NBER WORKING PAPER SERIES

INATTENTIVE CONSUMERS

Ricardo Reis

Working Paper 10883
http://www.nber.org/papers/w10883

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2004

I am grateful to N. Gregory Mankiw, Alberto Alesina, Robert Barro, and David Laibson for their guidance and to Andrew Abel, Susanto Basu, John Campbell, Larry Christiano, Mariana Colacelli, Benjamin Friedman, Jens Hilscher, Yves Nosbusch, David Romer, John Shea, Monica Singhal, Adam Szeidl, Bryce Ward, Justin Wolfers, and seminar participants at UC Berkeley, Board of Governors, Chicago GSB, Columbia GSB, ECB, Harvard, Johns Hopkins, Kansas City Fed, Kennedy School, New York Fed, Northwestern, Pompeu Fabra, Princeton, Stanford, and Yale for useful comments. The Fundacao Ciencia e Tecnologia, Praxis XXI and the Eliot Memorial fellowship provided financial support. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

Inattentive Consumers
Ricardo Reis
NBER Working Paper No. 10883
November 2004
JEL No. E2, D9, D1, D8

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 able to explain the excess sensitivity and excess smoothness puzzles. 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.

Ricardo Reis
Princeton University
Department of Economics
324 Bendheim Hall
Princeton, NJ 08544-1021
and NBER
rreis@princeton.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. After all, this is just the basic optimality condition from consumer choice that the marginal rate of substitution between two goods must equal their relative price. If the future is uncertain though, it is expected marginal utility that is relevant, and a crucial component of a model of consumption must specify how agents form their 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 expectations formation by consumers. With rational expectations, agents can costlessly absorb and process information on all the relevant characteristics of the economy, can costlessly think through this information, and can costlessly 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,

[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.

but are only sporadically updated. Following a new event, many agents will be unaware 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.

The model in this paper has further predictions beyond people's inattentiveness, excess sensitivity, and excess smoothness. It also predicts that while aggregate consumption moves sluggishly in response to shocks, the extent of this sluggishness is endogenously determined by the size of the information costs and income volatility, which may be different in different periods. Moreover, the model predicts that consumers only respond with a delay to a news that was not easily anticipated far in advance and that did not refer to some extraordinary event that captured everyone's attention. Finally, the model predicts that about one third of the U.S. population rationally chooses to never plan, live hand-to-mouth, and save very little.

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.[2] Moreover, I do not assume that agents infrequently adjust their plans, but rather I derive this behavior endogenously 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. 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.[3]

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

[3] A few other theoretical papers have explored consumption decisions with limited information: Goodfriend (1992) and Pischke (1995) assume that agents cannot distinguish between permanent and transitory income shocks, Ameriks et al. (2003b) model absent-minded consumers who cannot keep track of how much they have already consumed, and Mullainathan (2002), Bernheim and Thomadsen (2002), and Wilson (2003) model agents who have full information on the present but recall the past imperfectly.

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.

More generally, this paper is part of a recent wave of research rethinking how to model the process by which people form their expectations. Some have assumed that people instead use simple least squares learning algorithms to form their expectations of the future (see Evans and Honkapohja, 2001, for a survey). Others have studied models in which agents's expectations are consistent with the data while not using all of the available information (Kurz, 1997), and still others model agents as choosing between different simple mechanisms to form expectations according to their past performance (Brock and Hommes, 1997). Which is the right approach to model expectations is at this point still unclear (and many of these approaches are not mutually exclusive). One virtue of the inattentiveness model is that it remains firmly rooted in classical economics, in that agents are modelled as maximizing utility subject to constraints, the novelty being that the constraints also include costly information. One can therefore use the powerful tools of constrained optimization and rational expectations with limited information that economists have for long developed, to quickly get very far in terms of predictions.

This paper is organized as follows. Section 2 informally describes the model of inattentiveness and intuitively describes its predictions. 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.

Section 5 tests the implications of the model with aggregate and individual data. The model is also contrasted with models of rule-of-thumb behavior, habit formation, and state-contingent adjustment. Section 6 focusses on the informational assumptions of the model and discusses some extensions of the basic model. Section 7 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 directions for future research.

2 An informal description of inattentiveness and its predictions

Consider the problem facing a person who lives forever, earns a stochastic income and consumes every period, and maximizes utility subject to a standard budget constraint. The new assumption in this paper is that despite being fully rational and making optimal choices, this person must incur a cost whenever she acquires information and makes optimal decisions. This is the cost in money and time of obtaining information, processing and interpreting it, and deciding how to optimally act. It 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 the time to plan. While I model these costs as a monetary expense, they can be thought of as the wages foregone at times of planning, if planning takes time away from supplying labor at a market wage and leisure enters utility separately from consumption. Likewise, 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.

Facing these costs, a person setting a plan of action for consumption must choose not only what to consume, but also when to plan again. With regards to her consumption plan, between two periods which are in between planning dates, the person is not obtaining any new information. Therefore, the dynamics of consumption are as if the consumer was living under perfect certainty, with consumption following a pre-determined plan, irrespective of the news in the economy.[4] On the other hand, optimality with respect to consumption at two successive planning dates is determined by a stochastic Euler equation, just as in the Hall model. At planning dates, the consumer obtains new information and takes the random arrival of news into account in trading off current for future consumption.

Costly planning and inattentiveness affect not just the dynamics of consumption, but especially its level. The longer a person stays inattentive for, 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 in order to safeguard against a sequence of bad income shocks. Therefore, if a person faces higher planning costs, she plans less often, and saves more.

The optimal length of inattentiveness weights the costs of reacting with a delay to news against the costs incurred by planning. There are several interesting properties of optimal inattentiveness. First, a person who faces very small costs of planning can be inattentive for a long time. The reason is that being inattentive and reacting only with a delay to news is close to being optimal in the sense of implying only a small loss in welfare. The second property of optimal inattentiveness is that the lower is the risk faced by the person and the lower her aversion to this risk, the longer she will be inattentive for. The lower these are, the smaller is the effective cost of being inattentive in terms of exposure to risk, and thus the less frequently the desire to adjust plans. A third less intuitive property is that a larger interest rate lowers optimal inattentiveness. Inattentiveness leads

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

to sub-optimal savings and the larger is the interest rate, the larger is the impact of these inefficient savings on her future assets.

So far, I have been describing the problem of a person who chooses plans for consumption. 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. More concretely, an inattentive consumer is someone whose paycheck is deposited in her bank account, spends a planned amount, and leaves whatever remains in the bank. An inattentive saver is someone who receives her paycheck in her pocket, puts aside a planned amount in savings, and spends the rest until her pocket is empty.

One immediate implication of inattentive saver behavior is that consumption absorbs all of the income shocks. Therefore, the marginal propensity to consume out of current income is one, so the inattentive savers live hand-to-mouth. Another important feature of the behavior of an inattentive saver is that, as long as the costs of planning are not too small, it is optimal to never plan at all. To understand the intuition behind this result, consider the special case in which income shocks are serially uncorrelated. If the person does not update her information this period, next period's assets equal this period's assets plus savings and capital returns, all of which are not random. Therefore, the agent is facing exactly the same problem as in the previous period, and thus she must again choose not to update her plans. Iterating on this logic shows that the saver will either always be attentive, or never update her plans. If the cost of planning is not too small, she will choose to never plan. The inattentive saver is a rational non-planner.

Having characterized the behavior of inattentive consumers and inattentive savers, the next question is which do people choose to be. In the case where the inattentive savers rationally choose to never plan, they are unaffected by the costs of planning. The inattentive consumers on the other hand are worse off the larger are these costs. It then follows that if the costs of planning are above a certain threshold, people only choose savings plans, whereas if they are below, they will choose a consumption plan. 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 and never make plans. The other agents, who have lower planning costs, opt instead for following infrequently-updated plans on consumption.

In this inattentive economy, aggregate consumption responds gradually to a shock, with a reaction that builds up over time. The reason is that people only gradually update their plans and become aware of the news, which only slowly disseminates throughout the entire economy. If the shock affects income then in the inattentive economy one will find that aggregate consumption is excessively sensitive, since past income shocks affect current consumption growth. Moreover, since only a fraction of the agents react contemporaneously to changes in permanent income, consumption will be smoother than income. Slow dissemination of information can therefore solve both the excess sensitivity and excess smoothness puzzles.

Finally, note that the inattentiveness model refines the meaning of tests for excess sensitivity. In an inattentive economy, 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, it is reasonable to extend the baseline model to allow people to observe some extraordinary events when they take place. The defining features of these events is that they refer to changes in variables that only move infrequently, so the cost of monitoring them is very small, and which lead to large changes in the agent's income. For instance, if the agent suddenly becomes unemployed or wins the lottery, it is reasonable to suppose that the agent becomes immediately aware of these rare significant events, and responds to them. Past extraordinary events do not affect present individual consumption changes.

Summarized and simplified, these are the main features of the theory of inattentive consumption. The next two sections formalize this description, before turning to the evidence in Section 5.

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.[5] 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, who each instant consume an amount of goods c_t, which yields an amount of utility given by the function $u(c_t)$. This function is continuous, everywhere twice differentiable, increasing and concave, and future utility is discounted at the positive rate ρ.

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, $s_t = y(x_t) - c_t$. 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 σ-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

[5] An earlier version of this paper solved the model also in discrete time. Details are available from the author.

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 first makes a decision on whether to write a plan for consumption c_t or a plan for savings s_t. Each period, the agent visits a goods market and an asset market; she can choose to either follow a plan of conduct in one or in the other market. Having decided on the type of plan, the agent must then choose the content of the plan which consists of a sequence of actions until the next adjustment, $z(i) = [z_{D(i-1)}, z_{D(i)})$ where z equals c or s, and when to plan again $D(i)$. The restriction embodied in the existence of a plan is that these choices must be contingent on the information available at time $D(i-1)$: if $\{z, D\} = \{z(i), D(i)\}_{i=1}^{\infty}$ these 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. If the consumer enters period $D(i)$ with assets given by $a^-_{D(i)}$, her wealth then changes discontinuously to $a^+_{D(i)} = a^-_{D(i)} - K_{D(i)}$.[6] 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.

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

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

$$\text{s.t.} \quad : z(i) = c(i) \text{ or } s(i), \tag{2}$$

$$\{z, D\} \text{ are } \Im\text{-adapted}, \tag{3}$$

$$da_t = (ra_t - c_t + y_t)dt, \tag{4}$$

$$s_t = y_t - c_t, \tag{5}$$

$$a^+_{D(i)} = a^-_{D(i)} - K_{D(i)}, \text{for all } i \in \mathbb{N}_0, \tag{6}$$

$$\lim_{t \to \infty} e^{-rt} a_t \geq 0, \tag{7}$$

with initial conditions a_0, x_0. It is difficult to solve this problem both because it is hard to impose the measurability restriction (3) and because of the discontinuity in the level of assets at the planning dates (6). To make progress, the problem must be re-stated in a more convenient form.

Start by integrating the law of motion for assets in (4) between $D(i)$ and $D(i+1)$, and replace $a^-_{D(i)}$ by $a^+_{D(i)} + K_{D(i)}$ using (6). This gives:

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

[6] 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 (1)-(7) becomes:

$$
\begin{aligned}
V(a,x) &= \max\{V^c(a,x), V^s(a,x)\}, && (8)\\
V^c(a,x) &= \max_{c,d} \int_0^d e^{-\rho t} u(c_t)dt + e^{-\rho d} E\left[V(a',x')\right] \text{ and} && (9)\\
V^s(a,x) &= \max_{s,d} \int_0^d e^{-\rho t} E\left[u(y_t - s_t)\right] dt + e^{-\rho d} E\left[V(a',x')\right], && (10)\\
\text{subject to } a' &= e^{rd}\left(a + \int_0^d e^{-rt}(\underbrace{y_t - c_t}_{=s_t})dt\right) - K'. && (11)
\end{aligned}
$$

Focussing on the consumption problem, 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,the initial post-planning asset level is $a_0 - K_0$.[7]

The solution to the problem in (9)-(11) will be a pair of functions, $c_t(a,x)$ or $s_t(a,x)$ and $d(a,x)$, determining optimal consumption or savings from time 0 to time d and when the next planning will take place. Consumption or savings 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 (9)-(11) is a familiar dynamic programming problem. 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-from-below utility function, the problem of the inattentive consumer has one additional technical difficulty

[7] Some related problems have been studied in engineering under the headings: sampled-data control systems, and digital control. The two closest to this paper are control problems in which the state is observed at exogenously given infrequent dates (Franklin et al., 1990), and optimally choosing how often to sample a continuous time stochastic process to maximize the information content of the messages (Miller and Runggaldier, 1997).

relative to the full information problem. To illustrate it, consider the case in which income follows an arithmetic Brownian motion, which has infinite local variation. If the agent is inattentive for even an instant 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 bounded stochastic process that cannot fall extremely 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 is enough to guarantee a well-defined optimization problem. This is the approach that I will follow in section 4, and it is also the approach in the engineering literature which focusses on quadratic objective functions or H_∞ control (Chen and Francis, 1995).

3.2 Characterizing the solution

Taking the derivative of (9) with respect to d and setting it equal to zero gives:

$$E\left[u(c_d)\right] = \rho E\left[V(a', x')\right] - \frac{\partial}{\partial d} E\left[V(a', x')\right]. \tag{12}$$

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 flow of value from extending the interval of non-adjustment, which is the utility the agent would get if she kept to her outdated consumption plan. On the right-hand side is the value 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.

Consider first the case of an inattentive consumer. It is easy to show that the first-order conditions with respect to c_t and the envelope theorem condition imply that if the consumer is inattentive between times t and $s > t$, consumption between these periods obeys the deterministic

Euler equation:

$$u'(c_t) = e^{(r-\rho)(s-t)}u'(c_s). \tag{13}$$

If $D(i)$ and $D(i+1)$ are two successive planning dates, consumption between these periods obeys instead the stochastic Euler equation:[8]

$$u'(c_{D(i)}) = e^{(r-\rho)(D(i+1)-D(i))}E_{D(t)}\left[u'(c_{D(i+1)})\right]. \tag{14}$$

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 optimal choices incorporate it.

If the agent instead chooses a savings plan, then the most interesting result is that consumption moves one-to-one with income since $c_t = y_t - s_t$ and s_t does not respond to income news. The optimal path for s_t is determined by Euler equations similar to the ones above.

3.3 Aggregate consumption

There are many inattentive agents in the economy, individually behaving in the way described above. They 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 whether they choose consumption or savings plans, on how much they consume or save, and on how long they stay inattentive for.

First focus on the choices of inattentive consumers; at the end of this section I will consider savers as well. Following the literature, I work with linearized versions of the optimality conditions.[9] A first-order Taylor approximation of (13) around the point where $c_t = c_s$ and $r = \rho$ gives:

$$c_s = c_t + \frac{1}{\alpha}(r-\rho)t, \tag{15}$$

where $\alpha = -u''(c_t)/u'(c_t)$ is the coefficient of absolute risk aversion. A similar approximation of (14) leads to:

$$c_{D(i+1)} = c_{D(i)} + \frac{1}{\alpha}(r-\rho)t + e_{D(i+1),D(i)}, \tag{16}$$

where $e_{D(i+1),D(i)} \equiv c_{D(i+1)} - E_{D(i)}\left[c_{D(i+1)}\right]$, the innovation to consumption between $D(i)$ and $D(i+1)$, where $D(i) \leq s \leq t \leq D(i+1)$.

Consider then the change in consumption of any inattentive consumer between t and $t+1$. If she has not adjusted her plan between t and $t+1$, then her behavior is described by equation (15)

[8] To be rigorous, here $c_{D(i)}$ is the right-side time limit, whereas $c_{D(i+1)}$ refers to the left-side time limit.

[9] 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.

with $s = t+1$. If she has adjusted, then let j denote how long ago starting in $t+1$ did the agent last adjust, and similarly, let i denote how long starting from t one must go back to the last adjustment date for that same agent. Then, combining equation (15) and (16) establishes the relation between consumption choices at t and $t+1$ by these agents:

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

Summing over all of the inattentive consumers in the economy, it is then easy to see that

$$C_{t+1} - C_t = \text{constant} + u_{t+1},$$

where C_t is aggregate consumption by inattentive consumers and u_{t-1} is a sum of the $e_{t+1-j,t-i}$ of the different people in the economy. This has the property $E_{t-I}\left[u_{t+1}\right] = 0$, where I is the largest amount of time during which consumers remain inattentive. We therefore get the result:

Proposition 1 *Aggregate consumption growth by inattentive consumers 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 consumers remain inattentive.*

With full information ($I = 0$), Hall (1978) first showed 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, update their information and plans between t and $t+1$ and only then react to past events.

Assuming that there is a finite number of people in the world, and that the $e_{t+1-j,t-i}$ can be broken into independent homoskedastic increments, Appendix A shows that:

Proposition 2 *Aggregate consumption growth by inattentive consumers can be written as:*

$$C_{t+1} - C_t = \text{constant} + \Phi(0)e_{t+1} + \Phi(1)e_t + ... + \Phi(I)e_{t-I+1}, \tag{17}$$

with $\Phi(s) \geq \Phi(s+1) \geq 0$ for $s = 1, 2, ..., I$, while $E_{t-s}\left[e_{t+1-s}\right] = 0$ 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. The size of $\Phi(s)$ depends on the length of inattentiveness chosen by consumers, so to make equation (17) empirically testable with a time series, one must add the assumption that the economy has converged to a stationary distribution of inattentiveness.[10]

[10] Reis (2004) derives an interesting result regarding this distribution: if the decisions of when to adjust are mutually independent over time and across consumers, and the costs of planning are almost surely positive and such that inattentiveness is not always a constant multiple of some integer, then the stationary distribution will be exponential.

Equation (17) reveals another implication of the model for aggregate consumption. 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. With inattentiveness though, 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. The impulse response of aggregate consumption to a shock is therefore increasing for a few periods, and concave. A related implication from equation (17) is that consumption growth depends on past news with more recent news receiving a larger weight than older news does. Combining these two results:

Proposition 3 *Aggregate consumption by inattentive consumers 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 (17) implies that the change in aggregate consumption should follow an $MA(I)$ process with positive coefficients. Turning to the frequency domain emphasizes the difference between the two: the normalized power spectrum of aggregate consumption changes ($f_{\Delta C}(\omega)$) is horizontal in the Hall model, but has a shape determined by $\Phi(s)$ in the inattentiveness model. Moreover, Gali (1991), following Deaton (1987), showed that $\psi \equiv 1/\sqrt{2\pi f_{\Delta C}(0)}$ equals the excess smoothness ratio, that is the square root of the ratio between the variance of changes in consumption and the variance of changes in permanent income.[11] In the Hall model, this ratio equals one, since consumption reacts immediately one-for-one to changes in permanent income, so findings of $\psi < 1$ have been described as revealing excess smoothness of consumption. Appendix B shows that:

Proposition 4 *In the inattentiveness model:*
a) Changes in aggregate consumption have a normalized power spectrum given by:

$$f_{\Delta C}(\omega) = \frac{1}{2\pi}\left\{1 + 2\frac{\sum_{j=1}^{I}\sum_{k=0}^{I-j}\Phi(k)\Phi(k+j)\cos(\omega j)}{\sum_{k=0}^{I}\Phi(k)^2}\right\}. \tag{18}$$

[11] Heuristically, Gali's 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 2π 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.

b) The excess smoothness ratio is:

$$\psi = \sqrt{\frac{\sum_{i=0}^{I} \Phi(i)^2}{\left[\sum_{i=0}^{I} \Phi(i)\right]^2}}. \tag{19}$$

If some agents are inattentive for at least one period, 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. Proposition 4 shows the tight relation between excess sensitivity and excess smoothness in the inattentiveness model.[12] Any particular pattern of excess sensitivity coefficients ($\Phi(i)$) implies not just excess smoothness, but also an exact value for ψ. The model requires that the same set of parameters must fit these two related but distinct features of the data.

Finally, I turn to the behavior of inattentive savers. Following very similar steps, it is easy to show that for these agents

$$c_{t+1} - c_t = \text{constant} + y_{t+1} - y_t + e_{t+1-j,t-i}, \tag{20}$$

where, as before, $e_{t+1-j,t-i}$ captures news on savings to inattentive savers whose most recent planning date since $t+1$ was at $t+1-j$ and most recent planning date since t was at $t-i$. Note especially that if $j > 1$, then $j = i+1$ and $e_{t+1-j,t-i} = 0$. If, for instance, savers never plan, then all their $e_{t+1-j,t-i}$ are zero.

Aggregating over the inattentive savers then leads to a similar expression as in (17) but now with an added term involving the change in the aggregate income of inattentive savers. Aggregating over all of the consumers in the economy then leads to:

Proposition 5 *Aggregate consumption growth over all agents can be written as:*

$$\hat{C}_{t+1} - \hat{C}_t = \text{constant} + \lambda(Y_{t+1} - Y_t) + \hat{\Phi}(0)e_{t+1} + \hat{\Phi}(1)e_t + ... + \hat{\Phi}(I)e_{t-I+1}, \tag{21}$$

where λ is the share of aggregate income going to inattentive savers and the $\hat{\Phi}(s)$ and e_{t+1-s} have the same properties as in Proposition 2.

This result shows that regressing consumption growth on income growth, instrumenting the latter with information lagged at least I periods will give an estimate of the share of inattentive savers in the economy. This refines the prediction in proposition 1 that consumption growth is unpredictable I periods ahead: with inattentive savers, consumption will respond to movements in income predictable as of I periods ahead, but only through the behavior of inattentive savers.

[12] Campbell and Deaton (1989) link excess sensitivity and excess smoothness in the rational expectations model.

Propositions 1 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 (17) insofar as it turns the $MA(I)$ process into an $MA(I+1)$ with a new set of coefficients which are still non-increasing. All the propositions are likewise affected solely by replacing I by $I+1$.

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 income process 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:

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

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.[13]

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 σ_P^2 and Wiener increments dz_t^P. This corresponds for instance to changes in employment status or to changes 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 ϕ and independent Wiener impulses $\sigma_T dz_t^T$. Shocks to transitory income affect income only temporarily, and the larger is ϕ the more short-lived their effects are. For instance, these could stand for overtime payment, illness, or winning a prize.

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.[14] If ϕ is large,

[13] As is well-known, a caveat of the CARA model is that it lacks absolute wealth effects. These would lead, given a fixed K, to richer people planning more often. If the cost of planning is interpreted as a cost of time though, it is reasonable to expect that planning involves a higher opportunity cost for the wealthy, in which case they may plan more of less, depending on the precise assumptions made about K.

[14] MaCurdy (1982) finds that an MA(2) fits the data equally well, and his findings have been confirmed by Abowd and Card (1989), Pischke (1995), and Meghir and Pistaferri (2003).

income changes will be close to the MA(1) process originally proposed by Muth (1960).

4.1 Optimal consumption and inattentiveness

First, I solve the problem of an inattentive consumer. 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:

$$dw_t = (rw_t - c_t)dt + \frac{\sigma_P}{r}dz_t^P + \frac{\sigma_T}{r+\phi}dz_t^T. \tag{22}$$

Whereas generally the agent must keep track of a_t and y_t separately in order to assess how her constraints will evolve, (22) 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:

$$V(w) = \max_{c,d} \int_0^d e^{-\rho t}\left(-\frac{e^{-\alpha c_t}}{\alpha}\right)dt + e^{-\rho d}E\left[V(w')\right], \tag{23}$$

$$\text{subject to } w' = e^{rd}\left[w - \int_0^d e^{-rt}c_t dt + \int_0^d e^{-rt}\left(\frac{\sigma_P}{r}dz_t^P + \frac{\sigma_T}{r+\phi}dz_t^T\right)\right] - K. \tag{24}$$

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 inattentiveness intervals are given by:*

$$d^* = \frac{1}{r}\ln\left(1 + \sqrt{\frac{4K}{\alpha\sigma^2}}\right). \tag{25}$$

Optimal consumption between adjustments, for $D(i) < t < D(i+1)$, is:

$$c_t^* = rw_{D(i)} + \frac{(r-\rho)(t-D(i))}{\alpha} - \frac{(r-\rho)}{\alpha r} - \frac{rK}{e^{rd^*}-1} - \frac{\alpha r\sigma^2}{4}(e^{rd^*}+1) \tag{26}$$

$$= rw_{D(i)} + \frac{(r-\rho)(t-D(i))}{\alpha} - \frac{(r-\rho)}{\alpha r} - \frac{r\alpha\sigma^2}{2} - r\sqrt{\alpha\sigma^2 K}, \tag{27}$$

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:

$$c_{D(i)}^* = c_{D(i)}^A - \frac{rK}{e^{rd^*}-1} - \frac{\alpha r\sigma^2}{4}(e^{rd^*}-1). \tag{28}$$

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 (28). 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 (28). 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. Savings after expenditure on consumption and planning is therefore higher for precautionary reasons captured in the third term in (28), which increases in the length of inattention. Larger costs of planning lead to longer periods of inattentiveness thus strengthening the precautionary motive and raising savings.[15]

Inspecting the optimal inattentiveness in (25) establishes:

Corollary 2 *In the CARA-utility, ARMA-income case, inattentiveness by a consumer* (d^*)*:*
1. Falls with the volatility of the income shocks (σ^2)*;*
2. Falls with the coefficient of absolute risk aversion (α)*;*
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.

In a world that is quickly changing in which income is volatile, it is very costly to not pay attention to news so people avoid being inattentive for long. Similarly, if people are very averse to risk, they will want to lower the risk they face 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 σ^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. If policy can stabilize the economy, it will raise welfare by allowing people to be inattentive 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 final interesting property of inattentiveness is that even very small costs of planning can lead to considerable inattentiveness.[16] The intuition for this result is similar to that in Mankiw (1985), Akerlof and Yellen (1985) and Cochrane (1989). Inattentiveness leads to consumption

[15] 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.

[16] Further deviations from rationality may magnify this inertia. For instance, if agents have hyperbolic discount functions, costly planning can lead to procrastination (Akerlof, 1991).

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.[17]

Table 1 illustrates how large d^* can be using different parameter estimates. In the first column, are the estimates by Pischke (1995), who measures y_t^P as aggregate income and y_t^T as idiosyncratic income. His estimates of aggregate income variability, and of the serial correlation and standard deviation of income changes imply that $\sigma_P = \$45$, $\phi = 0.487$, and $\sigma_T = \$1,962$. I set the quarterly interest rate at 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 (25) 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.

Column 2 repeats the calculation with $r = 0.5\%$, which may be more appropriate since this is a riskless rate, while columns 3-5 follow Bound et al. (1994) by lowering the variance of income by 1/3, while changing σ_P by factors of 0.5, 1.5, and 2. Across these different parameter specifications, costs of planing between \$10 and \$50 still lead to 2 years of inattentiveness. Column 6 uses instead the estimates in Gourinchas and Parker (2002). In this case, a \$30 cost of planning leads to slightly lower inattentiveness at 2.5 quarters, and it takes now a cost of about \$80 to induce one year of inattentiveness. These calculations are solely meant to illustrate how large inattentiveness can be. They suggest that small costs can generate substantial inattentiveness.[18]

4.2 Optimal savings and inattentiveness

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 lives hand-to-mouth following a plan for savings. Her choice of inattentiveness $\hat{d} = +\infty$ if:*

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

[17] Note that the formula in equation (25) is scale-invariant, since K is in income units, σ^2 is in squared units of income, and α equals scale-free relative risk aversion divided by consumption.

[18] As discussed at the end of Section 3.1, tractability required allowing consumption to be negative. How often does this happen? Using the Pischke (1995) parameters and assuming that $\rho = r$, $K = \$30$, and 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, the probability that c_d is negative is essentially zero: it would take 8 successive quarters of negative wealth shocks equal to more than 10 times their standard deviation for c_d to be negative.

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

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

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. A person that chooses $\hat{d}=+\infty$ is a rational non-planer in the sense that she writes a plan once at time 0 and follows it forever. For the parameter estimates in Pischke (1995), she chooses to do so once the costs of planning exceed \$543.

Rational non-planners not only live hand-to-mouth, but also, 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.*

4.3 The choice between consumption and savings plans

Appendix E solves for the inattentive agent's optimal choice of which type of plan:

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 a savings plan otherwise. When the agent shifts from a consumption to savings plans, her inattentiveness rises discontinuously, and possibly to infinity.*

The condition in Proposition 8 likely holds for plausible parameter values. Most studies of individual income find that transitory shocks are the dominant source of income variation, so the condition is close 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%.

Proposition 8 shows 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 may be composed only of people who rationally choose to never plan:

Corollary 4 *As long as:*

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

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

For the parameter estimates of σ_P^2/σ_T^2 and ϕ 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:*

$$\frac{\sigma_P^2}{\sigma_T^2}(e^{rd^*}-1)+\left(\frac{r}{r+\phi}\right)^2 e^{rd^*}-\frac{r}{r+2\phi}<0. \tag{29}$$

While this condition involves an endogenous variable (d^*), it only requires knowledge of σ_P^2/σ_T^2 and ϕ from the earnings data, and no information on the degree of risk aversion. Using the benchmark estimates in Pischke (1995) for σ_P^2/σ_T^2 and ϕ, 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.

5 Evidence of inattentiveness

5.1 Slow adjustment to shocks, excess sensitivity, and excess smoothness in the aggregate data

The previous two sections stated a series of predictions of the inattentiveness model for the behavior of aggregate consumption. I test these using U.S. quarterly time series from 1953:1 to 2002:4 for aggregate consumption (C_t) measured as real consumption of non-durables and services per capita, and aggregate income (Y_t) measured as real disposable personal income per capita. I will also use data on real asset returns (r_t) using the value-weighted S&P500. All series are deflated using the price deflator for consumption of non-durables and services. I measure consumption in logs, since the series is closer to log-linear than linear. (The predictions in Section 3 could be re-stated in terms of log consumption by log-linearizing rather than linearizing the Euler equations.)[19]

[19] I tried using different consumption series, which exclude services and some components of non-durables that are arguably durable, and using other measures of returns, on different assets and using alternative adjustments for taxes. The results were robust.

Proposition 3 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 the impulse response to permanent shocks to consumption.

Figure 1 displays the impulse response of log consumption to a permanent shock together with a 90% point-wise confidence interval generated by a bootstrap. 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 concave shape predicted by the model is also visible in Figure 1.

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 2. Regressing consumption growth on several lags of y_t leads to a test of the model's predictions in equation (17).

The first possible model for income growth I consider is an AR(5). The results from regressing consumption growth on income news are in Panel A of Table 2. Panel B adds 5 lags of the log consumption income ratio to predict income. 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. Panel C further adds 5 lags of the real interest rate as predictors of future income growth, though these variables help very little in forecasting income growth.

The estimates are very similar across the three panels. As predicted by the inattentiveness model, lagged income surprises affect future consumption growth, with coefficients that are approximately unchanged across the different panels. 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 always strongly rejected at significance levels above 0.1%. Moreover, income surprises explain much of the variability of consumption growth; the adjusted R^2 of the regression is between 0.23 and 0.33. 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 produces the restricted estimates presented in the Table. 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 2, 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 statistical significance level below 37%.[20]

Figure 2 displays graphically these results. Since the estimates are so similar across panels, I display only the results in Panel B. The top panel of the Figure plots $\hat{\Phi}(j)/\sum_{i=0}^{I+1}\hat{\Phi}(i)$ for j from 0 to $I+1$, together with 95% confidence intervals. The inattentiveness model predicts a declining sequence of non-negative points, and this is consistent with the plot. In the bottom panel, I plot instead the cumulative dissemination of the news. It shows the increasing and concave shape that the model predicts, similar to that estimated earlier in Figure 1.

Proposition 4 makes sharp predictions on the shape of the power spectrum of aggregate consumption changes. Figure 3 plots estimates of the spectrum, constructed using a sample spectral density weighted over a 5-lag Bartlett window. Figure 3 also displays the spectrum for aggregate consumption growth predicted by the inattentiveness model using the weights $\hat{\Phi}(i)$ estimated in panel B of Table 2. The predicted spectrum matches the empirical spectrum well, despite somewhat more pronounced swings, and the fit is especially good using the theory-restricted estimates.

Table 3 displays different estimates of the excess smoothness ratio ψ. 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 estimated in Table 2 predicts an excess smoothness ratio between 0.47 and 0.66, well within what we observe in the data (or slightly below). The inattentiveness model is therefore able to simultaneously generate the extent of excess sensitivity and excess smoothness that we observe in the data.

5.2 The share of inattentive savers and inattentiveness versus Campbell-Mankiw

Proposition 5 stated that regressing consumption growth on income growth, instrumenting with variables lagged $I+1$ periods will give an estimate of the share of income attributed to inattentive savers. Table 4 presents these estimates, computed 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. 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 λ are quite low, between 0.05 and 0.15, and the null hypothesis that $\lambda = 0$ can never be rejected at conventional significance levels. This confirms the prediction in Section 4.3 that the share of aggregate consumption attributable to hand-to-mouth behavior should be small, with the bulk of aggregate consumption dynamics accounted for by inattentive consumers.

The instruments used in these regressions are weak though, as reflected by the low F-statistics: income growth is difficult to forecast 9 quarters in advance. With weak instruments, the IV esti-

[20] 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 2, so the inferences are robust to this econometric issue.

mates are biased towards the OLS estimates, so I report these in Panel B. Since the OLS estimates are higher than the IV estimates, the estimates of λ in Panel A are, if anything, too large. An alternative estimator is the limited information maximum likelihood (LIML) estimator, 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 different tests proposed in the literature on weak instruments to powerfully test the hypothesis that $\lambda = 0$: none of them rejects this hypothesis. While these tests likely suffer from lack of power, note that 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 λ to be small.

Equation (21) also 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 λ fraction is accounted for by irrational, myopic, hand-to-mouth people. The difference is that in their model, $\Phi(i) = 0$ for $i \geq 1$. 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 4 shows that doing so produces estimates of λ that are insignificant and much lower, between 5% and 15%, supporting instead the inattentiveness model.

It would be desirable to test the Campbell-Mankiw model against the inattentiveness model, having both stated as null hypotheses in order to ensure that lack of power does not bias the results in favour of the model that is stated as a null hypothesis. Note that $Y_{t+1} - Y_t$ can be written instead as $(E_t - E_{t-1})(Y_{t+1} - Y_t) + (E_{t-1} - E_{t-2})(Y_{t+1} - Y_t) + ... + E_{t-T}(Y_{t+1} - Y_t) + (Y_{t+1} - E_{tY t+1})$. Equation (21) can therefore instead be written in the form:

$$C_{t+1} - C_t = \beta_0 + \sum_{s=1}^{T} \beta_s (E_{t-s+1} - E_{t-s})(Y_{t+1} - Y_t) + \lambda E_{t-T}(Y_{t+1} - Y_t) + u_{t+1}, \tag{30}$$

In terms of this regression equation, the null hypothesis describing the Campbell-Mankiw model is[21]

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

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 is that

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

as long as $T \geq I + 1$, and the estimate of λ gives the share of inattentive savers. If there are only

[21] There is no restriction on β_1 because of time aggregation, and I do not impose the restriction $\lambda \geq 0$, which may bias the results in favor of the Campbell-Mankiw model.

inattentive consumers this leads to the stronger null hypothesis:

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

Finally, the Hall (1978) model predicts that

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

so that no lagged variables predict future consumption growth. 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.

I generated the regressors in (30) using the forecasts of income growth from a VAR with 5 lags on the change in log income, the log consumption-income ratio, and the real interest rate. Table 5 presents the point estimates of equation (30). They are somewhat discouraging for all four models since they do not seem to have the pattern described in either H_0^{CM} or H_0^{IN}, and several of them are individually large and statistically significant contrary to H_0^{RE}. Panel B of Table 5 formally tests the models using Wald tests for H_0^{CM}, H_0^{RE}, H_0^{IN}, and H_0^{ING}. 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 λ. 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 rather than to the model proposed by Campbell and Mankiw (1990). Moreover, as predicted by the theory, rational non-planning seems to have a small impact on aggregate consumption.

5.3 Inattentiveness versus habits

A popular alternative theory of consumption has stressed that consumers may develop habits over consumption, in which case they have a preference for sluggish consumption adjustments (e.g., Fuhrer, 2000). In the simplest case (see Deaton, 1992, pp. 31-33) this model implies that consumption growth will be an AR(1) process, with the AR coefficient equal to the preference parameter determining the habit. Since an $AR(1)$ is also an $MA(\infty)$ with declining coefficients a model with a representative consumer with a habit generates aggregate consumption observations close to the $MA(I+1)$ with declining coefficients predicted by the inattentiveness model. One can see the inattentiveness model as providing a "micro-foundation" for a representative consumer with a habit.[22]

Using other information aside from the stochastic process describing consumption, the models can be distinguished in several ways. For instance, since in the habit model sluggishness of consumption is a result of preferences, then this sluggishness should be constant across different periods in time. In the inattentiveness model on the other hand, slow adjustment is a result of inattentiveness, and this is optimally chosen by agents in response to, among other things, the volatility of income. 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 U.S. GDP after around 1984. Figure 4 repeats the calculations behind Figures 1 and 2, but now splitting the sample between before 1982 and after 1985. The model predicts that the latter period should show a more sluggish response to shocks, and this is what we observe.

Likewise, the representative consumer habit model predicts that consumption should respond sluggishly to any event. The inattentiveness model on the other hand predicts that consumption moves sluggish with respect to events towards which people are inattentive. It is reasonable to suppose that people become aware of extreme, very noticeable events, in which case, consumption should respond instantly.[23] 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, consistently with the inattentiveness model but not with the habit model.

Another approach to distinguishing inattentiveness form habits is to look at individual data. 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.

[22] 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.

[23] Section 6 incorporates these extraordinary events in the model.

5.4 Inattentiveness versus state-contingent adjustment and the micro evidence on inattention, slow dissemination of information, and planning

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 decide to adjust consumption at sporadic dates contingent on the current state of the economy. Consumption adjustment is now state-contingent.

Since 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 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 it should adjust consumption. 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, behind this model there must be some actual physical cost of adjusting consumption. It is difficult to find evidence for this cost in the consumption of non-durables.

There is an alternative way to compare state-contingent adjustment with inattentiveness though. It consists of 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.

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 the respondents were not aware of any change in withholding, and as many as 2/3 did not know whether withholding rates had increased or fallen.

The inattentiveness model further predicts that news disseminates slowly throughout the population. Carroll (2003) and Mankiw, Reis, and Wolfers (2003) use survey data on inflation expectations to test this prediction. Carroll (2003) finds that the expectations of the public lag those of professional forecasters and that when newspapers mention inflation more often, the public updates its expectations faster, consistent with an optimal choice of inattentiveness. Mankiw, Reis, and Wolfers (2003) find that a model with exogenous staggered updating of information matches the time-series of disagreement about expected inflation in the data well, and can explain the particularly large increase in disagreement that occurred in the early 1980s during the Volcker disinflation.

A separate piece of evidence supporting the inattentiveness model comes from recent empirical work on planning behavior. The inattentiveness model predicts that a fraction of the population choose to never make plans and save less than those who do. Lusardi (1999, 2002) finds that in the sample of people over 50 years old in the Health and Retirement Study (HRS), 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. Using these proxies of the costs of planning as instruments, she finds that planning strongly predicts accumulated wealth. Ameriks, Caplin and Leahy (2003a) perform a similar exercise in a sample of TIAA-CREF members, asking people whether they write financial plans, and using as measures of the costs of planning whether they are confident with their mathematical skills, and whether they usually plan their vacations. They find that approximately 25% of households report not having a financial plan, and that those who do not plan have significantly lower savings and accumulated wealth.

In addition, in their survey, Ameriks, Caplin, and Leahy (2003a) also asked those with plans, for how long they have had their plan in place. The inattentiveness model predicts that planners who update less frequently will save less. Ameriks, Caplin, and Leahy (2003a) support this prediction as they find that people who have had plans in place for longer accumulate significantly more wealth. In related work, Alessie, Kapteyn, and Lusardi (1999) use the Dutch CentER data-panel, and find that the longer is people's reported planning horizon, the larger are their savings, though the effect is not statistically significant.

Finally, Hurst (2003) uses data from the Panel Study of Income Dynamics (PSID) to find that people that reach retirement with low wealth, also have a larger drop in consumption at retirement (consistent with inadequate planning for retirement), and earlier in life had 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.

This evidence that, when asked, people report being unaware of important current economic events; that individual expectations are consistent with slow dissemination of information; that reported planning behavior is a determinant of accumulated wealth; and that there is a group in the population that does not plan, saves little, and lives hand-to-mouth, all support the inattentiveness model, but would not be predicted by a state-contingent adjustment model.

5.5 Inattentiveness and excess sensitivity at the micro level

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. The model predicts that people 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. Likewise, if the event is extraordinary in the sense of capturing people's attention (as will be formalized in section 6), the model also predicts an instantaneous reaction.

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.

5.6 Possible new tests of inattentiveness

The inattentiveness model makes a series of sharp predictions on the behavior of individual consumption. The discussion so far shows that these predictions are consistent with a wide set of known facts about individual behavior.

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 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.

6 Extensions

The main assumption behind the model is that it is costly to acquire, absorb, and process information. One might argue though that surely people observe some things, namely their income or the amount in their bank account. However, understanding what this information implies for their wealth, which includes forecasts of the entire path of future income, as well as figuring out the other relevant state variables that should guide people's optimal actions, are likely not easy and quite costly. Moreover, it is also not costless to be able to write down a mapping from these state variables to the person's choices, which include not just the cost of solving an often difficult optimization problem, but also the cost of discussing within the household what is the best course of action. Therefore, even if the costs of acquiring information may be small, the costs of absorbing and processing this information may be quite substantial.

6.1 Extraordinary events

There are some extraordinary circumstances though, that people likely keep an eye on. 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 δ 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 δ) an extraordinary event occurs which dramatically changes the person'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.

Appendix F solves the problem of an inattentive consumer with CARA utility and the income process in section 4, including these extraordinary events. (A similar calculation could be performed for the inattentive savers.) Under the convenient but inessential assumption that $r = \rho$, optimal inattentiveness minimizes the function $A(0)$, which solves the boundary value differential equation:

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

This differential equation can be solved numerically to find d^*. Panel A of Table 6 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 arrives.

Table 6 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.

6.2 Hybrid plans

In the United States, many workers enroll in fixed percentage contribution IRA plans, which every month save a fixed percentage of their income for retirement. The economy has provided a mechanism that saves people the cost of observing their income and deciding how much to save, by automatically saving a constant fraction every period.

With these hybrid consumption-savings plans, $c_t = \theta y_t + \tilde{c}_t$ 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 (θ). Appendix G solves the problem of this hybrid consumption-savings planner, with CARA utility and ARMA income. The optimal $\tilde{d}$ and θ are state-independent, and Table 7 displays them for different plausible values of the interest rate and the costs of planning. The ability to choose θ implies that relative to consumption-planning the agent is now inattentive for even longer. The optimal θ in turn are quite small, ranging from 0.02 to 0.19. Note that a world with hybrid planners would on aggregate be similar to a world partially populated by consumption and savings planners. Moreover, in both cases, the theory predicts that wither θ or λ should be quite small, which matches the empirical findings reported in section 5.2.

7 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 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 2

Treating the vector (i, j) as a random variable with distribution $\Psi(i, j)$, equation (17) 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 (17) 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 (17) 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 $E_{t-s}[e_{t,t-s}] = 0$. 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 $k > 0$.[24] 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:

$$\begin{aligned} \Delta C_{t+1} &= \sum_{k=0}^{N-1} \sum_{m=0}^{NI} \left[\int_{t}^{t+1-k/N} \varepsilon(v)dv + \int_{t-m/N}^{t} \varepsilon(v)dv \right] \Psi(m/N, k/N) \\ &= \sum_{k=0}^{N-1} \left[\int_{t}^{t+1-k/N} \varepsilon(v)dv \right] P^j(k/N) + \sum_{m=0}^{NI} \left[\int_{t-m/N}^{t} \varepsilon(v)dv \right] P^i(m/N). \end{aligned} \tag{31}$$

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 (31) into independent increments in intervals of length $1/N$:

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

[24] 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 Δ_1 and Δ_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:

$$\begin{aligned} \Delta C_{t+1} &= \sum_{k=0}^{N(I+1)-1} u_{t+1-k/N} F(k), \qquad (32) \\ \text{with } u_{t+1-k/N} &\equiv \int_{t+1-k/N}^{t+1-k/N+1/N} \varepsilon(v)dv, \end{aligned}$$

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 (32), 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 $Var[e_{t+1-s}] = \sigma_\varepsilon^2$. Moreover, 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 the Proposition.□

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 (32). 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:

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

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

Appendix B - Spectrum of consumption

Proof of Proposition 4

The power spectrum is defined as

$$h_{\Delta C}(\omega)=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j},$$

where $\gamma_j=E\left(\Delta C_t-E(\Delta C_t)\right)\left(\Delta C_{t-j}-E(\Delta C_t)\right)$, the j^{th} autocovariance of ΔC_t. The normalized power spectrum is $f_{\Delta C}(\omega)=h_{\Delta C}(\omega)/Var(\Delta C)$.

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 (17) is stationary, its autocovariance function is symmetric ($\gamma_j=\gamma_{-j}$). Then:

$$h_{\Delta C}(\omega)=\frac{1}{2\pi}\left[\gamma_0+2\sum_{j=1}^{\infty}\gamma_j\cos(\omega j)\right]. \tag{33}$$

For the process in (17) 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 (33) and dividing by γ_0, gives the expression in Proposition 5, which depends only on $\{\Phi(i)/\Phi(0)\}$ for i from 1 to I. Evaluating (18) at frequency zero and rearranging gives the excess smoothness ratio in (19).□

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

Proof of Proposition 6

The Euler equation in (13) with CARA utility implies:

$$c_t^* = \frac{(r-\rho)t}{\alpha} + c_0^*. \tag{34}$$

Using this to substitute out c_t in the budget constraint (24), 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:

$$E\left[w'\right] = e^{rd^*}\left[w - \frac{(r-\rho)(1-e^{-rd^*}-rd^*e^{-rd^*})}{\alpha r^2}\right] - K - \frac{e^{rd^*}-1}{r}c_0^*, \tag{35}$$

$$Var\left[w'\right] = \frac{\sigma^2}{2r}(e^{2rd^*}-1), \tag{36}$$

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 implies that:

$$-Bw = (r-\rho)d^* + \ln\left(E\left[e^{-Bw'}\right]\right). \tag{37}$$

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 (35)-(36) in (37), gives the solution for c_0^*:

$$c_0^* = rw - \frac{rK}{e^{rd^*}-1} - \frac{B\sigma^2}{4}(e^{rd^*}+1) - \frac{(r-\rho)}{e^{rd^*}-1}\left[\frac{rd^*}{B} + \frac{e^{rd^*}-1-rd^*}{\alpha r^2}\right]. \tag{38}$$

The marginal utility of consumption equals the marginal value of assets:

$$e^{-\alpha c_0^*} = ABe^{-Bw}. \tag{39}$$

If the guess of the value function is valid, (39) must hold for all possible realizations of w. Matching coefficients shows that $B = \alpha r$. Going back to (38) with this result gives:

$$c_0^* = rw - \frac{rK}{e^{rd^*}-1} - \frac{\alpha r\sigma^2}{4}(e^{rd^*}+1) - \frac{r-\rho}{\alpha r}. \tag{40}$$

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 (40) to obtain:

$$(e^{rd^*} - 1)^2 = \frac{4K}{\alpha \sigma^2}. \tag{41}$$

Solving this equation gives (25). Using the solution for d^* in (40) gives the solution for c_0^* in (27).□

Appendix D - The inattentive saver's problem

Proof of Proposition 7

The problem facing the agent can be written as:

$$\begin{aligned} W(w) &= \max_{d,\{s_t\}} E\left[\int_0^d e^{-\rho t} u(y_t - s_t)dt + e^{-\rho d} W(w')\right] && (42) \\ \text{s.t.} \quad da_t &= (ra_t + s_t)dt && (43) \end{aligned}$$

Integrating (43) 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:

$$\begin{aligned} E_0\left[w'\right] &= e^{rd}\left(w + \int_0^d e^{-rt} s_t dt\right) - K + \frac{1 - e^{rd}}{r} y^P + \frac{e^{-\phi d} - e^{rd}}{r+\phi} y^T, && (44) \\ Var_0\left[w'\right] &= \frac{\sigma_P^2}{r^2} d + \frac{\sigma_T^2 (1 - e^{-2\phi d})}{2\phi\,(r+\phi)^2} && (45) \end{aligned}$$

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

$$E_0\left[u'(y_t - s_t)\right] = e^{(r-\rho)(d-t)} E_0\left[W_w(w')\right], \quad \text{for } t \in [0, d). \tag{46}$$

Combining this equation for time t and for time 0:

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

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

$$s_t = s_0 - (1 - e^{-\phi t})y_0^T - \frac{\alpha}{2}\left(\sigma_P^2 t + \frac{\sigma_T^2(1 - e^{-2\phi t})}{2\phi}\right) - \frac{(r-\rho)\,t}{\alpha}. \tag{47}$$

The envelope theorem condition is:

$$W_w(w) = e^{(r-\rho)d} E\left[W_w(w')\right]. \tag{48}$$

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 (48), and using the properties of the log-normal distribution together with (44)-(45) 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 (47) to substitute out savings in this equation gives, after rearranging,:

$$s_0 = -rw + y + \frac{r-\rho}{\alpha r} + \frac{\alpha\sigma_P^2}{2r} + \frac{\alpha\sigma_T^2}{2(r+2\phi)} + \frac{rK}{e^{rd}-1} - \frac{\alpha r\phi\sigma_T^2(1-e^{-2\phi d})}{4(r+2\phi)(r+\phi)^2(e^{rd}-1)} \tag{49}$$

Combining the envelope theorem (48) with (46) 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 (49), this I can solve this equation A to obtain:

$$A = \frac{1}{\alpha r}\exp\left\{\frac{r-\rho}{r} + \frac{\alpha^2\sigma_P^2}{2r} + \frac{\alpha^2\sigma_T^2}{2(r+2\phi)} + \frac{\alpha r K}{e^{rd}-1} - \frac{\alpha^2 r\phi\sigma_T^2(1-e^{-2\phi d})}{4(r+2\phi)(r+\phi)^2(e^{rd}-1)}\right\}. \tag{50}$$

Given (50) 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:

$$\hat{A}(K,d) \equiv \frac{K}{e^{rd}-1} - \frac{\alpha\phi\sigma_T^2(1-e^{-2\phi\hat{d}})}{4(r+2\phi)(r+\phi)^2(e^{r\hat{d}}-1)}. \tag{51}$$

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

$$\frac{e^{(r-2\phi)d}}{\Xi\left(e^{rd}-1\right)^2}\underbrace{\left[re^{2\phi d}(1-K\Xi) + 2\phi e^{-rd} - 2\phi - r\right]}_{\equiv B(d)} = 0, \tag{52}$$

$$\text{where} : \Xi \equiv \frac{4(r+2\phi)(r+\phi)^2}{\alpha\phi\sigma_T^2}. \tag{53}$$

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 $W(w)$. It is easy to see that $V(w) = -\exp\left(-\alpha c_0^*\right)/\alpha r$, while $W(w)$ is in (50). The condition $V(w) > W(w)$ then 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 (41). Then, $sign\left\{H_K(.)\right\} = sign\left\{d^* - \hat{d}\right\}$, so I must compare optimal inattentiveness with consumption and savings plans.

Evaluating the function $B(d)$ defined in (52), 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 Ξ in (53) 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:

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

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 $\hat{d}$.□

<u>Proof of Corollary 4</u>

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 Ξ, the condition $\bar{K}>1/\Xi$ becomes the condition in Corollary 4.□

<u>Proof of Proposition 9</u>

Using the solutions for $V(w)$ and $W(w)$ with $\hat{d}=+\infty$, $V(w)>W(w)$ becomes

$$c_0^* > rw - \frac{r-\rho}{\alpha r} - \frac{\alpha\sigma_P^2}{2r} - \frac{\alpha\sigma_T^2}{2(r+2\phi)}. \tag{54}$$

Using the solution for c_0^* in (26) 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 (41) to replace for K and rearranging gives the condition in (29).□

<u>Proof of Corollary 3</u>

Using the fact that $\hat{c}_0 = y_0 - \hat{s}_0$ and (49) 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 (54) holding, which is true for the agent who chooses to be an inattentive saver.□

Appendix F - Extraordinary events

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 $\sigma^2 = \sigma_P^2/r^2 + \sigma_T^2/(r+\phi)^2$ 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:

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

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:

$$\begin{aligned}u'(c_{D+t}) \quad &= \quad J_w(\bar{w}_{D+t},t) \Leftrightarrow \\ c_{D+t} \quad &= \quad r\bar{w}_{D+t} - \frac{\ln(A(t))}{\alpha}.\end{aligned} \tag{55}$$

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 (55) to replace out consumption. Using the guess for the value function in this equation gives the differential equation in Proposition 10. 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. Finally, d^* can be found by minimizing $A(0)$.

Appendix G - Hybrid consumption-savings plans

The problem to solve is:

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

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

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

$$E\left[u'(\theta y_t+\tilde{c}_t)\right]=e^{(r-\rho)(d-t)}E\left[Z'(w')\right]. \tag{56}$$

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

$$\begin{aligned}u'(\theta y_0+\tilde{c}_0) &= e^{(r-\rho)t}E\left[u'(\theta y_t+\tilde{c}_t)\right]\Leftrightarrow\\ -\alpha\theta y_0-\alpha\tilde{c}_0 &= (r-\rho)t-\alpha\tilde{c}_t-\alpha\theta E[y_t]+\frac{\alpha^2\theta^2}{2}Var[y_t]\Leftrightarrow\\ \tilde{c}_t &= \tilde{c}_0+\theta(1-e^{-\phi t})y_0^T+\frac{(r-\rho)t}{\alpha}+\frac{\alpha\theta^2}{2}Var[y_t].\end{aligned} \tag{57}$$

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:

$$Z'(w)=e^{(r-\rho)d}E\left[Z(w')\right] \tag{58}$$

The first order condition (56) at time 0, combined with this condition, leads to:

$$\begin{aligned}e^{-\alpha(\lambda y_0+\tilde{c}_0)} &= \alpha r A e^{-\alpha r w}\Leftrightarrow\\ \tilde{c}_0 &= -\frac{\ln(\alpha r A)}{\alpha}+rw-\theta y_0\end{aligned} \tag{59}$$

Using the solutions for $\tilde{c}_t$ in (57) and $\tilde{c}_0$ in (59) to substitute for the consumption terms in the budget constraint and rearranging shows that w' is normally distributed with:

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

Using the envelope theorem condition (58), together with the guess for the value function, the normality of wealth, and the expression for $E[w']$, gives:

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

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 θ and d are then the solution to the problem:

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

Since none of the expressions in this objective function depend on the sate of the economy, the optimal d and λ 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. Solving this minimization numerically produces the results in Table 7.

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.

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.

Brock, William and Cars Hommes (1997) "A Rational Route to Randomness," *Econometrica*, vol. 65 (5), pp 1059-1095.

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.

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.

Evans, George and Seppo Honkapohja (2001) *Learning and Expectations*, Princeton, Princeton University Press.

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.

Fuhrer, Jeffrey (2000) "Habit Formation in Consumption and Its Implications for Monetary Policy," *American Economic Review*, vol. 90 (3), pp. 367-390.

Gabaix, Xavier and David Laibson (2002) "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.

Gourinchas, Pierre-Olivier, and Jonathan Parker (2002) "Consumption Over the Life-Cycle," *Econometrica*, vol. 70 (1), pp. 47-89.

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

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.

Kurz, Mordecai (1997) *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*, Springer series in Economic Theory, No. 6, Springer-Verlag.

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 (2004) "Disagreement in Inflation Expectations," *NBER Macroeconomics Annual 2003*.

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. 1464-76.

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.

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

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

Reis, Ricardo (2004) "Inattentive Producers," unpublished manuscript, Harvard University.

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.

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.

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 Inattentiveness Length

	Parameter combinations (σ_P, σ_T, ϕ, r, α)					
	Pischke	Riskless rate	Volatility 1	Volatility 2	Volatility 3	Gourinchas-Parker
d*	8	13	11	7	6	3
K*	$28	$12	$14	$36	$54	$324

Notes: In the row with d* is optimal inattentiveness with K = $30. In the row with K* are the costs of planning that would make optimal inattentiveness equal 8 quarters. The Pischke parameters are (45, 1962, 0.487, 0.015, 2/6926). In the riskless rate column, r is lowered to 0.005. In the volatility 1-3 columns (σ_P, σ_T) equal (23, 1602), (68, 1601), and (90, 1599). In the Gourinchas-Parker column, the parameter values are (1014, x, ∞, 0.0085, 0.51/6926).

Table 3: The Excess Smoothness Ratio

Panel A: Estimates

Method	Lags	ψ	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 Φ(i):	ψ
From news regressions in Table 2, 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 (ψ) 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 2: Regressing Consumption Growth on News on Income Growth

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
Panel A. Predictors of $\Delta\ln(Y_t)$										
$\Delta\ln(Y_{t-1}),\ldots, \Delta\ln(Y_{t-5})$	.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: .062 1st stage		W_{IN}: 4.71 (.701)	
Panel B. Predictors of $\Delta\ln(Y_t)$	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
$\Delta\ln(Y_{t-1}),\ldots, \Delta\ln(Y_{t-5})$, $\ln(C_{t-1}/Y_{t-1}),\ldots, \ln(C_{t-5}/Y_{t-5})$	.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: .196 1st stage		W_{IN}: 7.58 (.428)	
Panel C. Predictors of $\Delta\ln(Y_t)$	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
$\Delta\ln(Y_{t-1}),\ldots, \Delta\ln(Y_{t-5})$, $\ln(C_{t-1}/Y_{t-1}),\ldots, \ln(C_{t-5}/Y_{t-5})$, $r_{t-1},\ldots, r_{t-5}$	.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: .187 1st stage		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) $\Delta\ln(C_{t+1}) = const. + \beta_0 y_{t+1} + \beta_1 y_t + \ldots + \beta_9 y_{t-8} + \tilde{u}_t$. ***, ** 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: 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)
$\Delta\ln(Y_{t-9}),\ldots, \Delta\ln(Y_{t-12})$, Ln$(C_{t-9}/Y_{t-9}),\ldots, \ln(C_{t-12}/Y_{t-12})$	.166	.167	.64	2.33
	(.180)		(.743)	(.940)
$\Delta\ln(Y_{t-9}),\ldots, \Delta\ln(Y_{t-12})$, $\ln(C_{t-9}/Y_{t-9}),\ldots, \ln(C_{t-12}/Y_{t-12})$ $r_{t-9},\ldots, r_{t-12}$	.049	.073	.80	4.53
	(.139)		(.650)	(.952)

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)
$\Delta\ln(Y_{t-9}),\ldots, \Delta\ln(Y_{t-12})$, Ln$(C_{t-9}/Y_{t-9}),\ldots, \ln(C_{t-12}/Y_{t-12})$	.226	.148	2.542	.305	.164
			(.960)	(.950)	(.685)
$\Delta\ln(Y_{t-9}),\ldots, \Delta\ln(Y_{t-12})$, $\ln(C_{t-9}/Y_{t-9}),\ldots, \ln(C_{t-12}/Y_{t-12})$ $r_{t-9},\ldots, r_{t-12}$	.226	-.057	4.085	.086	.057
			(.982)	(.930)	(.812)

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. In Panel B, A-R is the Anderson-Rubin test, Moreira is the conditional likelihood ratio test, and LM is the conditional Lagrange multiplier test (see Moreira, 2004).

Table 5: Rational Expectations vs. Campbell-Mankiw 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 in parentheses.

Table 6: 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%, ϕ=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 7: 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 (θ)

	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 as in Table 6.

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

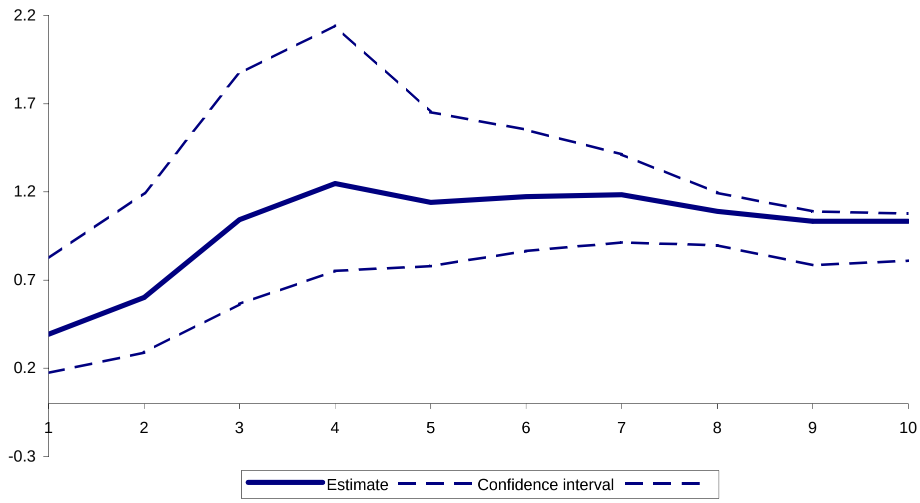

Figure 2. Estimates of the Inattentiveness Weights

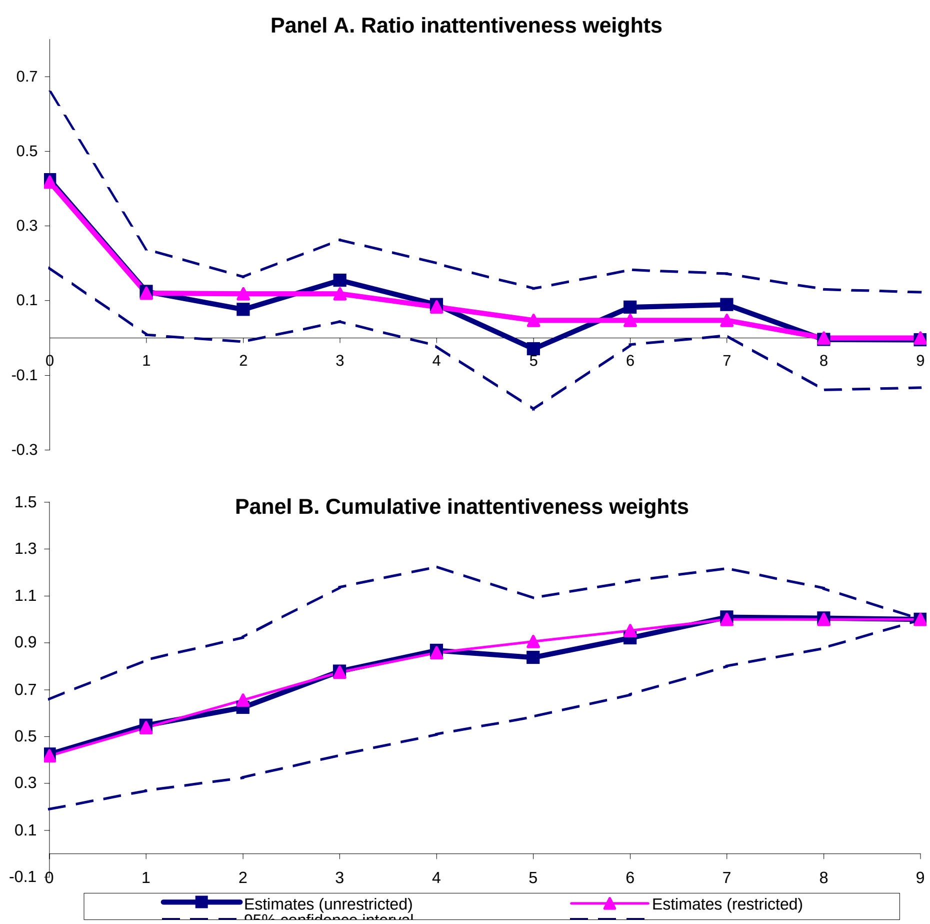

Figure 3. Predicted and Actual Normalised Spectral Densities

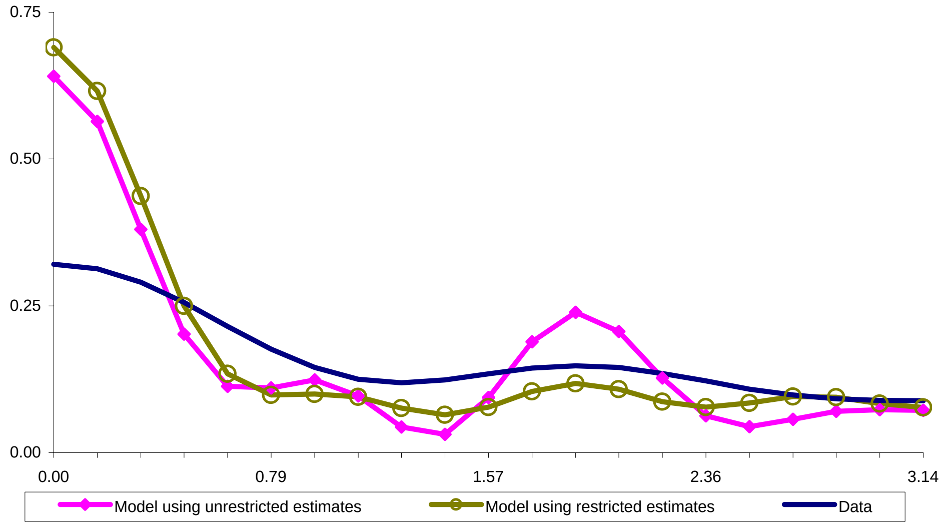

Figure 4. Sluggishness in Consumption pre and post 1982