# Inattentive Consumers

Ricardo Reis[*]
Harvard University

This draft: February 2004
First draft: September 2003

**Abstract**

This paper studies the consumption decisions of agents who face costs of acquiring, absorbing and processing information. These consumers rationally choose to only sporadically update their information and re-compute their optimal consumption plans. In between updating dates, they remain inattentive. This behavior implies that news disperses slowly throughout the population, so events have a gradual and delayed effect on aggregate consumption. The model predicts that aggregate consumption adjusts slowly to shocks and is excessively sensitive and excessively smooth relative to income. In addition, individual consumption is sensitive to ordinary and unexpected past news, but it is not sensitive to extraordinary or predictable events. The model further predicts that some people rationally choose to not plan, live hand-to-mouth, and save less, while other people sporadically update their plans. The longer are these plans, the more they save. Evidence using U.S. aggregate and microeconomic data generally supports these predictions.

[*]I am extremely grateful to N. Gregory Mankiw for mentoring my graduate studies, and to Alberto Alesina, Robert Barro, and David Laibson for their constant intellectual stimulation and priceless advice. Andrew Abel, Susanto Basu, John Campbell, Larry Christiano, Mariana Colacelli, Benjamin Friedman, Jens Hilscher, Yves Nosbusch, David Romer, Monica Singhal, Adam Szeidl, Bryce Ward, Justin Wolfers, and numerous seminar participants offered useful comments on this paper. The Fundacao Ciencia e Tecnologia, Praxis XXI and the Eliot Memorial fellowship provided financial support. Contact: reis@fas.harvard.edu.

*"Attention as the Scarce Resource. [...] Many of the central issues of our time are questions of how we use limited information and limited computational ability to deal with enormous problems whose shape we barely understand."*

*Herbert A. Simon (1978, page 13)*

*"Perhaps it is not surprising that many people do not report an expectation given the costs of it."*

*Sherwin Rosen (1990, page 284)*

# 1 Introduction

Most economists would agree that a rational consumer sets the marginal utility of consuming in the present equal to the discounted marginal utility of consuming in the future times the price of present relative to future consumption. If the future is uncertain, it is expected marginal utility that is relevant, and a crucial component of a model of consumption specifies how agents form expectations. In a pioneering contribution, Hall (1978) assumes that agents form expectations rationally in the Muth sense: they know the entire structure of the economy and have full information on all the relevant variables needed to form statistically optimal forecasts. Rational expectations leads to the prediction that consumption should be a martingale: consumption growth should not be predictable over time. Hall's finding that post-war U.S. aggregate consumption approximately follows a random walk was an early empirical success of rational expectations modelling.

Over the past 25 years though, many papers have found problems with the Hall model. Deviations of aggregate consumption from a martingale in the data have been convincingly established, taking the form of either excess sensitivity of consumption to past known information, or excess smoothness in response to permanent income shocks.[1] Campbell and Mankiw (1989, 1990) illustrate these failures by showing that if the world is partially populated by rational expectations agents, then there must be as many irrational consumers who consume their current income every period, in order to match the data on aggregate consumption.

This paper revisits the modelling of expectation formation by consumers. With rational expectations, agents have an unbounded ability to absorb and process information on all the relevant characteristics of the economy, and an unbounded ability to think through this information and calculate optimal forecasts and actions. I assume instead that it is costly for agents to acquire, absorb, and process information in forming expectations and making decisions. In a dynamic setting, while agents with rational expectations undertake these costly activities at every instant in time, in this paper agents rationally choose to update their information and plans infrequently: Expectations are rational, but are only sporadically updated. Following a new event, many agents will be unaware

[1] Consumption is excessively sensitive (Flavin, 1981) if future consumption growth depends on lagged information. It is excessively smooth (Deaton, 1987) if it does not respond one-to-one to shocks to permanent income, and thus is smoother than permanent income.

of the news for a while, and will continue following their outdated plans, only eventually updating their expectations. Agents are inattentive and the information in the economy is sticky, gradually dissipating over time to the entire population. Consumption in turn is excessively sensitive, since when agents adjust plans and consumption, they react to all the information (present and past) since their last adjustment date. Consumption is also excessively smooth, or insufficiently sensitive to permanent income shocks, since only a fraction of agents are attentive when there is a shock to permanent income and react to it instantly.

Beyond generating predictions that match the data on individual and aggregate consumption, a further contribution of this paper is that it provides a micro-foundation for time-contingent adjustment rules. Because inattentiveness emphasizes the costs of observing and processing the state of the economy, it naturally justifies agents adjusting to news at certain dates regardless of the state of the economy at those dates. However, when she plans, the inattentive agent optimally decides when to plan again, taking into account the state of the world at the *current* planning date. In traditional models with time-contingent adjustment rules, the adjustment dates are set regardless of the state of the world at *any* date. The inattentiveness model therefore implies recursively time-contingent adjustments, independent of the state of the economy at that date, but dependent on the state of the economy at the last adjustment date. In some special cases, this reduces to the standard time-contingent adjustment model, but even when it does not, the model retains the tractability that has made time-contingent adjustment so popular in the literature.

A few papers have recently explored the potential of modelling inattentiveness. Gabaix and Laibson (2001) assume that investors update their portfolio decisions infrequently, and show that this can explain the puzzling premium of equity over bond returns. Mankiw and Reis (2002, 2003) study inattentiveness on the part of price-setting firms and show that the resulting model of the Phillips curve matches well the dynamics of inflation and output that we observe in the data. Relative to these papers, this paper differs by focusing on consumption decisions and deriving predictions for individual and aggregate consumption, which are empirically tested. Moreover, I do not assume that agents infrequently adjust their plans, but rather I derive this behavior as the optimal response to explicitly modelled costs of planning.

Sims (2003) and Moscarini (2004) develop an alternative model of rational inattention. Both use Shannon's information theory to model the costs of obtaining information and solve for the optimal choice of which pieces of information to pay attention to, and how to use these to infer the current state of the world. Their approach is very complementary to the one in this paper, since the models differ more in focus than in substance. Whereas Sims and Moscarini focus on the information problem facing agents, at the cost of simplifying the study of their real actions, this paper focuses on these real decisions, their interaction with inattentiveness, and in deriving predictions to contrast with data, at the cost of simplifying the information acquisition problem.

Recent empirical work using microeconomic data has also emphasized that most people are inattentive and that this affects their behavior. Lusardi (1999, 2002) and Ameriks, Caplin, and

Leahy (2003a) find that a significant fraction of survey respondents make financial plans infrequently (if at all) and that their planning behavior has a statistically significant and sizeable effect on the amount of wealth they have accumulated. This paper contributes to this literature a theoretical model of costly and infrequent planning. Inattentiveness rationalizes these authors' findings and suggests further implications to test using observations of individual behavior.

This paper is organized as follows. Section 2 presents a simple model of inattentiveness that highlights the intuition behind the results. Section 3 rigorously sets up the general problem of an agent facing costs of planning, and derives the optimality conditions describing consumption and planning behavior. It aggregates individual consumption decisions over many such agents to obtain the predictions of the model for the time-series of aggregate consumption, which will later be tested in the data. Section 4 solves the inattentive agent's problem analytically for a particular specification of preferences and uncertainty. This provides further implications and intuition on the effects of costly planning on savings and optimal inattentiveness. Next, I show that the inattentiveness model predicts that if the agent's costs of planning are above a certain threshold, she rationally chooses to never plan, and to live hand-to-mouth, consuming her income every period less a pre-determined amount. The theory section of this paper concludes by examining the response of an inattentive agent to "extraordinary events," which occasionally induce large changes in the environment she faces.

The paper then turns to testing the implications of the model. In Section 5, I use U.S. aggregate consumption data. I examine whether these data exhibit the slow adjustment and the stickiness of information that the model predicts, and I show that the model can generate the extent of sensitivity and smoothness with respect to income that we observe. Furthermore, I show that the data favors the inattentiveness model over the model of Campbell and Mankiw (1989, 1990). Section 6 discusses studies which have used microeconomic data, and shows that the inattentiveness model matches the existing evidence on the sensitivity of individual consumption to past information, the expectations of individuals, and their planning and savings behavior.

Section 7 compares the inattentiveness model with alternatives in the literature. Section 8 concludes by collecting the many theoretical results and empirical estimates in the paper into a coherent description of individual and aggregate consumption in the United States, and by discussing some directions for future research.

## 2 A simple model of inattentiveness

Consider the problem of an agent living forever in discrete time who consumes $c_t$ each period, obtaining utility given by the function $u(c_t)$, which is of the constant absolute risk aversion (CARA) form. This agent discounts future utility by the factor $\beta$, and each period she receives stochastic income $y_t$, which is normally distributed with mean $\bar{y}$ and variance $\sigma^2$. She earns returns on her assets, $a_t$, at the gross interest rate $R$ which equals $1/\beta$, so the subjective discount rate equals the net real interest rate.

Despite being fully rational and making optimal choices, the consumer must pay a monetary cost $K$ whenever she acquires information and makes optimal decisions. This can be thought of as the cost in money and time of obtaining information, processing and interpreting it, and deciding how to optimally act. Facing this cost, the agent must then choose when to plan. Her decision dates are denoted by $D(i)$, and for any $t$ between $D(i)$ and $D(i+1)$ the consumer follows the plan set at time $D(i)$. The problem of the agent therefore is:

$$\begin{aligned} V(a_0) &= \max_{\{c_t, s_t\}_{t=0}^{\infty}, \{D(i)\}_{i=1}^{\infty}} E\left[\sum_{t=0}^{\infty} \beta^t \left(-\frac{e^{-\alpha c_t}}{\alpha}\right)\right] \\ a_{t+1} &= Ra_t - c_t + y_t - K\iota(t), \end{aligned}$$

where $\iota(t)$ is an indicator function that equals 1 if $t = D(i)$ and is zero otherwise. Note that the agent can choose either consumption or savings $s_t = y_t - c_t$. For now, I focus on picking optimal consumption, but the savings alternative will become relevant later in this Section.

The first-order condition with respect to consumption between two periods $t$ and $t+s$ which are in between planning dates ($D(i) \leq t < t+s < D(i+1)$) implies:

$$c_t = c_{t+s}. \tag{1}$$

With inattentiveness, consumption stays constant, or more generally, *consumption follows a pre-determined path in between planning dates.*[2] Since the consumer does not update her information between $t$ and $t+s$, consumption evolves deterministically between these two periods. The first-order condition with respect to consumption at two planning periods is:

$$e^{-\alpha c_{D(i)}} = E_{D(i)}\left[e^{-\alpha c_{D(i+1)}}\right]. \tag{2}$$

*Consumption at planning dates is determined by a stochastic Euler equation.* This is of the same form as in the problem without planning costs. Yet now it holds only between two planning dates rather than always, since only at these dates is new information observed by the consumer.

To proceed further, I guess that optimal consumption at adjustment dates is linear in the level of wealth $c_{D(i)} = A + Bw_{D(i)}$. Wealth includes not only her financial assets but also her human capital which includes current and expected future labor income: $w_t = a_t + (y_t - \bar{y})/R + \bar{y}/(R-1)$. The coefficients $A$ and $B$ are to be determined. Iterating on the budget constraint between time 0 and the decision date $D$:

$$w_D = R^D w_0 - c_0 \frac{1-R^D}{1-R} + \sum_{j=0}^{D-1} R^{j-1}\left(y_{D-j} - \bar{y}\right) - K. \tag{3}$$

[2] In the psychology literature, Bargh and Chartrand (1999) describe this as "the unbearable automaticity of being."

Since income is normally distributed, then so will be wealth, and given the guess that consumption is linear in wealth, consumption is also normally distributed. Then, $\exp(-\alpha c_D)$ is log-normally distributed so the log of its expectation equals $-\alpha E_0[c_D] + (\alpha^2/2)Var_0[c_D]$. Calculating these moments of consumption using (3), and simplifying the first-order condition (2) to solve for the unknowns in the consumption function, gives:

$$c_t = \underbrace{(R-1)}_{B} w_D \underbrace{-\frac{(R-1)K}{R^D-1} - \frac{\alpha\sigma^2}{2R^2}\frac{(R-1)\left(R^D+1\right)}{R+1}}_{A},$$

which is linear in wealth, confirming the initial guess. If she faces higher planning costs, *the agent plans less often and saves more.* The longer the agent remains inattentive, the larger is her exposure to risk, since she is not reacting to shocks as they occur. This larger risk leads in turn to higher precautionary savings by the agent to safeguard herself against a sequence of bad income shocks.

Given the solution for consumption, the value function of the agent is then given by:

$$V(a_0) = \max_D \left\{ -\frac{R}{\alpha(R-1)} e^{-\alpha A - \alpha(R-1)w_0} \right\},$$

Performing the maximization with respect to D gives:

$$\hat{D} = \frac{\ln\left(1 + \sqrt{\frac{2R^2(R+1)K}{\alpha\sigma^2}}\right)}{\ln(R)}. \tag{4}$$

The optimal discrete time inattentiveness interval, $D^*$, is the integer just before or after $\hat{D}$ that yields the higher value. Equation (4) shows that if the costs of planning are close to zero, $\hat{D}$ is of order $\sqrt{K}$, so *second-order costs of planning lead to first-order periods of inattentiveness.* An agent with even very small costs of planning can be inattentive for a long time since her behavior is close to the full information behavior, so the utility losses from inattentiveness are small. Equation (4) also shows that she will be *inattentive for longer the lower is risk aversion and the lower is income volatility.* The lower these are, the smaller is the effective cost of being inattentive driven by her exposure to risk, and thus the less frequently she adjusts her plans. Moreover *the larger is the interest rate $R$, the shorter the inattentiveness.* While the agent is inattentive, she is not adjusting her savings optimally. The larger the interest rate, the larger the impact of these inefficient savings on her future asset position. Facing a large interest rate, the agent will choose to update more often to avoid her asset position becoming severely sub-optimal.

So far, I have been solving the problem of an agent who chooses consumption plans. Yet, she could instead set plans for her savings. If the agent has full information or if there is no income uncertainty, then the two are indistinguishable. But if the agent is not monitoring her income every instant, she must choose to either set a plan for consumption and let savings adjust to the shocks, or to set a plan for savings and let consumption adjust.

An inattentive saver has $c_t = y_t - s_t$, so since $s_t$ follows a pre-determined path, *the inattentive saver behaves like a hand-to-mouth consumer*, every period consuming her income up to a pre-determined amount. As long as $K$ is not too small, *the inattentive saver rationally chooses to never plan*, every period just consuming her income less a constant amount. To see this, consider an agent who at some period decides not to update her plans. Since assets evolve deterministically and income is white noise, note that in the next period the agent is facing exactly the same problem and thus she must again choose not to update her plans. Iterating on this logic to infinity shows that the saver will either always be attentive, or never update her plans. If $K$ is not too small, she will choose to never plan.

The inattentive consumer is worse off the larger is $K$, whereas the inattentive saver never plans and so is unaffected by this cost. Therefore, the agent will only choose savings plans if the costs of planning are large enough. This gives the following characterization of behavior in an inattentive economy: some agents have high costs of planning and optimally choose to live hand-to-mouth consuming their income every period and never making plans. The other agents, who have lower planning costs, opt instead for having consumption follow infrequently-updated plans.

In an economy populated by many inattentive consumers, their consumption decisions can be aggregated to obtain implications for aggregate consumption. This requires a description of how agents differ in the economy, and for now I make the simplest assumption: I assume that all agents are identical so all choose the same $D^*$, but they are uniformly staggered with respect to their planning dates.[3] Between two successive periods, a fraction $(D^* - 1)/D^*$ of agents will not change their consumption, while a fraction $1/D^*$ updates consumption responding to the information that arrived over the last $D^*$ periods. If $C_{t+1}$ denotes aggregate consumption:

$$C_{t+1} - C_t = \frac{\mu}{D^*} e_{t+1} + \frac{\mu}{D^*} e_t + ... + \frac{\mu}{D^*} e_{t-D^*+1}, \tag{5}$$

where $e_t$ denotes the information that arrived at period $t$, and $\mu$ is the marginal propensity to consume out of that information. Since income is a relevant piece of information, equation (5) shows that *the change in aggregate consumption is sensitive to income up to $D^*$ periods earlier.* If the inattentiveness model describes the data, then conventional tests of excess sensitivity will find that consumption is excessively sensitive to past income, but only up to $D^*$ periods ago. Moreover, since only a fraction of the agents react contemporaneously to changes in income, *consumption will be smoother than income.* From the perspective of the Hall (1978) model, consumption will be excessively smooth.

Equation (5) also shows that the change in consumption should be a moving average process of order $D^*$. *Inattentive aggregate consumption adjusts slowly to shocks*, with a reaction that builds up over time. Given the arrival of a piece of news, only a few agents will be attentive and react instantly to it. The remaining consumers gradually update their plans and adjust consumption so

[3] Section 3 will allow for a general distribution of individual characteristics and optimal planning dates.

*information disseminates slowly*, and affects aggregate consumption gradually over time.

Finally, a consumer responds to present and past shocks only if she could not predict them when she last planned. *Past predictable events do not affect present individual consumption changes.* Moreover, if some variables only move infrequently, so the cost of monitoring them is very small, and if movement in these variables leads to large changes in the agent's income, she will pay attention to these variables and react to them instantly. For instance, if the agent suddenly becomes unemployed or wins the lottery, she responds immediately to these easily observed and significant shocks. *Past extraordinary events do not affect present individual consumption changes.*

Summarized and simplified, these are the main results from the theory of inattentive consumption. The next two Sections set up and solve the problem of the inattentive consumer more generally, deriving a set of predictions that Sections 5 and 6 contrast with the data.

# 3 The general inattentiveness model

## 3.1 The set-up of the problem

I model the problem of the inattentive consumer in continuous time, so that the planning dates are chosen from a continuous set. Time is indexed by $t$ on the positive real line while the decision periods are denoted by $D(i)$ where $i \in \mathbb{N}_0$ orders the decision times so that $D(i+1) \geq D(i)$ for all $i$ with $D(0) \equiv 0$. If $d(i)$ denotes the time until the next adjustment, defined recursively as $d(i) = D(i) - D(i-1)$, it is clearly equivalent for the agent to choose the calendar dates of planning, $D(i)$, or the inattentiveness intervals, $d(i)$.

The economy is populated by many infinitely-lived consumers, which can be interpreted as the result of altruistic links between generations. Each instant, an agent consumes an amount of goods $c_t$, which yields an amount of utility given by the function $u(c_t)$. This function is assumed to be continuous, everywhere twice differentiable, increasing and concave. Future utility is discounted at the positive rate $\rho$, either due to impatience or because the current generation discounts the well-being of future generations relative to its own.

Each instant, the agent receives an income flow $y(x)$, and her assets $a_t$ earn returns at the interest rate $r$. The flow budget constraint is $da_t = (ra_t - c_t + y(x_t))dt$, stating that at each instant, assets increase by the interest earned plus new savings. Borrowing is constrained by the condition that all debts must be repaid, so the agent cannot run Ponzi schemes rolling over debt forever: $\lim_{t\to\infty} e^{-rt} a_t \geq 0$. Income is a function of a state vector $x_t$, of potentially very large dimension, which is generated by a continuous time stochastic process defined on a standard filtered probability space $\{X, F, P\}$ where $X$ is the set of possible states, $F$ is the filtration $F = \{F_t, t \geq 0\}$ where $F_t$ is the $\sigma$-algebra through which information on $x_t$ is revealed, and $P$ is the probability measure on $F$. I will write $y(x_t)$ more compactly as $y_t$. The notation $E_k[.]$ will be used to denote the expectation conditional on information up until time $k$: $E_k[y_t] = \int y_t dP(F_k)$. I further assume that the state vector has the Markov property, and, without loss of generality, that it is arranged in such a way that it is first-order Markov. Therefore, a sufficient statistic for the probability of any state $y_t \in Y$ from the perspective of time $k < t$ is the state vector at time $k$: $P(y_t \mid F_k) = P(y_t \mid x_k)$.

The consumer's choice of planning dates defines a new filtration $\Im = \{\Im_t, t \geq 0\}$ such that $\Im_t = F_{D(i)}$ for $t \in [D(i), D(i+1))$. When the consumer writes a plan at time $D(i-1)$, she decides on a consumption sequence until the next adjustment, $c^i = [c_{D(i-1)}, c_{D(i)})$, and on when to plan again, $D(i)$. The restriction embodied in the existence of a plan is that both of these must be contingent on the information available at time $D(i-1)$: If $\{c, D\} \equiv \{c^i, D(i)\}_{i=1}^{\infty}$, then both $c$ and $D$ must be $\Im$-adapted processes.

Whenever she plans, the consumer incurs a fixed monetary cost given by $K_t \equiv K(x_t)$, which can be stochastic and time-varying. This cost can be interpreted as the money spent acquiring information and paying a financial advisor to interpret the information and compute the optimal financial plan, or it could stand for the opportunity cost of taking time to plan. If the utility function has an extra additively separable term that is a function of leisure, and planning takes time away from labor supply at a stochastic market wage, the wages foregone at times of planning will appear as a $K_t$ cost.[4] If the consumer enters period $D(i)$ with assets given by $a^-_{D(i)}$, her wealth then changes discontinuously to $a^+_{D(i)} = a^-_{D(i)} - K_{D(i)}$.[5] Formally, $a^-_{D(i)}$ is the left-hand side time limit of assets, while $a^+_{D(i)}$ is the right-hand side limit, and they differ by the fixed cost $K_{D(i)}$.

The problem of the consumer can then be compactly written as:

$$\max_{\{c,D\}} E_0 \left[ \sum_{i=0}^{\infty} \int_{D(i)}^{D(i+1)} e^{-\rho t} u(c_t) dt \right] \tag{6}$$

$$\begin{aligned} \text{s.t.} \quad &: \quad \{c, D\} \text{ are } \Im\text{-adapted}, && (7) \\ da_t &= (ra_t - c_t + y_t)dt, && (8) \\ a^+_{D(i)} &= a^-_{D(i)} - K_{D(i)}, \text{for all } i \in \mathbb{N}_0, && (9) \\ \lim_{t\to\infty} e^{-rt} a_t &\geq 0, && (10) \end{aligned}$$

with initial conditions $a_0$, $x_0$. It is difficult to solve this problem both because it is hard to impose the measurability restriction (7) and because of the discontinuity in the level of assets at the planning dates (9). To make progress, the problem must be re-stated in a more convenient form.

Start by integrating the law of motion for assets (8) between $D(i)$ and $D(i+1)$, and replace $a^-_{D(i)}$ by $a^+_{D(i)} + K_{D(i)}$, using (9). This gives:

$$a^+_{D(i+1)} = e^{rd(i)} \left( a^+_{D(i)} - \int_0^{d(i)} e^{-rt} c_{D(i)+t} dt + \int_0^{d(i)} e^{-rt} y_{D(i)+t} dt \right) - K_{D(i+1)},$$

[4]Modelling the costs of planning as additive reductions in utility, because some people may find the process annoying or frustrating, leads to similar results to the ones discussed in this paper.

[5]Implicit in this setup is the assumption that while it is costly to re-write new plans, this can be done in an instant of time. I could assume instead that it takes a fixed interval of time to devise a plan. While this would require some modifications to the analysis that follows, it would not affect the main conclusions.

thus eliminating the $a_t^-$ variables, so that only $a_t^+$'s are left. Moreover, realize that there is a recursive structure between planning dates so the cumbersome time indices can be dropped by denoting $a^+_{D(i)}$ by $a$ and $a^+_{D(i+1)}$ by $a'$, and similarly $x_{D(i)}$ by $x$ and $x_{D(i+1)}$ by $x'$. Next, let $V(a, x)$ be the value function associated with this problem. The state vector is $(a,\, x)$ since the law of motion for assets and the Markov assumption for the state vector imply that $(a, x)$ is a sufficient statistic for the uncertainty facing the agent until the next planning date.

With these changes, the problem in (6)-(10) becomes:

$$V(a,x) = \max_{c,d} \int_0^d e^{-\rho t} u(c_t)dt + e^{-\rho d} E\left[V(a', x')\right] \tag{11}$$

$$\text{subject to } a' = e^{rd}\left(a - \int_0^d e^{-rt} c_t dt + \int_0^d e^{-rt} y_t dt\right) - K'. \tag{12}$$

The measurability constraints are imposed by having passed the expectations operator through $\{c, d\}$, so that these choices are made conditional only on the information in $(a, x)$. The only unknown at this planning date is what assets and accumulated income will be by the next planning date. As for the initial conditions, note that since there is planning at time 0, a cost $K$ is incurred at this date so the initial post-planning asset level is $a_0 - K_0$.[6]

The solution to the problem in (11)-(12) will be a pair of functions, $c_t(a, x)$ and $d(a, x)$, determining optimal consumption from time 0 to time $d$ and when the next planning will take place. Consumption at any date between 0 and $d$ is inattentive since it is chosen regardless of the state of the world at that date. In turn, the date of the next adjustment does not depend on the state at that date – adjustment is not state-contingent. However, adjustment is also not purely time-contingent, since the date of the next adjustment depends on the state of the world at the last adjustment. For lack of better words, I describe adjustment with inattentiveness as *recursively time-contingent*: it occurs at a pre-set date which depends recursively on the state at the past planning date. In some cases, $d(a, x)$ might be independent of $(a, x)$, in which case the inattentiveness model leads to purely time-contingent adjustment.

The problem in (11)-(12) is a familiar dynamic programming problem, and if the utility function is bounded, arguments similar to those in Stokey et al. (1989) prove the existence of a solution and give the necessary restrictions for uniqueness of this solution. With an unbounded utility function, the problem of the inattentive consumer has one additional technical difficulty relative to the full information problem. To illustrate it, consider the case in which $u(c)$ is of the constant relative risk aversion form, so that marginal utility tends to $-\infty$ as $c$ approaches 0, and $y_t$ is an arithmetic Brownian motion, which has infinite local variation. If the agent is inattentive for even an instant

[6] Some related problems have been studied in the engineering literature on "sampled-data control systems" or "digital control." The two most related problems to the one in this paper, are optimal control problems in which the state is only observed at exogenously given infrequent dates (Franklin et al., 1990), and problems of determining how often to sample a continuous time stochastic process to maximize the information content of the messages (Miller and Runggaldier, 1997).

then with positive probability her income may fall to a very large negative number inducing her to borrow a very large amount. Satisfying the intertemporal budget constraint would then require setting consumption so low that utility would be unbounded from below, so being inattentive could never be optimal. On the other hand, being always attentive cannot be optimal since it involves an infinite expenditure of resources in planning, so the problem does not have a well-defined solution.

There are two ways to get around this problem. A reasonable solution is to simply assume that income follows a stochastic process that cannot fall to minus infinity with positive probability every instant. If the reader of this paper knew that in the instant it takes to read this sentence while being inattentive to her income, her life circumstances could change so suddenly as to throw her into a life of bondage, she would never read anything at all and would go through life doing nothing but monitoring income every instant. This is not the case for most people, so it is reasonable to assume it is also not the case for the inattentive consumer. A second solution to the problem is to retain the mathematical convenience of using Wiener processes for income, while specifying preferences that do not run into the problem. For instance, assuming that the utility function is of the constant absolute risk aversion form and allowing consumption to sometimes be negative, as I did in Section 2, is enough to guarantee a well-defined optimization problem.[7]

### 3.2 Characterizing the solution

Taking the derivative of (11) with respect to $d$ and setting it equal to zero gives:

$$u(c_d) = \rho E\left[V(a', x')\right] - \frac{\partial}{\partial d} E\left[V(a', x')\right].$$

This first-order condition states that the agent plans to adjust when the marginal cost of adjusting equals the marginal benefit of doing so. On the left-hand side is the marginal cost of adjusting, which is the utility the agent would get if she kept to her outdated consumption plan. On the right-hand side is the marginal benefit of adjusting at time $d$. The first term is the present flow value of having re-planned and obtained new information, while the second term is the benefit from acquiring this information at $d$ rather than in the next instant when this value has fallen. The cost $K$ enters the first-order condition on the right-hand side by lowering the benefits of planning through the fall in assets by $K$ to $a'$ at the planning date.

The first-order conditions with respect to $c_t$ are:

$$u'(c_t) = e^{(r-\rho)(d-t)} E\left[V_a(a', x')\right], \text{ for } t \in [0, d), \tag{13}$$

where $u'(.)$ is the first derivative of the utility function and $V_a(.)$ is the derivative of the value function with respect to its first argument. Using the fact that $e^{(r-\rho)d} E\left[V_a(a', x')\right]$ is independent

[7] The literature on digital control takes this second route by focusing on problems with a quadratic objective function or on $H_\infty$ control (Chen and Francis, 1995).

of time, take logs and derivatives with respect to time of equation (13) to find that for $t \in [0, d)$:

$$\frac{du'(c_t)/dt}{u'(c_t)} = -(r - \rho).$$

This is the famous Ramsey (1928) rule. The rate of change of the marginal utility of consumption equals the gap between the agent's impatience and the riskless rate of return.

A third optimality condition comes from the envelope theorem:

$$V_a(a, x) = e^{(r-\rho)d} E\left[V_a(a', x')\right]. \tag{14}$$

Combining (13) and (14) to substitute out the value function for the utility functions gives:

$$u'(c_0) = e^{(r-\rho)d} E\left[u'(c_d)\right].$$

This is also a familiar expression. It is the stochastic Euler equation that arises in the study of consumption under uncertainty. At time 0, a dollar can be used to consume goods yielding $u'(c_0)$ units of utility, or instead it can be saved returning $e^{rd}$ dollars in $d$ periods, which can be used to consume goods at $d$ giving $e^{-\rho d}u'(c_d)$ units of utility in time 0 utility units. Optimal behavior requires that these two uses of funds give the same benefit.

The dynamics of inattentive consumption over time are therefore simple to describe. During the intervals of inattentiveness, consumption evolves just like in the standard consumer problem with certainty. At adjustment dates, consumption evolves just like in the standard consumer problem with uncertainty. Intuitively, between adjustments the agent is not receiving new information so it is as if there is no uncertainty; at adjustments, information is revealed and her optimal choices incorporate it. This is summarized in the following result:

**Proposition 1** *If the consumer is inattentive between times $t$ and $s > t$, consumption between these periods obeys the deterministic Euler equation:*

$$u'(c_t) = e^{(r-\rho)(s-t)} u'(c_s). \tag{15}$$

*If $D(t)$ and $D(s)$ are two successive planning dates, consumption between these periods obeys the stochastic Euler equation:*

$$u'(c_{D(t)}) = e^{(r-\rho)(D(s)-D(t))} E_{D(t)}\left[u'(c_{D(s)})\right]. \tag{16}$$

A final optimality condition is the transversality condition, which requires that the present value of a unit of assets at infinity must be zero, for otherwise the agent could have used it to increase consumption and utility:

$$\lim_{t\to\infty}\left[e^{-\rho t} V_a(a_t, x_t) a_t\right] = 0. \tag{17}$$

### 3.3 Aggregate consumption

The economy is populated by many inattentive agents, whose individual behavior is described in Proposition 1. The different consumers have the same preferences but differ for instance in their realization of income shocks and in the costs of planning they face. They therefore differ in how long they stay inattentive and in how much they consume. Obtaining predictions for aggregate consumption is complicated by the non-linearities of the marginal utility function. Following the literature, I work instead with linearized versions of (15) and (16).[8] A first-order Taylor approximation of (15) around the point where $c_t = c_s$ and $r = \rho$ gives:

$$c_s = c_t + \frac{1}{\alpha}(r - \rho)t, \tag{18}$$

where $\alpha = -u''(c_t)/u'(c_t)$ is the coefficient of absolute risk aversion. A similar approximation of (16) leads to:

$$c_{D(s)} = c_{D(t)} + \frac{1}{\alpha}(r - \rho)t + e_{D(s),D(t)}, \tag{19}$$

where $e_{D(s),D(t)} \equiv c_{D(s)} - E_{D(t)}\left[c_{D(s)}\right]$, the innovation to consumption between $D(t)$ and $D(s)$.

Some form of indexing must be defined to keep track of the different agents. The following indexing turns out to be convenient: let $j$ denote how long, starting from $t+1$, one must go back to find the last date when the agent has adjusted, with $j$ lying in the interval $[0, J]$. Similarly, let $i \in [0, I]$ denote how long, starting from $t$, one must go back to the last adjustment date for that same agent. The change in consumption of an individual agent is denoted by $c_{t+1}(i,j) - c_t(i,j)$. For instance, $c_{t+1}(4, 0.75) - c_t(4, 0.75)$ is the change in the consumption of an agent whose last two adjustments were at $t + 0.25$ and at $t - 4$.

The population of consumers between two time periods always divides itself between two groups. On the one hand, there are the consumers with $j \geq 1$ (and so for whom $i + 1 = j$), which account for a fraction $\bar{\Psi}$ of aggregate consumption. These agents have not adjusted their consumption plans between $t$ and $t+1$, so using equation (18):

$$c_{t+1}(i,j) - c_t(i,j) = \frac{1}{\alpha}(r - \rho).$$

On the other hand, there is a $(1 - \bar{\Psi})$ fraction of agents who have adjusted at some time between $t$ and $t+1$ and so for whom $j < 1$. Equation (19) describes the consumption choices of these consumers between $t - i$ and $t + 1 - j$. Since they have not adjusted between $t + 1 - j$ and $t + 1$, equation (18) describes the change in consumption between these two periods. Likewise the definition of $i$ implies that equation (18) holds between $t - i$ and $t$. Combining these three equations

[8] This is not to say that these non-linearities are not important. Attanasio and Weber (1995) argue that they can significantly affect tests of the Hall model. Examining their effect on the inattentiveness model is left for future work.

gives the relation between consumption at $t$ and at $t+1$ for these agents:

$$c_{t+1}(i,j) - c_t(i,j) = \frac{1}{\alpha}(r-\rho) + e_{t+1-j,t-i}.$$

Summing over the two groups of consumers gives aggregate consumption:

$$C_{t+1} - C_t = \text{constant} + (1-\bar{\Psi}) \int_{j=0}^{1} \int_{i=0}^{I} e_{t+1-j,t-i} d\Psi(i,j), \tag{20}$$

where $\Psi(i,j)$ is the cumulative density function over the consumers for whom $j \in [0,1)$ and $i \in [0,I]$. I assume that this distribution takes only finitely many values, which matches the fact that there are a finite number of people in the world.

If $u_{t+1} \equiv (1-\bar{\Psi}) \int\int e_{t+1-j,t-i} d\Psi(i,j)$ is treated as the error term in a linear regression for consumption growth, the model predicts that $E_{t-I}[u_{t+1}] = 0$: consumption growth is unpredictable from the perspective of $t-I$ information. In the full information case ($J = I = 0$), Hall (1978) first derived the implication that any variable dated $t$ or before should not predict consumption growth between $t$ and $t+1$. With inattentive agents, events between $t-I$ and $t$ predict consumption growth, since some consumers who had been inattentive will update their information and plans between times $t$ and $t+1$ and will only then react to the past events.

**Proposition 2** *With inattentiveness, aggregate consumption growth between* $t$ *and* $t+1$ *should be unpredictable from the perspective of* $t-I$ *information, where* $I$ *is the largest amount of time during which agents remain inattentive.*

Breaking the $e_{t+1-j,t-i}$ news into independent increments and assuming that these are homoskedastic, Appendix A shows that:

**Proposition 3** *With inattentiveness, aggregate consumption growth can be written as:*

$$C_{t+1} - C_t = \Phi(0)e_{t+1} + \Phi(1)e_t + ... + \Phi(I)e_{t-I+1}, \tag{21}$$

*with* $\Phi(s) \geq \Phi(s+1) \geq 0$ *for* $s = 1, 2, ..., I$*, while* $E_{t-s}[e_{t+1-s}] = 0$ *defines the innovations.*

It is appropriate to call the $e_t$'s "news" since they are mutually uncorrelated and are unpredictable one period ahead. The $\Phi(s)$'s correspond approximately to the share of agents in the population that update their information between $t$ and $t+1$ and had last done so at or before $t-s$. Thus, they are non-increasing in $s$.

Equation (21) reveals another implication of the model for aggregate consumption. When news arrives, consumption rises immediately by $\Phi(0)$. The following period, consumption rises further but now by the smaller amount $\Phi(1)$, and the following period it rises further by the even smaller amount $\Phi(2)$, and so on until $I$ periods after. Aggregate consumption is thus characterized by slow adjustment. It reacts slowly and gradually to shocks, so the impulse response of aggregate consumption to a shock should be increasing for a few periods. Moreover, the impulse response should be concave since the change in consumption gets smaller over time. In contrast, with full information, consumption responds immediately to the news ($\Phi(0) = 1$ and $\Phi(s) = 0$ for $s \geq 1$), since all agents are attentive and so react immediately. A related implication of equation (21) is that consumption growth depends on past news with more recent news receiving a larger weight than older news does. Information disseminates slowly in the inattentive economy, as news gradually spreads and has an impact on consumption choices. Combining these two results:

**Proposition 4** *The inattentiveness model predicts that aggregate consumption exhibits:*
*a) Slow adjustment - the impulse response of consumption to shocks is increasing and concave.*
*b) Slow dissemination of information - consumption growth depends on current and past news and the estimates from regressing consumption growth on current and past news are non-increasing in how far in the past the news had arrived.*

While the Hall (1978) model predicts that aggregate consumption should follow a random walk, equation (21) implies that the change in aggregate consumption should follow an $MA(I)$ process with positive coefficients. The difference between the two is well illustrated by looking at their different predictions for the shape of the normalized power spectrum of aggregate consumption changes.[9] Appendix B derives a formula for the normalized power spectrum corresponding to equation (21), denoted by $f_{\Delta C}(\omega)$. In the full information case, consumption changes are white noise so the power spectrum is horizontal. Gali (1991) uses the power spectrum to test theories of consumption, focusing in particular on the spectrum at frequency zero. Following Deaton (1987), he examines the excess smoothness ratio $\psi \equiv 1/\sqrt{2\pi f_{\Delta C}(0)}$, interpreted as the square root of the ratio between the variance of changes in consumption and the variance of changes in permanent income.[10] In the Hall model, this ratio equals one, since consumption reacts immediately one-for-one to changes in permanent income. Findings of $\psi < 1$ have therefore been interpreted as suggesting that consumption is excessively smooth relative to income, whereas if $\psi > 1$ consumption is excessively volatile. With inattentive consumers, Appendix B shows that:

$$\psi = \sqrt{\frac{\sum_{i=0}^{I} \Phi(i)^2}{\left[\sum_{i=0}^{I} \Phi(i)\right]^2}}, \tag{22}$$

which is clearly smaller than one as long as some agents are inattentive for more than one pe-

[9] The power spectrum of a time series process $x_t$ is defined as $h_x(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j}$, where $\gamma_j = E(x_t - E(x_t))(x_{t-j} - E(x_t))$, the $j^{th}$ autocovariance of $x_t$. The normalized power spectrum is: $f_x(\omega) = h_x(\omega)/Var(x_t)$.

[10] The link between $\psi$ and $f_{\Delta C}(0)$ was rigorously established by Gali (1991). Heuristically, the argument goes as follows. The variance ratio of Deaton is $\psi = \sqrt{Var(\Delta C)/Var(\Delta Y^P)}$, where $Y^P$ denotes permanent income. Gali notes that since the agent faces a budget constraint, changes in permanent income must lead to changes in permanent consumption, so $Var(\Delta Y^P) = Var(\Delta C^P)$. But $2\pi$ times the normalized spectrum at frequency zero of consumption changes measures exactly the fraction of the variability of consumption changes driven by permanent movements: $2\pi f_{\Delta C}(0) = Var(\Delta C^P)/Var(\Delta C)$. That $\psi = 1/\sqrt{2\pi f_{\Delta C}(0)}$ then follows immediately.

riod, so that with inattentiveness, consumption is excessively smooth. Note that if there is excess smoothness, then it must be that $\Phi(i) \neq 0$ for some $i > 0$, so there is excess sensitivity. Yet, excess sensitivity per se does not necessarily imply excess smoothness. A virtue of the inattentiveness model is that it links these two concepts.[11] Any particular pattern of excess sensitivity coefficients ($\Phi(i)$) implies not just excess smoothness, but also an exact value for $\psi$. The model requires that the same set of parameters must fit these two related but distinct features of the data.

**Proposition 5** *In the inattentiveness model:*
*a) Changes in aggregate consumption have a normalized power spectrum given by:*

$$f_{\Delta C}(\omega) = \frac{1}{2\pi}\left\{1 + 2\frac{\sum_{j=1}^{I}\sum_{k=0}^{I-j}\Phi(k)\Phi(k+j)\cos(\omega j)}{\sum_{k=0}^{I}\Phi(k)^2}\right\}. \tag{23}$$

*b) Consumption is excessively smooth with an excess smoothness ratio, $\psi$, given by (22).*

Propositions 2 to 5 give a set of predictions that can be tested using aggregate data. Yet the available measurements of consumption do not give consumption at an instant in time, but rather as the sum over a time period. In other words, while the Propositions assert implications for $C_{t+1}$, the available observations are of $\bar{C}_{t+1} = \int_0^1 C_{t+1-s}ds$. Nevertheless, as Appendix A shows, this only affects equation (21) insofar as it turns the $MA(I)$ process into an $MA(I+1)$ with a new set of coefficients $\tilde{\Phi}(s)$ which are still non-increasing for $s = 1, ..., I+1$. Proposition 2 is modified to assert that measured consumption growth is unpredictable from the perspective of $t - I - 1$ information. The other Propositions are unchanged.[12]

# 4 Functional form assumptions and further predictions

The problem of optimal consumption over time with stochastic labor income even with full information only has a closed-form solution for particular forms of the utility function. In this Section, I derive further implications of the model making assumptions on the utility function, the stochastic process governing income, and the costs of planning, that lead to a closed-form solution while being roughly consistent with the data. I assume that the utility function is of the constant absolute risk aversion (CARA) form:

$$u(c) = -\frac{e^{-\alpha c}}{\alpha},$$

where $\alpha > 0$ is the coefficient of absolute risk aversion. It is well-known that this is one of the few utility functions for which the full information problem has an analytical solution. Also for tractability, I assume that the costs of planning are fixed at a constant $K$.

[11] Campbell and Deaton (1989) link excess sensitivity and excess smoothness in the rational expectations model.
[12] Christiano, Eichenbaum, and Marshall (1991) study the effect of time aggregation in the Hall model.

Following Friedman (1957), I assume that income is the sum of two independent components. The first component is permanent income, denoted by $y_t^P$, which is assumed to follow a driftless Brownian motion with variance $\sigma_P^2$ and Wiener increments $dz_t^P$. This corresponds for instance to changes in employment status or in experience, training or education. The second component is transitory income, $y_t^T$, which is assumed to follow an Orstein-Uhlenbeck process (a continuous time AR(1)), with mean reversion speed $\phi$ and independent Wiener impulses $\sigma_T dz_t^T$. Shocks to transitory income affect income only temporarily, and the larger is $\phi$ the more short-lived their effects are. For instance, these could stand for overtime payment or for occurrences such as illness or winning a lottery prize. If these transitory components are idiosyncratic to the agent, they will aggregate to zero, in which case $y_t^P$ is aggregate income in the economy, but this does not need to be the case. Aggregate but short-lived events, such as weather shocks to productivity, movements in the price level, or business cycles, affect disposable income through $y_t^T$.

If permanent income is observed at discrete points in time, it generates observations matching a discrete-time random-walk, while transitory income observed in discrete time is an AR(1). Income changes therefore follow an ARMA(1,1) process. MaCurdy's (1982) seminal study of annual earnings in the United States finds that this specification describes the data well.[13] If $\phi$ is large, income changes will be close to the MA(1) process originally proposed by Muth (1960).

### 4.1 Optimal inattentiveness and consumption

Defining the consumer's wealth, $w_t$, as the sum of her assets, $a_t$, and the present value of her expected income, $y_t^P/r + y_t^T/(r+\phi)$, the law of motion for wealth is:

$$dw_t = (rw_t - c_t)dt + \frac{\sigma_P}{r}dz_t^P + \frac{\sigma_T}{r+\phi}dz_t^T. \tag{24}$$

Whereas generally the agent must keep track of $a_t$ and $y_t$ separately in order to assess how her constraints will evolve, (24) shows that in this case $w_t$ is a sufficient statistic. I can then write the value function as $V(w_t)$, reducing the dimension of the state space. The agent solves the problem:

$$V(w) = \max_{c,d} \int_0^d e^{-\rho t}\left(-\frac{e^{-\alpha c_t}}{\alpha}\right)dt + e^{-\rho d}E\left[V(w')\right], \tag{25}$$

$$\text{subject to } w' = e^{rd}\left[w - \int_0^d e^{-rt}c_t dt + \int_0^d e^{-rt}\left(\frac{\sigma_P}{r}dz_t^P + \frac{\sigma_T}{r+\phi}dz_t^T\right)\right] - K. \tag{26}$$

Denoting the variance of wealth shocks by $\sigma^2 \equiv \sigma_P^2/r^2 + \sigma_T^2/(r+\phi)^2$, Appendix C proves:

**Proposition 6** *In the CARA-utility, ARMA-income, inattentive consumer problem, the optimal*

[13] MaCurdy (1982) also finds that an MA(2) fits the annual PSID earnings observations as well as an ARMA(1,1). His findings are confirmed by Hall and Mishkin (1982), Abowd and Card (1989), and Meghir and Pistaferri (2003). Pischke (1995) obtains similar results using the quarterly income observations in the 1984 SIPP.

*inattentiveness intervals are given by:*

$$d^* = \frac{1}{r}\ln\left(1+\sqrt{\frac{4K}{\alpha\sigma^2}}\right). \tag{27}$$

*Optimal consumption between adjustments is:*

$$\begin{aligned} c_t^* &= rw_{D(i)} + \frac{(r-\rho)(t-D(i))}{\alpha} - \frac{(r-\rho)}{\alpha r} - \frac{rK}{e^{rd^*}-1} - \frac{\alpha r\sigma^2}{4}(e^{rd^*}+1) && (28)\\ &= rw_{D(i)} + \frac{(r-\rho)(t-D(i))}{\alpha} - \frac{(r-\rho)}{\alpha r} - \frac{r\alpha\sigma^2}{2} - r\sqrt{\alpha\sigma^2 K}, && (29)\end{aligned}$$

*for $D(i) < t < D(i+1)$. The resulting value function is:*

$$V(w_{D(i)}) = -\frac{\exp(-\alpha c^*_{D(i)})}{\alpha r}. \tag{30}$$

If $c_t^A$ denotes the consumption decisions of an agent that has $K = 0$ and so is always attentive, then the consumption of an inattentive agent at a planning date equals:

$$c^*_{D(i)} = c^A_{D(i)} - \frac{rK}{e^{rd^*}-1} - \frac{\alpha r\sigma^2}{4}(e^{rd^*}-1). \tag{31}$$

**Corollary 1** *At time 0, in the CARA-utility, ARMA-income problem, inattentive agents consume less than attentive ones. The larger are the costs of planning, the longer they are inattentive for, and the more they save.*

The lower consumption is due to two reasons, captured by the two terms in (31). The first reason is that costly planning lowers the agent's wealth, since she must pay an amount $K$ every $d^*$ periods, and lower permanent income reduces optimal consumption. The present value of this periodic expense is given by the second term in the right-hand side of (31). The second reason for lower consumption is that the inattentive agent is more vulnerable to risk, since she only periodically adjusts her behavior to take account of the income shocks that are arriving every instant. The precautionary motive for savings is therefore larger by the third term in (31), which is increasing in the length of inattention. Larger costs of planning lead to longer periods of inattentiveness thus strengthening the precautionary motive and raising savings.[14]

Inspecting the optimal inattentiveness in (27) establishes:

**Corollary 2** *In the CARA-utility, ARMA-income case, inattentiveness ($d^*$):*
*1. Falls with the volatility of the income shocks ($\sigma^2$);*

[14] The inattentiveness model suggests a curious explanation for the decline in the U.S. personal savings rate in the last two decades. If advances in information technology have lowered the costs of obtaining and processing information, then agents should optimally respond by saving less.

*2. Falls with the coefficient of absolute risk aversion ($\alpha$);*

*3. Falls with the real interest rate ($r$);*

*4. Increases with the costs of planning ($K$);*

*5. Is first-order long with only second-order costs of planning.*

The intuition behind these results is as follows. The more volatile are income shocks, the more often the agent wants to re-plan so that she is able to adjust her behavior to the arrival of news. In a world that is quickly changing, it is very costly to not pay attention to news so the agent will avoid being inattentive for long. Similarly, if the agent is very risk averse, she will want to lower the risk she faces by updating information more often and responding to shocks faster. This does not imply that higher volatility is beneficial by inducing greater attentiveness. Quite on the contrary, a higher $\sigma^2$ unambiguously lowers welfare, since it increases uncertainty which the risk-averse agent dislikes, and moreover it forces her to spend more resources updating plans more frequently. Policy in a world with inattentive consumers should aim at stabilizing the economy. People can then be inattentive for long and direct their resources towards productive uses, rather than towards planning consumption.

Between planning dates the inattentive consumer (dis)saves all the unexpected changes in income, whereas the full-information consumer (dis)saves only a fraction of the new income. The larger is the interest rate, the larger is the repercussion that this inefficient (dis)saving will have on her future wealth. Facing a high interest rate, the agent will want to adjust more often to avoid past mistakes and to keep her assets under control.

The fourth property of inattentiveness is very intuitive: it states that the more costly it is to plan, the less often the agent plans. More interesting is the last property, which shows that even very small costs of planning can lead to considerable inattentiveness. The intuition for this result is similar to that in Mankiw (1985), Akerlof and Yellen (1985) and Cochrane (1989). Inattentiveness leads to consumption differing from its full information optimum. However, since the choices of the inattentive consumer are close to this optimum, this deviation only has a second-order effect on utility. Therefore, even a second-order cost of planning will induce the agent to tolerate the second-order costs of being inattentive for a first-order period of time.[15]

To illustrate how large $d^*$ can be, consider the parameter estimates by Pischke (1995). He identified $y_t^P$ with aggregate income and $y_t^T$ with idiosyncratic income, and measured them using aggregate and family income from the National Income and Product Accounts (NIPA) and the Survey of Income and Program Participation (SIPP). He estimates that $\sigma_P = \$45$ and his estimates of the autocorrelation of income changes imply that $\phi = 0.487$ (which implies an AR coefficient in the ARMA for income changes of 0.615). His estimate that income changes have an average standard deviation of $\$2,812$ then implies that $\sigma_T = \$1,962$. I set the quarterly interest rate at

[15] Further deviations from rationality may magnify this inertia. For instance, if agents have hyperbolic discount functions, costly planning can lead to procrastination (Akerlof, 1991; Laibson, 1997; O'Donoghue and Rabin, 1999).

1.5%, approximately its historical value in the United States, and $\alpha = 2/6926$, where \$6,926 is mean income in the Pischke sample, so the coefficient of relative risk aversion is about 2. Equation (27) implies that if the costs of updating plans are just \$30, the agent stays inattentive for over 2 years. Very small costs of planing can lead to considerable inattentiveness.[16]

## 4.2 What to plan

With rational expectations or if income is certain, seeing the agent as a consumer or as a saver are two equivalent ways of looking at the same problem. With inattentiveness though, they are no longer equivalent. The agent must choose whether to set a plan for consumption or a plan for savings, letting the other absorb the shocks to income.

An inattentive saver sets plans for savings $s_t$, subject to the constraint that this choice is conditional on the information at the last planning date. Appendix D solves for the optimal choices of this agent, proving the following:

**Proposition 7** *The CARA-utility, ARMA-income, inattentive saver sets consumption:*

$$\hat{c}_t = y_t - \hat{s}_t$$

*where optimal savings $\hat{s}_t$ are set conditional on information at the last adjustment date. The optimal inattentiveness $\hat{d} = +\infty$ if:*

$$K \geq \frac{\alpha\phi\sigma_T^2}{4(r+2\phi)(r+\phi)^2}.$$

*Otherwise, $\hat{d}$ is finite and is the unique solution of the equation:*

$$re^{2\phi\hat{d}}\left(1 - \frac{4(r+2\phi)(r+\phi)^2 K}{\alpha\phi\sigma_T^2}\right) = r + 2\phi(1 - e^{-r\hat{d}}).$$

Since consumption is just $c_t = y_t - s_t$ and $s_t$ is pre-determined, the inattentive saver consumes every period her total income less a pre-determined amount. She lives hand-to-mouth, with a marginal propensity to consume out of income equal to one.

[16] This calculation is solely meant to illustrate how large inattentiveness can be. Still, the reader may wonder how sensitive it is to the choice of parameters. First, since $r$ is a riskless rate, 0.5% may be a more appropriate value. If so, even smaller costs of planning lead to considerable inattentiveness: \$12 induces $d^* = 8$ quarters. Second, the variance of total earnings is probably over-estimated due to measurment error, and it is tricky to identify its permanent component. Following Bound et al. (1994), I lower the variance of toal income by 1/3, and experiment lowering $\sigma_P$ by 50%, and raising it by 50% and 100%. Then, the agent is inattentive for 2 years if the costs of planning are respectively \$14, \$36 and \$54.

Finally, as discussed at length at the end of Section 3.1, I allowed consumption to be negative. This was convenient but how often does it happen? Given the set of Pischke (1995) parameters, and further assuming that: $\rho = r$, $K = \$30$, and that $c_0$ equals 90% of median income (since the savings rate in the national accounts is about 10%) to infer a value for $w_0$, I can calculate the probability that by the next planning date, $c_d$ is negative. This probability is essentially zero. It would take 8 successive quarters of negative wealth shocks equal to more than 10 times their standard deviation to make $c_d$ negative.

The intuition for the $\hat{d} = +\infty$ result comes from realizing that while consumption reacts optimally (one-to-one) to permanent income shocks, it also responds one-to-one to transitory income shocks when the optimal reaction would be to consume only a fraction $r/(r+\phi)$ of these shocks. As the costs of planning and optimal inattentiveness rise, less remains of a transitory shock by the time the agent responds to it. The incentive to update her plans therefore falls as inattentiveness rises, and a small increase in the costs of planning leads to a large increase in inattentiveness. After a certain level, optimal inattentiveness becomes convex in the costs of planning, and shoots to infinity.

If the agent chooses $\hat{d} = +\infty$, she can be described as a *rational non-planner.* She writes a plan once at time 0 and then follows this plan forever. For the parameter estimates in Pischke (1995), she chooses to do so once the costs of planning exceed \$543. Moreover, as Appendix E shows:

**Corollary 3** *At time 0, in the CARA-utility, ARMA-income problem, rational non-planners save less than the consumption planners.*

If the agent is given the option of being either an inattentive consumer or an inattentive saver, which will she choose? Appendix E proves the following answer:

**Proposition 8** *If* $(\phi - r)/(\phi + r) > \sigma_P^2/\sigma_T^2$*, the CARA-utility, ARMA-income, inattentive agent prefers a consumption plan if her costs of planning are below a threshold* $\hat{K}$*, and prefers a savings plan if the costs of planning are above this threshold. When the agent shifts from consumption to savings plans, her inattentiveness rises discontinuously, and possibly to infinity.*

Since almost all studies of individual income find that transitory shocks are the dominant source of income variation, the parameter restriction in this Proposition reduces to assuming that $\phi > r$. With an annual interest rate of 6%, this requires that transitory income shocks have a half life of no more than 11.5 years. From the other perspective, if $\phi = 0.487$ as estimated by Pischke (1995), the annual interest rate must be lower than 601%. The constraint in Proposition 8 likely holds given plausible values of the interest rate and the persistence of transitory shocks.

Proposition 8 then states that the model predicts that there are two distinct groups in the population. On the one hand, are those who make financial plans for consumption, updating them sporadically. On the other hand, are those who are inattentive for longer, live hand-to-mouth and save less. This second group maybe composed only of people who rationally choose to never plan:

**Corollary 4** *As long as:*

$$\frac{\phi^3 - r^2(r+2\phi)}{(r+2\phi)(r+\phi)^2} > \sigma_P^2/\sigma_T^2,$$

*then agents who choose to be inattentive savers also choose to be rational non-planners.*

For the parameter estimates of $\sigma_P^2/\sigma_T^2$ and $\phi$ found by Pischke (1995), the condition in the corollary holds as long as the annual interest rate is below 232%. It is reasonable to expect that all inattentive savers are rational non-planners.

A convenient way to assess how likely it is to find rational non-planners in the economy is to use the following result, proven in Appendix E:

**Proposition 9** *If the conditions in Proposition 8 and Corollary 4 apply, then consumption plans are strictly preferred to rational non-planning if:*

$$\frac{\sigma_P^2}{\sigma_T^2}(e^{rd^*} - 1) + \left(\frac{r}{r+\phi}\right)^2 e^{rd^*} - \frac{r}{r+2\phi} < 0. \tag{32}$$

While this condition involves an endogenous variable ($d^*$), it only requires knowledge of $\sigma_P^2/\sigma_T^2$ and $\phi$ from the earnings data, and no information on the degree of risk aversion. Using the benchmark estimates in Pischke (1995) for $\sigma_P^2/\sigma_T^2$ and $\phi$, then if the agent would choose to be inattentive for 8 quarters under a consumption plan, she prefers this plan to being a rational non-planner as long as the quarterly real interest rate is below 12.5%. From a different perspective, if the quarterly interest rate is 1.5%, then only if the consumption-planning agent stays inattentive for more than 41 years would she prefer to become a rational non-planner. Some agents may face such high costs of planning and interest rates that they live hand-to-mouth, but these calculations suggest that the majority of the population follows consumption plans.

Aside from strict consumption or savings plans, another plausible form of planning sets a fixed percentage of income to be automatically consumed or saved. The agent still keeps herself from observing and calculating an optimal response to income shocks every instant, but now has savings and consumption absorbing the shock in fixed percentages. In the United States, many workers enroll in fixed percentage contribution IRA plans that resemble these hybrid consumption-savings plans. Consumption is then $c_t = \lambda y_t + \tilde{c}_t$ and at a planning date, the agent now sets a plan for consumption ($\tilde{c}_t$), for the next planning date ($\tilde{d}$), as well as for the fraction of income shocks so be absorbed by consumption ($\lambda$). Appendix F solves the problem of this hybrid consumption-savings planner, with CARA utility and ARMA income. The optimal $\tilde{d}$ and $\lambda$ are state-independent, and Table 1 displays them for different plausible values of the interest rate and the costs of planning. The ability to choose $\lambda$ implies that relative to consumption-planning the agent is now inattentive for even longer. The optimal $\lambda$ in turn are quite small, ranging from 0.02 to 0.19.

Considering savings or hybrid planning then qualifies the prediction in Proposition 2 as follows:

**Corollary 5** *If $I$ is the largest amount of time during which planning agents remain inattentive, regressing aggregate consumption growth between $t$ and $t+1$ on expected income as of $t-I$ periods ago, identifies either the fraction of aggregate consumption by hand-to-mouth inattentive savers, or the optimal percentage of income shocks consumed by hybrid planners. Either way, the estimated coefficient should be small.*

### 4.3 Extraordinary events

For most of the time in her ordinary life, a person is subject to random but small income shocks so it is not too costly to be inattentive. Occasionally though, big things happen in your life. You may lose your job, or win the lottery; your close family may be struck by a serious and expensive disease, or you may receive a sudden inheritance from a distant relative; an unexpected hyperinflation may eat up your purchasing power, or the shares in your small company may be worth a fortune after you come across a great invention. These things make you stop and think: the circumstances around you have changed so radically that old plans must be thrown out of the window and new plans made for the future.

A simple way to model these extraordinary events is by adding to the agent's income an independent Poisson term with arrival rate $\delta$ and jumps $u$ or $-u$ with equal probability. Most of the time (with probability $1-\delta$) no event takes place, but every so often (with probability $\delta$) an extraordinary event occurs which dramatically changes the agent's disposable income and to which she responds instantly. Because most of the time no event occurs, the computational cost of observing this variable is small so this is consistent with the underlying assumption that there are costs of absorbing and processing information. As long as the event is extraordinary (i.e., $u >> 0$), the agent responds to it by collecting information and setting a new plan. Further adding the convenient but inessential assumption that the interest rate equals the discount rate, Appendix G proves:

**Proposition 10** *With CARA preferences and income following an ARMA(1,1) process plus Poisson extraordinary events, optimal inattentiveness is the minimizer of the function $A(0)$, which solves the boundary value differential equation:*

$$\begin{aligned} A'(t) - rA(t)\ln(A(t)) - \delta A(t) &= -A(0)\frac{\delta}{2}e^{\alpha rK - \frac{\alpha^2 r\sigma^2}{4}}\left(e^{-\alpha ru} + e^{\alpha ru}\right)e^{\frac{\alpha^2 r\sigma^2}{4}e^{2rt}} \\ A(d^*) &= A(0)e^{\alpha rK - \frac{\alpha^2 r\sigma^2}{4}}e^{\frac{\alpha^2 r\sigma^2}{4}e^{2rd^*}} \end{aligned}$$

This differential equation can be solved numerically to find $d^*$. Panel A of Table 2 does so using the parameter estimates from Pischke (1995) in the cases when extraordinary events occur on average every 2, 5, or 10 years, and when they imply a change in income of \$500, \$2500, and \$5000. Panel B shows the probability that an extraordinary event occurs before the planning time is reached.

Table 2 shows that the larger is the size of the extraordinary event, the longer is inattentiveness. I hold the agent's total income variance constant over the different parameters, so as $u$ rises, a larger share of the variance is accounted for by extraordinary events. The variance of the small income shocks to which the agent is inattentive is then lower, so she stays inattentive for longer. Extraordinary events have a modest effect for these parameter values, at most raising inattentiveness by 3 quarters. Note also that as extraordinary events become more infrequent, the planning horizon approaches the solution without extraordinary events. If an extraordinary event occurs on average every 10 years, then the agent who stays inattentive for 2-year periods will adjust before the end of her plan only about 18% of the time.

Considering extraordinary events leads to the following prediction:

**Corollary 6** *Individual consumption responds immediately to extraordinary events, but only with a delay to ordinary shocks.*

## 5 Evidence on aggregate consumption

Propositions 2 to 5 stated a series of predictions of the inattentiveness model for the behavior of aggregate consumption. In this Section, I test these predictions using U.S. data. Section 5.1 examines Propositions 3 and 4 by studying whether aggregate consumption adjusts gradually to shocks in the data. Section 5.2 then contrasts the predictions in Proposition 5 on the spectrum and excess smoothness of consumption, with estimates from the data. Section 5.3 tests the prediction in Proposition 2, as refined by Corollary 5, by estimating the fraction of aggregate consumption attributable to hand-to-mouth behavior. This naturally leads to a test of the inattentiveness model against the alternative suggested by Campbell and Mankiw (1989, 1990).

I use U.S. quarterly time series data which come from the National Income and Product Accounts, and financial data from the Center for Research in Security Prices. Aggregate consumption, $C_t$, is measured as real consumption of non-durables and services per capita, while $Y_t$ denotes real disposable personal income per capita.[17] Both are deflated using the price deflator for consumption of non-durables and services. Data on real returns, $r_t$, will be used as a predictor of income growth. It is measured as the nominal return on the value-weighted S&P500 minus the inflation rate using the price deflator for non-durables and services.[18] The sample runs from 1953:1 through 2002:4.[19] One specification issue is whether to measure consumption in levels or logs. Past research has alternatively chosen one or the other, but I opt for log consumption since the series of observations on $C_t$ appears to be closer to log-linear than linear.[20] Note that while the Propositions in Section

[17] I have also experimented using two other measures of $C_t$: consumption only on non-durables, and Parker's (2001) consumption series, which excludes footwear, housing, medical care, education, and personal business expenditures from non-durables and services. Both led to similar results.

[18] I experimented with many alternative measures for $r_t$. Different assets were used (the New York Stock Exchange index, the 3-month Treasury Bill rate, and municipal bonds), and after-tax returns were computed using different measures of the tax rate (the ratio between the return on tax-free municipal bonds relative to taxable corporate bonds, and a fixed rate of 30%). The results are robust to using these alternatives.

[19] Data are available from 1947, but I follow Blinder and Deaton (1985) and Campbell and Mankiw (1990) in starting the sample in 1953:1 to avoid the effect of the Korean War and the unusual large spike in disposable income in 1950:1 due to the large payment of National Service Life Insurance benefits to World War veterans.

[20] Regressing the change in consumption on the level of consumption gives a coefficient of 0.003 with a t-statistic of 2.51, suggesting that the mean change rises with the level of the series. Regressing the squared change on the level of consumption gives a coefficient of 0.914 with a t-statistic of 4.88, which is a strong sign that the innovation variance of the series also increases with its level. Both point towards log-linearity.

3 concerned the level of consumption, they could equally well be stated for log consumption by log-linearizing rather than linearizing the Euler equations.

### 5.1 Slow adjustment to shocks and slow dissemination of information

Proposition 4 stated that the impulse response of consumption to shocks should be increasing and concave. A simple analysis of the adjustment of aggregate consumption to shocks comes from estimating a structural vector autoregression (VAR) on consumption and income growth. I set the lag length on the VAR at 5, as suggested by the use of the Schwartz's Bayesian information criterion and by examining the significance of the last lag included in the VAR. Following Blanchard and Quah (1989), I identify permanent shocks to consumption and, in Figure 1, I display the impulse response of log consumption to a permanent shock together with a 90% point-wise confidence interval generated by a bootstrap.

As shown in Figure 1, aggregate consumption adjusts with a delay to the shock, as the inattentiveness model predicts would be the case due to slow dissemination of information. Moreover, while consumption is sluggish, it is only moderately so: most of the adjustment is completed within one year of the shock. This is consistent with an inattentiveness model in which agents update their information approximately once a year. The model also predicts that the impulse response should be concave and this pattern is also present in Figure 1.

If the volatility of disposable income fell, the model predicts that agents would respond by staying inattentive for longer, so that consumption should adjust more sluggishly to shocks. Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) identify a large fall in the volatility of GDP in the US after around 1984.[21] Figure 2 shows the impulse responses of consumption estimated for the period between 1953 and 1982, and between 1985 and 2002. The model predicts that the latter period should show a more sluggish response to shocks; this is the case in Figure 2.

A sharper test of the slow adjustment of consumption to news comes from examining its response to news on a particularly important variable: income. Given a statistical model for income, surprises ($y_t$) can be constructed as one-step ahead forecast errors. By construction, these have mean zero and are uncorrelated, so they satisfy the properties that define the innovations $e_t$ in Proposition 3. Regressing consumption growth on several lags of $y_t$ leads to a test of the model's predictions in equation (21).

The first possible model for income growth I consider is an AR(5). The results from regressing consumption growth on income news are in Table 3. As predicted by the inattentiveness model, lagged income surprises affect future consumption growth. The F-statistic reported in the Table tests the null hypothesis of the Hall (1978) model that lagged income surprises do not affect current consumption growth. This hypothesis is strongly rejected with a p-value below 0.01%. Moreover, income surprises explain much of the variability of consumption growth; the adjusted $R^2$ of the

[21] Blanchard and Simon (2001) and Stock and Watson (2002) have confirmed this conclusion.

regression is 0.33.

The top panel of Figure 3 plots $\hat{\Phi}(j)/\sum_{i=0}^{I+1}\hat{\Phi}(i)$ for $j$ from 0 to $I+1$ using the estimates in Table 3, together with 95% confidence intervals. The inattentiveness model predicts a declining sequence of non-negative points. This is consistent with the plot in Figure 3. Another way to examine this property is to look at the cumulative dissemination of the news, which is plotted in Figure 4. It shows the increasing and concave shape that the model predicts, similar to that estimated earlier in Figure 1. Estimating a regression by least squares subject to the model's restriction that the coefficients on income surprises $s$ periods ago are declining in $s$ results in the estimates presented in the second to last row of Panel A in Table 3 and displayed in Figures 3 and 4. These restricted estimates are quite close to the unrestricted estimates supporting the validity of the model's restrictions. The null hypothesis of the model can be formally tested using Wolak's (1989) Wald test. In Table 3, $W_{IN}$ displays the value of the test statistic and the p-value of the test of the inattentiveness null. The model cannot be rejected at the 5% statistical significance level.

Panel B of Table 3 models income growth instead as depending on 5 lags of both past income growth and the log consumption income ratio. As did Campbell (1987), I find that these new regressors have significant predictive power for income growth: the p-value of an F-test on their significance is below 0.1% and the adjusted $R^2$ of the first stage regression rises by a factor of 3. The coefficients in the regression of consumption growth on current and past surprises are nevertheless little affected. They are still jointly very statistically significant and they broadly exhibit the non-increasing pattern predicted by the model. This can be visually apprehended in Figures 3 and 4, and is confirmed by the value of the $W_{IN}$ statistic, which does not reject the model at the 5% significance level. Panel C adds 5 lags of the after-tax real interest rate as further predictors of future income growth. These variables help very little forecasting income growth, but the estimates of equation (21) are again similar to those obtained before, and lead to similar inferences.[22]

As before, we can see whether these estimates changed with the fall in GDP volatility. Figure 6 plots the cumulative weights, before 1983 and after 1985. Again, we see that the period of lower volatility is associated with a more sluggish response of consumption to shocks.

## 5.2 Excess smoothness and the spectrum of consumption

Proposition 5 makes sharp predictions on the shape of the power spectrum of aggregate consumption changes. Figure 6 plots estimates of the spectrum, constructed using a sample spectral density weighted over a 5-lag Bartlett window. The normalized power spectrum of consumption growth generally declines with the frequency with a shape close to Granger's typical shape, but with a slight hump around the $\pi/2$ frequency. Figure 6 also displays the spectrum for aggregate consumption growth predicted by the inattentiveness model using the weights $\hat{\Phi}(i)$ estimated in panel A of Table

[22] Mishkin (1983) noted that the two-step econometric procedure that I used will not produce efficient estimates. I have estimated the system of two equations simultaneously using the iterative procedure suggested by Mishkin (1983). The results were very similar to those in Table 3, so the inferences are robust to this econometric issue.

3.[23] The predicted spectrum matches the empirical spectrum quite well, albeit with somewhat more pronounced swings. Also in Figure 6 is the predicted spectrum using the weights obtained after imposing the theoretical restrictions of the inattentiveness model. The fit to the data is even better with these restricted estimates.

Table 4 displays different estimates of the excess smoothness ratio ($\psi$) defined in Section 3.3. They lie between 0.52 and 0.7, and the full information rational expectations null hypothesis that they equal one is always rejected. The inattentiveness model using the weights from Panel A of Table 3 predicts an excess smoothness ratio of 0.66, well within the range of estimates. Table 4 shows the predictions of the model using different estimates of the weights. For all sets of weights, the inattentiveness model predicts a ratio of excess smoothness close to (or slightly below) that in the data. The inattentiveness model is therefore able to simultaneously generate the extent of excess sensitivity and excess smoothness that we observe in the data.

### 5.3 Excess sensitivity and hand-to-mouth behavior

The general inattentiveness model, in which there are both inattentive consumers and savers, or in which hybrid plans are formed, implies that

$$C_{t+1} - C_t = (1 - \lambda)e_{t+1} + \lambda(Y_{t+1} - Y_t), \tag{33}$$

with $E_{t-i}[e_{t+1}] = 0$ for any $i \geq I + 1$. Proposition 2 and Corollary 5 stated that $\lambda$ is the fraction of aggregate consumption accounted for by rational non-planners, or the share of income shocks absorbed by consumption in hybrid plans.

Table 5 estimates this equation, using as instruments for the change in income variables dated at least 9 quarters before. The estimates in Section 5.1 suggested that within one year most agents have updated their plans, so letting $I$ be 2 years is a conservative choice. I use as instruments 4 lags of income growth, and then successively add 4 lags of the log income-consumption ratio, and 4 lags of real returns. Since the model predicts that the residuals of this regression should be serially correlated, I compute the Hayashi and Sims (1984) nearly-efficient estimates, rather than the conventional (but inefficient) two-stage least squares estimates.

The estimates of $\lambda$ are quite low, between 0.05 and 0.15. This confirms the prediction in Corollary 5, that the share of aggregate consumption attributable to hand-to-mouth behavior should be quite small. The null hypothesis that $\lambda = 0$ cannot be rejected at conventional significance levels in any of the regressions, supporting Proposition 2. The data is consistent with a model in which inattentive consumers account for the bulk of aggregate consumption dynamics.

The instruments used in these regressions are weak, as reflected by the low F-statistics in the second to last column in Panel A of Table 5. Income growth is difficult to forecast 9 quarters in advance. With weak instruments, the IV estimates are biased towards the OLS estimates. These

[23] Using the estimates from panels B and C produces very similar plots.

are displayed in the second column of Panel B, and are higher than the IV estimates, suggesting that the estimates of $\lambda$ in Panel A are, if anything, too large. An alternative estimator is limited information maximum likelihood (LIML) and the third column shows that these estimates are slightly lower than those in Panel A. Columns 4 to 6 of Panel B present three tests proposed in the literature on weak instruments to powerfully test the Hall (1978) model: the Anderson and Rubin statistic, the Moreira (2003) conditional likelihood ratio statistic, and a conditional Lagrange multiplier statistic. All of them cannot reject the hypothesis that $\lambda$ is zero. While these tests likely suffer from lack of power, still both the IV and the LIML estimates are consistent (and the tests of the over-identifying restrictions implied by instrument validity are never rejected), and they consistently estimate $\lambda$ to be small.

One feature of equation (33) is that it is also the equation that describes aggregate consumption dynamics in the model proposed by Campbell and Mankiw (1989, 1990), in which a fraction $1-\lambda$ of consumption is accounted for by rational expectations agents, while the remaining $\lambda$ fraction is accounted for by irrational, myopic, hand-to-mouth people. The difference is that in their model, $E_{t-i}\left[e_{t+1}\right] = 0$ for any $i \geq 1$. Using variables lagged two quarters as instruments for income growth, Campbell and Mankiw (1989, 1990) found that hand-to-mouth agents account for 40-50% of aggregate consumption. According to their model though, it is equally valid to use instruments lagged nine quarters. However, Table 5 shows that doing so produces estimates of $\lambda$ that are insignificant and much lower, between 5% and 15%, supporting instead the inattentiveness model.

Given the concerns on the power of the tests, it would be desirable to test the Campbell-Mankiw against the inattentiveness model, having both models stated as null hypotheses, to avoid biasing the results towards the model that is stated as a null hypothesis due to lack of power. Note that I can expand the right-hand side of equation (33) to obtain:

$$\begin{aligned} C_{t+1} - C_t &= \lambda\left(E_t - E_{t-1}\right)\left(Y_{t+1} - Y_t\right) + \lambda\left(E_{t-1} - E_{t-2}\right)\left(Y_{t+1} - Y_t\right) + ... \\ &\quad + \lambda\left(E_{t-T+1} - E_{t-T}\right)\left(Y_{t+1} - Y_t\right) + \lambda E_{t-T}(Y_{t+1} - Y_t) + u_{t+1}, \end{aligned}$$

where $u_{t+1} \equiv (1-\lambda)e_{t+1} + \lambda\left(Y_{t+1} - E_t(Y_{t+1})\right)$ is uncorrelated with the other right-hand side variables. Estimating the regression equation

$$C_{t+1} - C_t = \beta_0 + \sum_{s=1}^{T} \beta_s\left(E_{t-s+1} - E_{t-s}\right)\left(Y_{t+1} - Y_t\right) + \lambda E_{t-T}(Y_{t+1} - Y_t) + u_{t+1}, \qquad (34)$$

the null hypothesis describing the Campbell-Mankiw model is[24]

$$H_0^{CM} : \beta_2 = ... = \beta_T = \lambda.$$

[24] There is no restriction on $\beta_1$ to allow for time aggregation. I do not impose the restriction that $\lambda \geq 0$, which biases the results in favor of the Campbell-Mankiw model.

Since $(E_{t-s+1} - E_{t-s})(Y_{t+1} - Y_t)$ has a zero expectation as of $t - s$, it fits into the definition of the news $e_{t-s+1}$, so the prediction of the inattentiveness model in Proposition 3 is that

$$H_0^{IN} : \beta_1 \geq \beta_2 \geq ... \geq \beta_T \geq 0,\ \lambda = 0.$$

If there are both inattentive savers and consumers, this leads to the weaker null hypothesis:

$$H_0^{ING} : \beta_1 \geq \beta_2 \geq ... \geq \beta_T \geq 0,$$

as long as $T \geq I + 1$, and the estimate of $\lambda$ gives the share of inattentive savers. Finally, the Hall (1978) model predicts that

$$H_0^{RE} : \beta_2 = ... = \beta_T = \lambda = 0,$$

so that no variable dated $t$ or before predicts consumption growth.

The different models are now expressed as different null hypotheses on the same regression, so a potential lack of power should not be expected to bias the inferences towards any one of them in particular. Intuitively, the difference between the Campbell-Mankiw and the inattentiveness models is that in the former consumption depends on lagged income news solely through hand-to-mouth consumption, so how far away in the past the news was revealed does not affect its impact. With inattentiveness instead, the longer the news has been known for, the more likely it is that agents have since updated their plans and so the smaller their current impact.

The variables on the right-hand side of (34) were generated by estimating a VAR with 5 lags on the change in log income, the log consumption-income ratio, and the real interest rate, and then using this VAR to construct $s$-step ahead forecasts of income growth. Table 6 presents the point estimates of equation (34), which are somewhat discouraging for all four models. The point estimates of the $\beta_s$'s do not seem to have the pattern described in either $H_0^{CM}$ or $H_0^{IN}$, and several of the estimated coefficients are individually large and statistically significant contrary to $H_0^{RE}$. Panel B of Table 6 formally tests the models using Wald tests for $H_0^{CM}$, $H_0^{RE}$, $H_0^{IN}$, and $H_0^{ING}$.[25] Consistent with the other results in this paper, the full information rational expectations model is decisively rejected even at a 0.01% significance level. The Campbell-Mankiw model is also rejected at the 5% significance level (but not at the 1% level), which is not surprising given the low estimate of $\lambda$. The null hypothesis of the general inattentiveness model on the other hand has a p-value of 12.8%, so it is not statistically rejected at conventional significance levels. Moreover, note that it is estimated that only 3.4% of consumption is done by inattentive savers, so hand-to-mouth behavior is economically and statistically insignificant. Consequently, the model with inattentive consumers alone is not statistically rejected at the 5% significance level.

These results suggest that the hand-to-mouth behavior detected in aggregate consumption data may be attributable to inattentiveness. Moreover, as predicted by the theory, rational non-planning

[25] The null hypotheses $H_0^{IN}$ and $H_0^{ING}$ are tested using the Wolak (1989) procedure.

(or hybrid planning) seems to have a small impact on aggregate consumption.

## 6 Microeconomic evidence on inattentiveness

The inattentiveness model makes a series of sharp predictions on the behavior of individual consumption. Over the last decade, much research has established a set of facts about individual behavior. As a first step, it is important to examine whether the main predictions are consistent with what we know from these studies. At the end of this Section, I then discuss possible new tests of inattentiveness using micro data.

### 6.1 Sensitivity to past information

There have been many tests of excess sensitivity using individual consumption data but the results so far are inconclusive: some studies find it, while others do not, and it is unclear what explains the different results. The inattentiveness model suggests an explanation. Proposition 1 established that agents are inattentive to ordinary unpredictable events and thus react with a delay to these shocks, only at their next planning date. If the event is easily predictable though, the agents will have reacted to it when they set their plans in the past. If the event is extraordinary, Corollary 6 established that reaction was instantaneous. The inattentiveness model therefore predicts that consumption is sensitive to past ordinary unpredictable events, but it is not sensitive to predictable or extraordinary events.

Two key papers that have found evidence that small past news on after-tax income affects consumption some time after are Parker (1999) and Souleles (1999). Parker (1999) looks at the patterns of Social Security tax withholding, while Souleles (1999) looks at income tax refunds. In both cases, the news were to ordinary components of income and were not especially noticeable and they were also unpredictable. Parker and Souleles findings that consumption is sensitive to these past news supports the inattentiveness model.

In turn, Browning and Collado (2001) and Souleles (2000) look at the response of consumption to large and easily predictable changes in income, and find that consumption does not react to these past news. Browning and Collado (2001) examine the reaction of Spanish households to well-known income fluctuations driven by the timing of bonus payments, while Souleles (2000) examines the impact of the easily predicted college tuition payments on parent's consumption. Hsieh (2003) studies the reaction of Alaskans to the extraordinary payments made to them by the Alaska's Permanent Fund associated with oil royalties. These payments were very large (on average $1,964 in 2000), infrequent, and amply discussed in the media. Hsieh (2003) finds that consumption does not respond to this past extraordinary news, supporting the inattentiveness model.

The inattentiveness model can therefore reconcile the apparently contradictory findings in the literature that has tested for excess sensitivity to past events.

## 6.2 Inattention

There is much evidence that people are inattentive, but one is particularly relevant to the model in this paper. In 1992, President George H. Bush announced a reduction in the standard rates of withholding for income taxes, which lowered employees' tax withholding by about $29 per month. Using a survey of 501 people 1-2 months after the announcement, Shapiro and Slemrod (1995) find that about half of respondents were not aware of any change in withholding, and as many as 2/3 did not even know whether withholding rates had increased or fallen.

The inattentiveness model further predicts that news disseminates slowly throughout the population. Carroll (2003) looks at survey data on inflation expectations and finds that the expectations of the public lag those of professional forecasters, supporting the slow dissemination of information. Moreover, he finds that when the number of references to inflation in the newspapers rises, the public updates its expectations faster, which is consistent with the endogenous determination of inattentiveness in this paper. Mankiw, Reis, and Wolfers (2003) examine the dynamics of disagreement in inflation expectations in three U.S. surveys. In the inattentiveness model, information and expectations are only updated infrequently so, at any point in time, agents will have different expectations determined by when they last updated. Mankiw, Reis, and Wolfers (2003) find that a model with exogenous staggered updating of information matches the time-series of disagreement well, and can explain the particularly large increase in disagreement that occurred in the early 1980s during the Volcker disinflation.

## 6.3 Planning

Proposition 8 together with Corollaries 3 and 4 predicted tat a fraction of the population should choose to never make plans and save less than those who do. Lusardi (1999, 2002) uses data from the Health and Retirement Study (HRS), which surveys people over 50 years old on their attitudes towards retirement, and finds that approximately one third have hardly thought about retirement. She finds that those who have not planned are more likely to be less educated, self-report lower cognitive abilities, be single, and do not have older siblings to use as a source of information. If these proxy for the costs of planning, they can be used as instruments for planning in determining whether it determines accumulated wealth. Lusardi (1999, 2002) estimates this regression and finds a strong significant effect. Ameriks, Caplin and Leahy (2003a) use a TIAA-CREF survey in which households were asked "Have you personally gathered together your household's financial information, reviewed it in detail, and formulated a specific financial plan for your household's long-term future?" The households were also asked further questions on their attitudes towards planning, namely whether they are confident with their mathematical skills, and whether they usually plan their vacations. If these responses provide a proxy for the costs of planning, they can be used as instruments for the household's planning decisions in assessing whether planning predicts savings and wealth. Ameriks, Caplin and Leahy (2003a) find that approximately 25% of households report not having a financial plan, and that these households seem to face higher costs of planning. Moreover, they find that those who do not plan have significantly lower savings and accumulated wealth.

These studies using microeconomic data therefore suggest that between one quarter and one third of the U.S. population does not make plans. Section 5 in turn found that about 5% of aggregate consumption can be attributed to rational non-planners. These two estimates are surprisingly consistent, since the theory predicts (and the data confirms) that the non-planners have lower savings and wealth, and so account for a small share of aggregate consumption.[26]

Hurst (2003) uses the data from the Panel Study of Income Dynamics (PSID) to distinguish between two groups of consumers, according to whether they have high or low wealth when they reach retirement. He finds that the low wealth group suffers a larger drop in consumption at retirement (consistent with inadequate planning for retirement), and has consumption growth responding to predictable changes in income. Moreover, he finds that this behavior cannot be accounted for by liquidity constraints, precautionary savings, or habit formation, but can be explained by hand-to-mouth behavior. The identification of a group in the population that simultaneously does not plan, saves less, and lives hand-to-mouth, supports the model in this paper.

In addition, in the Ameriks, Caplin, and Leahy (2003a) survey, agents who reported having a plan were further asked for how long they have had their plan in place. Proposition 6 and Corollary 1 stated that planners who update information less frequently due to larger costs of planning will save less. Ameriks, Caplin, and Leahy (2003a) regress accumulated wealth on the length of the agent's plan, using as instruments their proxies for the costs of planning, and find that agents who have had plans in place for longer accumulate significantly more wealth. This finding supports the inattentiveness model. In related work, Alessie, Kapteyn, and Lusardi (1999) use the Dutch CentER data-panel, which asks households for their planning horizon for expenditures and savings. Again in agreement with the inattentiveness model, they find that the longer is the planning horizon, the larger are savings (but the effect is not statistically significant at the 5% level).

### 6.4 Possible new tests of inattentiveness

The inattentiveness model is therefore able to explain many of the existing findings on individual consumption. The model also generates many novel predictions that can be tested using micro data. For instance, the model predicts that the longer it takes from the announcement of an income shocks to its realization, then the smaller should be a cross-sectional estimate of the response of consumption to the shock when it is realized. An alternative test of the model would be to use information that at some point in time (e.g., at tax-filing dates) some agents are more likely to be paying attention to their income than others (e.g., those that fill their fax forms on their own

[26] On the other hand, the Campbell and Mankiw (1989, 1990) estimate that hand-to-mouth consumers account for 40-50% of aggregate consumption implies that a large majority of U.S. households live hand-to-mouth, which is difficult to reconcile with the microeconomic evidence.

vs. those that use a tax-preparer), and see whether those that are inattentive are more likely to respond to the available information with a delay (e.g., change consumption when the income tax refund check arrives). Yet a third alternative could examine whether shocks that are common to all individuals raise the dispersion of consumption over households, as some react to it and others do not. These, and other tests, are beyond the scope of this already long paper, but hopefully can be undertaken in future work.

# 7 Alternative models of near-rationality and consumption

## 7.1 Other models of inattention

Gabaix and Laibson (2001), Sims (2003), and Moscarini (2004) are the closest papers to this one.

Gabaix and Laibson (2001) study the problem of an inattentive agent who allocates her savings between a risky asset (equity) and a riskless one (bonds). The focus of their paper was on the equity premium puzzle, and in particular on the correlation between consumption growth and equity returns. The focus of this paper is instead on the dynamics of consumption and its relation to income. Moreover, during most of their paper, Gabaix and Laibson (2001) set the inattentiveness intervals exogenously. In one section, they endogenize them by solving for optimal constant inattentiveness intervals through a series of approximations. The theory in this paper has focussed on determining optimal inattentiveness, and Sections 3.1 and 3.2 set up and solved this problem in great generality. Inattentiveness is constant only under special circumstances.[27]

Sims (2003) assumes that economic agents can at each point in time only obtain k bits of information. This limited transmission channel could be interpreted literally, since human senses have a finite capacity to absorb information, or instead as a metaphor for the limited ability to interpret this information. Using Shannon's (1948) entropy measure of the information in a signal and his results on optimal coding of messages, Sims (2003) determines which pieces of information the agent wishes to pay attention to, and solves the signal extraction problem faced by the agent to infer the current state of the world.

Moscarini (2003) instead assumes that agents must pay a cost for information that is increasing in the number of bits. He shows that agents optimally decide to be inattentive, only periodically turning on their channel to receive information on the state of the world. This result suggests that modelling the costs of information using the Sims-Moscarini approach or the approach in this paper leads to similar infrequent updating of plans. One significant difference is that when updating occurs, agents in the Sims-Moscarini model only obtain an imperfect signal on the state of the world, whereas consumers in the inattentiveness model obtain full information on this state. This feature of the inattentiveness model makes it significantly more tractable.

[27] Carrol and Sommer (in progress) also study the implications of slow dissemination of information for consumption.

### 7.2 Habit formation

A popular theory of consumption has stressed that consumers may develop habits over consumption. In its simplest form, this model assumes that utility at time $t$ depends on $c_t - \gamma c_{t-1}$, with $\gamma > 0$, so that higher consumption last period creates a habit that lowers utility this period. In this case, optimal consumption growth is: $\Delta c_{t+1} = \gamma \Delta c_t + e_{t+1}$ (see Deaton, 1992, pp. 31-33).

Taken as a model of a representative consumer, the habit theory predicts that aggregate consumption follows an AR(1). Since an $AR(1)$ is also an $MA(\infty)$ with declining coefficients, and if $\gamma$ is not too large the higher-order moving average coefficients are negligible, then a representative consumer with a habit model generates aggregate consumption observations very close to the $MA(I+1)$ with declining coefficients predicted by the inattentiveness model. The two models are close to indistinguishable using only the stochastic properties of aggregate consumption, and the inattentiveness model can be seen as providing a "micro-foundation" for a representative consumer with a habit.[28]

Using other information aside from the time-series properties of aggregate consumption, the two models can be distinguished. The representative consumer habit model predicts that consumption should respond sluggishly to any event. The inattentiveness model on the other hand predicts that in response to an event that is very noticeable and grabs the attention of the population, consumption should respond instantly. One notable such event is the end of hyper- and high-inflations, which usually occurs suddenly with the implementation of drastic and well-publicized stabilization programs. Fischer et al. (2002) examine 45 such episodes in 25 countries since 1960. They find that these noticeable disinflation programs have a large effect on real variables, and especially that aggregate consumption responds immediately. They write "...per-capita consumption growth: it is essentially zero in the year before stabilization and jumps to around 2 percent in the year of stabilization..." and moreover they find that after this immediate reaction, consumption growth is stable in the following two years. This immediate response of consumption to these noticeable events is consistent with the inattentiveness model, but not with a habit model that predicts a sluggish response with respect to all shocks.

It is easier to distinguish habit formation at the level of the individual from inattentiveness. Inattentiveness implies that individual consumption adjusts infrequently, so that all the sluggishness in aggregate consumption comes from aggregation, whereas habit formation predicts that individual consumption is serially correlated. Dynan (2002) uses data from the PSID to find that individual consumption growth is close to serially uncorrelated. Therefore, when she estimates the optimality conditions imposed by the habit model she finds no evidence for habits. Her findings are consistent with inattentiveness.

[28] With more flexible specifications of habits, the match between the two models may be even closer. Chetty and Szeidl (2003) show that, under some circumstances, a model with time-contingent consumption adjustment exactly mimics the aggregate consumption dynamics that would be chosen by a representative agent with a specific habit formation process.

### 7.3 State-contingent adjustment

Caballero (1995) proposes a model of non-durables consumption in which it is costless to obtain, acquire, and process information, but it is costly to implement the optimal consumption plan. Consumers are always attentive, but only adjust consumption when the deviation between current consumption and the consumption that would be optimal given the current state exceeds a certain threshold. Consumption adjustment is now *state-contingent,* occurring at sporadic dates contingent on the state of the economy at those dates.

Because both models imply a disconnect between available information and observed actions, given data on these two alone, it will generally be difficult to distinguish the two models. One way to do contrast the models is over the empirical realism of their assumptions. I have argued that it is costly to collect and process information and to compute an optimal solution. With state contingent adjustment though, every instant the agent *is* observing the full state of the economy, *is* processing this information to realize what is her wealth, and *is* performing costly computations to determine whether consumption is in the inaction region or not. State-contingent behavior is as complicated as following the full information rational expectations optimal plan, so it cannot be justified as describing "near-rational" behavior. Rather, it can only be justified by appealing to some actual physical cost of adjusting consumption. It is difficult to find evidence of any relevant such cost in the consumption of non-durables.

Another way to compare the models is by looking at an intermediate step in the disconnect between publicly available information and observed actions: the information that agents have. According to the inattentiveness model, agent's private information, expectations and future plans are only sporadically updated, whereas in the state-contingent model private and public information coincide at all instants. The evidence discussed in Section 7 that individual expectations are consistent with slow dissemination of information; that in surveys a large fraction of agents reports being unaware of important current economic events; and that reported planning behavior is a determinant of accumulated wealth, all support the inattentiveness model but would not be predicted by a state-contingent adjustment model.[29]

## 8 Conclusion

In his Nobel lecture, James Tobin (1982, page 189) wrote:

*"Some decisions by economic agents are reconsidered daily or hourly, while others are reviewed at intervals of a year or longer except when extraordinary events compel revisions. It would be*

[29] A few other papers worth mentioning are: Goodfriend (1992) and Pischke (1995) who assume that agents cannot distinguish between permanent and transitory income shocks, Ameriks et al.'s (2003b) model of absent-minded consumers who cannot keep track of how much they have already consumed, and models where agents have full information on the present but only recall the past infrequently (Mullainathan, 2002; Wilson, 2003; and Bernheim and Thomadsen, 2002).

*desirable in principle to allow for differences among variables in frequencies of change and even to make these frequencies endogenous. But at present, models of such realism seem beyond the power of our analytical tools."*

In this paper, I developed some of the tools that Tobin called for and examined the implications of modelling behavior in this way for the dynamics of aggregate consumption. I assumed (and justified) the existence of decision costs inducing agents to only sporadically update their decisions and characterized the decisions of these agents on how much to consume and how often to plan. This individual behavior implies that information should be sticky in the aggregate economy, only gradually dissipating throughout the population, so that aggregate consumption adjusts slowly to the arrival of news. I found that this prediction is confirmed in U.S. data and that the model also generates dynamics for aggregate consumption which have the "excess sensitivity" and "excess smoothness" with respect to income that had been previously identified in the data. For individual consumption, the model predicted that consumption changes should be sensitive to small and unpredictable past shocks, but should not be sensitive to past large or predictable changes. This dichotomy reconciles the disparate findings of the many microeconomic studies which have studied the excess sensitivity of consumption to shocks. The model further predicted that information and expectations are only sporadically updated, which has also been shown to be the case using inflation expectations surveys. Finally, the model predicted that a group of people do not plan and save less than those who plan, and that among planners, those who plan for longer, save more. Again, this has been confirmed in the data.

Beyond passing tests in the data, the set of theoretical results and empirical estimates in this paper offer a plausible description of consumption behavior. There are two types of agents in the United States. About one third of people face high costs of planning (e.g., because of lack of education) and so rationally choose to never plan, living hand-to-mouth and consuming their income less a predetermined amount every period. These people save less and accumulate less wealth. Because they are poorer, they account for only a small fraction of aggregate consumption, around 5%. The bulk of aggregate consumption is accounted for instead by the other two thirds of people who form plans for consumption regularly. Because they only sporadically update their plans, these people react to small unexpected income shocks only gradually over time. Aggregate consumption therefore reacts sluggishly to shocks, but not too sluggishly since people do update their plans within a year or so.

Because the model in this paper is a model of how dynamic decisions are made and expectations are formed, in principle it is widely applicable to different economic problems. Decisions on how much to invest in stocks or bonds, how often to change prices or revise contracts are some to which the inattentiveness approach can be applied. While it is difficult to know for sure how successful these applications will be, there is enough promise to justify paying some attention to inattention.

## Appendix A - The discrete-time representation of consumption

Proof of Proposition 3

Treating the vector $(i,j)$ as a random variable with distribution $\Psi(i,j)$, equation (20) shows that (up to a constant), aggregate consumption growth is the expected value of $e_{t+1-j,t-i}$. Because $(i,j)$ can only take finitely many values, it is a *simple random variable* (Billingsley, 1995, Section 5) so the integrals in (20) are Riemann integrals and can be represented as sums. Breaking each unit interval into $N$ parts, $j$ takes $N$ equidistant values from 0 to $1-1/N$ and $i$ takes $IN+1$ equidistant values from 0 to $I$. Equation (20) becomes:

$$\Delta C_{t+1} = \sum_{k=0}^{N-1} \sum_{m=0}^{NI} e_{t+1-k/N,t-m/N} \Psi(m/N, k/N).$$

Recall that $e_{t,t-s}$ is a random variable such that $E_{t-s}\left[e_{t,t-s}\right] = 0$. It can be broken into independent increments by writing: $e_{t,t-s} = \int_{t-s}^{t} \varepsilon(v)dv$, where $\varepsilon(v)$ is a continuous time "white noise" process with $E\left[\varepsilon(v)^2\right] = \sigma_\varepsilon^2$ but $E\left[\varepsilon(v)\varepsilon(v-k)\right] = 0$ for any $k > 0$.[30] Then:

$$\Delta C_{t+1} = \sum_{k=0}^{N-1} \sum_{m=0}^{NI} \left[ \int_{t-m/N}^{t+1-k/N} \varepsilon(v)dv \right] \Psi(m/N, k/N).$$

Separating the random variables occurring after $t$ from those before $t$:

$$\begin{aligned} \Delta C_{t+1} &= \sum_{k=0}^{N-1} \sum_{m=0}^{NI} \left[ \int_{t}^{t+1-k/N} \varepsilon(v)dv + \int_{t-m/N}^{t} \varepsilon(v)dv \right] \Psi(m/N, k/N) \\ &= \sum_{k=0}^{N-1} \left[ \int_{t}^{t+1-k/N} \varepsilon(v)dv \right] P^j(k/N) + \sum_{m=0}^{NI} \left[ \int_{t-m/N}^{t} \varepsilon(v)dv \right] P^i(m/N). \qquad (35) \end{aligned}$$

The last expression uses $P^j(.)$ to denote the marginal distribution of $j$, $P^j(k/N) = \sum_{m=0}^{NI} \Psi(m/N, k/N)$, as well as $P^i(.)$ to denote the marginal distribution over the $i$, $P^i(m/N) = \sum_{k=0}^{N-1} \Psi(m/N, k/N))$.

Breaking the integrals in (35) into independent increments in intervals of length $1/N$:

$$\begin{aligned} &\sum_{k=0}^{N-1} \left[ \int_{t}^{t+1/N} \varepsilon(v)dv + \int_{t+1/N}^{t+2/N} \varepsilon(v)dv + ... + \int_{t+1-k/N-1/N}^{t+1-k/N} \varepsilon(v)dv \right] P^j(k/N) \\ &+ \sum_{m=0}^{NI} \left[ \int_{t-1/N}^{t} \varepsilon(v)dv + \int_{t-2/N}^{t-1/N} \varepsilon(v)dv + ... + \int_{t-m/N}^{t-m/N+1/N} \varepsilon(v)dv \right] P^i(m/N) \\ = &\sum_{k=0}^{N-1} \left[ \sum_{n=0}^{N-k-1} \int_{t+n/N}^{t+n/N+1/N} \varepsilon(v)dv \right] P^j(k/N) + \sum_{m=0}^{NI} \left[ \sum_{n=0}^{m/N-1/N} \int_{t-n-1/N}^{t-n} \right] P^i(m/N). \end{aligned}$$

[30] More rigorously, $\varepsilon(v)dv = \zeta(dv)$, where $\zeta(dv)$ is a random measure defined on all subsets of the real line such that $E[\zeta(dv)] = 0$, $E[\zeta(dv)^2] = \sigma^2 dv$, and $E[\zeta(\Delta_1)\zeta(\Delta_2)] = 0$ for any disjoint sets $\Delta_1$ and $\Delta_2$(Rozanov, 1967).

Collecting all the terms corresponding to each $1/N$ length interval gives:

$$\Delta C_{t+1} = \sum_{k=0}^{N-1} \left[ \int_{t+1-k/N-1/N}^{t+1-k/N} \varepsilon(v)dv \right] G^j(k) + \sum_{m=1}^{NI} \left[ \int_{t-m/N}^{t-m/N+1/N} \varepsilon(v)dv \right] G^i(m),$$

where I defined $G^j(k) \equiv \sum_{p=0}^{k} P^j(p/N)$, which is increasing in $k$; and $G^i(m) \equiv \sum_{p=m}^{NI} P^i(p/N)$, which is decreasing in $m$. One can then re-write this expression as an $MA(N+NI)$ process with independent increments:

$$\begin{aligned} \Delta C_{t+1} &= \sum_{k=0}^{N(I+1)-1} u_{t+1-k/N} F(k), \\ \text{with } u_{t+1-k/N} &\equiv \int_{t+1-k/N}^{t+1-k/N+1/N} \varepsilon(v)dv, \end{aligned} \tag{36}$$

where $F(k) = G^j(k)$ for $k = 0, ..., N-1$, while $F(k) = G^i(k-N)$ for $k = N, ..., N(I+1)-1$. Clearly, $E_{s-1/N}[u_s] = 0$ and $E[u_s u_k] = 0$, while $F(k)$ is increasing from $k = 0$ to $N-1$, and decreasing from $N$ to $N(I+1)-1$.

Given (36), the process for aggregate consumption changes in discrete time is:

$$\begin{aligned} \Delta C_{t+1} &= \sum_{s=0}^{I} \left( \sum_{k=0}^{N-1} u_{t+1-k/N-s} F(sN+k) \right) \\ &= \Phi(0)e_{t+1} + \Phi(1)e_t + ... + \Phi(I)e_{t-I+1}, \\ \text{defining} &: \quad \Phi(s) \equiv \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} F(sN+k)^2}, \\ e_{t+1-s} &\equiv \frac{1}{\Phi(s)} \sum_{k=0}^{N-1} u_{t+1-k/N-s} F(sN+k). \end{aligned}$$

Clearly, $E_{t-s}[e_{t+1-s}] = 0$ and $Var[e_{t+1-s}] = \sigma_\varepsilon^2$. This proves the first part of Proposition 3. Since $F(k)$ is decreasing for $k \geq N$, then $\Phi(s)$ is also decreasing for $s = 1, 2..., I$, which completes the proof of Proposition 3.□

Time Aggregation

Using the sum representation of the Riemann integral, we observe:

$$\bar{C}_{t+1} - \bar{C}_t = \frac{1}{N} \sum_{p=0}^{N-1} \Delta C_{t+1-p/N} = \sum_{p=0}^{N-1} \sum_{k=0}^{N(I+1)-1} u_{t+1-p/N-k/N} F(k),$$

where the second equality follows from (36). I can collect terms to see that $\Delta\bar{C}_{t+1}$ equals:

$$\sum_{p=0}^{N-1}\left[u_{t+1-p/N}\left(\frac{\sum_{v=0}^{p}F(v)}{N}\right)\right]+\sum_{s=1}^{I}\left[\sum_{k=0}^{N-1}u_{t+1-k/N-s}\left(\frac{\sum_{v=N(s-1)+k+1}^{Ns+k}F(v)}{N}\right)\right]$$
$$+\sum_{p=N(I+1)}^{N(I+2)-1}\left[u_{t+1-p/N}\frac{F(p-N)}{N}\right].$$

This can then be written in discrete time as:

$$\Delta\bar{C}_{t+1} = \tilde{\Phi}(0)e_{t+1}+\tilde{\Phi}(1)e_t+...+\tilde{\Phi}(I)e_{t-I+1}+\tilde{\Phi}(I+1)e_{t-I},$$
$$\tilde{\Phi}(s) \equiv \begin{cases} \sqrt{\frac{1}{N^2}\sum_{p=0}^{N-1}\left(\sum_{v=0}^{p}F(v)\right)^2}, \text{ for } s=0 \\ \sqrt{\frac{1}{N^2}\sum_{k=0}^{N-1}\left(\sum_{v=N(s-1)+k+1}^{Ns+k}F(v)\right)^2}, \text{ for } s=1,...,I, \\ \sqrt{\frac{1}{N}\sum_{p=NI}^{N(I+1)-1}F(p)^2}, \text{ for } s=I+1 \end{cases}$$

Time aggregation therefore turns an $MA(I)$ process into an $MA(I+1)$. The non-increasing pattern of the $\tilde{\Phi}(i)$ is unaltered, and applies up to $I+1$.

# Appendix B - Spectrum of consumption

Proof of Proposition 5

Recall four results: (a) De Moivre's formula, $e^{-i\omega j}=\cos(\omega j)-i\cdot\sin(\omega j)$, (b) $\sin(-\omega j)=-\sin(\omega j)$, (c) $\cos(\omega j)=\cos(-\omega j)$, and (d) that since the MA(I) process in equation (21) is stationary, its autocovariance function is symmetric ($\gamma_j=\gamma_{-j}$). The formula for the power spectrum in footnote 9 then is:

$$h_{\Delta C}(\omega)=\frac{1}{2\pi}\left[\gamma_0+2\sum_{j=1}^{\infty}\gamma_j\cos(\omega j)\right]. \tag{37}$$

For the process in (21) the autocovariance function is:

$$\gamma_j=\begin{cases}\sigma^2\sum_{k=0}^{I-j}\Phi(k)\Phi(k+j), \text{ for } j=0,1,2,...,I \\ 0, \text{ for } j>I\end{cases}$$

Replacing this into (37) and dividing by $\gamma_0$, gives the expression in Proposition 5, which depends only on $\{\Phi(i)/\Phi(0)\}$ for $i$ from 1 to $I$. Evaluating (23) at frequency zero and rearranging gives the excess smoothness ratio in (22).□

# Appendix C - CARA-utility, ARMA-income, consumer problem

Proof of Proposition 6

The Ramsey rule in equation (18) with CARA utility implies:

$$c_t^*=\frac{(r-\rho)t}{\alpha}+c_0^*. \tag{38}$$

Using this to substitute out $c_t$ in the budget constraint (26), a little algebra shows that wealth at the next planning period is:

$$w' = e^{rd^*}\left[w - \frac{(r-\rho)(1-e^{-rd^*}-rd^*e^{-rd^*})}{\alpha r^2} + \int_0^{d^*} e^{-rt}\left(\frac{\sigma_P}{r}dz_t^P + \frac{\sigma_T}{r+\phi}dz_t^T\right)\right] - K - \frac{e^{rd^*}-1}{r}c_0^*$$

Since $w'$ is a linear combination of normally distributed variables, it is normally distributed with:

$$E\left[w'\right] = e^{rd^*}\left[w - \frac{(r-\rho)(1-e^{-rd^*}-rd^*e^{-rd^*})}{\alpha r^2}\right] - K - \frac{e^{rd^*}-1}{r}c_0^*, \tag{39}$$

$$Var\left[w'\right] = \frac{\sigma^2}{2r}(e^{2rd^*}-1), \tag{40}$$

Next, I make the (educated) guess that the value function is exponential: $V(w) = -A\exp(-Bw)$, where $A$ and $B$ are coefficients to be determined. The envelope theorem condition becomes:

$$-Bw = (r-\rho)d^* + \ln\left(E\left[e^{-Bw'}\right]\right). \tag{41}$$

Since $w'$ is normally distributed, from the properties of the log-normal distribution, $\ln\left[E\left[\exp(-Bw')\right]\right]$ equals $-BE\left[w'\right] + B^2Var\left[w'\right]/2$. Using this result and (39)-(40) in (41), gives the solution for $c_0^*$:

$$c_0^* = rw - \frac{rK}{e^{rd^*}-1} - \frac{B\sigma^2}{4}(e^{rd^*}+1) - \frac{(r-\rho)}{e^{rd^*}-1}\left[\frac{rd^*}{B} + \frac{e^{rd^*}-1-rd^*}{\alpha r^2}\right]. \tag{42}$$

Combining the first-order condition (13) at $t=0$ with the envelope theorem in (14) gives:

$$e^{-\alpha c_0^*} = ABe^{-Bw}. \tag{43}$$

If the guess of the value function is valid, (43) must hold for all possible realizations of $w$. Matching coefficients shows that $B = \alpha r$. Going back to (42) with this result gives:

$$c_0^* = rw - \frac{rK}{e^{rd^*}-1} - \frac{\alpha r\sigma^2}{4}(e^{rd^*}+1) - \frac{r-\rho}{\alpha r}. \tag{44}$$

The last optimality condition is the first-order condition with respect to $d$, which is just: $\partial V(w)/\partial d = 0$. Given the guess for the value function,

$$V(w) = \max_d\left\{-\frac{e^{-\alpha c_0^*}}{\alpha r}\right\},$$

the first-order condition is just $\partial c_0^*/\partial d = 0$, which I can evaluate using (44) to obtain:

$$(e^{rd^*}-1)^2 = \frac{4K}{\alpha\sigma^2}. \tag{45}$$

Solving this equation gives (27). Using the solution for $d^*$ in (44) gives the solution for $c_0^*$ in (29).

The final step is to verify that the guess for the value function accords with the Bellman equation and satisfies the transversality condition. This is left to the reader.□

# Appendix D - The inattentive saver's problem

Proof of Proposition 7

The problem facing the agent can be written as:

$$W(w) = \max_{d,\{s_t\}} E\left[\int_0^d e^{-\rho t} u(y_t - s_t)dt + e^{-\rho d} W(w')\right] \tag{46}$$

$$\text{s.t.} \quad da_t = (ra_t + s_t)dt \tag{47}$$

Integrating (47) between two decision dates, using the fact that $w' = w_d - K$, that $y_t = y_t^P + y_t^T$, and the definition $w_t = a_t + y_t^P/r + y_t^T/(r+\phi)$, leads to:

$$w' = e^{rd}\left(w + \int_0^d e^{-rt} s_t dt\right) - K + \frac{y^{P'} - e^{rd}y^P}{r} + \frac{y^{T'} - e^{rd}y^T}{r+\phi}.$$

Since permanent income follows a Brownian motion, $y^{P'}$ is normally distributed with mean $y^P$ and variance $\sigma_P^2 d$. Likewise, since $dy_t^T = -\phi y_t^T dt + \sigma_T dz_t^T$, then transitory income is normally distributed with mean $y^T \exp(-\phi d)$ and variance $\sigma_T^2(1-\exp(-2\phi d))/2\phi$. Therefore, $w'$ is normally distributed with:

$$E_0\left[w'\right] = e^{rd}\left(w + \int_0^d e^{-rt} s_t dt\right) - K + \frac{1-e^{rd}}{r} y^P + \frac{e^{-\phi d} - e^{rd}}{r+\phi} y^T, \tag{48}$$

$$Var_0\left[w'\right] = \frac{\sigma_P^2}{r^2} d + \frac{\sigma_T^2(1-e^{-2\phi d})}{2\phi\,(r+\phi)^2} \tag{49}$$

The first-order conditions determining the optimal choices of $s_t$ are:

$$E_0\left[u'(y_t - s_t)\right] = e^{(r-\rho)(d-t)} E_0\left[W_w(w')\right], \quad \text{for } t \in [0,d). \tag{50}$$

Combining this equation for time $t$ and for time 0:

$$u'(y_0 - s_0) = e^{(r-\rho)t} E_0\left[u'(y_t - s_t)\right] \Leftrightarrow$$
$$-\alpha y_0 + \alpha s_0 = (r-\rho)t + \alpha s_t + \ln\left(E_0\left[e^{-\alpha y_t}\right]\right).$$

Using the normality of $y_t$, it takes a few steps to obtain:

$$s_t = s_0 - (1 - e^{-\phi t}) y_0^T - \frac{\alpha}{2}\left(\sigma_P^2 t + \frac{\sigma_T^2(1-e^{-2\phi t})}{2\phi}\right) - \frac{(r-\rho)\,t}{\alpha}. \tag{51}$$

The envelope theorem condition is:

$$W_w(w) = e^{(r-\rho)d} E\left[W_w(w')\right]. \tag{52}$$

Again, I guess that the value function is exponential: $W(w) = -Ae^{-\alpha r w}$, where $A$ is a coefficient to be determined. Taking logs of (52), and using the properties of the log-normal distribution together with (48)-(49) gives leads to:

$$\begin{aligned} w(e^{rd}-1) \quad = \quad & \frac{(r-\rho)d}{\alpha r} - e^{rd}\int_0^d e^{-rt} s_t dt + K + \frac{e^{rd}-1}{r} y^P + \frac{e^{rd}-e^{-\phi d}}{r+\phi} y^T \\ & + \frac{\alpha \sigma_P^2}{2r} d + \frac{\alpha r \sigma_T^2 (1-e^{-2\phi d})}{4\phi (r+\phi)^2}. \end{aligned}$$

Using the solution for $s_t$ in (51) to substitute out savings in this equation gives, after rearranging,:

$$s_0 = -rw + y + \frac{r-\rho}{\alpha r} + \frac{\alpha \sigma_P^2}{2r} + \frac{\alpha \sigma_T^2}{2(r+2\phi)} + \frac{rK}{e^{rd}-1} - \frac{\alpha r \phi \sigma_T^2 (1-e^{-2\phi d})}{4(r+2\phi)(r+\phi)^2(e^{rd}-1)} \tag{53}$$

Combining the envelope theorem (52) with (50) gives the condition:

$$u'(y_0 - s_0) = W_w(w).$$

Using the form of the utility function, the guess for the value function, and the expression for $s_0$ in (53), this I can solve this equation $A$ to obtain:

$$A = \frac{1}{\alpha r} \exp\left\{ \frac{r-\rho}{r} + \frac{\alpha^2 \sigma_P^2}{2r} + \frac{\alpha^2 \sigma_T^2}{2(r+2\phi)} + \frac{\alpha r K}{e^{rd}-1} - \frac{\alpha^2 r \phi \sigma_T^2 (1-e^{-2\phi d})}{4(r+2\phi)(r+\phi)^2(e^{rd}-1)} \right\}. \tag{54}$$

Given (54) and the guess for the value function, to maximize $W(w)$ with respect to $d$ is equivalent to minimizing $A$ with respect to $d$, which in turn is equivalent to minimizing:

$$\hat{A}(K,d) \equiv \frac{K}{e^{rd}-1} - \frac{\alpha \phi \sigma_T^2 (1-e^{-2\phi \hat{d}})}{4(r+2\phi)(r+\phi)^2(e^{r\hat{d}}-1)}. \tag{55}$$

The first-order necessary condition for an interior minimum is:

$$\frac{e^{(r-2\phi)d}}{\Xi\left(e^{rd}-1\right)^2} \underbrace{\left[re^{2\phi d}(1-K\Xi) + 2\phi e^{-rd} - 2\phi - r\right]}_{\equiv B(d)} = 0, \tag{56}$$

$$\text{where} : \Xi \equiv \frac{4(r+2\phi)(r+\phi)^2}{\alpha \phi \sigma_T^2}. \tag{57}$$

If $K\Xi > 1$, $B(d)$ is always negative, which implies that $\hat{A}$ falls monotonically with $d$, and so the optimal $\hat{d}$ is $+\infty$. Otherwise, $\hat{d}$ is the zero of $B(d)$. Straightforward evaluation and differentiation of $B(d)$ shows that with strictly positive costs of planning: $B(0) < 0$, $B_d(0) < 0$, $B_{dd}(.) > 0$, and $\lim_{d \to +\infty} B(D) = +\infty$. Thus, there is a unique solution to $B(d) = 0$, where $B(d)$ cuts the horizontal axis from below, and therefore there is a unique optimal $\hat{d}$.□

## Appendix E - Consumption versus savings plans

Proof of Proposition 8

The agent prefers a consumption plan if the value from doing so, $V(w)$, is larger than the value from following a savings plan, $W(w)$. Using the expressions in (30) and (54), the condition $V(w) > W(w)$ becomes:

$$H(K) \equiv \frac{rK}{e^{r\hat{d}} - 1} - \frac{\alpha r \phi \sigma_T^2 (1 - e^{-2\phi \hat{d}})}{4(r + 2\phi)(r + \phi)^2 (e^{r\hat{d}} - 1)} - r\sqrt{\alpha \sigma^2 K} + \frac{\alpha \sigma_T^2 \phi^2}{2(r + 2\phi)(r + \phi)^2} > 0.$$

If $K = 0$, then $\hat{d} = 0$ and using L'Hopital's rule it follows that $H(0) = 0$: under full information rational expectations, consumption and savings plans are equivalent. Moreover, when $K > 1/\Xi$ and so $\hat{d} = +\infty$, then the first two terms in the definition of $H(K)$ are zero, so clearly $H(K)$ is declining in $K$ tending towards minus infinity. More generally, using the envelope theorem:

$$H_K(.) = \frac{r}{e^{r\hat{d}} - 1} - r\sqrt{\frac{\alpha \sigma^2}{4K}} = \frac{r}{e^{r\hat{d}} - 1} - \frac{r}{e^{rd^*} - 1},$$

where the second equality follows from (45). Then, $sign\{H_K(.)\} = sign\left\{d^* - \hat{d}\right\}$, so I must compare optimal inattentiveness with consumption and savings plans.

Evaluating the function $B(d)$ defined in (56), whose zero is the optimal inattentiveness with savings, at the optimal inattentiveness with consumption $d^*$, replacing for $K$, gives:

$$F(d^*) \equiv re^{2\phi d^*} \left(1 - \frac{\Xi \alpha \sigma^2 \left(e^{rd^*} - 1\right)^2}{4}\right) + 2\phi e^{-rd^*} - 2\phi - r.$$

Since I know that if $B(d)$ is negative it is to the left of its zero, and when it is positive it is to the right of its zero, then when $F(d^*)$ is positive it follows that $d^* > \hat{d}$. Conversely when $F(d^*)$ is negative, then $d^* < \hat{d}$, and at $\hat{d}$, $F(\hat{d}) = 0$.

Straightforward evaluation and differentiation of $F(.)$ shows that: $F(0) = 0$, $F_d(0) = 0$, and $F_{dd}(0) = 2r\phi(2\phi + r) - r^3 \Xi \alpha \sigma^2 / 2$. Using the definition of $\Xi$ in (57) shows that if the assumption in Proposition 8 holds, then $F_{dd}(0) > 0$. Thus, close to 0, $F(.)$ is positive and so $d^* > \hat{d}$.

Next, I will show that aside from the trivial intersection at 0, $d^* = \hat{d}$ only once. Note that the derivative of $F(.)$ at a point of intersection is:

$$
\begin{aligned}
F_d(\hat{d}) &= 2\phi r e^{2\phi\hat{d}}\left(1-\frac{\Xi\alpha\sigma^2\left(e^{r\hat{d}}-1\right)^2}{4}\right)-\frac{r^2 e^{(2\phi+r)\hat{d}}\Xi\alpha\sigma^2\left(e^{r\hat{d}}-1\right)}{2}-2\phi r e^{-r\hat{d}} \\
&= 2\phi\left(2\phi+r-2\phi e^{-r\hat{d}}\right)-\frac{r^2 e^{(2\phi+r)\hat{d}}\alpha\sigma^2\Xi\left(e^{r\hat{d}}-1\right)}{2}-2\phi r e^{-r\hat{d}} \\
&= 2\phi\left(r+2\phi\right)\left(1-e^{-r\hat{d}}\right)\left(1-\frac{r^2\alpha\sigma^2\Xi e^{2(\phi+r)\hat{d}}}{4\phi(r+2\phi)}\right),
\end{aligned}
$$

where the second line follows from replacing the first term using the condition $F(\hat{d})=0$, and the third line follows from rearranging. Then, it is clear that if $\hat{d}$ is small enough, $F_d(\hat{d})$ is positive, but once $\hat{d}$ rises above a certain threshold, it becomes negative forever. Now, since for small $K$, $F(d^*)$ is positive, this continuous function must intersect the horizontal axis first at a point where $F_d(\hat{d})<0$. Towards a contradiction, say that is intersects the horizontal axis again at some higher $d$. By continuity of the $F(d)$ function, it must cut the axis from below. Yet, we know that at any zero of the $F(d)$ function the slope must be negative, which leads to a contradiction. Therefore, $d^*=\hat{d}$ only once at some value of $K$, and if the costs of planning exceed this value then $\hat{d}>d^*$.

Returning back to the initial aim of studying $H(.)$ I conclude that starting from 0 when $K=0$, the function increases up to a certain $K$ (when $d^*=\hat{d}$). Then it declines monotonically towards minus infinity, intersecting the horizontal axis at a unique point $\hat{K}$. Therefore, if $K\in(0,\hat{K})$, then $H(K)>0$, so consumption plans are preferred. If $K>\hat{K}$, savings plans are preferred.

Finally, note that at $\hat{K}$ where $H(\hat{K})=0$, we know that $H_K(\hat{K})<0$, and so that $\hat{d}>d^*$; therefore when $K$ passes $\hat{K}$ and the agent shifts from consumption to savings plans, her inattentiveness takes a discontinuous jump from $d^*$ to $\hat{d}$.□

Proof of Corollary 4

If $K>1/\Xi$, then $\hat{d}=+\infty$. Also, consumption plans are preferred as long as $H(K)>0$, which if $\hat{d}=+\infty$ becomes:

$$
K<\bar{K}\equiv\frac{\alpha\sigma_T^4\phi^4}{4(r+2\phi)^2(r+\phi)^2\left((r+\phi)^2\sigma_P^2+r^2\sigma_T^2\right)}.
$$

Moreover, if $K>\hat{K}$, then savings plans are preferred. Combining these three facts, it follows that if $\bar{K}>1/\Xi$, then $\hat{K}=\bar{K}$. Using the definitions of $\bar{K}$ and $\Xi$, the condition $\bar{K}>1/\Xi$ becomes the condition in Corollary 4.□

Proof of Proposition 9

Using the solutions for $V(w)$ and $W(w)$ in (30) and (54) with $\hat{d}=+\infty$, $V(w)>W(w)$ leads to:

$$c_0^* > rw - \frac{r-\rho}{\alpha r} - \frac{\alpha\sigma_P^2}{2r} - \frac{\alpha\sigma_T^2}{2\,(r+2\phi)}. \tag{58}$$

Using the solution for $c_0^*$ in (28) gives, after cancelling terms:

$$\frac{4K}{e^{rd^*}-1} + \alpha\left(\frac{\sigma_P^2}{r^2} + \frac{\sigma_T^2}{(r+\phi)^2}\right)(e^{rd^*}+1) < 2\alpha\left(\frac{\sigma_P^2}{r^2} + \frac{\sigma_T^2}{r(r+2\phi)}\right).$$

Using (45) to replace for $K$ and rearranging gives the condition in (32).□

<u>Proof of Corollary 3</u>

Using the fact that $\hat{c}_0 = y_0 - \hat{s}_0$ and (53) with $\hat{d} = +\infty$, shows that:

$$\hat{c}_0 = rw_0 - \frac{r-\rho}{\alpha r} - \frac{\alpha}{2}\left(\frac{\sigma_P^2}{r} + \frac{\sigma_T^2}{r+2\phi}\right).$$

Then, for $\hat{s}_0 < s_0^*$, it must be that $\hat{c}_0 > c_0^*$, which using the expressions above is equivalent to condition (58) holding, which is true for the agent who chooses to be an inattentive saver.□

## Appendix F - Hybrid consumption-savings plans

The problem to solve is:

$$\begin{aligned}
Z(w) &= \max_{d,\lambda,\{\tilde{c}_t\}} E\left[\int_0^d e^{-\rho t} u(\lambda y_t + \tilde{c}_t)dt + e^{-\rho d} Z(w')\right] \\
\text{s.t.}\quad w' &= e^{rd}\left(w - \int_0^d e^{-rt}\tilde{c}_t dt\right) - K + e^{rd}(1-\lambda)\int_0^d e^{-rt} y_t dt + \frac{y^{P'} - e^{rd}y^P}{r} + \frac{y^{T'} - e^{rd}y^T}{r+\phi},
\end{aligned}$$

where the constraint is derived by combining the law of motion for assets, the definition of wealth, and the consumption rule $c_t = \lambda y_t + \tilde{c}_t$.

The first-order condition with respect to $\tilde{c}_t$ is:

$$E\left[u'(\lambda y_t + \tilde{c}_t)\right] = e^{(r-\rho)(d-t)} E\left[Z'(w')\right]. \tag{59}$$

Combining this condition at time 0 with that at some $t < d$ gives:

$$\begin{aligned}
u'(\lambda y_0 + \tilde{c}_0) &= e^{(r-\rho)t} E\left[u'(\lambda y_t + \tilde{c}_t)\right] \Leftrightarrow \\
-\alpha\lambda y_0 - \alpha\tilde{c}_0 &= (r-\rho)t - \alpha\tilde{c}_t - \alpha\lambda E[y_t] + \frac{\alpha^2\lambda^2}{2} Var[y_t] \Leftrightarrow \\
\tilde{c}_t &= \tilde{c}_0 + \lambda(1-e^{-\phi t})y_0^T + \frac{(r-\rho)t}{\alpha} + \frac{\alpha\lambda^2}{2} Var[y_t]. \qquad (60)
\end{aligned}$$

The second line follows from the CARA form of the utility function and the normality of income, and the third line from rearranging. I guess that the value function has the same exponential form as before: $Z(w) = -A\exp(-\alpha r w)$, with the coefficient $A$ to be determined. The envelope theorem condition is:

$$Z'(w) = e^{(r-\rho)d} E\left[Z(w')\right] \tag{61}$$

The first order condition (59) at time 0, combined with this condition, leads to:

$$\begin{aligned} e^{-\alpha(\lambda y_0 + \tilde{c}_0)} &= \alpha r A e^{-\alpha r w} \Leftrightarrow \\ \tilde{c}_0 &= -\frac{\ln(\alpha r A)}{\alpha} + rw - \lambda y_0 \end{aligned} \tag{62}$$

Using the solutions for $\tilde{c}_t$ in (60) and $\tilde{c}_0$ in (62) to substitute for the consumption terms in the budget constraint and rearranging shows that $w'$ is normally distributed with:

$$\begin{aligned} E[w'] &= w + \frac{\ln(\alpha r A)(e^{rd}-1)}{\alpha r} + \frac{\left(1 + dr - e^{rd}\right)(r-\rho)}{\alpha r^2} - \frac{\alpha \lambda^2 e^{rd}}{2} \int_0^d e^{-rt} Var[y_t] dt - K \\ Var[w'] &= Var\left[\frac{y^{P'} - e^{rd} y^P}{r} + \frac{y^{T'} - e^{rd} y^T}{r+\phi} + e^{rd}(1-\lambda) \int_0^d e^{-rt} \left(y_t^P + y_t^T\right) dt\right] \end{aligned}$$

Using the envelope theorem condition (61), together with the guess for the value function, the normality of wealth, and the expression for $E[w']$, gives:

$$\ln(\alpha r A) = \frac{r-\rho}{r} + \frac{\alpha^2 r \lambda^2 e^{rd}}{2(e^{rd}-1)} \int_0^d e^{-rt} Var[y_t] dt + \frac{\alpha r K}{(e^{rd}-1)} + \frac{\alpha^2 r^2 Var[w']}{2(e^{rd}-1)}.$$

The fact that $A$ does not depend on the state $w_t$ or on any component of income, validates the guess for the value function. The optimal $\lambda$ and $d$ are then the solution to the problem:

$$\min_{d,\lambda} \left\{ \frac{\alpha \lambda^2 e^{rd} \int_0^d e^{-rt} Var[y_t] dt + 2K + \alpha r Var[w']}{(e^{rd}-1)} \right\}$$

Since none of the expressions in this objective function depend on the sate of the economy, the optimal $d$ and $\lambda$ are independent of the state. Using the stochastic processes for $y^T$ and $y^P$, $Var[y_t]$ and $Var[w']$ can be easily (but tediously) evaluated. The results in Table 1 are found by solving this minimization numerically.

## Appendix G - Extraordinary events

Proof of Proposition 10

Define accumulated ordinary income shocks as:

$$\varepsilon_t = e^{rt} \int_0^t e^{-rs} \left( \frac{\sigma_P}{r} dz_s^P + \frac{\sigma_T}{r+\phi} dz_s^T \right).$$

It follows from the properties of Wiener processes that $\varepsilon_t \sim N(0, \sigma^2(e^{2rt}-1)/2r)$, where $\sigma^2 = \sigma_P^2/r^2 + \sigma_T^2/\left(r+\phi\right)^2$ and that $d\varepsilon_t = r\varepsilon_t + \left(\frac{\sigma_P}{r} dz_t^P + \frac{\sigma_T}{r+\phi} dz_t^T\right)$. Then, if I define $\bar{w}_t = w_t - \varepsilon_t$, the law of motion for $w_t$ implies that $d\bar{w}_t = (r\bar{w}_t - c_t)dt$.

Denote the value function in terms of $\bar{w}_{D+t}$, and in terms of how long has elapsed since the last planning date $t$ by $J(\bar{w}_{D+t}, t)$. This is an optimal stopping problem. The Bellman equation is:

$$\begin{aligned}(r+\delta)J(\bar{w}_{D+t}, t) &= \max_{c_{D+t}, d}\{u(c_{D+t}) + \frac{\delta}{2}E_0\left[J(\bar{w}_{D+d} + \varepsilon_{D+d} + u - K, 0) + J(\bar{w}_{D+d} + \varepsilon_{D+d} - u - K, 0)\right] \\ &\quad + J_w(\bar{w}_{D+t}, t)(r\bar{w}_{D+t} - c_{D+t}) + J_t(\bar{w}_{D+t}, t)\},\end{aligned}$$

and the value matching condition at the optimal stopping date is:

$$J(\bar{w}_{D+d^*}, d^*) = E_0\left[J(\bar{w}_{D+d^*} + \varepsilon_d - K, 0)\right].$$

To solve this problem, I guess that $J(\bar{w}, t) = -(A(t)/\alpha r)\exp(-\alpha r \bar{w})$, where $A(t)$ is a time varying function to be determined. The first-order condition for the optimal choice of $c_t$ is:

$$\begin{aligned}u'(c_{D+t}) &= J_w(\bar{w}_{D+t}, t) \Leftrightarrow \\ c_{D+t} &= r\bar{w}_{D+t} - \frac{\ln(A(t))}{\alpha}.\end{aligned} \tag{63}$$

The envelope theorem condition with respect to $\bar{w}_t$ is:

$$\begin{aligned}\delta J_w(\bar{w}_{D+t}, t) &= \frac{\delta}{2}E_0\left[J_w(\bar{w}_{D+d} + \varepsilon_{D+d} + u - K, 0) + J_w(\bar{w}_{D+d} + \varepsilon_{D+d} - u - K, 0)\right] \\ &\quad + J_{ww}(\bar{w}_{D+t}, t)\frac{\ln(A(t))}{\alpha} + J_{wt}(\bar{w}_{D+t}, t),\end{aligned}$$

where I used (63) to replace out consumption. Using the guess for the value function in this equation gives the differential equation in Proposition 9. Since $A(t)$ does not depend on $\bar{w}_{D+t}$, the guess of the value function was valid. Using it in the value matching condition gives the boundary condition in Proposition 9. Finally, $d^*$ can be found by minimizing $A(0)$.

## References

Abowd, John, and David Card (1989) "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, pp. 411-445.

Akerlof, George A. (1991) "Procrastination and Obedience," *American Economic Review Papers and Proceedings*, 81 (2), pp. 1-19.

Akerlof, George A., and Janet L. Yellen (1985) "A Near Rational Model of the Business Cycle with Wage and Price Inertia," *Quarterly Journal of Economics*, 100, pp. 823-838.

Alessie, Rob, Arie Kapteyn, and Annamaria Lusardi (1999) "Saving after Retirement: Evidence from Three Different Surveys," *Labour Economics*, 6 (2), pp. 277-310.

Ameriks, John, Andrew Caplin and John Leahy (2003a) "Wealth Accumulation and the Propensity to Plan," *Quarterly Journal of Economics*, 118 (3), pp. 1007-1047.

Ameriks, John, Andrew Caplin and John Leahy (2003b) "The Absent-Minded Consumer," unpublished, New York University.

Andrews, Donald W. K., and Christopher J. Monahan (1992) "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60 (4), pp. 953-966.

Attanasio, Orazio P. and Guglielmo Weber (1995) "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103 (6), pp. 1121-1157.

Bargh, John, and Tania Chartrand (1999) "The Unbearable Automaticity of Being," *American Psychologist*, 54, pp. 462-479.

Bernheim, B. Douglas and Raphael Thomadsen (2003) "Memory and Anticipations," unpublished manuscript, Stanford University.

Billingsley, Patrick (1995) *Probability and Measure*, 3rd edition, New York, Wiley & Sons.

Blanchard, Olivier J. and Danny Quah (1989) "The Dynamic Effects of Aggregate Demand and Aggregate Supply Disturbances," *American Economic Review*, 79, pp. 655-673.

Blanchard, Olivier J. and John Simon (2001) "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity*, pp. 135-174.

Blinder, Alan S., and Angus Deaton (1985) "The Time Series Consumption Function Revisited," *Brookings Papers on Economic Activity*, 2, pp. 465-511.

Bound, John, Charles Brown, Greg J. Duncan and Willard L. Rodgers (1994) "Evidence on the Validity of Cross-sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, vol. 12, pp. 345-368.

Browning, Martin and M. Dolores Collado (2001) "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates," *American Economic Review*, 91 (3) pp. 681-692.

Caballero, Ricardo J. (1995) "Near-Rationality, Heterogeneity, and Aggregate Consumption" *Journal of Money, Credit and Banking*, 27 (1), pp. 29–48.

Campbell, John Y. (1987) "Does Savings Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica*, 55, pp. 1249-1273.

Campbell, John Y. and Angus Deaton (1989) "Why is Consumption So Smooth?" *Review of Economic Studies*, 56, pp. 357-374.

Campbell, John Y. and N. Gregory Mankiw (1989) "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," *NBER Macroeconomics Annual*, 4, pp. 185-216.

Campbell, John Y. and N. Gregory Mankiw (1990) "Permanent Income, Current Income and Consumption," *Journal of Business and Economic Statistics*, 8 (3), pp. 265-279.

Carroll, Christopher D. (2003) "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of Economics*, 118 (1), pp. 269-298.

Carroll, Christopher D., and Martin Sommer (in progress) "Epidemiological Expectations and Consumption Dynamics," Johns Hopkins University, unpublished.

Chen, Tongwen, and Bruce Francis (1995) *Optimal Sampled-Data Control Systems*, Springer-Verlag, London.

Chetty, Raj, and Adam Szeidl (2003) "Consumption Commitments and Asset Prices," Harvard University, unpublished.

Christiano, Lawrence J., Martin Eichenbaum and David Marshall (1991) "The Permanent Income Hypothesis Revisited," *Econometrica*, 59 (2), pp. 397-423.

Cochrane, John H. (1989) "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives," *American Economic Review*, 79, pp. 319-337.

Deaton, Angus (1987) "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?" in *Advances in Econometrics, Fifth World Congress*, vol. 2, ed. Truman Bewley, Cambridge, Cambridge University Press.

Deaton, Angus (1992) *Understanding Consumption*, Oxford, Oxford University Press.

Dynan, Karen (2000) "Habit Formation in Consumer Preferences: Evidence from Panel Data," *American Economic Review*, 90, pp. 391-406.

Fischer, Stanley, Ratna Sahay, and Carlos A. Vegh (2002) "Modern Hyper- and High Inflations," *Journal of Economic Literature*, 150 (3), pp. 837-880.

Flavin, Marjorie A. (1981) "The Adjustment of Consumption to Changing Expectations about Future Income," *Journal of Political Economy*, 89, pp. 974-1009.

Franklin, Gene F., J. David Powell, Michael L. Workman (1990) *Digital Control of Dynamic Systems: second edition*, Addison-Wesley, Massachussets,

Friedman, Milton (1957) *A Theory of the Consumption Function*, Princeton, Princeton University Press.

Gabaix, Xavier and David Laibson (2001) "The 6D Bias and the Equity Premium Puzzle," *NBER Macroeconomics Annual 2001*, pp. 257-311.

Gali, Jordi (1991) "Budget Constraints and Time Series Evidence on Consumption," *American Economic Review,* 81, pp. 1238-1253.

Goodfriend, Marvin (1992) "Information-Aggregation Bias," *American Economic Review*, 82, pp. 508-519.

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

Hall, Robert E. and Frederick S. Mishkin (1982) "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 50, pp. 461-481.

Hayashi, Fumio and Christopher A. Sims (1983) "Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous, Instruments," *Econometrica*, 51 (3), pp. 783-798.

Hsieh, Chang-Tai (2003) "Do Consumers React to Anticipated Income Shocks? Evidence from the Alaska Permanent Fund," *American Economic Review*, 93 (1), pp. 397-405.

Hurst, Erik (2003) "Grasshoppers, Ants and Pre-Retirement Wealth: A Test of Permanent Income Consumers," University of Chicago, unpublished.

Kim, Chang-Jin and Charles R. Nelson (1999) "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *Review of Economics and Statistics*, 81 (4), 608-616.

Laibson, David (1997) "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 62, pp. 443-477

Lusardi, Annamaria (1999) "Information, Expectations, and Savings," in *Behavioral Dimensions of Retirement Economics*, ed. Henry Aaron, Brookings Institution Press/Russell Sage Foundation, New York, pp. 81-115.

Lusardi, Annamaria (2002) "Explaining Why So Many Households Do Not Save," Dartmouth College, unpublished.

MaCurdy, Thomas E. (1982) "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18, pp. 83-114.

Mankiw, N. Gregory (1985) "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *Quarterly Journal of Economics*, 100, pp. 529-537.

Mankiw, N. Gregory and Ricardo Reis (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117 (4).

Mankiw, N. Gregory and Ricardo Reis (2003) "Sticky Information: A Model of Monetary Nonneutrality and Structural Slumps," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, eds. Phillipe Aghion, Romain Frydman, Joseph Stiglitz and Michael Woodford, Princeton, Princeton University Press.

Mankiw, N. Gregory, Ricardo Reis and Justin Wolfers (2003) "Disagreement in Inflation Expectations," *NBER Macroeconomics Annual*, forthcoming.

McConnell, Margaret and Gabriel Perez Quiros (2000) "Output Fluctuations in the United States: What has Changed Since the Early 1980s?" *American Economic Review*, 90 , pp. 1464-1476.

Meghir, Costas, and Luigi Pistaferri (2003) "Income Variance Dynamics and Heterogeneity," *Econometrica*, forthcoming.

Miller, Boris M. and Wolfgang J. Runggaldier (1997) "Optimization of Observations: a Stochastic Control Approach," *SIAM Journal of Control and Optimization*, 35 (3), pp. 1030-1052.

Mishkin, Frederick (1983) *A Rational Expectations Approach to Macroeconometrics: Testing Policy Ineffectiveness and Efficient-Markets Models*, Chicago, University of Chicago Press.

Moreira, Marcelo J. (2003) "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71 (4), pp. 1027-1048.

Moscarini, Giuseppe (2003) "Limited Information Capacity as a Source of Inertia," *Journal of Economic Dynamics and Control*, forthcoming.

Mullainathan, Sendhil (2002) "A Memory-Based Model of Bounded Rationality", *Quarterly Journal of Economics*, 117 (3), pp. 735-774.

Muth, John (1960) "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 55, pp. 290-306.

Newey, Whitney and Kenneth West (1987) "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, pp. 703-708.

O'Donoghue, Ted and Matthew Rabin (1999) "Doing it Now or Later," *American Economic Review*, 89 (1), pp. 103-124

Parker, Jonathan (1999) "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes," *American Economic Review,* 89 (4), pp. 959-973.

Parker, Jonathan (2001) "The Consumption Risk of the Stock Market," *Brookings Papers on Economic Activity*, 2, pp. 279-348.

Pischke, Jorn-Steffen (1995) "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica*, 63 (4), pp. 805-840.

Ramsey, Frank (1928) "A Mathematical Theory of Saving," *Economic Journal*, 38, pp. 543-559.

Rosen, Sherwin (1990) "Comment," in *Issues in the Economics of Aging*, David Wise, ed., Chicago, The University of Chicago Press.

Rozanov, Yu A. (1967) *Stationary Random Processes*, San Francisco, Holden Day.

Shannon, Claude (1948) "A Mathematical Theory of Communication," *Bell Systems Technical Journal*, 27, pp. 379-423 and 623-656.

Shapiro, Matthew D. , and Joel Slemrod (1995) "Consumer Response to the Timing of Income: Evidence from a Change in Tax Withholding," *American Economic Review*, 85 (1), pp. 274-283.

Simon, Herbert A. (1978) "Rationality as Process and as Product of Thought," *American Economic Review Papers and Proceedings*, 68 (2), pp. 1-16.

Sims, Christopher A. (2003) "Implications of Rational Inattention," *Journal of Monetary Economics*, 50 (3), pp. 665-690.

Souleles, Nicholas S. (1999) "The Response of Household Consumption to Income Tax Refunds," *American Economic Review,* 89 (4), pp. 947-958.

Souleles, Nicholas S. (2000) "College Tuition and Household Savings and Consumption," *Journal of Public Economics*, 77 (2), pp. 185-207.

Stock, James H. and Mark Watson (2002) "Has the Business Cycle Changed and Why?" *NBER*

*Macroeconomics Annual*, forthcoming.

Stokey, Nancy L., Robert E. Lucas Jr., with Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*, Cambridge, Harvard University Press.

Tobin, James (1982) "Money and Finance in the Macroeconomic Process," *Journal of Money, Credit and Banking*, May, 14 (2), pp. 171-204.

Wilson, Andrea (2003) "Bounded Memory and Biases in Information Processing," unpublished, Princeton University.

Wolak, Frank A. (1989) "Testing Inequality Constraints in Linear Econometric Models," *Journal of Econometrics*, 41 pp. 205-235.

**Table 1: Optimal Hybrid Consumption-Savings Plans**

*Panel A: Inattentiveness(d)*

| | K = \$30 | K = \$100 | K = \$250 | K = \$500 | K = \$1000 |
|---|---|---|---|---|---|
| r = 0.5% | 13 | 24 | 36 | 50 | 67 |
| r = 1.5% | 10 | 16 | 23 | 31 | 41 |
| r = 4% | 6 | 9 | 12 | 15 | 20 |

*Panel B: Optimal share of income shocks consumed (λ)*

| | K = \$30 | K = \$100 | K = \$250 | K = \$500 | K = \$1000 |
|---|---|---|---|---|---|
| r = 0.5% | 0.02 | 0.03 | 0.03 | 0.04 | 0.04 |
| r = 1.5% | 0.06 | 0.06 | 0.07 | 0.08 | 0.09 |
| r = 4% | 0.13 | 0.14 | 0.16 | 0.17 | 0.19 |

Notes: The remaining parameters were set at the benchmark values: $\phi$=0.487, α=2/6926, $\sigma_P$=45, $\sigma_T$=1962.

**Table 2: Extraordinary Events and the Length of Inattentiveness**

*Panel A: Inattentiveness*

| | u = \$500 | u = \$2,500 | u = \$5,000 |
|---|---|---|---|
| δ = 1/8 | 10 | 10 | 11 |
| δ = 1/20 | 9 | 9 | 9 |
| δ = 1/40 | 8 | 9 | 9 |

*Panel B: Probability of planning in response to an extraordinary event*

| | u = \$500 | u = \$2,500 | u = \$5,000 |
|---|---|---|---|
| δ = 1/8 | 71% | 71% | 75% |
| δ = 1/20 | 36% | 36% | 36% |
| δ = 1/40 | 18% | 20% | 20% |

Notes: The remaining parameters were set to match the benchmark values: r=1.5%, $\phi$=0.487, α=2/6926, $\sigma^2=(45/r)^2+[1962/(r+\phi)]^2-\delta u^2$. The costs of planning K were set at \$30 so that without extraordinary events, the agent plans every 8 quarters.

**Table 3: Regressing Consumption Growth on News on Income Growth**

| | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ | $\beta_6$ | $\beta_7$ | $\beta_8$ | $\beta_9$ |
|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A*. Predictors of $\Delta\ln(Y_t)$ | | | | | | | | | | |
| $\Delta\ln(Y_{t-1}),\dots, \Delta\ln(Y_{t-5})$ | .288*** | .077*** | .072*** | .104*** | .029 | -.034 | .032 | .006 | -.035 | -.035 |
| | (.042) | (.032) | (.027) | (.029) | (.035) | (.038) | (.032) | (.022) | (.032) | (.028) |
| | Restricted Least Squares estimates | | | | | | | | | |
| | .287 | .084 | .084 | .084 | .023 | .001 | .001 | .001 | 0 | 0 |
| | F-test: 7.20*** (.000) | | Adj. $R^2$: .334 | | F-test 1st stage 3.55*** (.004) | | Adj. $R^2$: 1st stage .062 | | $W_{IN}$: 4.71 (.701) | |
| *Panel B*. Predictors of $\Delta\ln(Y_t)$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ | $\beta_6$ | $\beta_7$ | $\beta_8$ | $\beta_9$ |
| $\Delta\ln(Y_{t-1}),\dots, \Delta\ln(Y_{t-5})$, $\ln(C_{t-1}/Y_{t-1}),\dots, \ln(C_{t-5}/Y_{t-5})$ | .279*** | .082*** | .050* | .102*** | .059 | -.019 | .054 | .059*** | -.003 | -.003 |
| | (.049) | (.034) | (.026) | (.030) | (.037) | (.044) | (.033) | (.033) | (.039) | (.037) |
| | Restricted Least Squares estimates | | | | | | | | | |
| | .278 | .080 | .079 | .079 | .055 | .032 | .032 | .032 | 0 | 0 |
| | F-test: 5.35*** (.000) | | Adj. $R^2$: .262 | | F-test 1st stage 5.69*** (.000) | | Adj. $R^2$: 1st stage .196 | | $W_{IN}$: 7.58 (.428) | |
| *Panel C*. Predictors of $\Delta\ln(Y_t)$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ | $\beta_6$ | $\beta_7$ | $\beta_8$ | $\beta_9$ |
| $\Delta\ln(Y_{t-1}),\dots, \Delta\ln(Y_{t-5})$, $\ln(C_{t-1}/Y_{t-1}),\dots, \ln(C_{t-5}/Y_{t-5})$, $r_{t-1},\dots, r_{t-5}$ | .265*** | .075** | .044 | .089*** | .059 | -.019 | .058* | .066** | -.001 | -.005 |
| | (.050) | (.037) | (.027) | (.033) | (.033) | (.046) | (.033) | (.033) | (.042) | (.038) |
| | Restricted Least Squares estimates | | | | | | | | | |
| | .263 | .073 | .070 | .070 | .053 | .035 | .035 | .035 | .002 | 0 |
| | F-test: 4.66*** (.000) | | Adj. $R^2$: .229 | | F-test 1st stage 3.96*** (.000) | | Adj. $R^2$: 1st stage .187 | | $W_{IN}$: 8.11 (.374) | |

Notes: These are the estimates of the system of two equations: (first stage) $y_t = \Delta\ln(Y_t) - E_{t-1}[\Delta\ln(Y_t)]$, and (second stage) $\Delta\ln(C_{t+1}) = \text{const.} + \beta_0 y_{t+1} + \beta_1 y_t + \dots + \beta_9 y_{t-8} + \tilde{u}_t$. ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively. In brackets below the estimates are Newey-West standard errors corrected for heteroskedasticity and autocorrelation up to 8 lags. The F-test is on the significance of the regression, and $W_{IN}$ tests the inattentive consumers model. In brackets below the test statistics are the p-values.

**Table 4: The Excess Smoothness Ratio**

*Panel A: Estimates*

| Method | Lags | $\psi$ | Standard Errors |
|---|---|---|---|
| Bartlett window | 5 | .704 | .065 |
| | 10 | .662 | .088 |
| | 20 | .671 | .129 |
| AR-HAC | 2 | .679 | .088 |
| | 5 | .651 | .115 |
| | 10 | .643 | .159 |
| Andrews-Monahan | 5 | .515 | .047 |
| | 10 | .559 | .073 |
| | 20 | .584 | .107 |

*Panel B: Predictions of the inattentiveness model*

| Estimates of the weights $\Phi(i)$: | $\psi$ |
|---|---|
| From news regressions in Table 6, with predictors: | |
| - lagged income | .660 |
| (restricted coefficients) | .570 |
| - lagged income and savings | .498 |
| (restricted coefficients) | .480 |
| - lagged income, savings and interest rates | .494 |
| (restricted coefficients) | .473 |

Notes: The estimates of the excess smoothness ratio ($\psi$) use data on the change of log aggregate consumption from 1954 to 2002. The different methods used to obtain estimates of the spectrum at frequency zero were: a Bartlett kernel estimator with window length 5, 10 and 20; a parametric AR-HAC estimate using an AR with lags 2, 5 and 10; a Andrews-Monahan (1992) estimator which pre-whitens the data using an AR(1) and then uses a Bartlett kernel with window lengths 5, 10 and 20. Standard errors are obtained by the delta method, and using the result that asymptotically $Var(h_{\Delta C}(\omega))=(4/3)*(M/N)*h_{\Delta C}(\omega)$ for the Bartlett kernel, where M is the window length, and N is the number of observations (see Priestley, 1981, pages 457-461).

**Table 5: Excess Sensitivity and Hand-to-Mouth Behavior in the Inattentiveness Model**

| *Panel A. IV regressions* Instruments for $\Delta\ln(Y_{t+1})$: | Estimates (standard errors) | Adj. $R^2$ | F-stat. 1st stage | J-stat. (p-value) |
|---|---|---|---|---|
| $\Delta\ln(Y_{t-9}),\ldots,\Delta\ln(Y_{t-12})$ | .157 | .165 | .80 | .81 |
| | (.229) | | (.525) | (.848) |
| $\Delta\ln(Y_{t-9}),\ldots,\Delta\ln(Y_{t-12})$, | .166 | .167 | .64 | 2.33 |
| $\text{Ln}(C_{t-9}/Y_{t-9}),\ldots,\ln(C_{t-12}/Y_{t-12})$ | (.180) | | (.743) | (.940) |
| $\Delta\ln(Y_{t-9}),\ldots,\Delta\ln(Y_{t-12})$, | .049 | .073 | .80 | 4.53 |
| $\ln(C_{t-9}/Y_{t-9}),\ldots,\ln(C_{t-12}/Y_{t-12})$ | (.139) | | (.650) | (.952) |
| $r_{t-9},\ldots,r_{t-12}$ | | | | |

| *Panel B. Weak Instruments* Instruments for $\Delta\ln(Y_{t+1})$: | Estimates | | Test statistics (p-values) | | |
|---|---|---|---|---|---|
| | OLS | LIML | A-R | Moreira | LM |
| $\Delta\ln(Y_{t-9}),\ldots,\Delta\ln(Y_{t-12})$ | .226 | .147 | 1.040 | .252 | .186 |
| | | | (.904) | (.887) | (.667) |
| $\Delta\ln(Y_{t-9}),\ldots,\Delta\ln(Y_{t-12})$, | .226 | .148 | 2.542 | .305 | .164 |
| $\text{Ln}(C_{t-9}/Y_{t-9}),\ldots,\ln(C_{t-12}/Y_{t-12})$ | | | (.960) | (.950) | (.685) |
| $\Delta\ln(Y_{t-9}),\ldots,\Delta\ln(Y_{t-12})$, | .226 | -.057 | 4.085 | .086 | .057 |
| $\ln(C_{t-9}/Y_{t-9}),\ldots,\ln(C_{t-12}/Y_{t-12})$ | | | (.982) | (.930) | (.812) |
| $r_{t-9},\ldots,r_{t-12}$ | | | | | |

Notes: The dependent variable in all regressions is $\Delta\ln(C_{t+1})$. $^{***}$, $^{**}$ and $^{*}$ denote statistical significance at the 1%, 5% and 10% levels respectively. The estimates use the Hayashi and Sims (1983) procedure with an estimated MA(9) to forward-filter the data. In Panel A, the J-stat. refers to the Hansen-Sargan statistic for testing the over-identifying restrictions associated with the validity of the instruments.

**Table 6: Rational Expectations vs. Hand-to-mouth vs. Inattentiveness**

*Panel A. Regression Estimates*

| Const. | $E_t$-$E_{t-1}$ | $E_{t-1}$-$E_{t-2}$ | $E_{t-2}$-$E_{t-3}$ | $E_{t-3}$-$E_{t-4}$ | $E_{t-4}$-$E_{t-5}$ | $E_{t-5}$-$E_{t-6}$ | $E_{t-6}$-$E_{t-7}$ | $E_{t-7}$-$E_{t-8}$ | $E_{t-8}$-$E_{t-9}$ | $E_{t-9}$-$E_{t-10}$ | $E_{t-10}$-$E_{t-11}$ | $E_{t-11}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $\Delta \ln(Y_{t-1})$ | | | | | | |
| .005$^{**}$ | .320$^{*}$ | .620$^{**}$ | .521$^{***}$ | .289$^{*}$ | .104 | .536 | .790 | .816$^{***}$ | .680 | 1.010 | -.498 | .034 |
| (.002) | (.163) | (.255) | (.179) | (.166) | (.161) | (.367) | (1.023) | (.315) | (.688) | (.543) | (1.063) | (.453) |
| Restricted Estimates | | | | | | | | | | | | |
| .005 | .394 | .394 | .394 | .314 | .314 | .314 | .314 | .314 | .314 | .314 | 0 | 0 |

Unrestricted Adjusted $R^2$: .090 Restricted Adjusted $R^2$: .055

*Panel B. Tests of the alternative models*

| Model | Test statistics (p-values) | Accept/Reject (5% significance level) |
|---|---|---|
| Rational Expectations (Hall): | 72.60 (.000) | Reject |
| Hand-to-mouth (Campbell-Mankiw): | 18.80 (.043) | Reject |
| Inattentive consumers: | 18.10 (.080) | Accept |
| Inattentive consumers and savers: | 15.09 (.128) | Accept |

Notes: $^{***}$, $^{**}$ and $^{*}$ denote statistical significance at the 1%, 5% and 10% levels respectively. All standard errors are corrected for heteroskedasticity and autocorrelation using a Newey-West procedure. Panel B displays Wald test statistics and asymptotic p-values.

Figure 1. Impulse Response of Aggregate Consumption to a Permanent Shock

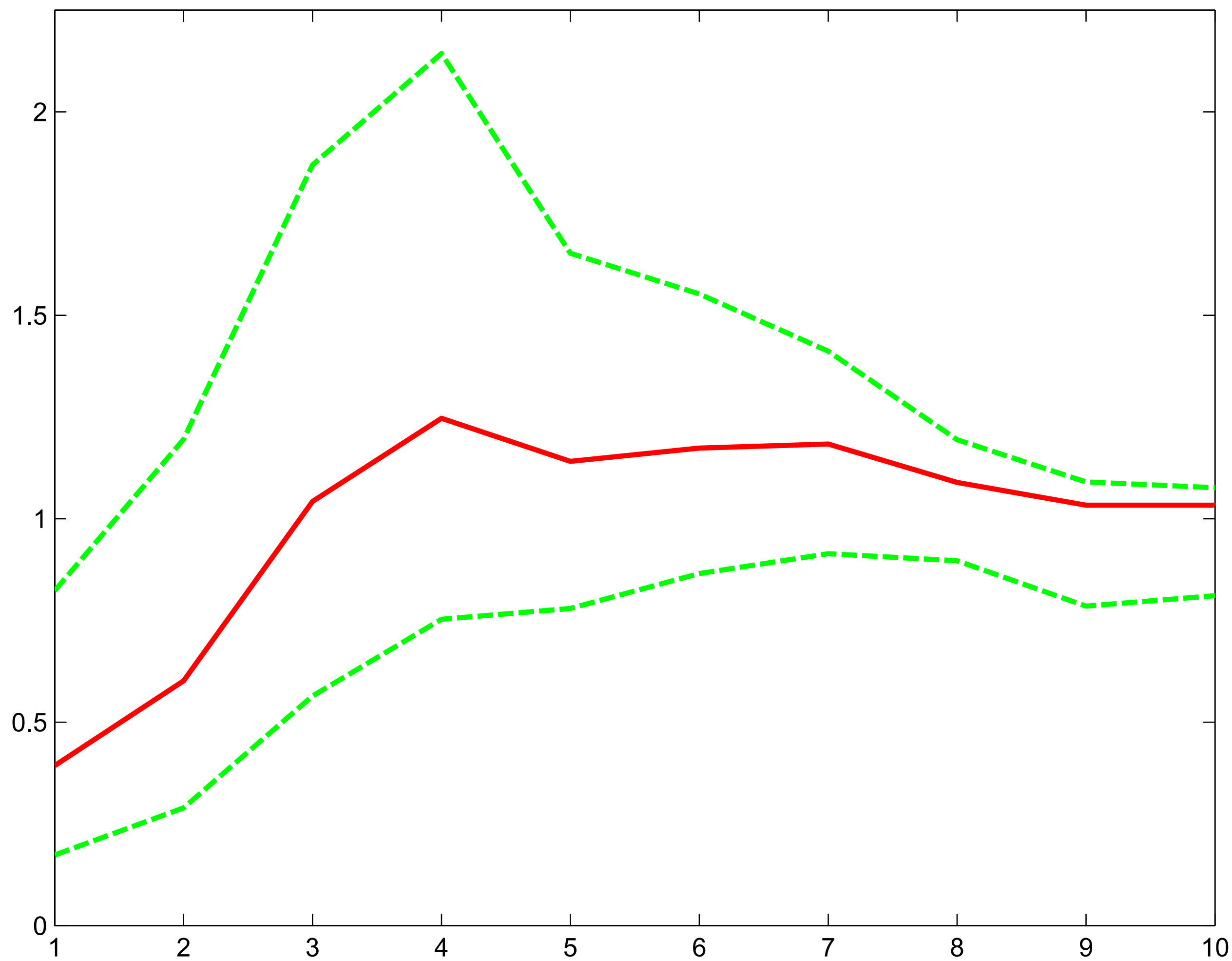
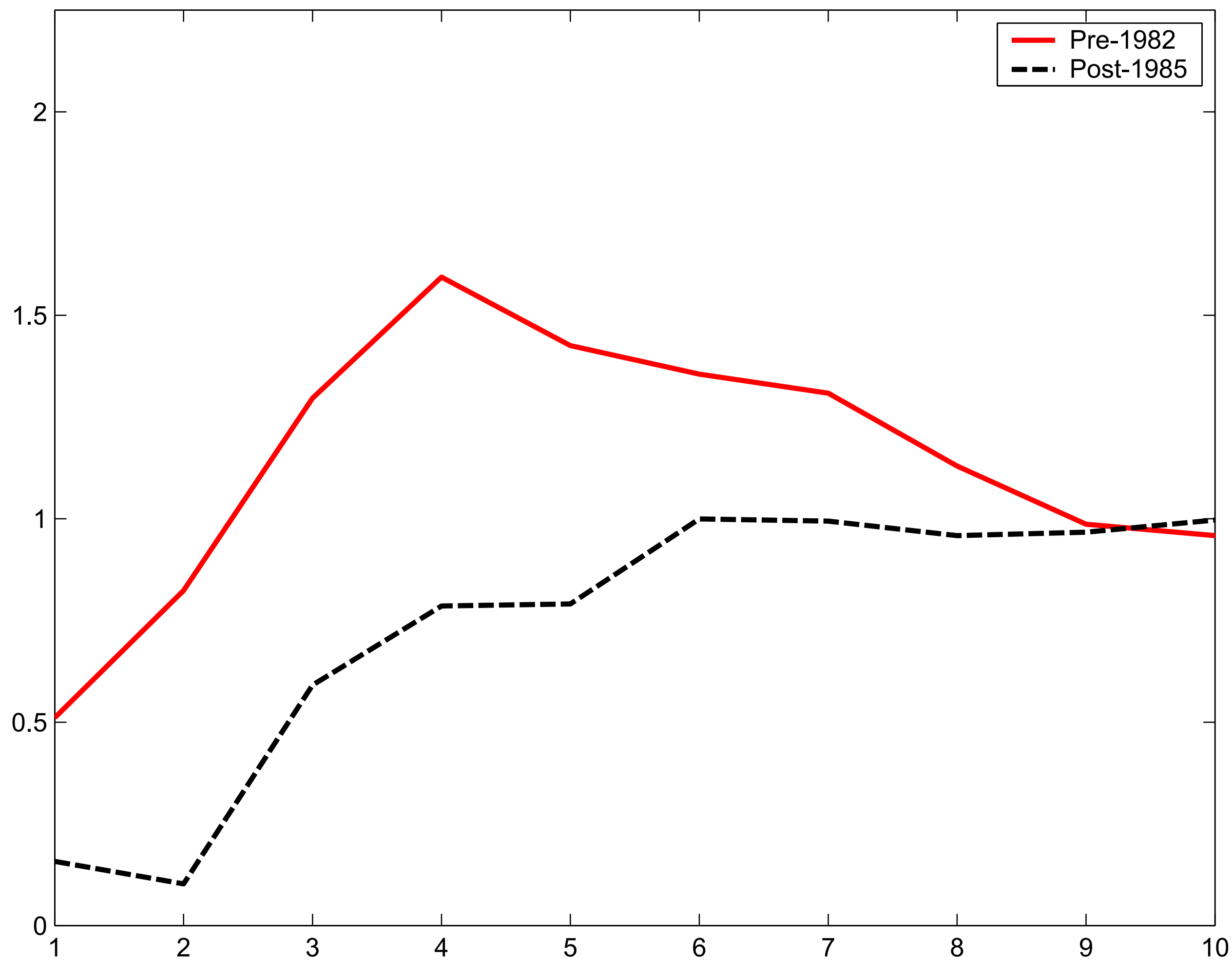
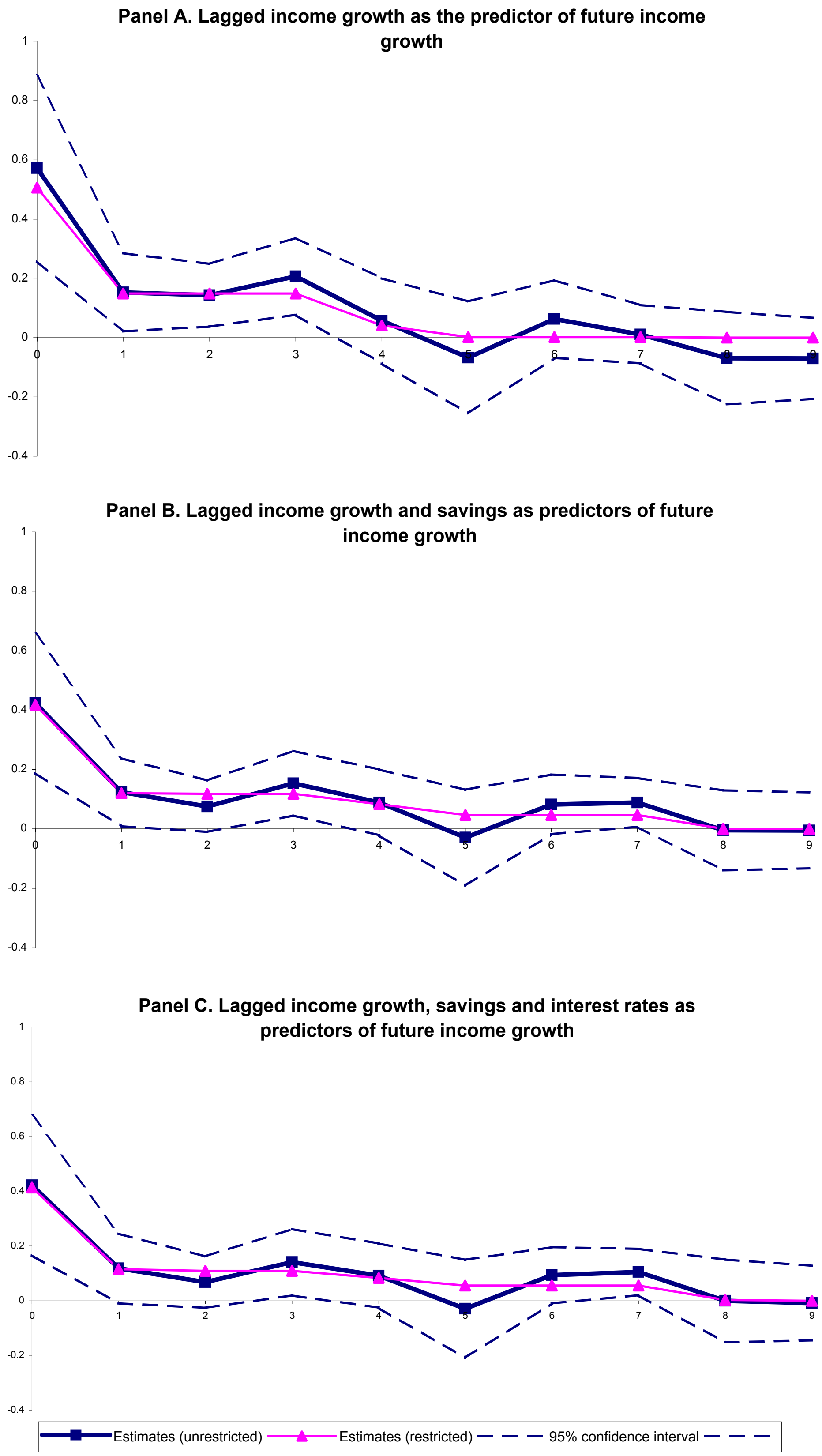


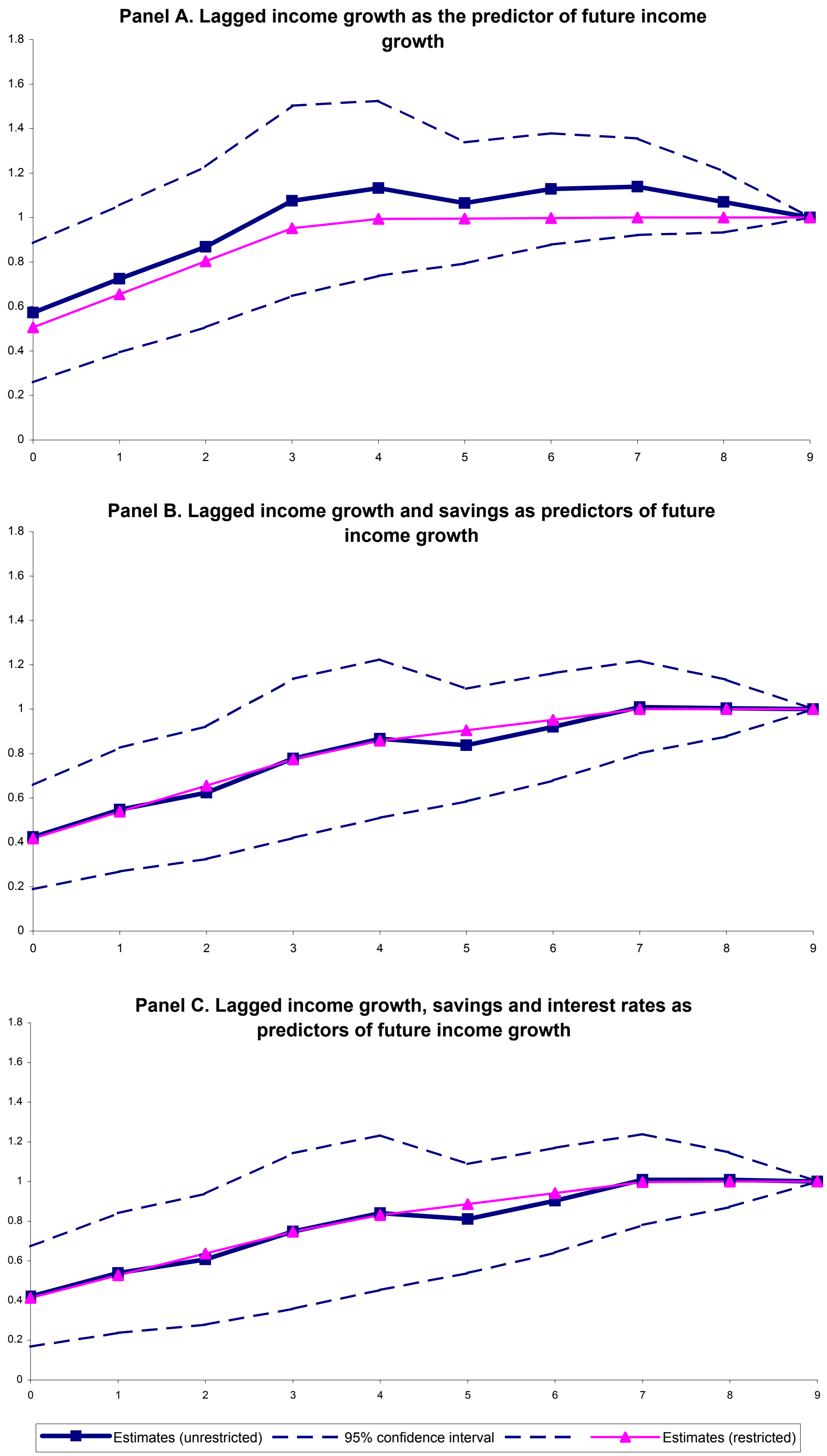
Figure 2. Impulse Response of Aggregate Consumption to a Permanent Shock, pre-1982 and post-1985



# Figure 3. Estimates of the (ratio) Inattentiveness Weights

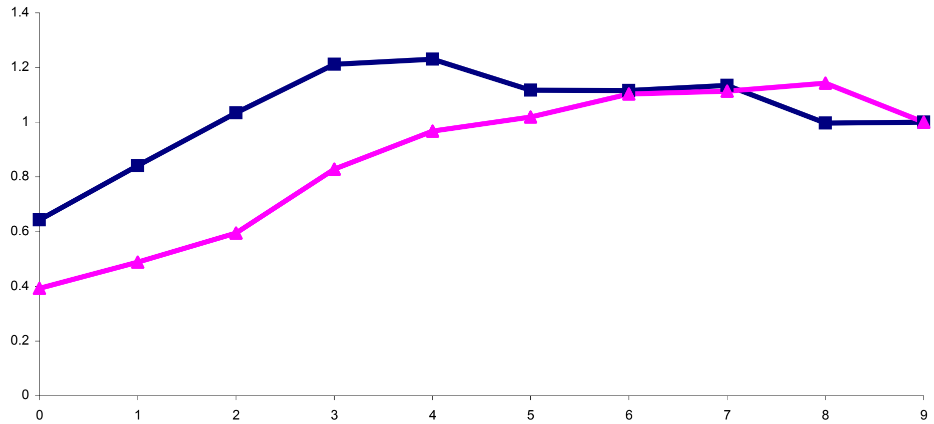


## Figure 4. Estimates of the Cumulative Inattentiveness Weights

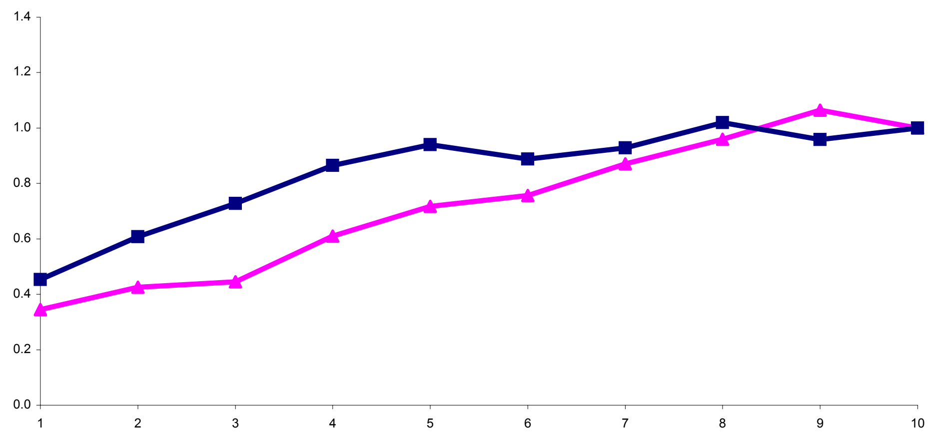


## Figure 5. The Cumulative Inattentiveness Weights, pre-1982 and post-1985

**Panel A. Lagged income growth as the predictor of future income growth**




**Panel B. Lagged income growth and savings as predictors of future income growth**




**Panel C. Lagged income growth, savings and interest rates as predictors of future income growth**

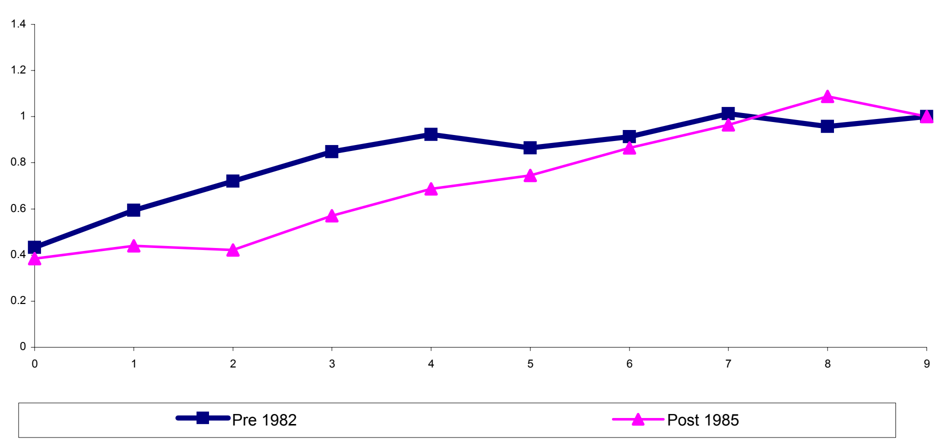


Figure 6. Model-Predicted and Actual Normalised Spectral Densities