Costly Information Processing and Consumption Dynamics

Jérémy Boccanfuso*

November 18, 2019

[Job Market Paper]

Latest version here

Abstract

This paper studies the consumption-saving problem of a consumer who faces a fixed cost for paying attention to noisy information and whose attention strategy, i.e., whether or not she pays attention, can be a function of the underlying information. At the optimum, consumers chose to be attentive when evidence accumulates far from their prior beliefs. The model provides an explanation for four puzzling empirical findings on consumption and expectations. First, consumers' attention depends on the information content. Second, aggregate information rigidities vary over the business cycle. Third, consumers only react to large anticipated shocks and neglect the impact of small ones. Fourth, aggregate consumption dynamics vary over the business cycle.

Keywords: Consumption dynamics, information frictions and inattentive consumers.

JEL Classification: E21, E70, C61.

*Paris School of Economics – EHESS. I am extremely grateful to Xavier Ragot, Philippe Andrade, Olivier Compte, Lena Drager, Antoine Ferey, Gaetano Gaballo, Cars Hommes, Ricardo Reis, Mirko Wiederholt, anonymous referees and participants at the CEF and EEA annual meetings, the Second Behavioral Macroeconomics Workshop, the CeDNEF seminar at the University of Amsterdam, the Paris School of Economics and the University Paris 1 for helpful comments and suggestions. The Labex OSE fellowship provided financial support. The paper was awarded at the 2016 CEF Graduate Student Paper Contest. Contact: jeremy.boccanfuso@psemail.eu.

1 Introduction

Theories of consumption play an important role in business cycle models and recently many authors have proposed changes to the modeling of consumption in DSGE models. In Heterogeneous Agents New Keynesian (HANK) models, hand-to-mouth consumers, precautionary savings, and heterogeneity of marginal propensities to consume shape the response of aggregate consumption to shocks. In models with information frictions on the household side, slow updating of beliefs over time generates hump-shaped responses of aggregate consumption to shocks (Mankiw and Reis,2006; Maćkowiak and Wiederholt,2015; Carroll et al., forthcoming).

However, several empirical findings on consumption and expectations are not explained by these models. First, consumers' attention depends on the information content. Using information experiments, Armantier et al. (2016), Abe and Ueno (2015) and Khaw et al. (2017) find that individuals are more likely to incorporate new information when it contradicts their prior beliefs. Second, information rigidities are not constant over time. Dräger and Lamla (2012) and Coibion and Gorodnichenko (2015) find that information rigidities drop persistently in the aftermath of a recession. Third, consumers only react to large anticipated shocks and neglect the impact of small ones. Jappelli and Pistaferri (2010) term this mechanism the *magnitude hypothesis* and empirical evidence in favor of the hypothesis is accumulating (Browning and Collado, 2001; Scholnick, 2013; Kueng, 2018).[1] Fourth, aggregate consumption dynamics vary over the business cycle. Caballero (1995) finds that "in good times, consumers respond more promptly to positive than to negative wealth shocks, while the opposite is true in bad times" and Kumar and Jia (2019) report systematic decreases in consumption growth persistence during recessions.

This paper proposes a model that can match these four empirical findings, while preserving the aforementioned prediction of information friction models regarding the sluggish response of aggregate consumption to shocks. In the model, there is a source which conveys noisy information about shocks to one's permanent income. Consumers face a fixed cost for paying attention to this information and their attention strategy, i.e., whether or not an agent pays attention, can be a function of the underlying information. Apart from this novel information structure, the consumption-saving problem considered in the paper is left as transparent as possible and coincides with Hall's (1978) random walk model with quadratic utility. The model

[1] Each of these papers controlled for the presence of credit constraints and concluded that it cannot explain the observed features of excess sensitivity. A priori, the magnitude hypothesis could arise as a consequence of consumption adjustment costs. In these models, consumption adjusts sporadically based on a state-dependent rule. Accordingly, Chetty and Szeidl (2016) demonstrate that in the presence of commitment costs, excess sensitivity and smoothness vanish for large shocks. However, empirical studies on the magnitude hypothesis have mainly focused on non-durable consumption goods, that is, those consumption goods which are the less likely to exhibit commitment costs.

predictions are thus isolated from other refinements of the textbook consumption theory.

The main difference to Reis (2006a) microfounded model of inattentive consumers is that consumers can choose information-dependent attention strategies and do not necessarily access perfect information when attentive. These two deviations from Reis (2006a) allow to match the aforementioned empirical evidence that consumers do not update on a purely time-dependent basis and, also, that expectations are heterogenous among attentive consumers (e.g. Armantier et al. (2016)).

The main findings of the paper are as follows. First, consumers' attention depends on the information content. At the optimum, the consumer faces an inattention region where she disregards new information and does not adjust her consumption plan. It is only when evidence against her prior beliefs accumulates that the consumer is willing to pay attention to new information releases and revises her consumption plan accordingly.

Second, information rigidities are time-varying. Consumers adopt an information-dependent attention strategy. Therefore, and starting from the stationary cross-sectional distribution of consumers, an aggregate income shock prompts more consumers to be attentive.[2] Information being noisy, the impact of an aggregate shock disseminates slowly in the economy and the increase in the share of attentive consumers is persistent.

Third, the model predicts a positive correlation between the size of an income shock and the marginal propensity to consume. It thus provides a rationale to explain *the magnitude hypothesis*. The intuition is as follows. For a consumer who was last attentive one period ago, a small permanent income shock (in absolute value) is unlikely to be significant enough to trigger her attention. However, as the size of the shock gets larger, the shock becomes more likely to trigger the consumer's attention and to prompt her to revise her consumption path. More generally, the expected marginal propensity to consume out of a shock to permanent income is found to be both history-dependent and shock-dependent.

Fourth, aggregate consumption dynamics vary over the business cycle. The model predicts that during economic busts (respectively booms) – that is periods when lagged consumption growth was below (resp. above) its steady state level – a positive (resp. negative) shock generally lowers the share of attentive consumers and generates a smoother response, while

[2]Consumers are ex-post heterogenous because of their idiosyncratic income shocks, information noises and attention histories. To apprehend this multi-dimensional heterogeneity, I show that the cross-sectional distribution of consumers may be characterized by a function of the information that each consumer disregards when being inattentive. Starting from the stationary cross-sectional distribution, an aggregate income shock shifts the distribution of consumers and increases the average information that consumers would disregard by remaining inattentive. Thereby, more consumers are willing to pay attention and the share of attentive consumers increases.

a negative (resp. positive) shock generates a sharper response. Moreover, the persistence of aggregate consumption growth depends on the endogenously time-varying share of attentive consumers. In normal times, information rigidities are near their steady state level and consumption persistence relatively constant. However, during unusual times such as recessions, information rigidities decrease a lot and so does aggregate consumption persistence.

In sum, the model provides an explanation to the aforementioned four facts indicating that consumers' attention depends on the information content, information rigidities are time-varying, consumers only react to large anticipated shocks, and aggregate consumption dynamics vary over the business cycle.

The consumption theory proposed in this paper also has some other attractive properties. First, in models of consumption adjustment costs (Caballero, 1993, 1995; Chetty and Szeidl, 2016) household consumption is constant between adjustments and consumption adjustments are large; while here consumption is not constant whilst the consumer is inattentive and the consumption change that triggers attention is modest.[3,4] Second, the model is tractable and a simple iterative method makes it possible to track the evolution of the cross-sectional distribution of consumers. It therefore enables us to analyze multiple aspects of consumption nonlinearities. Moreover, it offers a natural benchmark to asses the approximation loss from relying on time-invariant information rigidities. In the absence of aggregate shocks, the cross-sectional distribution of consumers is stationary and information rigidities constant. Predictions from time-invariant information-rigidity models however become less accurate as the variance of aggregate shocks increases relatively to that of idiosyncratic shocks. Finally, the model accurately matches the persistence of aggregate consumption growth and provides an explicit mapping between aggregate information frictions and consumption growth persistence.

This paper belongs to a growing literature analyzing the implications of consumers' information rigidities for consumption dynamics. This literature generally opposes two forms of information rigidities: rational inattention and sticky expectations. Papers building on Sims (2003) rational inattention generally consider linear-quadratic Gaussian frameworks (Luo, 2008; Luo and Young, 2014) or assume ex-post Gaussian distributions of the true state and noise (Luo

[3]See Reis (2006a) or Carroll et al. (forthcoming) for a discussion of consumption adjustment cost models. Unattractive predictions of these models are that (i) we do not observe adjustment costs for non-durable consumption goods, (ii) consumption must be constant between adjustments and (iii) consumption adjustments must be large in these models.

[4]I find that the consumption change that triggers attention is less than half the consumption change that triggers an adjustment in Caballero (1995).

et al., 2017).[5] Sticky expectation models fall into one of two categories. The first corresponds to the Calvo-like models in which agents have a constant probability to update their expectations, e.g., Mankiw and Reis (2002), Carroll (2003) and Carroll et al. (forthcoming), and the second category builds on microfounded model of sticky expectations (Gabaix and Laibson, 2001; Reis, 2006a,b) where agents update every nth period. Consequently, both rational inattention and sticky expectations imply that information rigidities are constant over time. In comparison, this paper proposes a model of rationally inattentive consumers that nests both forms of information rigidities as limiting cases (within a LQG framework) and that naturally generates time-varying information rigidities. These variations in information rigidities are found to be large and to hold important implications for consumption dynamics at both the household and aggregate levels.

A few papers have already proposed mechanisms to generate time-varying aggregate information rigidities. Gorodnichenko (2008) and Woodford (2009) consider the price setting problem of firms and respectively find that information externalities and inattention between price reviews may result in time-varying information rigidities. Similarly, Cheremukhin and Tutino (2016) highlight that time variations in firms' exit rates and markups lead to countercyclical information rigidities. These explanations are based on information rigidities on the side of firms. Nimark (2014) and Larsen et al. (2019) argue that the media coverage of economic events results in time-varying information rigidities. However, most of the uncertainty faced by consumers is the consequence of idiosyncratic shocks, for which the media are unlikely to be the main source of information. This paper thus adds to this literature by simultaneously analyzing the joint dynamics of aggregate information rigidities and consumption. Doing so, it provides a novel mechanism to generate time-varying information rigidities. This mechanism is internal and does not arise as the consequence of an aggregate externality. It is thus more likely to be transferable to other settings.

Finally, this paper relates to the large literature studying the consumption response to income changes (see Jappelli and Pistaferri (2017)). In particular, it provides novel explanations to some puzzling consumption nonlinearities previously identified in the data. These nonlinearities are shown to be the consequence of household level adjustments in expectations. Hence, the paper also relates to the large literature analyzing the implications of microeconomic adjustments for aggregate dynamics, e.g., Caballero and Engel (2007), and Chetty and Szeidl (2016).[6]

[5]A notable exception is Tutino (2013) who numerically solves for the optimal discrete distribution of actions when utility is CRRA. In comparison, the setup considered here is substantially different, the model is solved analytically and the focus is on aggregate consumption.

[6]Similarly to Burstein (2006) where firms may choose a pricing plan at revision dates, I allow consumers to choose a consumption plan when attentive. In both models, a plan can be contingent on the information available at the adjustment date but cannot be contingent upon future realizations of economic shocks.

The paper is organized as follows. The consumption problem is presented and solved in Section 2. Section 3 further discusses consumers' attention strategy and derives implications for inattention lengths. Section 4 provides a quantitative application with an ARMA income process. Sections 5 and 6 respectively derive implications for consumption dynamics at the household and aggregate levels.

2 Optimization with costly information processing

2.1 The consumer problem

This section presents the consumer's problem. Importantly, it discusses the information structure which is the main innovation of this paper and explains how it interacts with the consumer's problem.

I consider the problem of a rationally inattentive consumer with memory who lives for T periods, consumes c_t each period and whose utility $u(c_t)$ is quadratic.[7] This agent discounts future utility by the factor $\beta \in (0,1)$ and can borrow and lend freely at the gross interest rate $1+r$. At each period, she receives an exogenous stochastic income y_t which follows from a multivariate linear state space model with Gaussian white innovations. The consumer's budget constraint therefore writes $a_{t+1} = (1+r)a_t - c_t + y_t$ where a_t are the consumer's assets at time t.

For ease of exposition, I follow the literature (e.g. Luo (2008)) and reformulate the consumer's problem in terms of permanent income $s_t \equiv a_t + \mu_t$ where $\mu_t \equiv \sum_{k=t}^{T-1} E_t[(1+r)^{(t-k)} y_k]$ is the expected flow of actualized incomes. The period budget constraint may accordingly be written in terms of permanent income and thus becomes $s_{t+1} = (1+r)s_t - c_t + \zeta_{t+1}$ with $\zeta_{t+1} = \mu_{t+1} - (1+r)\mu_t$ a Gaussian white noise with variance σ_ζ^2.

Because of information frictions, the consumer cannot perfectly observe the economic state s_t. Nevertheless, there exists a source that continuously conveys information about the evolution of s_t. This information channel stands for instance for the news from newspapers, TV and cheap talk agents get every day and that may potentially reflect an evolution of the economic environment. At any moment, the consumer may either be attentive to this information – in

[7]The quadratic utility assumption allows to derive an analytical solution and is a widely used framework for the study of rationally inattentive consumers (e.g. Sims (2003), Luo and Young (2014)).

return for a fixed utility cost denoted λ – or remain inattentive and put this information aside for a later use.

Formally, I follow the signal extraction literature (e.g. Luo and Young (2014) for a recent application to consumer theory) and model the continuous information using an additive noisy signal $z_t = s_t + \vartheta_t$ where ϑ_t is an i.i.d. Gaussian white noise with variance σ_ϑ^2 at each period. The smaller the variance of the noises σ_ϑ^2 is, the more informative these signals are. Moreover, I define a latent information set, denoted $\mathcal{I}_t$, that contains all signals since the beginning of time and past actions of the consumer. These actions consist in her consumption choices (c_k) and whether she was attentive ($\tau_k = 1$) or not ($\tau_k = 0$) at each period $k \in [0, 1, \ldots, t-1]$. That is, $\mathcal{I}_t$ is the σ-algebra defined as

$$\mathcal{I}_t \quad \equiv \quad \Big\{z_0, \tau_0, c_0, \ldots, z_{t-1}, \tau_{t-1}, c_{t-1}, z_t\Big\} \tag{1}$$

This information is not observable by the consumer and is used as a modeling tool. Consequently, and in opposition to this latent information set, let $\overline{\mathcal{I}}_t$ be the information set in the hands of the consumer. $\overline{\mathcal{I}}_t$ and $\mathcal{I}_t$ are generally different and they coincide only when the consumer is constantly attentive to each new signal release. More specifically, we have

$$\overline{\mathcal{I}}_t \quad \equiv \quad \Big\{\bar{z}_0, \tau_0, c_0, \ldots, \bar{z}_{t-1}, \tau_{t-1}, c_{t-1}, \tau_t, \bar{z}_t\Big\} \tag{2}$$

The consumer's information set also contains past actions. However, it is not necessarily incremented at each period by the signal z_t as the consumer may rationally prefer not to pay attention to the information channel. Instead, let $\bar{z}_t$ be the novel information that the consumer gets at period t. Then, by definition, we have that $\bar{z}_t = \emptyset$ is empty whenever the consumer is inattentive ($\tau_t = 0$). I do not impose a specific form for $\bar{z}_t$ at periods when the consumer is attentive. It could for example be a truncated sequence of signals or a filtration of these signals. The only restriction I impose is that at any period t, the consumer's information $\overline{\mathcal{I}}_t$ may be retrieve from the latent information set $\mathcal{I}_t$. That is, $\sigma(\{\bar{z}_k\}_{k=0}^t) \subseteq \sigma(\{z_k\}_{k=0}^t)$ where $\sigma(.)$ denotes a σ-algebra. Following Molin and Hirche (2010), I will refer to this property as the nestedness of the information structure. Intuitively, this condition implies that the signals z_t are the only source of information for the consumer. Moreover, the consumer may catch up with any information she previously ignored.

The above-described consumer problem may be expressed as a discounted linear-quadratic

Gaussian problem in the following form:

$$\begin{aligned}
\min_{\{c_t,\tau_t\}_{t=0}^{T-1}\in\mathbb{R}^T\times\{0,1\}^T} \quad & E_0\left(\sum_{t=0}^{T-1}\beta^t\Big((c_t-\bar{c})^2+\lambda\tau_t\Big)+\beta^T q_T s_T^2\middle|\left\{\mathcal{I}_t,\overline{\mathcal{I}}_t\right\}\right) && (3)\\
\text{s.t.} \quad & s_{t+1}=(1+r)s_t-(c_t-\bar{c})+\zeta_{t+1}\\
& c_t=f_t(\overline{\mathcal{I}}_t);\tau_t=g_t(\mathcal{I}_t)\\
& s_0|\overline{\mathcal{I}}_0\sim\mathcal{N}(\bar{s}_0,\sigma_{s_0}^2)\\
& \bar{s}_0=a_0+\mu_0+\frac{1-(1+r)^{-T}}{1-(1+r)^{-1}}\bar{c}
\end{aligned}$$

Problem (3) states that the consumer maximizes her intertemporal utility given the aforementioned period budget constraint and information structure. Her instantaneous utility is quadratic $u(c)=-(c-\bar{c})^2$ with $\bar{c}\in\mathbb{R}^+$ the bliss point. The control variables are consumption c_t and attention $\tau_t\in\{0,1\}$ and λ is a fixed utility cost from being attentive. In the term $\beta^T q_T s_T^2$, q_T is an arbitrary large constant which is used to impose the standard terminal condition $s_T=0$ in expectation. In the last condition, I normalize the initial state $\bar{s}_0$ by substracting an intertemporal consumption stream equal to $\bar{c}$ at each period to s_0.[8] The period budget constraint is amended accordingly (hence the term $(c_t-\bar{c})$). Finally, the third constraint imposes that the initial uncertainty in the state variable is Gaussian.

Following from the information structure, the policies $f_t(.)$ and $g_t(.)$, which respectively refer to the consumption and attention choices, are Borel-measurable functions with respect to $\overline{\mathcal{I}}_t$ and $\mathcal{I}_t$. Hence, the consumption choice c_t depends on the consumer's information at period t. Alternatively, I consider more sophisticated attention strategies. Here, the consumer bases her attention strategy on the information contained in the latent information set $\mathcal{I}_t$. That is, the consumer has the opportunity to select the type of information she will be attentive to. A real-life illustration of this strategy could for example to be attentive to economic news only when they contain keywords e.g. crises, recession or low interest rate.

The above formulation encompasses some well-known models of inattentive consumers as limiting cases. When the attention cost λ is nil, the consumer is always attentive to the Gaussian signals. Hence, the consumer's information set coincides with the latent information set and problem (3) collapses to the textbook linear-quadratic Gaussian problem with incomplete state information.

Closer in spirit is the sticky information model of Reis (2006a). He also considers the problem

[8]This novel state variable represents the consumer net wealth given that she will consume $\bar{c}$ at each period. Therefore, the period budget constraint is amended to account for the deviation between actual consumption c_t and the consumption bliss point $\bar{c}$.

of a rationally inattentive consumer who must pay a fixed cost to observe information. In his model, the consumer perfectly observes the state variable s_t when she is attentive. Transposed to problem (3), it implies that the signals are noiseless $z_t = s_t$. Further, he restricts the attention strategies to depend only on the consumer's information. Using the above notation, he focuses on solutions in the form $\tau_t = g_t(\overline{\mathcal{I}}_t)$. In particular, this assumption implies that the consumer's attention strategy must be independent of the economic conditions between information updates. As a result, Reis (2006a) finds that the optimal updating behavior is purely time-dependent.[9]

2.2 Consumption and Information

This section highlights important intermediary results. In particular, it characterizes the optimal consumption policy and treatment of signals.

Problems related to (3) have recently been studied in engineering. The closest paper is Molin and Hirche (2010) who study an undiscounted discrete-time LQG setup with a similar information structure. In particular, they show that the certainty equivalence holds in their setup (Lemma 2, Molin and Hirche (2010)). Appendix A.1 shows that the introduction of discounting does not affect this conclusion. Consequently, the consumption policy function coincides with the one we would have obtained under full-information rational expectations and is recalled in Lemma 1.

Lemma 1 (Certainty equivalence)**.** *The optimal consumption is*

$$c_t = L_t E[s_t|\overline{\mathcal{I}}_t] + \bar{c} \qquad \forall t \in 0, \dots, T-1 \tag{4}$$

where $L_t \equiv (1+r)\beta p_{t+1}/(1+\beta p_{t+1})$ *and* p_t *follows from iterating on the backward Riccati equation* $p_t = (1+r)^2 \beta p_{t+1}/(1+\beta p_{t+1})$ *with terminal condition* $p_T = q_T$*.*

Proof. See Appendix A.1. □

Because the certainty equivalence holds here, the consumption function is not affected by the information structure. We therefore retrieve the well-established conclusions that the consumption function is linear and there are no precautionary savings when the consumer's utility is quadratic. More specifically, the consumption deviation from the bliss-point $\bar{c}$ depends on

[9] Reis (2006a) also develops a behavioral extension for extreme events where it is assumed that the latter are fully observable and instantaneously internalized by consumers. By construction, problem (3) directly accounts for these extreme events and allows to study their implications within a fully microfounded framework.

the discount rate β, the interest rate r, the terminal condition q_T and the perceived permanent income given the information that has been processed at period t. The constant L_t measures the change in consumption following a marginal increase in expected permanent income.

Equation (4) in Lemma 1 states that the consumption choice depends on the expected permanent income given the consumer's information set $\overline{\mathcal{I}}_t$. Characterizing the latter expectation $E[s_t|\overline{\mathcal{I}}_t]$ is not trivial. It first requires to determine the optimal estimator $E[s_t|\mathcal{I}_t]$ with respect to the latent information set $\mathcal{I}_t$. Following common practice in the literature (e.g. Sims (2003), Luo et al. (2017) and Maćkowiak et al. (2018)), I assume that the initial uncertainty surrounding the state variable $\sigma^2_{s_0}$ is at its steady state value. Consequently, $E[s_t|\mathcal{I}_t]$ is the linear least squares estimator given by the Kalman filter in Lemma 2.

Lemma 2 (Latent Kalman filter)**.** *The optimal estimate of s_t given the latent information set is*

$$E[s_t|\mathcal{I}_t] \quad = \quad (1+r)E[s_{t-1}|\mathcal{I}_{t-1}] - c_{t-1} + \bar{c} + K(z_t - (1+r)E[s_{t-1}|\mathcal{I}_{t-1}] + c_{t-1} - \bar{c}) \qquad (5)$$

where K is the steady state Kalman gain defined in Appendix A.2

Proof. See Appendix A.2. □

Now that we have identified $E[s_t|\mathcal{I}_t]$ in Lemma 2, we can characterize $E[s_t|\overline{\mathcal{I}}_t]$. When the consumer is attentive ($\tau_t = 1$), she may access the information contained in $\mathcal{I}_t$ to form an updated estimate $E[s_t|\overline{\mathcal{I}}_t, \tau_t = 1]$. Therefore, $E[s_t|\overline{\mathcal{I}}_t, \tau_t = 1] = E[s_t|\mathcal{I}_t]$ since the latter estimator is optimal given $\mathcal{I}_t$ and the nestedness property of the information structure implies that there is no other source of information.

Deriving the optimal estimator $E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0]$ at non-updating periods ($\tau_t = 0$) is more complex because it depends on the attention strategy of the consumer. To illustrate this dependance, realize that $E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0] = E\Big[E[s_t|\mathcal{I}_t]\Big|\overline{\mathcal{I}}_t, \tau_t = 0\Big]$. Then, using equation (5) we have

$$\underbrace{E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0]}_{\text{estimate when inattentive}} \quad = \quad \underbrace{E[s_t|\overline{\mathcal{I}}_{t-1}]}_{\text{update}} + \underbrace{E\Big[(1+r)e_{t-1} + K(z_t - E[s_t|\mathcal{I}_{t-1}])\Big|\overline{\mathcal{I}}_t, \tau_t = 0\Big]}_{\text{corrective term accounting for inattention}} \qquad (6)$$

where

$$e_t \equiv E[s_t|\mathcal{I}_t] - E[s_t|\overline{\mathcal{I}}_t] \qquad (7)$$

is the perceived forecast error given the information in $\mathcal{I}_t$. The last term in the right-hand side of equation (6) represents a correction that the consumer may be willing to integrate in her consumption path when she decides to remain inattentive. Intuitively, if the choice to be attentive results from the occurence of a predetermined event, then being inattentive indicates that this event did not occur. For example, suppose that the consumer is extremely loss averse and willing to be attentive to permanent income drops only. Implicitly, this attention strategy would imply that no permanent income drops occur when she is inattentive. She would therefore infer that, on average, her permanent income is higher than what would be implied by a mechanical update $E[s_t|\overline{\mathcal{I}}_{t-1}] = (1+r)E[s_{t-1}|\overline{\mathcal{I}}_{t-1}] - c_t + \bar{c}$ at periods when she is not attentive. Consequently, she would revise her estimate $E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0]$ upward and, following from Lemma 1, increase her consumption during these periods.[10]

Appendix A.3 shows that the corrective term accounting for inattention in equation (6) is nil at the optimum. We can therefore characterize the optimal permanent income estimate of the consumer in Lemma 3.

Lemma 3 (Perceived permanent income)**.** *The optimal estimate of s_t given the consumer's information set is*

$$E[s_t|\overline{\mathcal{I}}_t] = \begin{cases} E[s_t|\mathcal{I}_t] & \text{if } \tau_t = 1 \\ (1+r)E[s_{t-1}|\overline{\mathcal{I}}_{t-1}] - c_{t-1} + \bar{c} & \text{if } \tau_t = 0 \end{cases}$$

Proof. See Appendix A.3 . □

Lemmas 1-3 together characterize the consumption behavior of the consumer. At times when she is attentive, she chooses the consumption path which maximizes her expected intertemporal utility. The dynamics of this path is similar to the one she would have selected in the absence of information frictions. Therefore and unsurprisingly, it depends on the interest rate r, the discounting factor β, time t and the horizon T. Then, at times when she is inattentive, she does not revise her consumption path and behaves as if no shock to permanent income occurred since the period she was last attentive. However, at times when she is attentive, she catches up with the information she previously ignored, revises her permanent income estimate and selects a novel consumption path.

[10]This mechanism, referred to as negative information, is central to event-based state estimation (see the book of Shi et al. (2016)). It may hold implications for consumption dynamics when one relaxes the assumption of Gaussian shocks to permanent income and/or quadratic utility. Extrapolating from Molin and Hirche (2010), it might result in a predetermined adjustment in consumption that would depend on time and the duration of inattention. See Nimark (2014) for an application of negative information in macroeconomics.

2.3 Attention strategy

This section focuses on the consumer's attention strategy. Building on the previous results, it demonstrates that the complex task of identifying the optimal attention strategy directly from the full consumer's problem (3) collapses to a much simpler discrete choice control problem that can be solved using standard dynamic programming tools.

The perceived forecast error e_t defined in equation (7) measures the expected permanent income discrepancy between the latent information – i.e. the information the consumer would access if attentive – and the consumer's information when she remains inattentive. Using Lemmas 2 and 3, this error follows a Markov process given by

$$e_{t+1} = (1-\tau_t)(1+r)e_t + K\Big(z_{t+1} - E[s_{t+1}|\mathcal{I}_t]\Big) \tag{8}$$

That is, the error is incremented at each period by the innovation from the latent Kalman filter and previous errors are augmented by the gross interest rate $1+r$ whilst the consumer remains inattentive. Because e_t depends only on the signals, consumption and attention choices, it is observable given the latent information $\mathcal{I}_t$. Moreover, Appendix A.4 indicates that e_t is a sufficient statistics to apprehend the consumer's disutility from being inattentive at time t. As a result, the difficult task of characterizing the optimal attention policy in problem (3) collapses to the much simpler task of computing the solution to the following discrete choice optimal control problem with perfect state observation e_t.

Lemma 4 (Attention problem)**.** *The optimal attention strategy of the consumer is the solution to the following Bellman equation*

$$J_t(e_t) = \min_{\tau_t\in\{0,1\}} (1-\tau_t)L_t^2(1+\beta p_{t+1})e_t^2 + \tau_t\lambda + \beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t] \tag{9}$$

$$\text{s.t.} \quad e_{t+1} = (1-\tau_t)(1+r)e_t + K\Big(z_{t+1} - (1+r)E[s_t|\mathcal{I}_t] + c_t - \bar{c}\Big)$$

Proof. See Appendix A.4 . □

The loss function in equation (9) represents the expected intertemporal utility costs associated to the attention choice once we account for Lemmas 1-3. When the consumer is inattentive ($\tau_t = 0$), her consumption choice is suboptimal given $\mathcal{I}_t$. In terms of intertemporal utility, this translates into an instantaneous misoptimization cost $L_t^2(1+\beta p_{t+1})e_t^2$ and a discounted future misoptimization cost $\beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t]$. When the consumer is attentive today ($\tau_t = 1$), she observes the perceived forecast error e_t (Lemma 3) and adjusts her consumption plan accordingly

(Lemma 1). Therefore, the consumer does not suffer from the aforementioned misoptimization costs but must pay a utility cost λ to be attentive.

The problem in Lemma 4 is standard and can be solved using a DP algorithm. More specifically, the following Proposition holds.

Proposition 1 (Attention strategy)**.** *The optimal attention policy $g_t(e_t)$ is symmetric and such that $g_t(e_t) = 1 \iff |e_t| \geq \pi_t$ and 0 otherwise. The threshold $\pi_t \in \mathbb{R}^+$ follows from Lemma 4 and solves $\forall t \in \{0, \ldots, T-1\}$*

$$\lambda + \beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t, e_t = 0] \quad = \quad L_t^2(1 + \beta p_{t+1})\pi_t^2 + \beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t, e_t = \pi_t] \tag{10}$$

Proof. See Appendix A.4. □

There is therefore a symetric inattention region, such that $|e_t| < \pi_t$, where the consumer disregards new information and adopts a wait-and-see consumption strategy. Sporadically, the absolute value of the perceived forecast error gets larger than the threshold π_t. The occurence of this event triggers the consumer's attention.

2.4 Stationary policies

As a final step in characterizing the model solution, I consider the infinite horizon limit of problem (3). Appendix A.5 demonstrates that when the horizon T is infinite, problem (3) converges to stationary policies $f(.)$ and $g(.)$. These stationary policies are reported in Proposition 2.

Proposition 2. *When the horizon is infinite, the policy functions converges to stationary policies $f(.)$ and $g(.)$. Consequently, and assuming it exists, the consumption function is*

$$c_t = \frac{\beta(1+r)^2 - 1}{\beta(1+r)} E[s_t|\overline{\mathcal{I}}_t] + \bar{c} \tag{11}$$

and the consumer updates the information set $\overline{\mathcal{I}}_t$, that is $\tau_t = 1$, whenever $|e_t| \geq \pi$ where $e_t \equiv E[s_t|\mathcal{I}_t] - E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0]$ and

$$\pi \quad = \quad \frac{\sqrt{\beta(1+r)\Big(\lambda + \beta(E[J(e_{t+1})|\mathcal{I}_t, e_t = 0] - E[J(e_{t+1})|\mathcal{I}_t, e_t = \pi])}}{\beta(1+r)^2 - 1} \tag{12}$$

$J(.)$ is the functional fixed-point solution to the infinite horizon reformulation of the Bellman equation (9).

Proof. See Appendix A.5. □

The stationary consumption policy (11) is standard and we retrieve the well-known result that the consumption path is constant over time when $\beta^{-1} = (1 + r)$. Further, the inattention region becomes time-independent as well when the horizon tends to infinity.

Unless stated otherwise, I consider the infinite horizon formulation in the rest of the paper to avoid the unnecessary burden of indexing each variable with a time t index. Extending the results to the finite horizon formulation can nevertheless be easily done using Proposition 1.

3 Consumer's inattention

Following from Proposition 1, there exists an inattention region that depends on the discrepancy between the consumer's information and the latent information stemming from the continuous signals. This section further analyses how this inattention region affects the joint dynamics of the latent perceived forecast error and consumers' attention. It then derives implications for the distribution of inattention lengths.

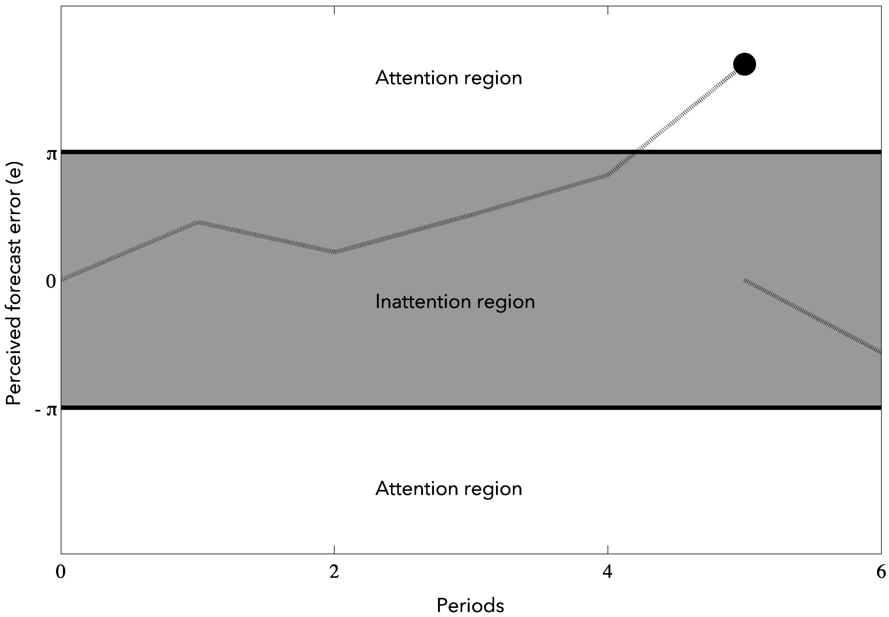
3.1 Attention dynamics

The consumer's attention dynamics is driven by the latent perceived forecast error e_t whose law of motion is given in equation (8). Further noticing that $z_{t+1} - E[s_{t+1}|\mathcal{I}_t]$ is the latent Kalman filter innovation, this law of motion equivalently writes $e_{t+1} = (1 - \tau_t)(1 + r)e_t + \omega_{t+1}$ where ω_{t+1} is a Gaussian white noise with variance $\sigma_\omega^2 = K^2(p_+ + \sigma_\eta^2)$. Consequently, the perceived forecast error follows an AR(1) process with a reseting at 0 when the consumer is attentive ($\tau_t = 1$).

Figure 1 illustrates the joint dynamics of the perceived forecast error and attention. Starting from a period 0 when the consumer was attentive, the perceived forecast error smoothly incorporates the continuous information arising from the Gaussian signals. This evolution is represented by the grey line on the graph. As long as the latent forecast error remains in the inattention region, the consumer does not observe it. However, when it exceeds the lower $(-\pi)$ or upper (π) threshold, the consumer becomes attentive and observes e_t. In the illustration from Figure 1, this event occurs at the fifth period and is pictured with the black dot in the upper attention region. Because the consumer observes e_t when being attentive, she catches up with the latent information and the dynamics of the perceived forecast error restarts from zero.

The attention dynamics depicted in Figure 1 is coherent with the observation that consumers sometimes rationally prefer to ignore new information. Using a controlled information experiment, Armantier et al. (2016) investigates how US consumers' inflation expectation are

Figure 1: Inattention region



NOTE: An illustration of the joint dynamics of the latent perceived forecast error (e_t) and attention. The grey line represents the evolution of the latent perceived forecast error over time. The latter is unknown to the consumer. The only information available to her corresponds to black lines (i.e. the threshold π) and the black dot in the upper attention region. The threshold π follows from the infinite horizon problem whose solution is time-invariant and reported in proposition 2.

affected by the provision of novel information. They show that while the provision of information – about the average inflation forecast in the Survey of Professional Forecasters – increases the probability that a consumer is attentive to this information, the probability to remain inattentive to it remains large (42% in their study). Further, they find a nonlinear relation between the perception gap – namely the difference between a consumer's prior and the information she is provided with (i.e. the latent Kalman filter innovation in the framework of this paper) – and consumers' average revision, thus suggesting that consumers are more likely to be attentive to the information they are provided with when it contradicts their prior. Abe and Ueno (2015) run a similar information experiment on Japanese consumers and reach similar conclusions. Moreover, they provide direct evidence of an inattention region with respect to the perception gap.

3.2 Inattention lengths

How long will a consumer remain inattentive? A consumer's inattention duration is stochastic. Answering this question therefore requires to derive the distribution of inattention lengths.

This is the purpose of this section. It shows that this distribution is the solution to a first passage problem. It then provides a method to characterize this distribution and an easily implementable approximation procedure.

The attention dynamics discussed in Section 3.1 may be apprehended as resulting from a first time passage problem (with reseting).[11] That is, how long will it take for the latent perceived forecast error e_t to reach one of the attention regions? Formally, let $l_t \equiv \sup\{i : \tau_i = 1, i \leq t\}$ be the most recent period when the consumer was attentive and the first passage time be defined as $d \equiv \inf\{i : \tau_{l_{t-1}+i} = 1, i \in \mathbb{N}\}$. The associated probability density function is thus

$$q(k) \equiv P(d = k) = P(\tau_{l_t+k} = 1 | \cap_{i=1}^{k-1} \tau_{l_t+i} = 0) \qquad \forall k \in \mathbb{N} \tag{13}$$

where $q(k)$ is the probability that a consumer remains inattentive for k consecutive periods. Following from Jaskowski and van Dijk (2016), a first passage time always exists here as $P(d = \infty) = 0$ at the limit. Similarly, a finite average inattention length $\bar{d}_t \equiv \sum_{i=1}^{T-1-t} i q_{i,t+i}$ exists as well.

It is well-known that directly computing the probabilities $q(k)$ is difficult. Therefore, I use the relation between these probabilities and the hazard rates, denoted $\Lambda(k)$, which are easier to compute. By definition, we have that

$$\Lambda(k) \equiv 1 - \int_{\Xi} f(e|k) de \qquad \forall k \in \mathbb{N} \tag{14}$$

where $\Xi \equiv [-\pi, \pi]$ and $f(e|k)$ is the distribution of the latent perceived forecast error e_t given that the consumer was inattentive for k consecutive periods. Equation (14) thus states that the hazard rate $\Lambda(k)$ is equal to the probability that the latent perceived forecast does not belong to the inattention region after k periods of inattention.

In Section 3.1 we have seen that the latent perceived forecast error follows an AR(1) process with a reseting at 0 when the consumer is attentive. Therefore, we have from Bayes law

$$f(e|k) \propto \int_{\Xi} f_{ar}(e|\bar{e}) f(\bar{e}|k-1) d\bar{e} \qquad \forall k \in \mathbb{N} \tag{15}$$

where $f_{ar}(e|\bar{e}) = \frac{1}{\sigma_\omega}\phi\left(\frac{e-(1+r)\bar{e}}{\sigma_\omega}\right)$ and the initial condition $f(e|0) = \delta(e)$ with $\delta(.)$ the dirac

[11] There are, at least, two complementary views to analyze the stochasticity of the attention behavior. One may either consider tracking the evolution of e_t and analyze an update as a situation such that $|e_t| \geq \pi$. In this case, it is best to apprehend the updating as resulting from a first time passage problem (in discrete time) with reseting. On the other hand, one may only value the time dimension without regard for the specifics of the dynamics of e_t. In this case, one may use tools from survival analysis. In the rest of the paper, I will rely on tools from both approaches depending on which is the most convenient. See Aalen et al. (2001) for a discussion.

distribution. Equation (15) illustrates how the uncertainty from not using the extra information acquired through the latent information set $\mathcal{I}_t$ evolves whilst the consumer is inattentive. On the one hand, if the consumer was not attentive at the previous period, she did not observe e_{t-1}. Consequently, she did not adjust her consumption path accordingly and the uncertainty surrounding her current permanent income increases mechanically by a factor indexed on the gross interest rate $(1+r)$. This mechanical increase in uncertainty is captured through $f_{ar}(e|\bar{e})$ in equation (15). On the other hand, she knows that being inattentive at the previous period implies that $\tau_{t-1} = g(e_{t-1})$ was equal to zero. Since, at the optimum, she chooses a triggering law such that $g(e_t) = 1 \iff |e_t| \geq \pi$ and zero otherwise, she knows that e_{t-1} belonged to Ξ. Hence, this negative information leads to truncate the integration of the distribution $f(\bar{e}|k-1)d\bar{e}$ in equation (15).

In order to provide quantitative predictions, it is necessary to compute the distribution $f(e|k)$ from equation (15). The latter distribution is not standard and explicitly iterating on equation (15) may lead to large numerical errors (Shi et al., 2016). I therefore rely on the approximation procedure presented in Lemma 5 which provides closed-form approximations. This approximation relies on a truncation of histories, a procedure which is well-suited for realistic calibrations of the problem under consideration. Indeed, the average inattentiveness length being generally of a few periods, the share of consumers who will encounter a long duration without being attentive should be small. Therefore, a good approximation method here should be close to exact for small k. As is highlighted in Lemma 5, the proposed method is exact when k is equal to one or two periods.

Lemma 5. *For $k = 1$, we have*

$$f(e|1) = \frac{1}{\sigma_\omega}\phi\Big(\frac{e}{\sigma_\omega}\Big)$$

and for $k = 2$,

$$f(e|2) \propto \phi\Bigg(\frac{e}{\sqrt{1+(1+r)^2}\sigma_\omega}\Bigg)\Bigg[\Phi\Bigg(\frac{\pi - \frac{(1+r)}{1+(1+r)^2}e}{\frac{\sigma_\omega}{\sqrt{1+(1+r)^2}}}\Bigg) - \Phi\Bigg(-\frac{\pi + \frac{(1+r)}{1+(1+r)^2}e}{\frac{\sigma_\omega}{\sqrt{1+(1+r)^2}}}\Bigg)\Bigg]$$

For higher $k \in \{3, 4, \ldots, \infty\}$, the distribution $f(e|k)$ is approximated by truncating the histories and we have

$$f^{\text{app}}(e|k) \propto \phi\Bigg(\frac{e}{\sqrt{z(k)}\sigma_\omega}\Bigg)\Bigg[\Phi\Bigg(\frac{\pi - \frac{(1+r)u(k)e}{z(k)}}{\sqrt{\frac{u(k)}{z(k)}}\sigma_\omega}\Bigg) - \Phi\Bigg(-\frac{\pi + \frac{(1+r)u(k)e}{z(k)}}{\sqrt{\frac{u(k)}{z(k)}}\sigma_\omega}\Bigg)\Bigg]$$

where $z(k) = \sum_{i=0}^{k-1}(1+r)^{2i}$ and $u(k) = \sum_{i=0}^{k-2}(1+r)^{2i}$.

Proof. See Appendix A.9. □

In conclusion, the distribution of inattention lengths is driven by three parameters: the interest rate r which captures the propagation of past errors over time, the triggering threshold π which characterizes the shape of the inattention region and the variance σ_ω^2 reproducing the volatility of the perceived forecast error.

4 Application to ARMA Income

This section calibrates the model parameters with an ARMA(1,1) income change process and derives quantitative implications for optimal inattentiveness.

4.1 Income process and calibration

The setup introduced in Section 2 requires that the income process follows from a multivariate linear state space model with Gaussian white innovations.[12] Following Friedman (1957), and more recently Reis (2006a) and Luo (2008), I assume that income is the sum of two independent components y_t^P and y_t^T. The first component is the permanent part of income and follows a random walk with variance σ_P^2. It captures permanent variations in income that may arise for instance from changes in employment status, experience, education or severe health shocks. The second component is transitory income and follows an AR(1) with parameter ρ and variance σ_T^2. Shocks to transitory income have a temporary effect on income and the larger ρ is the less persistent their effects are. These transitory shocks may represent for instance fluctuations in overtime labor supply, bonuses, lottery prizes and bequests. MaCurdy (1982) finds that such income process fits the US data well.

Following the methodology presented in Section 2, I reformulate this income process in terms of permanent income s_t. Shocks to permanent income are thus equal to

$$\zeta_t \quad = \quad \frac{1+r}{r}\varepsilon_t^P + \frac{1+r}{1+r-\rho}\varepsilon_t^T \tag{16}$$

where ε_t^P is the shock to the permanent part of income and ε_t^P the shock to the transitory part. In the following, I directly focus on the impact of shocks to permanent income. It allows me to characterize the impact of income shocks independently of their type. Equation (16) nevertheless permits to retrieve the impact of each shock separately.

[12] It is worth mentioning that these linear state space models are unable to capture important nonlinearities in the income process that have recently been identified (Meghir and Pistaferri, 2011; Arellano et al., 2017). Examining how these nonlinearities interact with consumers' inattention and, ultimately, consumption dynamics is left for future work.

The income process is calibrated following Pischke (1995) and such that $\sigma_P = \$45$, $\bar{y} = \$6,926$, $\rho = 0.487$, $\sigma_T = \$1,962$ and $r = 0.015$. The time period is a quarter and the observational unit a household. The discount rate is $\beta = 0.99$. Regarding the updating behavior, estimates at the macro-level indicate that individuals update once a year on average (Carroll, 2003; Mankiw et al., 2003; Reis, 2006a). I set the attention cost λ accordingly.

The remaining parameter σ_ϑ^2 stands for the signal informativeness. Everything else being equal, σ_ϑ^2 determines the latent Kalman filter gain from Lemma 2 and hence the rate at which income shocks are incorporated. I therefore calibrate σ_ϑ^2 to match the impulse-response function of US consumption estimated in Reis (2006a). He finds that about 40% of the consumption response to an income shock arises on impact. Given the steady state dynamics of the model, this implies that the latent Kalman filter gain is equal to 55%.

4.2 Optimal inattentiveness

Table 1 reports the threshold π normalized by the permanent income standard deviation. At the benchmark calibration, households update whenever their perceived forecast error e_t is larger than 1.40 σ_ζ. Using the approximation procedure from Lemma 5 it implies that consumers update their expectations once a year on average.

Table 1: Optimal inattentiveness

	Benchmark	Impact of a 5% decrease						
		r	β	ρ	σ_T	σ_P	λ	σ_ϑ
$\bar{\pi}$	1.40	1.45	0.82	1.43	1.43	1.42	1.37	1.40
$\bar{d}$	4.00	4.19	2.25	4.09	4.10	4.07	3.90	4.01

NOTE: Optimal normalized threshold $\bar{\pi} = \pi/\sigma_\zeta$ and implied average duration between updates $\bar{d}$ in quarters. The first column is for the benchmark calibration. Subsequent columns evaluate the impact of decreasing one of the parameters by 10% while keeping others constant.

In order to asses the sensitivity of consumers' optimal inattentiveness to the model parameters, Table 1 also displays the implied change in the normalized threshold and average duration between updates when one parameter decreases by 5% while others remain at the benchmark calibration. The information provided in Table 1 could thus be used to compute the updating threshold and average duration elasticities with respect to each of these parameters.[13] The updating threshold and average duration are decreasing in the interest rate r, the persistence of transitory shocks ρ, the standard deviation of permanent and transitory shocks (resp. σ_P and

[13] For example, the elasticity of the average duration with respect to the interest rate r is $-\frac{4.19-4}{0.05\times4} = -0.95$. These elasticities are only valid to locally asses the impact of a five percent decrease in the parameter values.

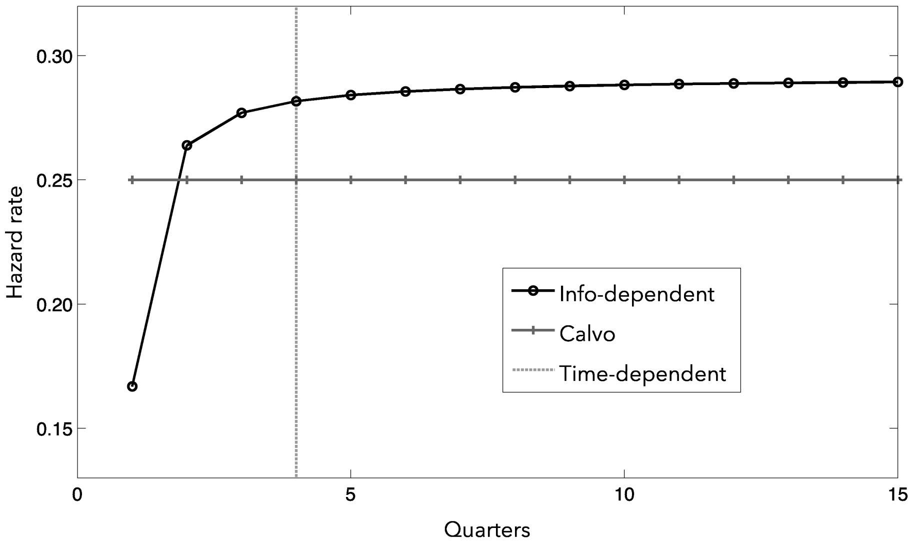
σ_T). These results are not surprising as an increase in any of the latter parameters ultimately rises the ex ante standard deviation of permanent income, thus making the consumer willing to be relatively more attentive to changes in permanent income. The updating threshold and average duration increase with the discount rate β. This is because an individual smoothes consumption less when she values tomorrow less. Therefore, the share of permanent income that she consumes (i.e. the variable L in Lemma 1) increases and so does the instantaneous cost of misoptimization. Trivially, the threshold and average inattentiveness length increase when the attention cost λ increases. Finally, the attention behavior is relatively unaffected by the signal informativeness σ_ϑ. This is because, as we saw, the consumer relies on a Kalman filter to smoothly incorporate new information from the noisy signals. Consequently, when these signals are noisier, the consumer's optimal strategy is essentially to adjust her estimator with respect to $\mathcal{I}_t$ from Lemma 2 in order to smooth these signals even more.

Figure 2 reports the hazard rates for the benchmark calibration. As can be seen, they increase with the inattention length. In particular, the probability that a consumer updates after only one quarter is small, less than 17%. These probabilities, conditionally on not updating in between, then smoothly converge to a value of 29% as time passes.

The literature offers different models of sticky expectations that lead to incompatible predictions regarding agents' inattention. The first generation of sticky information models (e.g. Mankiw and Reis (2002), Carroll (2003) and recently Carroll et al. (2020)) assumes that agents have a constant probability to update their expectations, independently of their attention history. I label these models as Calvo sticky expectations models in analogy to the Calvo price setting model (Calvo, 1983). As is reported in Figure 2 these models imply that the hazard rates are constant and equal to 25% when the average inattention length is of a year. Following these initiatives, Reis (2006a) develops a microfounded model of sticky expectations and reach the opposite conclusion that the updating behavior depends exclusively on the attention history. That is, assuming away any form of ex ante heterogeneity, each consumer should update her expectations each year (for the average inattention length to be of a year). In Reis' [2006a] model the attention strategy is therefore purely time-dependent and the associated hazard rates – reported in Figure 2 as well – are systematically nil excepted at the fourth quarter where it is equal to 1. The model considered in this paper follows Reis' [2006a] attempt to provide microfoundations to sticky expectations. By allowing for more general attention strategies, it generates an information-dependent attention[14] with increasing hazard rates. In this regard, it

[14]In analogy to the pricing literature, one may be willing to label the present model as one of state-dependent attention. I nevertheless believe that this terminology might be misleading as the state of interest to the consumer here is permanent income s_t. However, when the information source is imperfect, this state is

Figure 2: Inattention in sticky expectation models



NOTE: The figure reports the hazard rates for different models of sticky expectations. For the information-dependent model considered in this paper, the distributions is computed for the benchmark calibration and relies on the approximation procedure from Lemma 5. The Calvo model assumes that there is a constant probability to update at each period. It corresponds to the sticky expectation model in e.g. Mankiw and Reis (2002), Carroll (2003) and Carroll et al. (2020). The time-dependent model assumes that consumers update on a purely time dependent basis. It corresponds to the sticky expectation model in e.g. Reis (2006a). For the latter, the vertical line indicates that the hazard rate jumps from zero to one, and return back to zero afterward. All models are calibrated so that the average inattention length is equal to a year.

is more flexible as it allows the probability that a consumer becomes attentive to evolve with her inattention history but does not impose that inattention histories are the sole determinant of attention.

4.3 Welfare cost

Appendix A.7 decomposes the overall welfare cost from costly information processing. It may be apprehended as the sum of three independent terms: the utility cost from paying λ at each update (updating cost), the misoptimization cost from being inattentive to signals (latent information cost), and the misoptimization cost from observing noisy signals instead of perfect information (noisy information cost).

Table 2 reports these welfare costs for different values of the coefficient of relative risk aversion (CRRA) under the benchmark calibration. The costs induced by costly information processing are small. When the CRRA is equal to one, the overall welfare cost represents 0.04%

never observable and the consumer bases her attention behavior according to the information content of the information she will observe (i.e. the latent perceived forecast error e_t).

Table 2: Welfare Cost

Coeff. of relative risk aversion	1	2	4	10
Bliss point $\bar{c}$	$13,852	$10,389	$8,658	$7,619
Overall welfare cost (% consumption)	0.04%	0.07%	0.12%	0.29%
Welfare decomposition ($/quarter)	1.28	2.57	5.13	12.83
Updating cost	*0.70*	*1.40*	*2.81*	*7.02*
Latent information cost	*0.41*	*0.81*	*1.62*	*4.05*
Noisy information cost	*0.18*	*0.35*	*0.71*	*1.76*
Consumption change to update	2.92%	2.32%	2.10%	1.99%

NOTE: The welfare cost refers to misoptimization cost induced by costly information processing (equation **??**). The coefficient of relative risk aversion is equal to $\mu/(\bar{c}-\mu)$. The consumption change to update measures the threshold change in perceived consumption that will prompt the consumer to internalize new information at period 0. These results were obtained under the benchmark calibration for the infinite horizon problem.

of a household consumption (at period 0). That is, less than $1.3 per quarter. These costs remain negligible even when one considers extremely risk averse consumers. When the CRRA is 10, the welfare cost increases to only 0.29% of consumption. Decomposing the welfare cost, we find that more than half of it is attributable to the fixed utility cost λ that the consumer pays when attentive. When the CRRA is equal to one (resp. 10), the monetary equivalent from paying λ when being attentive is ¢70 (resp. $7.02) per quarter. Given that she will update once a year on average, the monetary equivalent for λ is $2.1 (resp. $21.06).

Consumption models with fixed costs of adjustment (e.g. Caballero (1995) and Chetty and Szeidl (2016)) predict that we should observe (i) large consumption jumps and (ii) long lasting periods when a household consumption remains constant.[15] The model presented in this paper does not share these unappealing predictions. First, the predicted consumption jumps are modest. Given the model calibration, the consumption change which prompts the consumer to update ($L \times \pi$) is relatively small, equal to $133. When normalized by period 0 consumption, the last row in Table 2 reports an order of magnitude of a few percentage points. In contrast Caballero (1995) estimates that in order to capture the stickiness of US aggregate consumption data, the implied jump in an adjustment consumption model would be of almost 6% – namely twice as large than what we find here. Second, and as is further discussed in the subsequent section, consumption is not constant here whilst the consumer is inattentive. Intuitively, this is because the consumer is inattentive around a consumption path and not a consumption level.

[15]See Reis (2006a) and Carroll et al. (2020) for a critical discussion.

5 Households' consumption dynamics

This section analyses the consumption change following a shock to permanent income at the household level. It shows that while consumption changes are partially predictable, they are not serially correlated. Moreover, the expected marginal propensity to consume out of an income shock depends on the perceived forecast error and the permanent income shock.

5.1 Consumption changes

At the micro level, consumption changes are conditional on the updating behavior. When inattentive, the consumer follows a committed consumption path. As such, the consumption change solely reflects the trend in this consumption path and consumption growth is constant. More specifically, we have from equation (1) that

$$\Delta c_{t+1}|(\tau_{t+1} = 0) = (r - L)c_t \tag{17}$$

As a consequence, consumption growth at non-updating periods is predetermined and orthogonal to permanent income shocks and information noises. However, at updating periods the consumer updates her information set and the consumption change

$$\Delta c_{t+1}|(\tau_{t+1} = 1) = Le_{t+1} + \Delta c_{t+1}|(\tau_{t+1} = 0) \tag{18}$$

is a Borel-measurable function with associated σ-statistics $\mathcal{I}_{t+1}$. As such, the change in consumption at updating periods depends on the complete history $\{\{\zeta_i\}_{i=1}^{t+1}, \{\vartheta_i\}_{i=1}^{t+1}\}$ and is therefore partially forecastable using past information about income shocks. In comparison, the sticky expectation models of Carroll (2003) and Reis (2006a) predict that, in an otherwise similar setup, a household consumption growth would be unpredictable using information prior to her last update at time l_t. Furthermore, the following proposition about serial correlation holds

Proposition 3. *Consumption growth is not serially correlated at the household level.*

Proof. Equations (8), (17) and (18) together imply that Δc_t is orthogonal to e_t when $\tau_t = 0$ and that e_{t+1} is orthogonal to e_t when $\tau_t = 1$ so that e_{t+1} is also independent from Δc_t in that case. □

Many studies have tested wether household level consumption growth is serially correlated. The recent meta-analysis of Havranek et al. (2017) considers 190 estimates from these studies on micro data and reports a median (resp. mean) estimate of 0.0 (resp. 0.1). These findings are in line with Proposition 3. They are however hardly reconcilable with the use of consumption habits as a mechanism to generate smoothness in aggregate consumption dynamics: "If habits are a true structural characteristic of people's utility functions, we should see their effects in microeconomic data as well as macroeconomic aggregates. But empirical studies using household-level data strongly reject the existence of habits of the magnitude necessary to explain aggregate consumption dynamics." (Carroll et al., 2020)

5.2 Asymmetry and magnitude

Now, consider a translation in the distribution of permanent income shocks at time $t+1$ such that $\zeta_{t+1} \sim \mathcal{N}(\bar{\zeta}, \sigma_\zeta^2)$. Let $\text{MPC}(e, \bar{\zeta}) \equiv \partial E[c_{t+1}|e = e_t]/\partial\bar{\zeta}$ be the expected instantaneous marginal propensity to consume for a consumer whose latent perceived forecast error is e_t. Then,

$$\text{MPC}(e_t, \bar{\zeta}) = L(1+r)\left[\frac{\partial q_1(e_t + \nu\bar{\zeta})}{\partial\bar{\zeta}}(e_t + \nu\bar{\zeta}) + \nu q_1(e_t + \nu\bar{\zeta})\right] \tag{19}$$

where $\nu = K/(1+r)$ and $q_1(x) = 1 + \Phi(-(\pi + (1+r)x)/\sigma_\omega) - \Phi((\pi - (1+r)x)/\sigma_\omega)$ is the probability to update in one period given an initial latent forecast error x. Equation (19) underlines the two margins affecting the expected consumption response to an income change. On the one hand, and taking the probability to update as given, the consumer will internalize a proportion K of the shock when updating. Accordingly, the expected marginal propensity to consume increases by LK times the probability of an update. On the other hand, the shock $\bar{\zeta}$ also affects the consumer's probability to update. The latter probability is $q_1(e_t + \nu\bar{\zeta})$ where the term ν is used to account for the fact that the consumer will only internalize a fraction K of the shock on average and to express the impact of the shock on the probability from a change in the initial condition at period t.

Equation (19) further reveals that the expected marginal propensity to consume is both history- and shock-dependent. To highlight the mechanisms behind these two dependences, first realize that the function $q_1(x)$ is symmetric around its minimum at 0, attains its maximum 1 at $\pm\infty$ and is monotonically decreasing on $[-\infty, 0]$ and monotonically increasing on $[0, \infty]$. As a consequence, the marginal change in the probability to update at the next period given a history leading to e_t and a shock $\bar{\zeta}$ is positive if and only if $e_t + \nu\bar{\zeta} \geq 0$ and negative otherwise. When the latter term is positive, the consumer is more likely to internalize the impact of the income shock and its expected marginal propensity to consume out of this income news is larger. More specifically, equation (19) implies the following behavior

Proposition 4. *The one-period ahead expected marginal propensity to consume out of an income shock is such that*

- $\mathrm{MPC}(e_t, \bar{\zeta})$ *increases with respect to* $|e_t + \nu\bar{\zeta}|$.
- *Given that* e_t *belongs to* Ξ_t, *there always exists a finite and large enough income shock* $\varkappa$ *such that*

$$\mathrm{MPC}(e_t, \varkappa) > \mathrm{MPC}(e_t, 0) \quad \text{and} \quad \mathrm{MPC}(e_t, -\varkappa) > \mathrm{MPC}(e_t, 0)\ \forall e_t \in \Xi_t$$

- *Let* $\bar{\zeta} > 0$, *then*

$$\begin{aligned} \mathrm{MPC}(e_t, \bar{\zeta}) &= \mathrm{MPC}(e_t, -\bar{\zeta}) \iff e_t = 0 \\ \mathrm{MPC}(e_t, \bar{\zeta}) &> \mathrm{MPC}(e_t, -\bar{\zeta}) \iff e_t > 0 \end{aligned}$$

According to the magnitude hypothesis, the consumption response to an income shock depends on its size. Such relation is expected to hold here as is apparent from the first two bullets in Proposition 4. The literature review in Section 1 mentions studies providing evidence to support the magnitude hypothesis. A few authors postulated that rational inattention may offer an explanation since the costs from not smoothing consumption increase with the size of the shock. For example, Hsieh (2003) states that "households will not bother to change their consumption paths when the computational costs involved are large relative to the utility gains". A similar argument is made in Browning and Collado (2001): households "do not bother to adjust optimally to small income changes since the utility cost [...] is small". The consumption model developed in this paper offers a microfounded framework confirming this guess. Accordingly, the magnitude hypothesis arises as a consequence of the joint dynamics of information rigidities and consumption. Large income shocks are more likely to prompt consumers to revise their expectations and, consequently, to adjust their consumption path to account for this shock. It it is worth mentioning that this conclusion holds independently of the sign of the income shock and therefore differs from any explanation based on the presence of credit constraints or risk aversion.[16] Proposition 4 however reveals that the size of an income shock is not a sufficient statistics to apprehend the magnitude hypothesis because other perceived income shocks since the last update matter as well. Alternatively saying, the consumption response to an

[16]Following an income decline, a credit constraint does not affect the marginal propensity to consume. See Jappelli and Pistaferri (2010) for a discussion. Regarding risk aversion, Tutino (2013) shows that the consumption response to negative income shocks is higher in a framework with a CRRA utility and inattention à la Sims (2003).

income shock is both history (through e_t) and shock-dependent (through $\bar{\zeta}$). One shall therefore simultaneously account for both dependences to derive implications for the consumption response.

The first and last bullets in Proposition 4 indicate that the consumption response to an income shock is asymmetric with respect to the sign of the shock. This asymmetry results from two complementary forces, the history-dependence and the shock-dependence. The consumption model predicts that the expected marginal propensity to consume at the household level is large for positive shocks when the perceived change in permanent income that has not been internalized yet is positive (and large) and for negative shocks when the perceived change in permanent income that has not been internalized yet is negative (and large in absolute value). On the other hand, the expected MPC decreases after a positive shock when the perceived change in permanent income that has not been internalized yet is negative (and large in absolute value) and after a negative shock when the perceived change in permanent income that has not been internalized yet is positive (and large). These asymmetries, which cannot be explained by standard extensions of the permanent income model, such as credit constraints or habits, have been identified by Caballero (1995) using data on US aggregate consumption. Up to the best of my knowledge, no study has analyzed the potential asymmetric reaction to negative and positive shocks during periods of income increases and declines at the household level.

Finally, equation (19) also indicates that the expected marginal propensity to consume out of an income shock is bounded by the rate at which the estimator with respect to $\mathcal{I}_t$ incorporates new information. It is easily seen from computing $\lim_{\bar{\zeta}\mapsto\pm\infty} \text{MPC}(e_t, \bar{\zeta}) = L(1+r)K$ where $0 < K < 1$ is the Kalman gain from equation (5).

6 Aggregate consumption dynamics

This section focuses on the implications for aggregate consumption. It shows that costly information processing generates impulse response functions for consumption which depend on both the state of the economy and the size of the shock. Moreover, aggregate consumption growth is found to be highly persistent at the steady state. This persistence however depends on the endogenously time-varying share of attentive consumers.

6.1 Cross-sectional distribution

To asses the aggregate dynamics of the economy, I assume it is composed of a unit mass continuum of infinitely-lived consumers who are ex ante identical and whose problem is given by (3). For simplicity, I further assume that the discount rate is equal to $\beta = (1+r)^{-1}$ in the following. Because they have different idiosyncratic shocks, signals and updating histories, households differ in terms of their latent perceived forecast error. Analyzing the aggregate dynamics of information frictions and consumption therefore requires to keep track of the distribution of households. To this end, let $a_t(e)$ denote the cross-sectional distribution of consumers at the end of period t and before the resetting. The share of updating consumers in the economy at period t is therefore $\lambda_t \equiv \int_{e \notin \Xi} a_t(e) de$. In the absence of aggregate shocks, the cross-sectional distribution dynamics is

$$a_t(e) \propto \frac{1}{\sigma_\omega} \Bigg[\underbrace{\int_{\tilde{e} \in \Xi} \phi\left(\frac{e-(1+r)\tilde{e}}{\sigma_\omega}\right) a_{t-1}(\tilde{e}) d\tilde{e}}_{\text{Non updaters at } t-1} + \underbrace{\phi\left(\frac{e}{\sigma_\omega}\right) \int_{\tilde{e} \notin \Xi} a_{t-1}(\tilde{e}) d\tilde{e}}_{\text{Updaters at } t-1} \Bigg] \tag{20}$$

The dynamics of the cross-sectional distribution directly follows from the dynamics of the latent perceive forecast error in equation (8). Proposition 5 characterizes the steady state cross-sectional distribution in the absence of aggregate shocks.

Proposition 5. *Equation (20) admits a stationary distribution*

$$a^\star(e) = \sum_{k=1}^{\infty} \lambda^\star(k) f^\star(e|k) \tag{21}$$

where

$$\lambda^\star(k) = \lambda^\star(1) S^\star(k-1) \quad \forall k \geq 2 \tag{22}$$

and $\lambda^\star(1) = 1/\sum_{k=0}^{\infty} S^\star(k)$ *with* $S^\star(k)$ *the unconditional survival function and* $\lambda^\star(k)$ *the steady-state share of consumers whose last update was* k *periods ago.*

Proof. Appendix A.6. □

I now introduce aggregate shocks in the economy and make the following assumption.

Assumption 1. *Shocks to permanent income* ζ_t *are the sum of an aggregate shock* χ_t *common to each consumer and an independent idiosyncratic shock* $\iota_{i,t}$*. Both shocks are i.i.d. Gaussians with zero mean and respective variance* σ_χ^2 *and* σ_ι^2*.*

Under assumption 1, the consumption and triggering policies from Proposition 2 are unaffected. As a result, the only impact of aggregate shocks is to continuously and persistently translate the cross-sectional distribution of consumers. Indeed, when a shock occurs, it is only observable through the noisy information channel $\mathcal{I}_t$ and is therefore gradually perceived by consumers. Therefore, an aggregate shock does not only disturb the cross-sectional distribution when it occurs, but does so persistently. Given that consumers rely on a latent Kalman filter to incorporate new information (Lemma 2), the translation in the cross-sectional distribution at period t is given by a weighted sum of past aggregate income shocks $\{\bar{\chi}_i\}_{i=1}^{t}$

$$S_t = \sum_{s=0}^{t-1} \upsilon_s \bar{\chi}_{t-s} \tag{23}$$

where

$$\upsilon_s = K(1+r)^s(1-K)^s \tag{24}$$

is the share of the aggregate shock – augmented by its returns – that is internalized on average at period $t+s$. Accounting for this new state variable, it is possible to derive the impact of an aggregate shock on consumption dynamics for any history of aggregate shocks.

Proposition 6. *The dynamics of the economy in the presence of aggregate shocks is characterized by the following system of dynamic equations*

$$a_{t+1}(e) \propto \frac{1}{\sigma_\omega}\left[\int_{\tilde{e}\in\Xi} \phi\left(\frac{e - S_{t+1} - (1+r)\tilde{e}}{\sigma_\omega}\right) a_t(\tilde{e})d\tilde{e} + \int_{\tilde{e}\notin\Xi} \phi\left(\frac{e - S_{t+1}}{\sigma_\omega}\right) a_t(\tilde{e})d\tilde{e}\right] \tag{25}$$

$$S_{t+1} = (1-K)(1+r)S_t + K\bar{\chi}_{t+1} \tag{26}$$

along with initial conditions $a_0(e)$ *and* $S_0 = \int_{\mathbb{R}} e a_0(e)de$.

Proposition 6 indicates that S_t follows a markovian process. In general, the dynamics of S_t is stationary since r should be small in comparison to K. This state variable characterizes the business cycle in this economy: a positive S_t coincides with an expansionary cycle while a negative S_t results in a contraction.

6.2 Consumption dynamics at the steady state

To illustrate the dynamics of the economy starting from the steady-state cross-sectional distribution, I first analyze a one time only aggregate shock. Formally, consider a shift $\bar{\chi}$ in the distribution $\chi_t \sim \mathcal{N}(\bar{\chi}, \sigma_\chi^2)$ at time t. Because the economy is initially at its steady-state and there is a one time only aggregate shock, Proposition 6 implies that the impact of the aggregate shock on the cross-sectional distribution follows from iterating on

$$a_{t+s}(e)|\bar{\chi} \propto \frac{1}{\sigma_\omega}\left[\int_{\tilde{e}\in\Xi}\phi\left(\frac{e-\upsilon_s\bar{\chi}-(1+r)\tilde{e}}{\sigma_\omega}\right)a_{t-1+s}(\tilde{e})d\tilde{e}+\phi\left(\frac{e-\upsilon_s\bar{\chi}}{\sigma_\omega}\right)\int_{\tilde{e}\notin\Xi}a_{t-1+s}(\tilde{e})d\tilde{e}\right] \tag{27}$$

which has for initial condition $a_{t-1}(e) = a^\star(e)$ so that $S_{t-1} = 0$. Therefore, the impact of an aggregate shock is first to shift the cross-sectional distribution of consumers. The stationary distribution $a^\star(e)$ being symmetric, unimodal and centered around zero, an aggregate shock increases the share of agents who update and the magnitude of the average consumption change of those who would have updated in the absence of the shock.[17]

Not everyone will update (a.s.) after the aggregate shock. Hence, the cross-sectional distribution at the next period must account for the impact of the initial shock for those who have not updated (whence the term $\int_{\tilde{e}\in\Xi}\phi\left(\frac{e-(1+r)\tilde{e}}{\sigma_\omega}\right)a_{t-1+s}(\tilde{e})d\tilde{e}$ in the right-hand side of equation (27)). Moreover, because information is imperfect, the cross-section of consumers continues to be translated at each period in order to account for new pieces of information regarding the initial shock after it occurred. This new information depends on the rate at which agents rely on their signals z_t to form $E[s_t|\mathcal{I}_t]$. Consequently, the cross-sectional distribution is translated a second time at period $t+1$ by a factor $\upsilon_1\bar{\chi}$. As time passes, υ_s tends to zero and the cross-sectional distribution is not disturbed anymore. It then smoothly converges back to the stationary distribution $a^\star(e)$ as consumers update to account for their idiosyncratic shocks.

Aggregating equations (17) and (18) over consumers when $\beta = (1+r)^{-1}$, the impulse response function (IRF) following an aggregate shock is thus

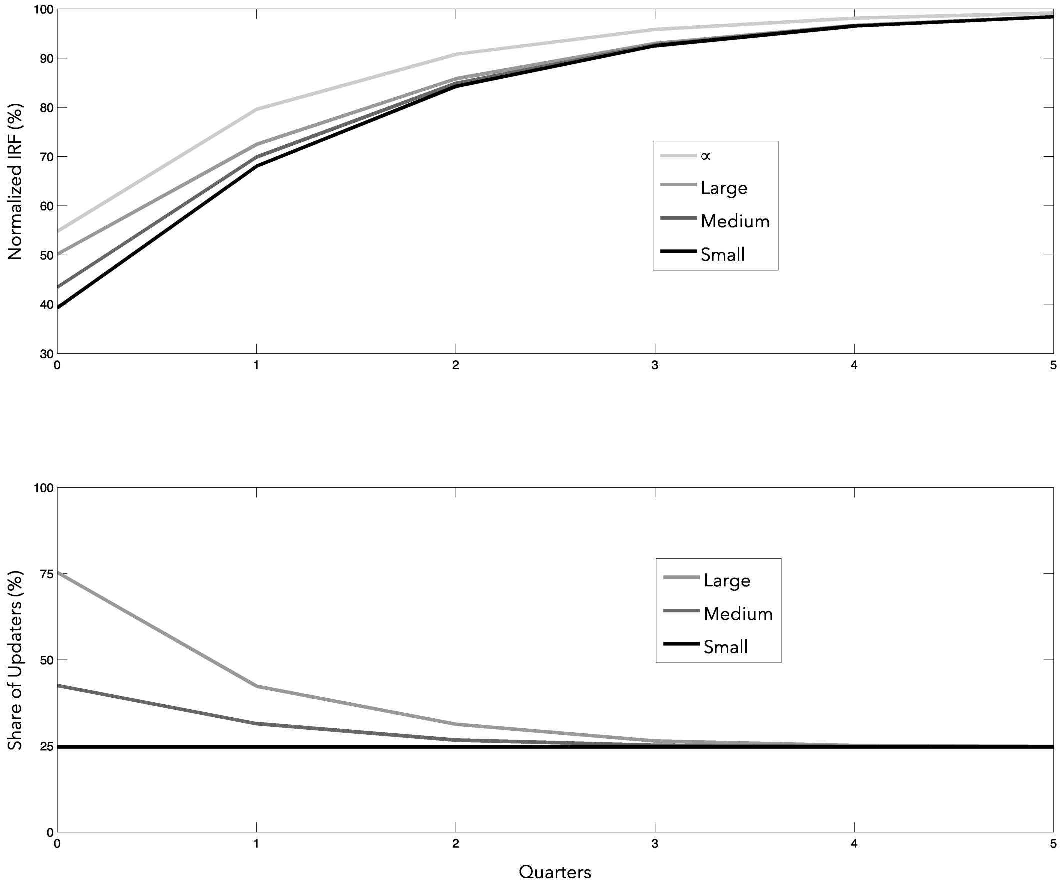
$$\Delta C_{t+s}|_{\chi=\bar{\chi}} - \Delta C_{t+s}|_{\chi=0} = L\int_{e\notin\Xi} e\ \left(a_{t+s}(e)|\bar{\chi} - a^\star(e)\right)\, de \tag{28}$$

Equation (28) highlights the dependence between consumption and information rigidities dynamics. Indeed, $(a_{t+s}(e)|\bar{\chi} - a^\star(e))$ stands for the change in the perceived forecast error distribution stemming from the aggregate shock. It thus captures the change in the share of updating consumers – as well as the change in their average e.

Figure 3 plots the impulse response function for different aggregate shocks. As is apparent from the top panel, aggregate information rigidities and consumption dynamics are shock-dependent. When the shock size is small, it barely affects the updating behavior which remains at its steady state dynamics. The shock is then smoothly accounted for by consumers as they update to internalize the impact of idiosyncratic shocks. However, when the shock is large, the

[17] In the state-dependent pricing literature, these responses are respectively labeled the extensive and intensive margins (Caballero and Engel (2007)).

Figure 3: IRF at the steady state



NOTE: The figure reports the impulse response function of aggregate consumption (top panel) and shares of updates (middle panel) for different values of aggregate income shocks. The small shock is a one dollar shock, the medium is a one standard deviation σ_ζ shock and the large shock a two standard deviations shock. The bottom panel plots the evolution of the cross-sectional distribution of consumers following the medium shock. The large plain line is the steady-state distribution.

share of agents updating jumps on impact and persistently remains above its steady state level. As a consequence, the short run response of aggregate consumption is much sharper following a large shock. Irrespectively of the size of the initial shock, the normalized IRFs then converge as consumers update to internalize the impact of idiosyncratic shocks.[18] Figure 3 also reports the dynamics at the limit when $\bar{\chi} \mapsto \infty$. After such shock, the share of updaters jumps to 100%. The instantaneous normalized response is thus equal[19] to $1-(1-K)(1+r)$. At the

[18]Note that the long run impact of the shocks are not identical because income news that haven't been processed grow at a constant rate r. See Luo (2008) for a related discussion. However, the normalized IRF being defined as the response at time $t+s$ divided by the long run response, it must converge to one.

[19]When $(1-K)(1+r) < 1$. This value is computed as follows. The long run response is $K/(1-(1-K)(1+r))$,

second period, the new piece of information prompts each consumer to update again a.s. and the normalized response is equal to $(v_0 + v_1)(1 - (1 - K)(1 + r))/K$.

6.3 Consumption dynamics during booms and busts

The previous section studies the consumption dynamics assuming the economy was initially at its steady state and concludes that it depends only on the size of the aggregate shock. However, and similarly to what we found at the household level, aggregate consumption dynamics is both history-dependent and shock-dependent. The present section thus considers the role of previous aggregate shocks for consumption dynamics. Importantly, it shows that the dynamics of aggregate consumption depends on the state of the economy when the shock occurs.

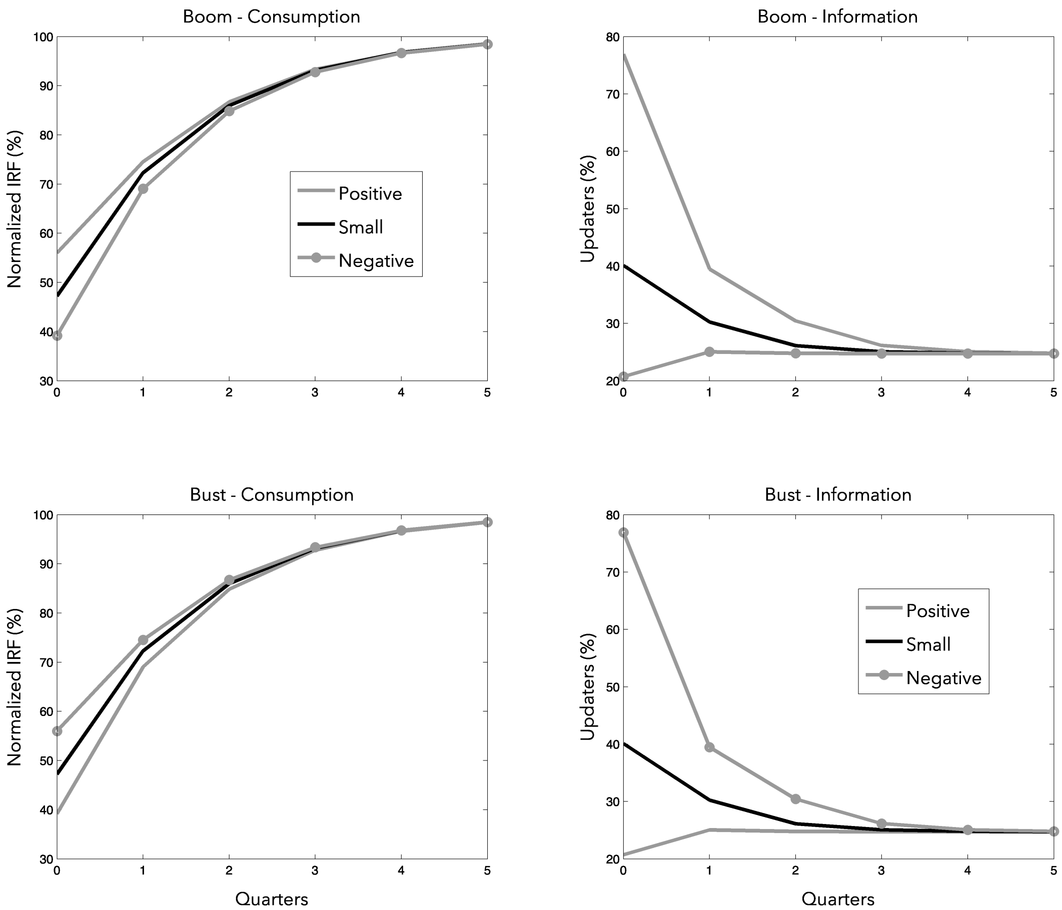
The situation studied in section 6.2 was such that $a_0(e)$ was equal to $a^\star(e)$. The mean of the latter being nil, we had $S_0 = 0$. It was thus a special case of proposition 6. The impulse response function follows from equation (28) where the counterfactual distribution is now $a_{t+s}(e)|\bar{\chi}_t = 0$, that is the cross-sectional distribution in the absence of aggregate shock at time t. Figure 4 reports the implied impulse response functions when the economy is initially either in a booming period ($S_0 > 0$) or in a busting period ($S_0 < 0$). The implied dynamics are fundamentally different. During a boom (resp. bust), a positive (resp. negative) aggregate shock increases the share of updates and consumers rapidly revise their consumption plan accordingly. In this case, the consumption response is brisker when the shock size is larger. However, a positive (resp. negative) shock may have an ambiguous effect on the updating behavior during recessions (resp. expansions). If the size of the shock is small relatively to $|S_0|$, then the share of updates decreases after the shock and consumption dynamics are more sluggish. However, and similarly to proposition 4, there always exists a large enough income shock such that consumers update more and consumption dynamics are abrupter. This latter observation may easily be understood by realizing that the impact of an infinitely large shock (computed in section 6.2) is independent of the initial state of the economy.

Caballero (1995) finds that US aggregate consumption dynamics display asymmetries during booms and busts as those described above: "in good times, consumers respond more promptly to positive than to negative wealth shocks, while the opposite is true in bad times". Similarly, Ocal and Osborn (2000) recently concluded that the dynamics of aggregate consumption in the UK depends on the the state of the economy and the sign of the shock.

---

and the response on impact is $v_0 = K$ as everyone update. The normalized response being the ratio of the latter over the former, it is equal to $1 - (1 - K)(1 + r)$.

Figure 4: IRF During Booms and Busts



NOTE: The figure reports the impulse response functions of aggregate consumption (left panels) and the shares of updates (right panels) following a one and two standard deviations positive permanent income shocks and a one and two standard deviations negative shocks when the economy is initially in a boom (top panels) or a bust (bottom panels).

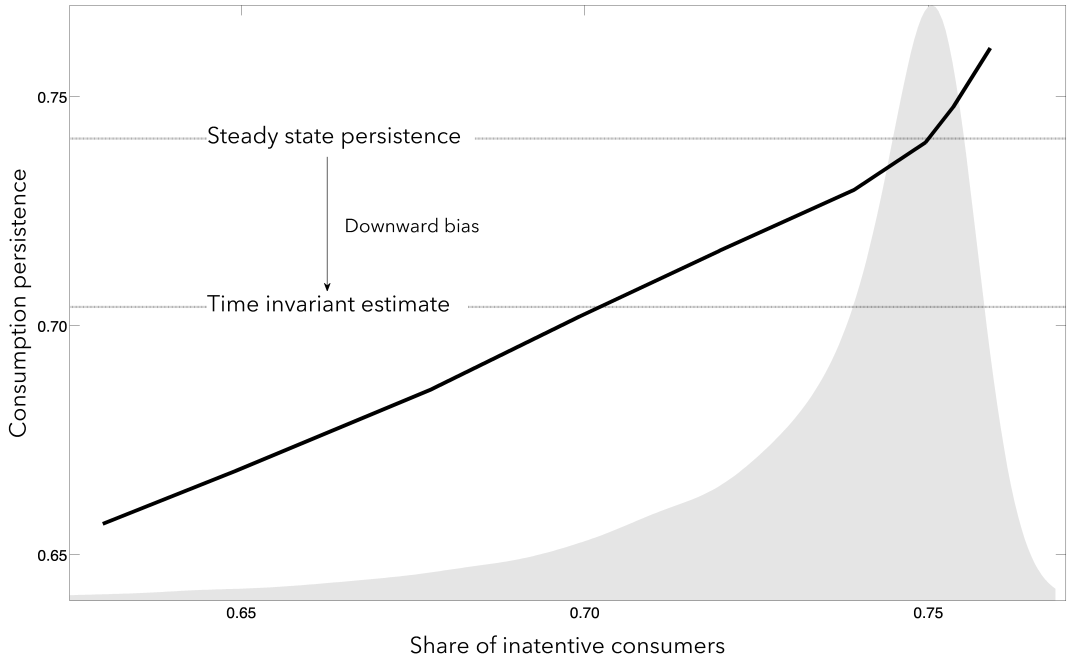
6.4 Implications for the persistence of aggregate consumption

[*Methodology to be written*]

The persistence of consumption growth depends on the endogenously time-varying share of attentive consumers. Figure 5 displays the predicted mapping between consumption persistence and consumers' information rigidities. In normal times, information rigidities are near their steady state level and consumption persistence relatively constant. However, during unusual times such as recessions, information rigidities decrease a lot and so does aggregate consumption persistence.

These findings are in line with Kumar and Jia (2019) who report systematic decreases in

Figure 5: Aggregate consumption persistence and attention



NOTE: Model predictions for the relation between consumption persistence and information frictions. Aggregate consumption change is approximatively equal to $\Delta C_t = \gamma_t \Delta C_{t-1} + \beta X_t$ where γ_t is a measure of the time-varying persistence of consumption growth and X_t a set of controls discussed in Section 6. The black line relates the evolution of the estimated time-varying persistence γ_t to the observed share of inattentive consumers from simulating the model over 200,000 periods. The grey area is the distribution of the share of inattentive consumers. The steady state persistence is the persistence when the share of attentive consumers is at its steady state level (i.e. 75%) and the time-variant estimate is obtained from an OLS estimation when the persistence is assumed to be constant. The model is calibrated at the household level (see Section 4).

consumption growth persistence during recessionary periods. An innovation of this paper is to illustrate how these drops in aggregate consumption persistence are fundamentally related to Dräger and Lamla (2012) and Coibion and Gorodnichenko (2015) findings that information rigidities decrease during these periods. Furthermore and as is depicted in Figure 5, time-invariant estimates reported in the literature will generally underestimate the steady state persistence of consumption.

7 Conclusions

This paper proposes a novel model to explain the state-dependence of information rigidities. Consumers must pay a fixed cost to observe noisy signals on the state of the economy. They face an inattention region where they temporarily ignore new signal releases and do not update their expectations. Accounting for the interaction between variations in information rigidities and consumption choices allows to reproduce non-linear features of consumption dyhnamics that have been observed in the data.

This paper stresses potentially important and novel policy implications that would be interesting to investigate. At the aggregate level, the dynamics of consumption is governed by the dynamics of information rigidities. As a consequence, households tend to pay more attention to negative income shocks than to positive income shocks during recessions. Although there is no explicit fiscal or monetary authorities in the present model, this result may hold significant consequences for stabilization policies. By leaning against the wind, these policies may reduce households' incentive to update their expectations. As a result, feeble policy interventions aiming to stabilize consumption may in fact increase the persistence of consumption growth in the aftermath of a recession, and thereby turn out to be even more destabilizing. It would therefore be interesting to enrich the model with fiscal and monetary authorities in future work. Such work would be useful in any attempt to understand the interaction between economic policies, information frictions, and macroeconomic outcomes.

References

Aalen, O. O., Gjessing, H. K., et al. (2001). Understanding the shape of the hazard rate: A process point of view (with comments and a rejoinder by the authors). *Statistical Science*, 16(1):1–22.

Abe, N. and Ueno, Y. (2015). Measuring Inflation Expectations: Consumers' Heterogeneity and Nonlinearity. Technical report, Research Center for Economic and Social Risks, Institute of Economic Research, Hitotsubashi University.

Arellano, M., Blundell, R., and Bonhomme, S. (2017). Earnings and consumption dynamics: a nonlinear panel data framework. *Econometrica*, 85(3):693–734.

Armantier, O., Nelson, S., Topa, G., Van der Klaauw, W., and Zafar, B. (2016). The price is right: Updating inflation expectations in a randomized price information experiment. *Review of Economics and Statistics*, 98(3):503–523.

Browning, M. and Collado, M. D. (2001). The response of expenditures to anticipated income changes: panel data estimates. *The American economic review*, 91(3):681–692.

Burstein, A. T. (2006). Inflation and output dynamics with state-dependent pricing decisions. *Journal of Monetary Economics*, 53(7):1235–1257.

Caballero, R. J. (1993). Durable goods: An explanation for their slow adjustment. *Journal of Political Economy*, 101(2):351–384.

Caballero, R. J. (1995). Near-rationality, heterogeneity, and aggregate consumption. *Journal of Money, Credit and Banking*, 27(1):29–48.

Caballero, R. J. and Engel, E. M. (2007). Price stickiness in ss models: New interpretations of old results. *Journal of monetary economics*, 54:100–121.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, 12(3):383–398.

Carroll, C. (2003). Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics*.

Carroll, C. D., Crawley, E., Slacalek, J., Tokuoka, K., and White, M. N. (2020). Sticky expectations and consumption dynamics. *American economic journal: macroeconomics*.

Cheremukhin, A. and Tutino, A. (2016). Information rigidities and asymmetric business cycles. *Journal of Economic Dynamics and Control*, 73:142–158.

Chetty, R. and Szeidl, A. (2016). Consumption commitments and habit formation. *Econometrica*, 84(2):855–890.

Cochrane, J. H. (1989). The sensitivity of tests of the lntertemporal allocation of consumption to near-rational alternatives. *The American Economic Review*, 79(3):319–337.

Coibion, O. and Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review*, 105(8):2644–78.

Dräger, L. and Lamla, M. J. (2012). Updating Inflation Expectations: Evidence from Micro-data. *Economics Letters*, 117(3):807–810.

Friedman, M. (1957). Introduction to" a theory of the consumption function". In *A theory of the consumption function*, pages 1–6. Princeton university press.

Gabaix, X. and Laibson, D. (2001). The 6d bias and the equity-premium puzzle. *NBER macroeconomics annual*, 16:257–312.

Gorodnichenko, Y. (2008). Endogenous Information, Menu Costs and Inflation Persistence. NBER Working Papers 14184, National Bureau of Economic Research, Inc.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of political economy*, 86(6):971–987.

Havranek, T., Rusnak, M., and Sokolova, A. (2017). Habit formation in consumption: A meta-analysis. *European Economic Review*, 95:142–167.

Hsieh, C.-T. (2003). Do consumers react to anticipated income changes? evidence from the alaska permanent fund. *The American Economic Review*, 93(1):397–405.

Jappelli, T. and Pistaferri, L. (2010). The consumption response to income changes. *Annu. Rev. Econ.*, 2(1):479–506.

Jappelli, T. and Pistaferri, L. (2017). *The economics of consumption: theory and evidence*. Oxford University Press.

Jaskowski, M. and van Dijk, D. (2016). First-passage-time in discrete time and intra-horizon risk measures. *Forthcoming*.

Khaw, M. W., Stevens, L., and Woodford, M. (2017). Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*, 91:88–103.

Kueng, L. (2018). Excess sensitivity of high-income consumers. *The Quarterly Journal of Economics*, 133(4):1693–1751.

Kumar, S. and Jia, P. (2019). Financial crisis and persistence: evidence from sticky expectations consumption growth model. *Applied Economics*, 51(17):1799–1807.

Larsen, V. H., Thorsrud, L. A., and Zhulanova, J. (2019). News-driven inflation expectations and information rigidities. Technical report, BI Norwegian Business School.

Ljungqvist, L. and Sargent, T. J. (2004). *Recursive macroeconomic theory (Second edi% tion)*. Cambridge, Mass: MIT Press.

Luo, Y. (2008). Consumption Dynamics under Information Processing Constraints. *Review of Economic Dynamics*, 11(2):366–385.

Luo, Y., Nie, J., Wang, G., and Young, E. R. (2017). Rational inattention and the dynamics of consumption and wealth in general equilibrium. *Journal of Economic Theory*, 172:55–87.

Luo, Y. and Young, E. R. (2014). Signal Extraction and Rational Inattention. *Economic Inquiry*, 52(2):811–829.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2018). Dynamic rational inattention: Analytical results. *Journal of Economic Theory*, 176:650–692.

Maćkowiak, B. and Wiederholt, M. (2015). Business Cycle Dynamics under Rational Inattention. *The Review of Economic Studies*, 82(4):1504–1532.

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of econometrics*, 18(1):83–114.

Mankiw, N. G. and Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics*, 117(4):1295–1328.

Mankiw, N. G. and Reis, R. (2006). Pervasive stickiness. *American Economic Review*, 96(2):164–169.

Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about Inflation Expectations. NBER Working Papers 9796, National Bureau of Economic Research, Inc.

Meghir, C. and Pistaferri, L. (2011). Earnings, consumption and life cycle choices. In *Handbook of labor economics*, volume 4, pages 773–854. Elsevier.

Molin, A. and Hirche, S. (2010). Structural characterization of optimal event-based controllers for linear stochastic systems. In *Proceedings of the IEEE International Conference on Decision and Control (CDC 2010)*.

Molin, A. and Hirche, S. (2012). An iterative algorithm for optimal event-triggered estimation. In *4th IFAC Conference on Analysis and Design of Hybrid Systems (ADHS'12)*.

Nimark, K. P. (2014). Man-Bites-Dog Business Cycles. *American Economic Review*, 104(8):2320–67.

Ocal, N. and Osborn, D. R. (2000). Business Cycle Non-Linearities in UK Consumption and Production. *Journal of Applied Econometrics*, 15(1):27–43.

Pischke, J.-S. (1995). Individual income, incomplete information, and aggregate consumption. *Econometrica: Journal of the Econometric Society*, pages 805–840.

Reis, R. (2006a). Inattentive Consumers. *Journal of Monetary Economics*, 53(8):1761–1800.

Reis, R. (2006b). Inattentive producers. *The Review of Economic Studies*, 73(3):793–821.

Scholnick, B. (2013). Consumption smoothing after the final mortgage payment: Testing the magnitude hypothesis. *Review of Economics and Statistics*, 95(4):1444–1449.

Shi, D., Shi, L., and Chen, T. (2016). Event-based state estimation. *Stud. in Syst. Decis. & Control*, pages 47–75.

Sims, C. A. (2003). Implications of Rational Inattention. *Journal of Monetary Economics*, 50(3):665–690.

Tutino, A. (2013). Rationally Inattentive Consumption Choices. *Review of Economic Dynamics*, 16(3):421–439.

Woodford, M. (2009). Information-Constrained State-Dependent Pricing. *Journal of Monetary Economics*, 56(S):S100–S124.

A Appendix

A.1 Proof of Lemma 1

We have

$$\beta^T q_T s_T^2 \quad = \quad p_0 s_0^2 + \sum_{t=0}^{T-1} \left(\beta^{t+1} p_{t+1} s_{t+1}^2 - \beta^t p_t s_t^2 \right) \tag{29}$$

since $p_T = q_T$. Moreover,

$$p_{t+1} s_{t+1}^2 = ((1+r)s_t - u_t + \zeta_{t+1})^2 p_{t+1} \tag{30}$$

where $u_t \equiv c_t - \bar{c}$. Further noticing from the definition of L_t and the Riccati equation for p_t presented on page 8 that $p_{t+1}(1+r)s_t u_t = \frac{1+\beta p_{t+1}}{\beta} L_t s_t u_t$ and $p_{t+1} u_t^2 = \frac{1+\beta p_{t+1}}{\beta} u_t^2 - \frac{1}{\beta} u_t^2$, it holds

$$\begin{aligned} E[\beta^{t+1} p_{t+1} s_{t+1}^2 | \mathcal{I}_0] \quad = \quad & E\Big[\beta^t (u_t - L_t s_t)^2 (1 + \beta p_{t+1}) + \beta^{t+1} p_{t+1} \zeta_{t+1}^2 \\ & + \beta^{t+1} (1+r)^2 p_{t+1} s_t^2 - \beta^t (1 + \beta p_{t+1}) L_t^2 s_t^2 - \beta^t u_t^2 \Big| \mathcal{I}_0 \Big] \end{aligned} \tag{31}$$

because ζ_{t+1} is independent with respect to u_t and s_t. Moreover,

$$\beta^t p_t s_t^2 \quad = \quad \beta^t (1+r) L_t s_t^2 \tag{32}$$

so that equation (29) writes in expectation

$$E[\beta^T q_T s_T^2|\mathcal{I}_0] \;=\; E\Big[p_0 s_0^2 + \sum_{t=0}^{T-1} \beta^t (u_t - L_t s_t)^2 (1+\beta p_{t+1}) + \beta^{t+1} p_{t+1} \zeta_{t+1}^2 - \beta^t u_t^2 \Big| \mathcal{I}_0\Big] \quad (33)$$

Consequently, the objective function $V_0 \equiv E\Big[\sum_{t=0}^{T-1} \beta^t \Big(u_t^2 + \lambda \tau_t\Big) + \beta^T q_T s_T^2 \Big| \mathcal{I}_0\Big]$ is

$$V_0 \;=\; E\Big[p_0 s_0^2 + \sum_{t=0}^{T-1} \beta^t \lambda \tau_t + \beta^{t+1} p_{t+1} \zeta_{t+1}^2 + \beta^t (u_t - L_t s_t)^2 (1+\beta p_{t+1}) \Big| \mathcal{I}_0\Big] \quad (34)$$

Therefore, guessing that the triggering choices τ_t are independent of the control law – a guess that will hold – it is optimal to set $u_t = L_t E[s_t|\overline{\mathcal{I}}_t]$ (Q.E.D. Lemma 1).

A.2 Proof of Lemma 2

Using this result, the last term in equation (34) writes $\beta^t L_t^2 (E[s_t|\overline{\mathcal{I}}_t] - s_t)^2 (1+\beta p_{t+1})$ where $s_t - E[s_t|\overline{\mathcal{I}}_t] = s_t - E[s_t|\mathcal{I}_t] + e_t$ and $e_t \equiv E[s_t|\mathcal{I}_t] - E[s_t|\overline{\mathcal{I}}_t]$. Thus,

$$E[(s_t - E[s_t|\overline{\mathcal{I}}_t])^2|\mathcal{I}_0] \;=\; E[(s_t - E[s_t|\mathcal{I}_t])^2|\mathcal{I}_0] + E[e_t^2|\mathcal{I}_0] \quad (35)$$

as the estimation error from $E[s_t|\mathcal{I}_t]$ is independent from e_t. Hence,

$$V_0 \;=\; E\Big[p_0 s_0^2 + \sum_{t=0}^{T-1} \beta^t \lambda \tau_t + \beta^{t+1} p_{t+1} \zeta_{t+1}^2 + \beta^t L_t^2 \Big((s_t - E[s_t|\mathcal{I}_t])^2 + e_t^2\Big)(1+\beta p_{t+1}) \Big| \mathcal{I}_0\Big] \quad (36)$$

Thanks to the additivity of the above equation, we can now characterize the optimal estimator $E[s_t|\mathcal{I}_t]$. Make the educated guess that it is a Kalman filter which admits a steady state variance. Then, the steady state posterior variance solves the algebraic Riccati equation $p_+ = (1+r)^2 \Big(p_+ - p_+^2/(p_+ + \sigma_\vartheta^2)^{-1}\Big) + \sigma_\zeta^2$. The Kalman gain is $K = p_+(p_+ + \sigma_\eta^2)^{-1}$ and the posterior steady state variance is $p_- = (1-K)p_+$.

We focus on situations such that the initial uncertainty surrounding the state variable is initially at its steady state value. That is, we impose $\sigma_{s_0}^2 = p_-$. Then, $E[(s_t - E[s_t|\mathcal{I}_t])^2|\mathcal{I}_0] = p_-$ is at its steady state. Therefore, the optimal estimator must minimize this steady state variance given the linear law of motion for s_t. By definition, this estimator is the Kalman filter, thus confirming our guess when it admits a steady state variance (Q.E.D. Lemma 2).

A.3 Proof of Lemma 3

The corrective term in equation (6) is a predetermined bias that depends on the information in the hands of the consumer when inattentive ($\overline{\mathcal{I}}_t, \tau_t = 0$). It therefore depends on time t, the inattention length and the triggering law $g_t(.)$ (see Lemma 4 in Molin and Hirche (2010)). Let $l_t \equiv \sup\{k : \tau_k = 1, k \leq t\}$ be the most recent period when the consumer was attentive. Equation (6) then writes

$$E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0] \;\; = \;\; E[s_t|\overline{\mathcal{I}}_{t-1}] + \alpha(t, l_t) \tag{37}$$

where $\alpha(t, l_t) \equiv E\Big[(1+r)e_{t-1} + K(z_t - E[s_t|\mathcal{I}_{t-1}])\Big|\overline{\mathcal{I}}_t, \tau_t = 0\Big]$. Using the definition for e_t in (7) we have

$$e_{t+1} = (1-\tau_t)(1+r)e_t - \alpha(t, l_t) + K\Big(z_{t+1} - (1+r)E[s_t|\mathcal{I}_t] + c_t - \bar{c}\Big) \tag{38}$$

Moreover, note that only the second and fourth terms in (36) depend on the triggering law $g_t(.)$. Hence, the triggering law solves the following problem

$$\begin{aligned} \min_{g_.(.),\alpha(.)} \quad & E\Big[\sum_{t=0}^{T-1} \beta^t \lambda \tau_t + \beta^t L_t^2 e_t^2 (1 + \beta p_{t+1})\Big|\mathcal{I}_0\Big] \\ \text{s.t.} \quad & e_{t+1} = (1-\tau_t)(1+r)e_t - \alpha(t, l_t) + \omega_{t+1} \end{aligned} \tag{39}$$

where $\omega_{t+1} \equiv K\Big(z_{t+1} - (1+r)E[s_t|\mathcal{I}_t] + c_t - \bar{c}\Big)$ is the innovation from the latent Kalman filter and is an i.i.d Gaussian white noise with variance $\sigma_\omega^2 = k^2(\bar{p}_+ + \sigma_\eta^2)$. The difficulty in solving problem (39) is that $\alpha(.)$ depends on $g_.(.)$ and vice versa.

Molin and Hirche (2012) develop an iterative algorithm to solve a similar problem when $\beta = 1$. They show that when the distributions of e_0 and $\{\omega_t\}$ are symmetric and unimodal, $\alpha(.) = 0$ is a globally asymptotically stable fixed-point of the algorithm (Theorem 5, Molin and Hirche (2012)). The Online Appendix A.8 shows that this theorem still holds for (39). (Q.E.D. Lemma 3)

A.4 Optimal Triggering Rule

Note that only the second and fourth terms in (36) depend on the triggering law $g(.)$. Hence, the updating behavior solves the following problem

$$\min_{\{\tau_t\}_{t=0}^{T-1}\in\{0,1\}^T} E\Big[\sum_{t=0}^{T-1}\beta^t\lambda\tau_t+\beta^t L_t^2 e_t^2(1+\beta p_{t+1})\Big|\mathcal{I}_0\Big] \tag{40}$$

along with a transitory dynamics for e_{t+1}. Following from Lemma 3, the law of motion for e_{t+1} is at the optimum

$$e_{t+1}=(1-\tau_t)(1+r)e_t+K\Big(z_{t+1}-(1+r)E[s_t|\mathcal{I}_t]+c_t-\bar{c}\Big) \tag{41}$$

Problem (40) along with the law of motion (41) is a standard optimal control problem with perfect state observation e_t and could therefore be solved using a DP algorithm.

From problem (40) the choice to update depends on the state variable $e_t\equiv E[s_t|\mathcal{I}_t]-E[s_t|\overline{\mathcal{I}}_t]$. Let the cost associated to the terminal condition q_T be arbitrarily large. Then, at period $T-1$, the consumer updates almost surely as $\lim_{q_T\mapsto\infty} L_{T-1}^2(1+\beta q_T)=\infty$. Therefore, $g_{T-1}(e_{T-1})=1\iff|e_{T-1}|>0$. Let π_t^+ and π_t^- denote the thresholds such that the consumer updates at period t if and only if $e_t\le\pi_t^-$ or $e_t\ge\pi_t^+$. Then, $\pi_{T-1}^+=-\pi_{T-1}^-=0$ and the triggering law is indeed symmetric at period $T-1$. Writing problem (40) in its Bellman form, we have

$$\begin{aligned}J_t(e_t) &= \min_{\tau_t\in\{0,1\}} L_t^2(1+\beta p_{t+1})e_t^2+\tau_t\lambda+\beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t] \\ \text{s.t.} &\quad e_{t+1}=(1-\tau_t)(1+r)e_t+K\Big(z_{t+1}-(1+r)E[s_t|\mathcal{I}_t]+c_t-\bar{c}\Big)\end{aligned} \tag{42}$$

We thus have $E[J_{T-1}(e_{T-1})|\mathcal{I}_{T-2}]=\lambda$ and the consumer updates if and only if

$$|e_{T-2}|\ge\frac{1}{L_{T-2}}\sqrt{\frac{\lambda}{1+\beta p_{T-1}}} \tag{43}$$

Again, the thresholds are symmetric $-\pi_{T-2}^-=\pi_{T-2}^+$. This symmetry arises because $J_{T-1}(e_{T-1})=J_{T-1}(-e_{T-1})$ is symmetric, the expectation is taken over an unimodal and symmetric distribution so that $E[J_{T-1}(e_{T-1})|\mathcal{I}_{T-2},e_t]=E[J_{T-1}(e_{T-1})|\mathcal{I}_{T-2},-e_t]$ and $L_t^2(1+\beta p_{t+1})e_t^2$ is symmetric as well. Therefore, iterating backward and using the same argument, $J_t(e_t)=J_t(-e_t)\ \forall t\in\{0,\dots,T-1\}$ so that $-\pi_t^-=\pi_t^+\ \forall t\in\{0,\dots,T-1\}$ thus confirming the guess that the optimal triggering law $g_t(e_t)$ is symmetric. Moreover, at any time $t\in\{0,\dots,T-1\}$, $\pi_t\in\mathbb{R}^+$ solves

$$\lambda+\beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t,e_t=0] = L_t^2(1+\beta p_{t+1})\pi_t^2+\beta E[J_{t+1}(e_{t+1})|\mathcal{I}_t,e_t=\pi_t] \tag{44}$$

(Q.E.D. Lemma 4 and Proposition 1)

A.5 Stationary Policies

I now demonstrate that the triggering law converges to a stationary policy when T tends to ∞. The first step is to show that p_t converges to a stationary solution. To this end, consider the following infinite horizon deterministic linear quadratic control problem:

$$\begin{aligned} \min_{\{c_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t (c_t - \bar{c})^2 \\ \text{s.t.} \quad x_{t+1} &= (1+r)x_t - c_t + \bar{c} \end{aligned} \tag{45}$$

where $\beta \in (0,1)$. Assuming it exists,[20] it is well-known that the stationary control law is $\bar{L} = (1+r)\frac{\beta\bar{p}}{1+\beta\bar{p}}$ where $\bar{p}$ is the solution to algebraic Ricatti equation

$$\bar{p} = (1+r)^2 \frac{\beta\bar{p}}{1+\beta\bar{p}} \tag{46}$$

that is, $\bar{p} = \frac{\beta(1+r)^2-1}{\beta}$.[21] Thanks to the certainty equivalence of the orignal problem, the consumption policy (4) admits a stationary solution $L = (1+r)\frac{\beta\bar{p}}{1+\beta\bar{p}} = \frac{\beta(1+r)^2-1}{\beta(1+r)}$. As a result, $L_t^2(1+\beta p_{t+1})$ converges to $\frac{[\beta(1+r)^2-1]^2}{\beta(1+r)}$ so that the reward function in the Bellman equation (9) is stationary. The latent Kalman filter being at its steady state, $z_{t+1} - E[s_{t+1}|\mathcal{I}_t]$ follows a gaussian distribution with mean zero and variance $p_+ + \sigma_\eta^2$ which is time-invariante. Hence, problem (9) is an infinite horizon discrete time Markov decision problem where the reward, transition, constraint and shock distribution are independent of time. As such, the problem is stationary and the Bellman equation takes the form of a functional fixed-point equation

$$\begin{aligned} J(e_t) &= \min_{\tau_t \in \{0,1\}} \frac{\Big(\beta(1+r)^2-1\Big)^2}{\beta(1+r)} e_t^2 + \tau_t\lambda + \beta E[J(e_{t+1})|\mathcal{I}_t] \\ \text{s.t.} \quad & e_{t+1} = (1-\tau_t)e_t + k\Big(z_{t+1} - (1+r)E[s_{t+1}|\mathcal{I}_t] + c_t - \bar{c}\Big) \end{aligned} \tag{47}$$

We thus have $\tau_t = g(e_t)$ where $g(e_t) = 1 \iff |e_t| \geq \pi$ and 0 otherwise. π solves

$$\lambda + \beta E[J(e_{t+1})|\mathcal{I}_t, e_t = 0] = \frac{\Big(\beta(1+r)^2-1\Big)^2}{\beta(1+r)}\pi^2 + \beta E[J(e_{t+1})|\mathcal{I}_t, e_t = \pi] \tag{48}$$

[20]See Ljungqvist and Sargent (2004) section 5.4.1 for a discussion on stability. Section 5.2.2 characterizes the solution to the problem under consideration.

[21]These results also follow directly from Lemma 1 since the certainty equivalence holds.

and $J(.)$ is the stationary value function from (47).

A.6 Cross-sectional Stationary Distribution

Take the limit when $T \mapsto \infty$ and let $\lambda_t(k)$ be the share of consumers who did not update for $k \in \{1, 2, \dots, \infty\}$ periods. Then,

$$a_t(e) = \sum_{k=2}^{\infty} \lambda_{t-1}(k-1)(1-\Lambda_t(k-1))f_t(e|k) + f_t(e|1)\Big(\sum_{k=1}^{\infty} \lambda_{t-1}(k)\Lambda_t(k)\Big) \tag{49}$$

The stationary distribution being time independent, the $\lambda^\star(k)$ must solve

$$\lambda^\star(1) = \sum_{k=1}^{\infty} \lambda^\star(k)\Lambda^\star(k) \tag{50}$$

$$\lambda^\star(k) = \lambda^\star(k-1)(1-\Lambda^\star(k-1)) \quad \forall k \geq 2 \tag{51}$$

$$1 = \sum_{k=1}^{\infty} \lambda^\star(k) \tag{52}$$

where $\Lambda^\star(k) = 1 - \int_{\Xi} f(e|k)de$ and Ξ is the time invariant non-updating set following from corollary 1 and $f(e|k)$ the corresponding time invariant[22] distribution obtained from iterating on equation (15). Iterating backward, equation (51) writes

$$\lambda^\star(k) = \lambda^\star(1) \prod_{i=1}^{k-1} (1-\Lambda^\star(i)) \quad \forall k \geq 2 \tag{53}$$

Noting that $S^\star(k-1) = \prod_{i=1}^{k-1}(1-\Lambda^\star(i))$ where $S^\star(k)$ is the time invariante survival function and $S^\star(0) = 1$, we may introduce this expression in (52) to get

$$\lambda^\star(1) = \frac{1}{\sum_{k=1}^{\infty} S^\star(k-1)} \tag{54}$$

Further noticing that equation (50) holds independently of $\lambda^\star(1)$ because $q_k^\star \equiv S(k-1)\Lambda^\star(k)$ and $\sum_{k=1}^{\infty} q_k^\star = 1$, equations (54), (53) and (15) fully characterize the stationary cross-sectional distribution

$$a^\star(e) = \sum_{k=1}^{\infty} \lambda^\star(k) f^\star(e|k) \tag{55}$$

[22] It is straightforward to see that the latter distribution is time invariant when the non-updating set is time invariante.

As a weighted sum of unimodal and symmetric distributions centered around zero, the stationary cross-sectional distribution of consumers is itself symmetric, unimodal and centered around zero. In the simulations, the (evenly discretized) stationary distribution is first computed using lemma 5 as a first guess and then iterated on a few times following (20) to achieve convergence.

A.7 Welfare Decomposition

Taking the infinite horizon version of equation (36), the value function at period 0 writes

$$V_0 = E\Big[ps_0^2 + \frac{p}{1-\beta}\sigma_\zeta^2 + \frac{L^2(1+\beta p)}{1-\beta}p_- + \sum_{t=0}^{\infty}\beta^t\Big(L^2(1+\beta p)e_t^2 + \lambda\tau_t\Big)\Big|\mathcal{I}_0\Big] \quad (56)$$

The above expression provides a straightforward decomposition of the welfare costs from imperfect information. The first term stands for the expected value function of a consumer facing a deterministic linear quadratic control problem. Taking the first two terms leads to the value function of a consumer facing a stochastic control problem with perfect state observation – that is the standard permanent income model with full-information rational expectation. The third term measures the welfare cost from the noisy state observation. Finally, the remaining sum stands for the cost of processing information.

Equation (56) is conditional on $\mathcal{I}_0$ and therefore imposes that period 0 is an updating period. To avoid such restriction and consider an initial period that does not rely on the specifics of the consumer behavior, I instead compute the expected value function unconditionally on the updating behavior at period 0. Let $E[V_0(e_0)]$ be this expected value function. Further, assume that the consumer has initially already lived for a long time. Accordingly, the pdf associated to e_0 is given by the cross-sectional stationary distribution $a^\star(.)$ from Proposition 5. Now, realize that $e_t = e_0$ if $e_0 \in \Xi$ and zero otherwise. Consequently, the relevant distribution for e_t is the transformation of $a^\star(.)$ which accounts for the resetting at zero when e is outside the boundaries. Given that $\int ea^\star(e)de = 0$ and denoting $\sigma_a^2 = \int_\Xi e^2 a^\star(e)de$, we find that $E_{a^\star(.)}[e_t^2] = \sigma_a^2$ is time invariante. Furthermore, let $\bar{\lambda}^\star \equiv \int_{\notin\Xi} a^\star(e)de$ be the share of updates at the stationary distribution. Then, $\lambda\sum_{t=1}^{\infty}\beta^t E_{a^\star(.)}[\tau_t] = \frac{\lambda\bar{\lambda}^\star}{1-\beta}$. Therefore,

$$E_{a^\star(.)}[V_0(e_0)] = p(\bar{s}_0^2 + p_-) + \frac{p}{1-\beta}\sigma_\zeta^2 + \frac{L^2(1+\beta p)}{1-\beta}(p_- + \sigma_a^2) + \frac{\lambda\bar{\lambda}^\star}{1-\beta} \quad (57)$$

Following Cochrane (1989), I use a money metric to measure the welfare cost of deviating from the full information rational expectation solution. Dividing the expected welfare loss marginal utility of consumption and converting it to quarterly rates, we get a welfare cost converted in dollars per period:

$$\mathrm{WC} \;=\; \frac{r(1-\beta)^{-1}}{2(\bar{c}-\mu)(1+r)}\left[\frac{[\beta(1+r)^2-1]^2}{\beta(1+r)}(p_- + \sigma_a^2) + \lambda\bar{\lambda}^\star\right] \tag{58}$$

A.8 On the optimality of $\alpha(.) = 0$ (Not for publication)

In this appendix, I show that Theorem 5 in Molin and Hirche (2012) applies to the problem considered in Appendix A.3. The theorem is meant to derive the optimal design of event-triggered estimation for first-order linear stochastic systems with an identical information structure. The general requirements for the theorem are that the distributions of the initial state e_0 and $\{w_t\}$ are symmetric and unimodal. This is the case in our setup since these distributions are Gaussian. The difference from their problem is with regard to the objective function. They consider the sum of square errors, whereas we are interested in a weighted and discounted sum of these errors here.

In the following, I recast the problem in Appendix A.3 using the notation used in their proof.[23] Let

$$\hat{e}_t \equiv E[s_t|\mathcal{I}_t] - E[s_t|\overline{\mathcal{I}}_t, \tau_t = 0] + \alpha(t, l_t) \tag{59}$$

Accordingly, problem (39) can be written as

$$\begin{aligned} \min_{g.(.),\alpha(.)} \quad & E\Big[\sum_{t=0}^{T-1} \beta^t \Big((1-\tau_t)\Gamma_t(\hat{e}_t - \alpha(t, l_t))^2 + \lambda\tau_t\Big)\Big|\mathcal{I}_0\Big] \\ \text{s.t.} \quad & \hat{e}_{t+1} = (1-\tau_t)(1+r)\hat{e}_t + \omega_{t+1} \end{aligned} \tag{60}$$

where $\Gamma_t \equiv L_t(1 + \beta p_{t+1})$. Moreover, let

$$\begin{aligned} \hat{y}_t &= \frac{\hat{e}_t}{R^t}, \quad t = 0, \dots, N-1 \\ \varrho_{t,l_t} &= \frac{\alpha(t, l_t)}{R^t}, \quad t = 0, \dots, N-1, l_t = 0, \dots, t \end{aligned}$$

where $R \equiv (1 + r)$. Given this transformation, the running cost is

$$\hat{c}_t^{\varrho_t}(\hat{y}_t, l_t, \tau_t) = \beta^t \Big((1-\tau_t)R^{2t}\Gamma_t(\hat{y}_t - \varrho_{t,l_t})^2 + \lambda\tau_t\Big) \tag{61}$$

The optimization problem for Molin and Hirche (2012) iterative procedure is thus given by

$$\min_{\hat{g},\varrho} \hat{J} \tag{62}$$

[23]The original proof of Theorem 5 in Molin and Hirche (2012) can be found here: https://arxiv.org/pdf/1203.4980.pdf .

with

$$\hat{J}(\hat{g},\varrho) = E_{\hat{g}}\left[\sum_{t=0}^{N-1} \hat{c}_t^{\varrho_t}(\hat{y}_t, l_t, \tau_t)\right] \tag{63}$$

where the subscript $\hat{g}$ emphases that the expectation is taken with respect to the triggering policy. The proof in Molin and Hirche (2012) requires that, for a fixed vector ϱ^i of all combinaisons ϱ_{t,l_t}, the following symmetry and monotonicity properties hold for the running cost:

$$\begin{aligned} \hat{c}_t^{\varrho_t^i}(\varrho_{t,l_t}^i + \Delta, l_t, \tau) &= \hat{c}_t^{\varrho_t^i}(\varrho_{t,l_t}^i - \Delta, l_t, \tau) \qquad (64) \\ & \quad \forall \Delta \in \mathbb{R}, l_t \in \{0,\ldots,t-1\}, \tau \in \{0,1\} \end{aligned}$$

and

$$\begin{aligned} 0 \leq \Delta_1 \leq \Delta_2 \Longrightarrow \quad \hat{c}_t^{\varrho_t^i}(\varrho_{t,l_t}^i + \Delta_1, l_t, \tau) &\leq \hat{c}_t^{\varrho_t^i}(\varrho_{t,l_t}^i + \Delta_2, l_t, \tau) \qquad (65) \\ & \quad \forall l_t \in \{0,\ldots,t-1\}, \tau \in \{0,1\} \end{aligned}$$

It is straightforward to see that these properties hold here given (61). As a result, the subsequent results in the proof in Molin and Hirche (2012) are valid and their Theorem 5 applies.

A.9 Approximating the distribution $f_t(e|k, e_{t-k})$ (Not for publication)

To avoid confusion in the following, let the realization $e_{t-k} = a$ and recall that $\sigma_\omega^2 = k^2(\bar{p}_+ + \sigma_\eta^2)$ where k and p_+ are respectively the gain and the one period ahead error variance at the Kalman filter steady state. Define $f_t(e|0,a) = \delta(e-a)$. Then from iterating on (15), we have for k=1:

$$f_t(e|1,a) = \frac{1}{\sigma_\omega}\phi\left(\frac{e-(1+r)a}{\sigma_\omega}\right) \tag{66}$$

When $k=2$,

$$\begin{aligned} f_t(e|2,a) &\propto \int_{\Xi_{t-1}} \frac{1}{\sigma_\omega^2}\phi\left(\frac{e-(1+r)\bar{e}}{\sigma_\omega}\right)\phi\left(\frac{\bar{e}-(1+r)a}{\sigma_\omega}\right)d\bar{e} \qquad (67) \\ &\propto \int_{\Xi_{t-1}} \frac{1}{2\pi\sigma_\omega^2}\exp\left\{-\frac{e^2 - 2(1+r)\bar{e}e + (1+r)^2\bar{e}^2 + \bar{e}^2 - 2(1+r)\bar{e}a + (1+r)^2a^2}{2\sigma_\omega^2}\right\}d\bar{e} \end{aligned}$$

Focusing on the numerator in the exponential and using the shortcut notation $R = 1 + r$

$$\begin{aligned}
& (1+R^2)\Big[\bar{e}^2 - \frac{2R}{1+R^2}\bar{e}(e+a) + \frac{R^2}{1+R^2}a^2 + \frac{1}{1+R^2}e^2\Big] \\
= \;& (1+R^2)\Big[\bar{e}^2 - \frac{2R}{1+R^2}\bar{e}(e+a) + \Big(\frac{R}{1+R^2}\Big)^2(e+a)^2 + \frac{R^2}{1+R^2}a^2 + \frac{1}{1+R^2}e^2 - \Big(\frac{R}{1+R^2}\Big)^2(e+a)^2\Big] \\
= \;& (1+R^2)\Big(\bar{e} - \frac{R}{1+R^2}(e+a)\Big)^2 + (1+R^2)\Big[\frac{R^2}{1+R^2}a^2 + \frac{1}{1+R^2}e^2 - \Big(\frac{R}{1+R^2}\Big)^2(e+a)^2\Big]
\end{aligned}$$

Where

$$\begin{aligned}
& \frac{R^2}{1+R^2}a^2 + \frac{1}{1+R^2}e^2 - \Big(\frac{R}{1+R^2}\Big)^2(e+a)^2 \\
= \;& \frac{R^2}{1+R^2}a^2 + \frac{1}{1+R^2}e^2 - \Big(\frac{R}{1+R^2}\Big)^2(e^2+a^2+2ea) \\
= \;& \frac{1}{(1+R^2)^2}e^2 - 2\Big(\frac{R}{1+R^2}\Big)^2 ea + \Big(\frac{R^2}{1+R^2}\Big)^2 a^2 + \Big[\frac{R^2}{1+R^2} - \Big(\frac{R}{1+R^2}\Big)^2 - \Big(\frac{R^2}{1+R^2}\Big)^2\Big]a^2 \\
= \;& \Big(\frac{1}{1+R^2}e - \frac{R^2}{1+R^2}a\Big)^2
\end{aligned}$$

Therefore, (67) writes

$$\begin{aligned}
f_t(e|2,a) \;&\propto\; \int_{\Xi_{t-1}} \frac{1}{2\pi\sigma_\omega^2}\exp\Big\{-\frac{(1+R^2)}{2\sigma_\omega^2}\Big[\Big(\bar{e} - \frac{R}{1+R^2}(e+a)\Big)^2 + \Big(\frac{1}{1+R^2}e - \frac{R^2}{1+R^2}a\Big)^2\Big]\Big\}d\bar{e} \\
&\propto\; \int_{\Xi_{t-1}} \frac{\sqrt{1+R^2}}{\sqrt{2\pi}\sigma_\omega}\exp\Big\{-\frac{(\bar{e} - \frac{R}{1+R^2}(e+a))^2}{2\frac{\sigma_\omega^2}{1+R^2}}\Big\}\frac{1}{\sqrt{2\pi}\sqrt{1+R^2}\sigma_\omega}\exp\Big\{-\frac{(e-R^2a)^2}{2(1+R^2)\sigma_\omega^2}\Big\}d\bar{e} \\
&\propto\; \int_{\Xi_{t-1}} \frac{\sqrt{1+R^2}}{\sigma_\omega}\phi\Big(\frac{\bar{e} - \frac{R}{1+R^2}(e+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\Big)\frac{1}{\sqrt{1+R^2}\sigma_\omega}\phi\Big(\frac{e-R^2a}{\sqrt{1+(1+r)^2}\sigma_\omega}\Big)d\bar{e} \\
&\propto\; \frac{1}{\sqrt{1+R^2}\sigma_\omega}\Big[\Phi\Big(\frac{\pi_{t-1} - \frac{R}{1+R^2}(e+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\Big) - \Phi\Big(-\frac{\pi_{t-1} + \frac{R}{1+R^2}(e+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\Big)\Big]\phi\Big(\frac{e-R^2a}{\sqrt{1+R^2}\sigma_\omega}\Big) \quad (68)
\end{aligned}$$

When $k = 3$,

$$\begin{aligned}
f_t(e|3,a) \;&\propto\; \int_{\Xi_{t-1}} \frac{1}{\sigma_\omega}\phi\Big(\frac{e-R\bar{e}}{\sigma_\omega}\Big)f_{t-1}(\bar{e}|2,a)d\bar{e} \\
&\propto\; \int_{\Xi_{t-1}} \frac{1}{\sqrt{1+R^2}\sigma_\omega^2}\phi\Big(\frac{e-R\bar{e}}{\sigma_\omega}\Big)\phi\Big(\frac{\bar{e}-R^2a}{\sqrt{1+R^2}\sigma_\omega}\Big)\Big[\Phi\Big(\frac{\pi_{t-2} - \frac{R}{1+R^2}(\bar{e}+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\Big) - \Phi\Big(-\frac{\pi_{t-2} + \frac{R}{1+R^2}(\bar{e}+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\Big)\Big]d\bar{e}
\end{aligned}$$

Again, I develop and reduce the product of the two gaussian pdfs. To do so, I first focus on the numerator within the exponential.

$$
\begin{aligned}
& \frac{1+R^2}{1+R^2+R^4}\left[(e-R\bar{e})^2+\left(\frac{\bar{e}-R^2a}{\sqrt{1+R^2}}\right)^2\right] \\
= \; & \bar{e}^2-2\frac{(1+R^2)R}{1+R^2+R^4}\bar{e}e+\frac{1+R^2}{1+R^2+R^4}e^2-2\frac{R^2}{1+R^2+R^4}a\bar{e}+\frac{R^4}{1+R^2+R^4}a^2 \\
= \; & \left(\bar{e}-\frac{(1+R^2)Re+R^2a}{1+R^2+R^4}\right)^2+\frac{1+R^2}{1+R^2+R^4}e^2+\frac{R^4}{1+R^2+R^4}a^2-\left(\frac{(1+R^2)Re+R^2a}{1+R^2+R^4}\right)^2
\end{aligned}
$$

Where

$$
\begin{aligned}
& \frac{1+R^2}{1+R^2+R^4}e^2-\left(\frac{(1+R^2)Re+R^2a}{1+R^2+R^4}\right)^2 \\
= \; & \frac{(1+R^2)(1+R^2+R^4)-(1+R)^2R^2}{(1+R^2+R^4)^2}e^2-2\frac{(1+R^2)R^3}{(1+R^2+R^4)^2}ea-\left(\frac{R^2}{1+R^2+R^4}\right)^2a^2 \\
= \; & \left(\frac{\sqrt{1+R^2}}{1+R^2+R^4}\right)^2e^2-2\frac{(1+R^2)R^3}{(1+R^2+R^4)^2}ea+\left(\frac{\sqrt{1+R^2}R^3}{1+R^2+R^4}\right)^2a^2-\left(\frac{R^2}{1+R^2+R^4}\right)^2a^2-\left(\frac{\sqrt{1+R^2}R^3}{1+R^2+R^4}\right)^2a^2 \\
= \; & \frac{1+R^2}{(1+R^2+R^4)^2}(e^2-R^3a)^2-\frac{R^4(1+R^2+R^4)}{(1+R^2+R^4)^2}a^2
\end{aligned}
$$

so that

$$
(e-R\bar{e})^2+\left(\frac{\bar{e}-R^2a}{\sqrt{1+R^2}}\right)^2 = \frac{1+R^2+R^4}{1+R^2}\left(\bar{e}-\frac{(1+R^2)Re+R^2a}{1+R^2+R^4}\right)^2+\frac{(e^2-R^3a)^2}{1+R^2+R^4}
$$

Introducing back this expression in (69), I obtain

$$
\begin{aligned}
f_t(e|3,a) \; \propto \; & \frac{1}{\sqrt{1+R^2}\sigma_\omega^2}\phi\left(\frac{e-R^3a}{\sqrt{1+R^2+R^4}\sigma_\omega}\right)\times \\
& \int_{\Xi_{t-1}}\phi\left(\frac{\bar{e}-\frac{R(1+R^2)e+R^2a}{1+R^2+R^4}}{\sqrt{\frac{1+R^2}{1+R^2+R^4}}\sigma_\omega}\right)\left[\Phi\left(\frac{\pi_{t-2}-\frac{R}{1+R^2}(\bar{e}+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\right)-\Phi\left(-\frac{\pi_{t-2}+\frac{R}{1+R^2}(\bar{e}+a)}{\frac{\sigma_\omega}{\sqrt{1+R^2}}}\right)\right]d\bar{e}
\end{aligned}
\tag{69}
$$

The above expression may not be expressed in other terms to simplify the computation for $k=4$. As a consequence, the computational cost from conditioning on past histories grows exponentially and will likely generate large approximation error when k increases. Therefore, I approximate the above expression by not accounting for the impact of histories before $t-1$. The approximated distribution is thus

$$
\begin{aligned}
f_t^{\text{app}}(e|3) \; \propto \; & \frac{1}{\sqrt{1+R^2}\sigma_\omega^2}\phi\left(\frac{e-R^3a}{\sqrt{1+R^2+R^4}\sigma_\omega}\right)\int_{\Xi_{t-1}}\phi\left(\frac{\bar{e}-\frac{R(1+R^2)e+R^2a}{1+R^2+R^4}}{\sqrt{\frac{1+R^2}{1+R^2+R^4}}\sigma_\omega}\right)d\bar{e} \\
\propto \; & \frac{1}{\sqrt{1+R^2+R^4}\sigma_\omega}\phi\left(\frac{e-R^3a}{\sqrt{1+R^2+R^4}\sigma_\omega}\right)\left[\Phi\left(\frac{\pi_{t-1}-\frac{R(1+R^2)e+R^2a}{1+R^2+R^4}}{\sqrt{\frac{1+R^2}{1+R^2+R^4}}\sigma_\omega}\right)-\Phi\left(-\frac{\pi_{t-1}+\frac{R(1+R^2)e+R^2a}{1+R^2+R^4}}{\sqrt{\frac{1+R^2}{1+R^2+R^4}}\sigma_\omega}\right)\right]
\end{aligned}
$$

For $k = 4$,

$$
\begin{aligned}
f_t^{\text{app}}(e|4) &\propto \int_{\Xi_{t-1}} \frac{1}{\sqrt{1+R^2+R^4\sigma_\omega^2}}\phi\Big(\frac{e-R\bar{e}}{\sigma_\omega}\Big)\phi\Big(\frac{\bar{e}-R^3a}{\sqrt{1+R^2+R^4}\sigma_\omega}\Big)d\bar{e} \\
&\propto \frac{1}{\sqrt{1+R^2+R^4\sigma_\omega^2}}\phi\Big(\frac{e-R^4a}{\sqrt{1+R^2+R^4+R^6}\sigma_\omega}\Big)\int_{\Xi_{t-1}}\phi\Bigg(\frac{\bar{e}-\frac{R(1+R^2+R^4)e+R^3a}{1+R^2+R^4+R^6}}{\sqrt{\frac{1+R^2+R^4}{1+R^2+R^4+R^6}}\sigma_\omega}\Bigg)d\bar{e} \\
&\propto \frac{1}{\sqrt{1+R^2+R^4+R^6}\sigma_\omega}\phi\Big(\frac{e-R^4a}{\sqrt{1+R^2+R^4+R^6}\sigma_\omega}\Big) \\
&\qquad \Bigg[\Phi\Bigg(\frac{\pi_{t-1}-\frac{R(1+R^2+R^4)e+R^3a}{1+R^2+R^4+R^6}}{\sqrt{\frac{1+R^2+R^4}{1+R^2+R^4+R^6}}\sigma_\omega}\Bigg)-\Phi\Bigg(-\frac{\pi_{t-1}+\frac{R(1+R^2+R^4)e+R^3a}{1+R^2+R^4+R^6}}{\sqrt{\frac{1+R^2+R^4}{1+R^2+R^4+R^6}}\sigma_\omega}\Bigg)\Bigg]
\end{aligned}
$$

Using forward iteration, it holds

$$
f_t^{\text{app}}(e|k) \propto \frac{1}{\sqrt{z(k)}\sigma_\omega}\phi\Big(\frac{e-R^ka}{\sqrt{z(k)}\sigma_\omega}\Big)\Bigg[\Phi\Bigg(\frac{\pi_{t-1}-\frac{Ru(k)e+R^{k-1}a}{z(k)}}{\sqrt{\frac{u(k)}{z(k)}}\sigma_\omega}\Bigg)-\Phi\Bigg(-\frac{\pi_{t-1}+\frac{Ru(k)e+R^{k-1}a}{z(k)}}{\sqrt{\frac{u(k)}{z(k)}}\sigma_\omega}\Bigg)\Bigg] \quad \forall k\in\{3,\dots,T-t\} \quad (7
$$

where $z(k) = \sum_{i=0}^{k-1}(1+r)^{2i}$ and $u(k) = \sum_{i=0}^{k-2}(1+r)^{2i}$.
The Online Appendix discusses the implications to the proposed approximation procedure. It shows that the main drawback of this procedure is to potentially over estimate the hazard rates for large k by an order of magnitude of about one to two percentage points. The impact on the survival function and distribution of updates is however negligible as the proportion of agents who encounters a large k is small.