{\partial u}{\partial c_t^k} = (\beta\delta)^{-t}\lambda\rho_t^k \qquad \forall k$$

where $\rho_t^k = p_t^k / \prod_{i=1}^t (1 + r_i)$.

The following proposition establishes the necessary and sufficient conditions for the naive case.

Proposition 2 *The following statements are equivalent.*
(1) The data $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$ *satisfy the naive quasi-hyperbolic discounting model.*
(2) There exist constants $\delta \in [0, 1]$,$\beta \in [0, 1]$ *such that*

$$0 \leq \sum_{\forall s,t\ \in\sigma} \frac{1}{d^t}\boldsymbol{\rho}_t'\left(\mathbf{c}_s - \mathbf{c}_t\right) \qquad \forall\sigma \subseteq \tau \tag{RN1}$$

where $\rho_t^k = p_t^k / \prod_{i=1}^t (1 + r_i)$ *and* $d = \beta\delta$.

Once more the Proposition is an equivalence result: if there exists a constant d such that the inequality in $(RN1)$ holds then the data are precisely replicable by the naive hyperbolic model; if such a constant does not exist then the data, as they stand are not rationalisable. The computational problem is again straightforward: conditional on a choice of d the condition is simply a matter of checking a finite number of inequalities or an equivalent linear program.

Both Lemma 2 and Definition 2 are, as noted above, very similar to the corresponding results for the exponential individual. In particular, in contrast to the self aware hyperbolic consumer where the effective discount rate is endogenous, here it is not and the effective discount rate is just $\beta\delta$. This delivers the following corollary.

Corollary 4 *The following statements are equivalent.*
(1) The data $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$ *satisfy the naive quasi-hyperbolic discounting model.*
(2) The data $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$ *satisfy the exponential discounting model.*

The corollary shows that the nonparametric empirical implications of naive hyperbolic behaviour for the data $\{r_t, \mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$ are identical to those of the exponential model: if the data are consistent with exponential behaviour they are consistent with naive hyperbolic behaviour too and it is not possible to distinguish β and δ in (RN1) with standard observational consumption data. In fact the only difference between exponential discounters and naive hyperbolic discounters is one of degree rather than of kind: the naive hyperbolic consumer behaves exactly like an exponential consumer with a high discount rate ($\beta\delta < \delta$). It might be a little surprising therefore that an unsophisticated hyperbolic individual who fails to recognise the inconsistency in their own preferences will end up behaving in a way which, to the researcher, looks perfectly time-consistent Nonetheless the intuition is fairly clear: the naive consumer lays a series of consumption plans but can only enact the first period's consumption decision; they therefore effectively make a series of one-period decisions. The extension to allow for uncertainty simply follows Browning (1989) and replaces $1/(\beta\delta)^t$ in (RN1) with $\lambda_t/(\beta\delta)^t$ where λ_t follows a martingale process.

2.3 Summary

To summarise: we have presented revealed preference conditions for the self-aware and the naive versions of the quasi-hyperbolic consumption model. These are necessary and sufficient conditions for the agreement between the theory and data recording interest rates, spot prices and consumption decisions. With further assumptions on the precise form of the model one could derive further restrictions, but the ones presented in Propositions 1 and 2 and the corollaries presented exhaust the nonparametric content of the theory, i.e. those implications which do not follow from specific functional form assumptions.

The next section investigates the ideas discussed above using a household level consumption panel dataset and looks at the performance of the hyperbolic model compared to an atemporal model in which the budget in each period is parametric and the exponential discounting model (using the characterisation in Browning (1989)).

3 Empirical Illustration

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, 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 six commodity groups: food consumed at home, alcohol, tobacco, clothing and footwear, restaurants and bars and domestic energy. The discounted price data are calculated from published prices aggregated to correspond to the expenditure categories and the average interest rate on consumer loans. 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.[17]

3.1 Consistency Results

We examined the consistency of the data with each of there models: atemporal utility maximisation (i.e. a GARP test), the exponential consumption model and the sophisticated quasi-hyperbolic consumption model. We examine the data for consistency for each of these models using the data for each individual household, one at a time[18]. It is important to emphasise that the data are never pooled across households so we therefore allow for complete preference heterogeneity within the classes of models studied: households may differ with respect to whether they are rationalisable by a given model, their discount rates and their instantaneous utilities. Since we only have a handful of observations on each household a nonparametric statistical test of the martingale/sub-martingale hypothesis (for example along the lines of Phillips and Jin (2013)) would have no power to reject. We therefore focus on the perfect certainty version of the models in our empirical work. Table 1 reports the pass rates for each model considered.

Table 1: Rationalisability results

Model	Atemporal	Exponential discounting	Quasi-hyperbolic discounting
Pass Rate	0.9569	0.0045	0.5281
(Std. Error)	(0.0036)	(0.0012)	(0.0089)

From Table 1 we see that the atemporal model[19] rationalises the behaviour of 95% of the sample. The exponential discounting/naive model by contrast rationalises very few households

[17] For a discussion of the issues raised by collective models of households in a revealed-preference framework see Cherchye *et al* (2007) and especially Adams *et al* (2012) who consider the coexistence of exponential discounters with differing discount rates within a household as another potential source of time-inconsistency in aggregate (household) behaviour.

[18] For the intertemporal model we conducted a grid search for δ over the unit interval with a grid spacing of 0.001 - this means that formally the pass rates for these models are lower bounds.

[19] $\max_{\mathbf{q}} u(\mathbf{q})$ subject to $\mathbf{p}_t'\mathbf{q} = \mathbf{p}_t'\mathbf{q}_t$

(less than half of one percent). Given that within-period consistency with utility maximisation (GARP) is a necessary but not sufficient condition for the exponential model (see Browning (1989)) this disparity must therefore be due, in the main, to the intertemporal behaviour displayed by the households in the sample. The hyperbolic model falls between these extremes - just over half of the sample display behaviour which is precisely rationalisable by hyperbolic preferences. The standard errors for each of these statistics is given in brackets in Table 1. In each case the effects of sampling variation appear to be quite modest and we would therefore expect the proportions of households consistent with each model in another random sample of a similar size from this population to be close to those in Table 1.

3.2 Model comparison

The results from the consistency tests are fairly stark and are statistically significant, but are they economically significant in the sense of providing useful information about the various models considered? Can we conclude that the atemporal model is the best model followed by the quasi-hyperbolic model and that the exponential model stands condemned?

It is important to remember when looking at the pass rates that these models are not equally demanding. The atemporal model of utility maximisation does not consider intertemporal planning at all and takes the budget allocated to each period as parametric. It only requires that within-period preferences are rational and stable. It has no restrictions whatsoever concerning how spending is allocated across time and therefore, as long as within-period preference satisfy GARP, *any* intertemporal allocation is, in that sense, rationalisable with the atemporal model. The exponential model is much more demanding. It constrains both within-period and intertemporal choices: within-period choices must be rational and stable, and intertemporal choices must be time-consistent. The hyperbolic model is, in a sense, intermediate. It too requires within-period rationality and stability (as shown in Corollary 1), but whilst it does not require time-consistency, intertemporal behaviour is constrained by the form of the quasi-hyperbolic Euler equation. To view the comparison of the exponential with the quasi-hyperbolic model another way; compared to the exponential model, quasi-hyperbolic preferences introduce a free parameter - the β discount rate. Thus, whilst much is sometimes made of the ability of the hyperbolic model to explain observed behaviour which the exponential model cannot[20], this should come as little surprise. How, then, should we make sense of the results in Table 1?

In their revealed preference guise, stripped of special functional form assumptions, all three models generate restrictions in the form of *sets* of choices which are consistent with the model of interest (for example, given any collection of budget constraints there will be a *set* of demands which satisfy GARP). To investigate the performance of models which predict sets, it is useful to think about two objects. The first is the feasible outcome space (for example, the set of choices which satisfy the intertemporal budget constraint) which we will denote F. The second is the set of model-consistent choices (for example, that set of feasible choices which

[20]See Frederick *et al* (2001) for example.

satisfy the restrictions of the model), denoted P. Clearly the model-consistent choices are a subset of the feasible choices: $P \subseteq F$. When one conducts a particular empirical revealed preference test one is, in essence, checking to see whether the observed choices lie within P.

With this in mind it becomes clear that it is necessary to allow for the relative size of the predicted/theoretically consistent subset. The essential idea - which is due to Selten and Krischker (1983) and Selten (1991) - is that if the set of observations consistent with the model(P) is very large relative to the set of behaviours which the consumer could possibly display (F) then we should be little surprised if we find that many of the observed choices lie in P - they could hardly have done otherwise. For example, if we are testing the atemporal model and the collection of budget constraints never cross then all feasible choices, necessarily, satisfy GARP: it would be impossible to make an observation which wasn't in P because $P = F$. This means that empirical fit alone (the proportion of the sample which passes the relevant test) is not a sufficient basis for ranking the performance of alternative theories: if it were then no theory could out-perform a meaningless theory which can rationalise any behaviour. A better approach would be to consider the trade-off between the pass rate and some sort of measure of how demanding the theory is. Following Selten (1991) let a denote the size of the theory-consistent subset P, *relative* to the outcome space F for the model of interest. The relative area of the empty set is zero and the relative area of all outcomes is one so $a \in [0, 1]$. Now suppose that we have some choice/outcome data. Let r denote the pass rate; this is simply the proportion of the data that lies in P and hence satisfies the restrictions of the model of interest (i.e. the numbers in Table 1). Selten (1991) argues that demanding theories are characterised by small values for a; and empirically successful theories combine small values of a with a high degree of agreement between the data and theory (large r). He also argues that the trade-off between the ability to fit the data and the restrictiveness of the theory should be the difference measure[21]: $r - a$. In what follows we take a somewhat different approach.

Since both the relative size of the theory-consistent set and the empirical pass rate satisfy all of the necessary properties of probabilities[22] we can be justified in thinking about the problem of comparing them as the problem of comparing probability distributions. Here $\{a, 1 - a\}$ is a bivariate probability distribution describing the probability that random uniform choice over the feasible set will satisfy the restrictions of the model. Similarly $\{r, 1 - r\}$ is a bivariate probability distribution describing the probability that a subject drawn at random with uniform probability from the population will exhibit behaviour consistent with the model. A simple way in which to make a comparison between these two distributions in units which are meaningful is to use the Kullback-Leibler divergence

$$KL(r, a) = \sum_i r_i \log_2 \left(\frac{r_i}{a_i} \right)$$

[21]See also Beatty and Crawford (2010) for an application of this to revealed preference test.

[22]They are non-negative, sum to one and the area of two none-overlapping subsets is the sum of their individual areas.

where i indexes possible outcomes. The Kullback-Leibler divergence measures the information (in bits) conveyed by the empirical results relative to what one would expect from the theoretical restrictions - in a well-defined formal sense it measures how "surprising" the outcome of the test is given the restrictiveness (or otherwise) of the theory of interest[23]. For example, suppose that the pass rate was 0.7 and this matched the size of theory-consistent set predicted by the model precisely (i.e. the proportion of all possible observed choices which satisfy the theory was also 0.7) then the observed pass rate would convey no surprise at all and the data would generate zero bits of information about the empirical performance of the model. If, on the other hand, the set of outcomes which are rationalisable represent very little of the outcome space ($a \to 0$) (i.e. the model makes very precise predictions) and most of the data generally satisfy the theoretical restrictions ($r \to 1$) or even if the reverse were true ($r \to 0$) then the outcome of the empirical test is extremely informative and $KL(r, a)$ is large: in the former case it is informative of the spectacular success of the theory, in the latter it is informative about equally spectacular failure.

Table 2: Kullback-Leibler Divergence

Model	Atemporal	Exponential discounting/naive	Quasi-hyperbolic discounting
$KL(r, a)$	0.940 (43.876)	10.342 (5.014)	240.691 (26.639)

Table 2 presents the Kullback-Leibler results for the three models. The numbers represent the number of bits of information conveyed by the results. To give a sense of scale it might be helpful to bear in mind that one flip of a coin believed to be fair contains one bit of information, 3134 flips of a coin which was believed to be fair and which all came up heads would (i) be very surprising and (ii) would contain 0.39 kilobytes of information about the true probability of a head, and any number of flips of a coin which contained equal numbers of heads and tails would contain zero bits of information. The very high pass rate of the atemporal model is revealed to be rather like this last example and quite uninformative about the success or otherwise of that model - and indeed the standard error indicates that the non-zero value is very likely to be due to sampling variation: another sample could easily produce zero information about the model. The KL divergence being essentially zero indicates that the predictive success of the model is due, almost entirely, to its permissiveness. Turning to the comparison of interest, that between the exponential and the hyperbolic model, we see that even allowing for the less restrictive nature of the hyperbolic model compared to the exponential model, the performance of the hyperbolic model is far better than the exponential model. Furthermore this difference between the performances of the models which we observe in this sample of households appears to be unlikely to have arisen simply by chance through random sampling. The better performance of the quasi-hyperbolic model does not appear to be a simple mechanical artifact associated with a free parameter.

[23]The surprisal of an event with probability p_i is proportional to $\log_2(1/p_i)$ so the KL divergence is interpretable as the average surprisal of the empirical performance of the model compared to the performance of a uniform random number generator acting over the set of feasible outcomes.

4 Conclusion

We provide the a revealed preference characterisation of quasi-hyperbolic consumption behaviour by a self-aware consumer who recognises the tension between their short-term and long-term selves and a naive consumer who does not. This class of models may allow for time-inconsistent behaviour but we show that, despite this, there is enough consistency in their behaviour that the model is *meaningful* in the Samuelsonian sense - it is refutable on the basis of observables alone. Furthermore the empirical conditions are quite easy to apply since they require nothing more complicated than linear programming. In an empirical application we consider in some detail how to interpret the revealed preference performance of alternative models which differ in their restrictiveness. We suggest the Kullback-Leibler divergence as a means of doing this. This seems to be quite a fruitful way of describing the performance of the models in question - and possibly economic models in general. We find that, for these data, the hyperbolic model provides the best description of behaviour of the models considered here.

References

[1] Adams, A., Cherchye, L., De Rock, B and E. Verriest, (2012) "Consume Now or Later? Time Inconsistency, Collective Choice and Revealed Preference" mimeo.

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

[3] Angeletos, G-M, D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg. (2001). "The hyperbolic consumption model: Calibration, simulation, and empirical evaluation." *Journal of Economic Perspectives*, **15**:3, 47-68.

[4] Beatty T. and I. Crawford (2011), "How Demanding Is the Revealed Preference Approach to Demand?", *American Economic Review* 101,2782–2795

[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] Benabou, R. and J. Tirole. (2000). "Self-confidence: Intrapersonal strategies.". Princeton University Discussion Paper #209: Princeton.

[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] Carrillo, D. (1999). "Self-control, moderate consumption, and craving.". CEPR Discussion Paper 2017.

[14] Carrillo, J. D. and T. Mariotti. (2000). "Strategic ignorance as a self-disciplining device." *Review of Economic Studies*, 67:3, 529-44.

[15] Cherchye, L., De Rock, L., and F. Vermeulen (2007), "The Collective Model of Household Consumption: A Nonparametric Characterization", *Econometrica*, 75.2 p553-574.

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

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

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

[19] Elster, J, (1979). *Ulysses and the Sirens : studies in rationality and irrationality*. Cambridge, UK: Cambridge University Press.

[20] Fischer, C., (1999). "Read this paper even later: Procrastination with time-inconsistent preferences." Resources for the Future Discussion Paper 99-20.

[21] 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

[22] Gruber, J. and B. Koszegi. (2000). "Is addiction 'rational'? Theory and evidence.".NBER Working Paper 7507.

[23] Hall, R. E. (1978), "Stochastic implications of the life cycle-permanent income hypothesis". *Journal of Political Economy,* **86** (6), pp. 971-87.

[24] Harris C. and D. Laibson (2001), "Dynamic Choices of Hyperbolic Consumers" *Econometrica*, **69**, 935-957

[25] Laibson, D. (1997) "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics*, 112, pp. 443-77.

[26] Laibson, D. (1998) "Life-Cycle Consumption and Hyperbolic Discount Functions", *European Economic Review,*. **42**: 861-871.

[27] O'Donoghue, T. and M. Rabin. (1999). "Incentives for procrastinators." *Quarterly Journal of Economics*, 114:3, pp. 769-816.

[28] O'Donoghue, T. and M. Rabin. (2001). "Choice and procrastination." *Quarterly Journal of Economics*, 116:1, pp. 121-60.

[29] Phelps, E.S. and R. Pollak. 1968. "On Second-Best National Saving and Game-Equilibrium Growth." *Review of Economic Studies*, **35**, 185-199.

[30] Rockafellar, R.T (1970) *Convex Analysis*, Princeton, Princeton University Press.

[31] Rubinstein, A.(2003), "Economics and Psychology? The case of hyperbolic discounting", *International Economic Review*, **44**(4) 1207-1261.

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

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

[34] Selten, R. and W. Krischker. (1983). "Comparison of two theories for characteristic function experiments." In *Aspiration Levels in Bargaining and Economic Decision Making,* ed. R.Tietz, 259-264. Berlin: Springer.

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

[36] Strotz, R. H. (1956). "Myopia and inconsistency in dynamic utiilty maximization." *Review of Economic Studie*s, 23:3, 165-80.

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

[38] Varian, H., (1983), "Nonparametric tests of consumer behaviour" *Review of Economic Studies*, 50, 99-110

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

Appendix - Proofs

Proof of Lemma 1 and Lemma 2

Let β denote the additional discount factor associated with hyperbolic discounting. Denote savings as S_t and income as Y_t. In period t, person t maximises:

$$V(A_t) = \max_{c_t \ldots c_T} u(c_t) + \beta \sum_{i=1}^{T-t} \delta^i u(c_{t+i}) \tag{1}$$

$$\begin{aligned} s.t. \quad \mathbf{p}_j' \mathbf{c}_j + S_j &= Y_j + A_j \\ A_j &= (1 + r_j) S_{j-1} \qquad j = t, ..., T \end{aligned}$$

but believes that persons $\tau = t{+}1...T$ (his future selves) will use a different quasi hyperbolic discounting factor, $\widehat{\beta}$, and thus will maximise

$$W(A_\tau) = \max_{c_\tau \ldots c_T} u(\mathbf{c}_\tau) + \widehat{\beta} \sum_{i=1}^{T-\tau} \delta^i u(\mathbf{c}_{\tau+i}) \tag{2}$$

$$\begin{aligned} s.t. \quad \mathbf{p}_j' \mathbf{c}_j + S_j &= Y_j + A_j \\ A_j &= (1 + r_j) S_{j-1} \qquad j = t+1, ..., T \end{aligned}$$

We can re-write the value function in equation (1) as

$$\begin{aligned} V(A_t) &= \max_{\mathbf{c}_t \ldots \mathbf{c}_T} u(\mathbf{c}_t) + \delta u(\mathbf{c}_{t+1}) + \beta \sum_{i=2}^{T-t} \delta^i u(\mathbf{c}_{t+i}) + \beta \delta u(\mathbf{c}_{t+1}) - \delta u(\mathbf{c}_{t+1}) \\ &= \max_{c_t} u(c_t) + \delta \left[\frac{\beta}{\widehat{\beta}} W(A_{t+1}) - \beta \left(\frac{1}{\widehat{\beta}} - 1 \right) u(c_{t+1}) \right] \\ s.t. \quad A_{t+1} &= (1 + r_t) \left(Y_t + A_t - \mathbf{p}_t' \mathbf{c}_t \right) \end{aligned} \tag{3}$$

and that in equation (2) as

$$\begin{aligned} W(A_\tau) &= \max_{\mathbf{c}_\tau \ldots \mathbf{c}_T} u(\mathbf{c}_\tau) + \delta u(\mathbf{c}_{\tau+1}) + \widehat{\beta} \sum_{i=2}^{T-\tau} \delta^i u(\mathbf{c}_{\tau+i}) + \widehat{\beta} \delta u(\mathbf{c}_{\tau+1}) - \delta u(\mathbf{c}_{\tau+1}) \\ &= \max_{c_\tau} u(c_\tau) + \delta \left[W(A_{\tau+1}) - \left(1 - \widehat{\beta} \right) u(c_{\tau+1}) \right] \\ s.t. \quad A_{\tau+1} &= (1 + r_\tau) \left(Y_\tau + A_\tau - \mathbf{p}_\tau' \mathbf{c}_\tau \right) \end{aligned} \tag{4}$$

The first order condition from equation (3) is

$$\frac{\partial u_t}{\partial c_t^k} - \delta (1 + r_{t+1}) p_t^k \left[\frac{\beta}{\widehat{\beta}} W_{A_{t+1}} - \beta \left(\frac{1}{\widehat{\beta}} - 1 \right) \sum_{k=1}^{K} \frac{\partial u}{\partial c_{t+1}^k} \frac{\partial c_{t+1}^k}{\partial A_{t+1}} \right] = 0 \qquad \forall k \tag{5}$$

and the envelope theorem gives

$$V_{A_t} = \delta\left(1 + r_{t+1}\right)\left[\frac{\beta}{\widehat{\beta}} W_{A_{t+1}} - \beta\left(\frac{1}{\widehat{\beta}} - 1\right)\sum_{k=1}^{K}\frac{\partial u}{\partial c_{t+1}^k}\frac{\partial c_{t+1}^k}{\partial A_{t+1}}\right] \qquad \forall k \tag{6}$$

Equations (5) and (6) give

$$p_t^k V_{A_t} = \frac{\partial u_t}{\partial c_t^k} \qquad \forall k \tag{7}$$

From beliefs about the future (equation (4)) we get first order conditions that person t believes will be used at $t+1$

$$\frac{\partial u}{\partial c_{t+1}^k} - \delta\left(1 + r_{t+2}\right) p_{t+1}^k\left[W_{A_{t+2}} - \left(1 - \widehat{\beta}\right)\sum_{k=1}^{K}\frac{\partial u}{\partial c_{t+2}^k}\frac{\partial c_{t+2}^k}{\partial A_{t+2}}\right] = 0 \qquad \forall k \tag{8}$$

and the envelope theorem gives

$$W_{A_{t+1}} = \delta\left(1 + r_{t+2}\right)\left[W_{A_{t+2}} - \left(1 - \widehat{\beta}\right)\sum_{k=1}^{K}\frac{\partial u}{\partial c_{t+2}^k}\frac{\partial c_{t+2}^k}{\partial A_{t+2}}\right] \tag{9}$$

Equations (8) and (9) give that person t believes that

$$p_{t+1}^k W_{A_{t+1}} = \frac{\partial u}{\partial c_{t+1}^k} \qquad \forall k \tag{10}$$

Substituting equation (10) back into equation (5) gives:

$$\frac{\partial u}{\partial c_t^k} = \delta\left(1 + r_{t+1}\right)\frac{p_t^k}{p_{t+1}^k}\left[\frac{\beta}{\widehat{\beta}} - \beta\left(\frac{1}{\widehat{\beta}} - 1\right)\sum_{k=1}^{K} p_{t+1}^k\frac{\partial c_{t+1}^k}{\partial A_{t+1}}\right]\frac{\partial u}{\partial c_{t+1}^k} \qquad \forall k \tag{11}$$

For the sophisticated agent, $\widehat{\beta} = \beta$ and hence equation (11) reduces to the Euler equation in Lemma 1.

Define

$$\mu_{t+1} = \sum_{k=1}^{K} p_{t+1}^k\frac{\partial c_{t+1}^k}{\partial A_{t+1}}$$

Note that this is the period $t+1$ marginal propensity to spend out of wealth. Given consumption in each period is normal then it follows that

$$\mu_t \in [0, 1] \qquad \forall t$$

Solving for $\partial u / \partial c_{t+1}^k$ gives

$$\frac{\partial u}{\partial c_{t+1}^k} = \frac{\partial u}{\partial c_t^k}\frac{1}{\delta\left(1 + r_{t+1}\right)}\frac{p_{t+1}^k}{p_t^k}\left[\frac{\beta}{\widehat{\beta}} - \beta\left(\frac{1}{\widehat{\beta}} - 1\right)\mu_{t+1}\right]^{-1} \qquad \forall k$$

and solving recursively gives

$$\frac{\partial u}{\partial c_t^k} = \frac{\partial u}{\partial c_0^k} \frac{p_t^k}{p_0^k} \frac{1}{\delta^t} \prod_{i=1}^{t} \frac{1}{(1+r_i)} \prod_{i=1}^{t} \frac{1}{\left[\frac{\beta}{\widehat{\beta}} - \beta\left(\frac{1}{\widehat{\beta}} - 1\right)\mu_i\right]} \qquad \forall k$$

Using (7) dated in $t = 0$ and denoting $V_{A_0} = \lambda$

$$\lambda = \frac{\partial u}{\partial c_0^k} \frac{1}{p_0^k} \qquad \forall k$$

On substitution this gives the condition

$$\frac{\partial u}{\partial c_t^k} = \lambda \frac{1}{\delta^t} \rho_t^k \prod_{i=1}^{t} \frac{1}{\left[\frac{\beta}{\widehat{\beta}} - \beta\left(\frac{1}{\widehat{\beta}} - 1\right)\mu_i\right]} \qquad \forall k \tag{12}$$

For the sophisticated agent, $\widehat{\beta} = \beta$ and hence equation (12) reduces to

$$\frac{\partial u}{\partial c_t^k} = \lambda \frac{1}{\delta^t} \rho_t^k \prod_{i=1}^{t} \frac{1}{[1 - (1-\beta)\mu_i]} \qquad \forall k$$

which is Definition 1. Reinserting the expression for μ_i gives the first equation of Lemma 1.

For the naive agent, $\widehat{\beta} = 1$ and hence equation (12) reduces to

$$\frac{\partial u}{\partial c_t^k} = \lambda (\beta\delta)^{-t} \rho_t^k \qquad \forall k$$

which is the first equation of Lemma 2.■

Proof of Proposition 1

$(T) \Rightarrow (R)$

Definition 1 and Lemma 1 state that

$$\frac{\partial u}{\partial c_t^k} = \lambda \frac{1}{\delta^t} \rho_t^k \prod_{i=1}^{t} \frac{1}{[1 - (1-\beta)\mu_i]} \qquad \forall k$$

Denote

$$\psi_t = \prod_{i=1}^{t} \frac{1}{[1 - (1-\beta)\mu_i]}$$

Because $\beta \in [0,1]$ and $\mu_t \in [0,1]$ this imples that

$$\frac{1}{[1 - (1-\beta)\mu_i]} \geq 1$$

Therefore

$$\psi_t \geq \psi_{t-1} \geq 1$$

which is condition $(RS2)$.

Rewriting the first order condition on the equilibrium path in vector notation gives

$$\nabla u\left(\mathbf{c}_t\right) = \lambda \frac{1}{\delta^t} \boldsymbol{\rho}_t \psi_t$$

Concavity of the instantaneous utility function gives

$$u\left(\mathbf{c}_s\right) \leq u\left(\mathbf{c}_t\right) + \nabla u\left(\mathbf{c}_t\right)' \left(\mathbf{c}_s - \mathbf{c}_t\right) \qquad \forall s, t \in \tau$$

Substituting in the first order conditions means there exist real numbers $\{u_t\}_{t\in\tau}$ such that

$$u_s \leq u_t + \lambda \frac{\psi_t}{\delta^t} \boldsymbol{\rho}_t' \left(\mathbf{c}_s - \mathbf{c}_t\right)$$

Sum over any subset of observations $\sigma \subseteq \{1, ..., T\}$ gives

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

Since $\lambda > 0$ this implies

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

which is condition $(RS1)$

$(R) \Rightarrow (T)$

Restriction $(RS1)$ is a cyclical monotonicity condition (Rockafellar, 1970, Theorem 24.8). Cyclical monotonicity for the data $\left\{\frac{\psi_t}{\delta^t}\rho_t, \mathbf{c}_t\right\}_{t\in\tau}$ implies that there exists a concave, strictly increasing (utility) function $u\left(\mathbf{c}\right)$ such that

$$\frac{\partial u}{\partial c_t^k} = \frac{\rho_t^k}{\delta^t} \psi_t \qquad \forall k, t$$

(the non-differentiable case is similar but involves super-gradients). Setting $\lambda = 1$ we also have

$$\frac{\partial u}{\partial c_t^k} = \lambda \frac{\rho_t^k}{\delta^t} \psi_t \qquad \forall k, t$$

for all $t \in \tau$. We therefore have a suitable utility function and the parameters δ and λ. We now need to show that we can take the sequence

$$1 \leq \psi_t \leq \psi_{t+1}$$

and find $\beta \in [0, 1]$ and a sequence of $\mu_t \in [0, 1]$ such that

$$\psi_t = \prod_{i=1}^{t} \frac{1}{\left[1 - \left(1 - \beta\right) \mu_i\right]}$$

Given the sequence of ψ_t terms we can take the ratio

$$\frac{\psi_{t-1}}{\psi_t} = 1 - (1-\beta)\,\mu_t$$
$$1 - \frac{\psi_{t-1}}{\psi_t} = (1-\beta)\,\mu_t$$

and for $t = 1$ we have

$$\psi_1 = \frac{1}{[1-(1-\beta)\,\mu_1]}$$
$$1 - \frac{1}{\psi_1} = (1-\beta)\,\mu_1$$

Since $\mu_t \in [0,1]$this implies that we must have

$$0 < \beta < \frac{\psi_{t-1}}{\psi_t}$$

Since $\psi_{t-1} \leq \psi_t$ we know that this implies

$$(1-\beta)\,\mu_t \in [0,1]$$

This provides $T-1$ equations with T unknowns so we can always find a suitable β and μ_t for $t = 2, ..., T$. For $t = 1$ we have

$$\psi_1 = \frac{1}{[1-(1-\beta)\,\mu_1]}$$

which we can solve for μ_1 by substitution. There therefore exists a concave, strictly increasing (utility) function $u(\mathbf{c})$ and positive constants $\lambda, \beta, \{\mu_t\}_{t=1,...,T}$ such that

$$\frac{\partial u}{\partial c_t^k} = \lambda \frac{1}{\delta^t} \rho_t^k \prod_{i=1}^{t} \frac{1}{[1-(1-\beta)\,\mu_i]} \qquad \forall k, t$$

which is the definition of *rationalise* in in Definition 1■

Proof of Corollary 1

Restriction $(R1)$ is a cyclical monotonicity condition (Rockafellar, 1970, Theorem 24.8). By Afriat's Theorem cyclical monotonicity for the data $\left\{\frac{\psi_t}{\delta^t}\rho_t, \mathbf{c}_t\right\}_{t\in\tau}$ implies that there exists a concave, strictly increasing (utility) function $u(\mathbf{c})$ such that the data $\left\{\frac{\psi_t}{\delta^t}\rho_t, \mathbf{c}_t\right\}_{t\in\tau}$ satisfies GARP. Since

$$\frac{\psi_t}{\delta^t}\boldsymbol{\rho}_t' c_t \geq \frac{\psi_t}{\delta^t}\boldsymbol{\rho}_t' \mathbf{c}_t \Rightarrow \mathbf{p}_t' \mathbf{c}_t \geq \mathbf{p}_t' \mathbf{c}_t$$

it follows that the data $\left\{\frac{\psi_t}{\delta^t}\rho_t, \mathbf{c}_t\right\}_{t\in\tau}$ satisfies GARP so do the data $\{\mathbf{p}_t, \mathbf{c}_t\}_{t\in\tau}$.■

Proof of Corollary 2

Immediate by comparison with Browning (Proposition 1, 1989) (augmented to allow for time discounting which Browning does not explicitly consider).■

Proof of Corollary 3

From the proof of sufficiency in Proposition we show that we can take the sequence of constants $1 \leq \psi_t \leq \psi_{t+1}$and derive a set of $T-1$ equations with T unknowns such that

$$1 - \frac{\psi_{t-1}}{\psi_t} = (1-\beta)\,\mu_t$$

Thus β is under-determined.■

Proof of Proposition 2

$(T) \Rightarrow (R)$
Definition 2 and Lemma 2 state that

$$\frac{\partial u}{\partial c_t^k} = \lambda \frac{1}{(\beta\delta)^t} \rho_t^k \qquad\qquad \forall k$$

Define $d = \beta\delta$. Rewriting this first order condition in vector notation gives

$$\nabla u\left(\mathbf{c}_t\right) = \lambda \frac{1}{d^t} \rho_t$$

Concavity of the instantaneous utility function

$$u\left(\mathbf{c}_s\right) \leq u\left(\mathbf{c}_t\right) + \nabla u\left(\mathbf{c}_t\right)'\left(\mathbf{c}_s - \mathbf{c}_t\right) \qquad \forall s,t \in \tau$$

Substituting in the first order conditions means there exist real numbers $\{u_t\}_{t\in\tau}$ such that

$$u_s \leq u_t + \lambda \frac{1}{d^t} \boldsymbol{\rho}_t'\left(\mathbf{c}_s - \mathbf{c}_t\right)$$

Sum over any subset of observations $\sigma \subseteq \{1, ..., T\}$ gives

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

Since $\lambda > 0$ this implies

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

which is condition $(RN1)$
$(R) \Rightarrow (T)$
Restriction $(RN1)$ is a cyclical monotonicity condition (Rockafellar, 1970, Theorem 24.8).

Cyclical monotonicity for the data $\left\{\frac{1}{d^t}\rho_t, \mathbf{c}_t\right\}_{t\in\tau}$ implies that there exists a concave, strictly increasing (utility) function $u(\mathbf{c})$ such that

$$\frac{\partial u}{\partial c_t^k} = \frac{\rho_t^k}{d^t} \qquad \forall k, t$$

(the non-differentiable case is similar but involves super-gradients). Setting $\lambda = 1$ we also have

$$\frac{\partial u}{\partial c_t^k} = \lambda\frac{\rho_t^k}{d^t} \qquad \forall k, t$$

for all $t \in \tau$. We therefore have a suitable utility function and the parameters d and λ. Given any d we can find two real numbers β and δ such that

$$\beta\delta = d$$

There therefore exists a concave, strictly increasing (utility) function $u(\mathbf{c})$ and positive constants λ, β and δ such that

$$\frac{\partial u}{\partial c_t^k} = \lambda\frac{1}{(\beta\delta)^t}\rho_t^k \qquad \forall k, t$$

which is the definition of *rationalise* in in Definition 2.■

Proof of Corollary 3.

Setting $d = \beta\delta$ this is Immediate by comparison with Browning (Proposition 1, 1989) (augmented to allow for time discounting which Browning does not explicitly consider).■